Anais da

XXX Semana Acadêmica da Matemática 2016

unioeste
Universidade Estadual do Oeste do Paraná

Anais da

XXX Semana Acadêmica da Matemática

19/09/2016 a 23/09/2016

unioeste
Universidade Estadual do Oeste do Paraná

Realização
Curso de Matemática

Apoio
CCET
Centro de Ciências Exatas e Tecnológicas

Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UNIOESTE

Semana Acadêmica da Matemática (30.: 2016: Cascavel - PR)
S471a Anais da XXX Semana Acadêmica da Matemática. / Organização de Francieli Cristina Agostineto Antunes... [et al.]. -- Cascavel: Unioeste, 2016.
Online

ISSN: 2526-0804
Disponível em: http://midas.unioeste.br/sgev/eventos/sam/anais
Evento realizado no Campus de Cascavel, no período de 19 a 23 de setembro de 2016.

1. Matemática – Estudo e ensino. 2. Ensino superior - Congressos. I. Universidade Estadual do Oeste do Paraná. II. Antunes, Francieli Cristina Agostineto, (org.). III. Título.

CDD 20. ed.– 510.63

Sandra Regina Mendonça CRB 9/1090

Apresentação

A Semana Acadêmica da Matemática está na sua XXX edição. Este é o evento de extensão mais tradicional promovido pelo Curso de Matemática, da UNIOESTE campus de Cascavel. É um evento com periodicidade anual.

Na programação da XXX Semana Acadêmica de Matemática figuram palestras, minicursos e comunicações orais. As comunicações orais resultam da inscrição dos participantes na modalidade de apresentadores de trabalhos.

Nesta edição da Semana Acadêmica de Matemática, 26 trabalhos foram inscritos. São em geral trabalhos resultados das pesquisas de Iniciação Científica e de Monografia desenvolvidos por alunos do curso de Matemática. Registramos também alguns trabalhos realizados por professores do Curso de Matemática da UNIOESTE - Cascavel, e de alunos de outros cursos que desenvolveram suas pesquisas com teor matemático. A apresentação destes trabalhos no evento tem o objetivo de compartilhar com os colegas os conhecimentos adquiridos pelos alunos e professores nos seus respectivos projetos. O registro destes trabalhos nestes anais servirá para que os futuros alunos possam também fazer uso deste conhecimento.

A comissão organizadora agradece aos autores pelo envio dos trabalhos e também à comissão científica pelas contribuições dadas durante o processo de avaliação e correção dos trabalhos.

A comissão organizadora.

Cascavel, Setembro de 2016.

Índice de trabalhos

Modelo de otimização multiobjetivo em planejamento radioterápico

Andressa Aparecida de Lima
Universidade Estadual do Oeste do Paraná
Discente do Curso de Matemática
andressaaplima4@gmail.com

Simone Aparecida Miloca
Universidade Estadual do Oeste do Paraná
Docente do Curso de Matemática
smiloca@gmail.com

Resumo: Neste trabalho apresentamos conceitos introdutórios de otimização multiobjetivo e um métodos de busca de solução, denominado Método da Soma Ponderada (Método dos Pesos). Apresentamos também uma aplicação envolvendo um problema da área da saúde, relacionado ao planejamento de tratamento radioterápico, mostrando sua formulação matemática e testes numéricos.

Palavras-chave: Otimização; Método dos Pesos; Radioterapia.

1 Introdução

Em diversas áreas do conhecimento, existe a necessidade de se fazer a leitura de um determinado problema utilizando-se de estruturas que permitam auxiliar a busca de sua solução. A literatura apresenta várias técnicas e algoritmos que podem ser destinados a estruturar e solucionar tais problemas, fazendo parte destas, as técnicas de Pesquisa Operacional (GOLDBARG; LUNA, 2005).

Segundo Lopes, Rodrigues e Steiner (2013), a Pesquisa Operacional é uma área voltada ao desenvolvimento de modelos matemáticos e algoritmos para resolução de problemas, como planejamento florestal, avaliação genética de espécies, apoio a decisão para planejamento de empresas, alocação de recursos, planejamento de produção, roteamento de veículos, planejamento hidrelétrico, planejamento de redes elétricas, planejamento de tratamento em áreas da saúde, problemas de transporte, dentre outros. Os métodos desenvolvidos tem sido utilizados com sucesso na obtenção de soluções em problemas de otimização, destacando-se os algoritmos do campo da Programação Matemática (GOLDBARG; LUNA, 2005; TAHA, 2008; ZIONTS, 2006; HILLIER F. S. ; LIEBERMAN, 2005).

A ideia geral em um problema de otimização é construir uma função denominada função objetivo, que será maximizada (ou minimizada) e está sujeita a restrições de igualdade ou

desigualdade. Matematicamente escreve-se:

$$Max \qquad f(x) \tag{1}$$

$$sujeito\ a \qquad h_i(x) = 0,\ i = 1, ..., m \tag{2}$$

$$g_j(x) \leq 0,\ j = 1, ..., n \tag{3}$$

$$x \in \mathbb{R}^n \tag{4}$$

onde, $f : \mathbb{R}^n \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}$, $h : \mathbb{R}^n \to \mathbb{R}$ são funções contínuas, geralmente diferenciáveis (RIBEIRO; KARAS, 2013).

Conforme as características do conjunto de restrições e as propriedades das funções objetivo, teremos os diferentes problemas de otimização. Por exemplo, as funções envolvidas no problema podem ser contínuas ou não, diferenciáveis ou não, lineares ou não. O caso particular em que a função objetivo e as funções que definem o conjunto de restrições são funções lineares é conhecido como um Problema de Programação Linear (PPL) e pode ser resolvido por métodos específicos, como o Método *Simplex* (RIBEIRO; KARAS, 2013).

Por outro lado, um problema de otimização multiobjetivo é caracterizado por um vetor p-dimensional de funcões objetivo (7) e uma região viável X, como em (6). No entanto, ao invés de buscar uma única solução, procura-se um conjunto S de soluções, denominadas soluções não-dominadas (ou Pareto-Ótimas).

$$z\left(x\right) = \left[z_1\left(x\right), z_2\left(x\right), \ldots, z_p\left(x\right)\right] \tag{5}$$

$$x \in X.$$

Aborda-se neste trabalho, problemas da natureza da Programação Linear. Fazemos uma introdução a conceitos de otimização multiobjetivo e um dos métodos de busca de solução denominado Método dos Pesos é descrito. Um modelo de otimização multiobjetivo é apresentado considerando um problema sobre planejamento de tratamento radioterápico, apresentado por Craft(2014).

A técnica de radioterapia é uma alternativa para o tratamento de câncer. Este tipo de tratamento se fundamenta no bloqueio ou destruição da divisão celular das moléculas de DNA que compõe o tumor e consiste em irradiar o tumor de forma a maximizar o efeito de radiação sobre os tecidos afetados, minimizando os impactos nocivos sobre os demais tecidos do organismo. O problema é modelado com de otimização multiobjetivo e tem natureza conflitante na tomada de decisão pois uma vez que se maximiza a intensidade de radiação no tumor pode-se afetar muitas células de tecidos saudáveis, mas minimizar a intensidade de dose em tecidos saudáveis pode não destruir as células de tecido de tumor.

2 Otimização Multiobjetivo

Um problema de otimização mono-objetivo é caracterizado pela procura de uma solução ótima (valor máximo ou mínimo) de uma função, chamada Função Objetivo como em (1), que satisfaça a um conjunto de restrições escrito na forma (2) e (3), apresentados na introdução.

O conjunto X dado em (6), formado pelas restrições do problema, é denominado conjunto viável (região viável ou ainda região factível).

$$X = \{x : x \in \mathbb{R}^n, g_i(x) \leq 0, x_j \geq 0 \text{ para todo i e j}\}. \tag{6}$$

Um problema de otimização procura um elemento x^* na região viável X que é o maior valor para $z(x)$, ou seja, $\max z(x) = z(x^*)$.

Dentre os algoritmos existentes para obtenção de solução, estão os que pertencem ao campo da Programação Matemática. Os algoritmos são desenvolvidos dependendo de características dos modelos. Os da Programação Linear (PL) são algoritmos utilizados em modelos cujas variáveis são contínuas e apresentam comportamentos lineares tanto em relação às restrições como à função objetivo. Os algoritmos de Programação Não Linear (PNL), destinam-se a problemas que exibem qualquer tipo de não linearidade, seja nas restrições ou na função objetivo. Os da Programação Inteira (PI), para modelos que consideram variáveis que não podem assumir valores contínuos, ficando condicionadas a assumir valores discretos. Deixamos as referências Goldbarg e Luna (2005), Taha (2008), Zionts (2006) e Hillier F. S. ; Lieberman (2005) para maiores informações.

Por outro lado, um problema de otimização multiobjetivo é caracterizado por um vetor p-dimensional de funcões objetivo dado em (7) e uma região viável X, como em (6).

$$z(x) = [z_1(x), z_2(x), \ldots, z_p(x)] \tag{7}$$

$$x \in X.$$

No entanto, ao invés de buscar uma única solução, procura-se um conjunto S de soluções, denominadas soluções não-dominadas, que é um subconjunto de X construído como na Definição 1.

Definição 1. Dado um conjunto de soluções viáveis X, o conjunto de soluções não dominadas é denotado por S e definido como

$$S = \{x : x \in X, \text{ onde não exista nenhum outro } x' \in X \text{ tal que}$$
$$z_q(x') > z_q(x) \text{ para algum } q \in \{1, 2, ..., p\} \text{ e } z_k(x') \geq z_k(x)\ \forall k \neq q\}.$$

Uma das diferenças entre as funções mono-objetivo e multiobjetivo é que na otimização multiobjetivo as funções objetivo constituem um espaço multidimensional, que denominaremos espaço objetivo, Z. Para cada solução viável x no espaço de decisão X, existe um ponto no espaço objetivo Z, denotado por $f(x) = z = (z_1, z_2, ..., z_M)^T$. Os algoritmos nesta área tem a tarefa de encontrar as soluções não dominadas.

O conceito de soluções não dominadas aparece na literatura com os nomes Pareto Ótima ou Solução Eficiente. A otimização multiobjetivo é, por vezes, chamada de otimização vetorial, porque um vetor de objetivos é otimizado em vez de um único objetivo.

2.1 Método dos Pesos

O método dos pesos (ou soma ponderada) consiste em combinar as funções objetivos em uma única função objetivo por meio de um vetor de pesos w, e as suas restrições iniciais. Em outras palavras, o problema multiobjetivo é transformado em um problema mono-objetivo. Assim, um método de otimização mono-objetivo é utilizado para gerar soluções.

Matematicamente, a função objetivo segundo o método dos pesos, pode ser apresentada como em (8).

$$\begin{aligned} max \quad & z(x) = \sum_{i=1}^{p} w_i z_i(x) \\ s.a. \;\; a \quad & x \in X. \end{aligned} \tag{8}$$

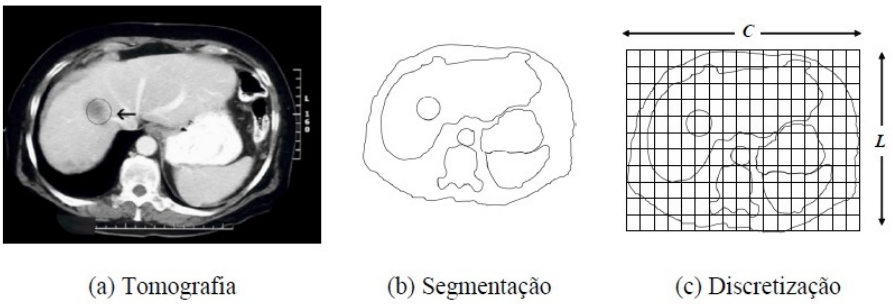
Em problemas com conjuntos convexos, o algoritmo gera diferentes retas suportes definidas pelos pesos w_i. Já em conjuntos não-convexos, nem todos os pontos não dominados as admitiriam, e assim, não geraria todas a soluções não dominadas.

3 Modelo Multiobjetivo de Planejamento Radioterápico

Um importante problema de aplicação envolvendo múltiplos objetivos é relacionado ao tratamento radioterápico, o qual consiste determinar a quantidade da dose de radiação x, cuja unidade de medida é o Gray (Gy), a ser emitida em um tumor de modo que seja a mais próxima da prescrita pelo médico e que atinja minimamente os demais tecidos (Ehrgott; Holder, 2009).

O problema é modelado considerando uma região do corpo humano obtido de um corte de imagem tomográfica, que é segmentada e discretizada. As regiões são representadas por uma rede de pixels, no caso de uma imagem bidimensional, ou uma grade de voxels, quando se tem

uma imagem em 3D, onde cada pixel ou voxel é considerado parte de tecido saudável, nobre ou do tumor, como mostrado na Figura 1.



(a) Tomografia (b) Segmentação (c) Discretização

Figura 1: Segmentação e Discretização (Goldbarg,2009)

Para a contrução do modelo parte-se do fato de que k ângulos (feixes) de radiação são previamente definidos, bem como alguns subângulos (subfeixes). Também assume-se conhecida a matriz de dose (D), construída considerando cortes de imagens tomográficas, contendo informações de atenuação de radiação nos tecidos, que é emitida pelos feixes.

Uma formulação matemática genérica para o problema é apresentado por Craft *et. al.* (2014). O modelo geral considera a matriz de absorção de dose D, cujas linhas contem informação da absorção de dose em cada voxel segundo um conjunto de subfeixes (colunas da matriz), oriundo de todos os feixes considerados no problema. Ou seja, a matriz D é uma concatenação das matrizes individuais de cada feixe. Seja x o vetor das fluências de dose dos j subfeixes e d o vetor de dose nos voxels. O conjunto C representa um conjunto de restrições relacionadas a limites de dose para os diferentes tecidos. O problema escrito na forma matricial é dado em (9).

$$Min \quad f(d) \tag{9}$$

$$s.a. \quad Dx = d \tag{10}$$

$$d \in C$$

$$x \geq 0$$

A função objetivo pode ser escrita de diversas maneiras, uma delas é apresentada por Holder(2003) que considera a minimização de desvios de dose nos tecidos. Exemplos numéricos foram estudados, elaborados e implementados usando os softwares Scilab e Lingo. Os resultados apresentados em Lima (2015).

Um modelo específico considerando uma função objetivo linear, poderia ser escrito escolhendo $f(d)$ como sendo a dose média para uma determinada estrutura (orgãos nobres, saudáveis ou de tumor). Considere como função objetivo, a minimização da dose média que atingirá os

tecidos nobres (d_N) e os tecidos saudáveis (d_S) e maximização da dose média recebida pelo tumor (d_T), o problema multiobjetivo poderia ser escrito, segundo o Método dos Pesos como:

$$Min \quad d_N x \tag{11}$$

$$Min \quad d_S x \tag{12}$$

$$Max \quad d_T x \tag{13}$$

$$s.a. \quad D_T x > d_{it} \tag{14}$$

$$D_N x \leq d_{sn} \tag{15}$$

$$D_S x \leq d_{ss} \tag{16}$$

$$x \geq 0$$

Onde,

(a) d_N, d_S e d_T são vetores que indica a atenuação média de dose nos pixels nobres, saudáveis e de tumor, que são atingidos por cada subfeixe.

(b) x é o vetor que quantifica a emissão de radiação de cada subfeixe j.

(c) $D_{N_{(q\times j)}}$ é a matriz de de atenuação de dose para os tecidos nobres, $D_{T_{(r\times j)}}$ é a matriz de de atenuação de dose para os tecidos de tumor.

(d) d_{it} é o vetor que indica os limites de dose inferior para tecidos de tumor, d_{sn} é o vetor que indica os limites de dose superior para tecidos nobres e d_{ss} é o vetor que indica os limites de dose superior para tecidos saudáveis.

As Restrições (14) indicam que os tecidos de tumor não podem receber dose inferior a d_{it}. As Restrições 15 e 16 indicam os limites superior de dose para tecidos nobres e saudáveis.

Este tipo de problema pode admitir várias soluções pareto-ótimas e por isto, é necessário a análise do decisor que terá a opção de escolher uma solução que seja intermediária.

4 Resultados Numéricos

Os testes numéricos que apresentamos a seguir refere-se a um problema de planejamento de tratamento radioterápico para um caso fictício de câncer descrito por Craft *et.al.* (2014). Em seu artigo, ele chama este caso de TG119.

O modelo implementado é apresentado em (17), já escrito segundo o método dos pesos.

$$Min \quad f(d) = w_1 d_N x + w_2 d_S x - w_3 d_T x \tag{17}$$

$$s.a. \qquad D_T x > d_{it} \tag{18}$$
$$x \geq 0$$

onde $w_i, i = 1, 2, 3$, indicam os pesos atribuídos para cada objetivo. Os valores mínimos para dose em tecidos de tumor foram $d_{it} = 1$. Os dados para este caso constam na Tabela 1.

Tabela 1: Dados de entrada para o caso TG119

Número de feixes	5
Número de subfeixes	418
Número de voxels de tumor	7429
Número de voxels de tecidos nobres	1280
Número de voxels de tecidos saudáveis	599440

Este modelo um total de 418 variáveis e 7429 restrições. A Fronteira de Pareto gerada para diferentes vetores de peso, considerando os valores para as funções objetivo relacionada a tecidos nobres (eixo x) e tecidos de tumor (eixo y) é mostrada na Figura 2. Observa-se o que ocorre em problemas multiobjetivos, um ponto bom para um problema de minimização não o é para o problema de maximização. Daí a importância do decisor no processo de escolha de solução.

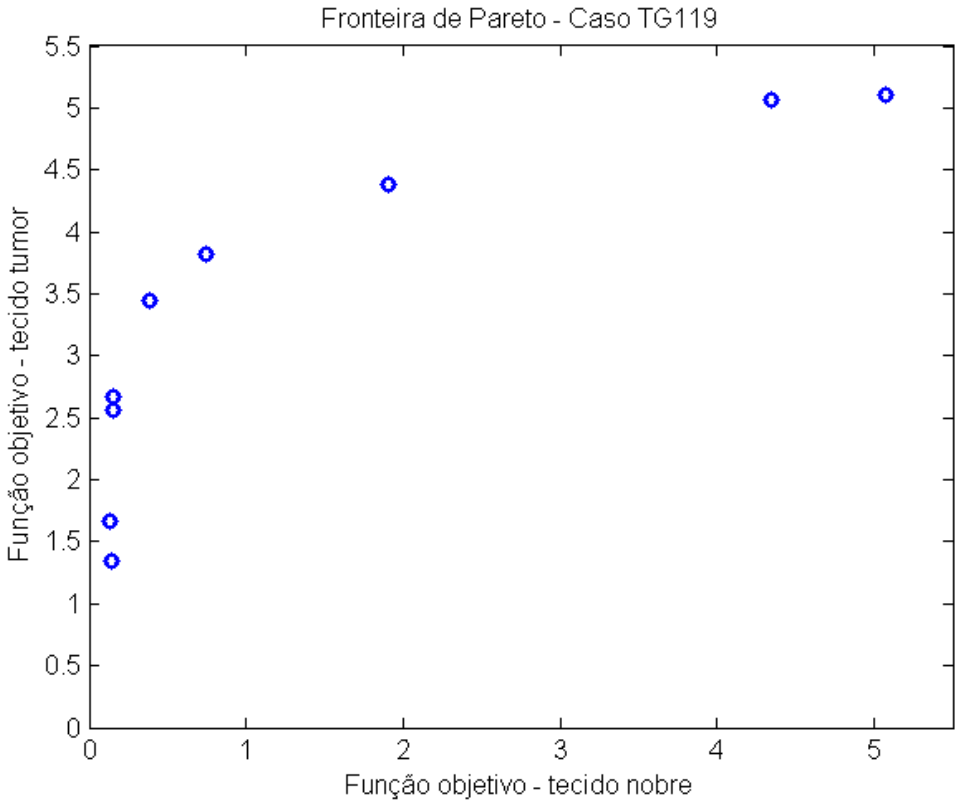


Figura 2: Fronteira de Pareto

A Figura 3 mostra a região tomográfica (CT) para o problema otimizado segundo Teste 1, que considera pesos iguais na função objetivo, e a Figura 4 mostra a região tomográfica (CT) para o Teste 2, que considera peso igual a zero para a função objetivo relacionada a tecidos de tumor e peso igual para os demais tecidos. Observe que quando se dá o mesmo peso para os três objetivos, a região de tumor é atingida e tecidos nobres e saudáveis procuram ser preservados, mas quando se retira a função objetivo (não há prioridade) relacionada a tecidos de tumor, a dose entregue nos voxels de tumor é bem menor.

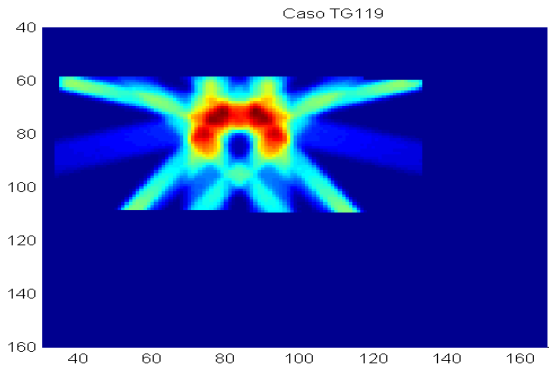


Figura 3: Região CT - Teste 1

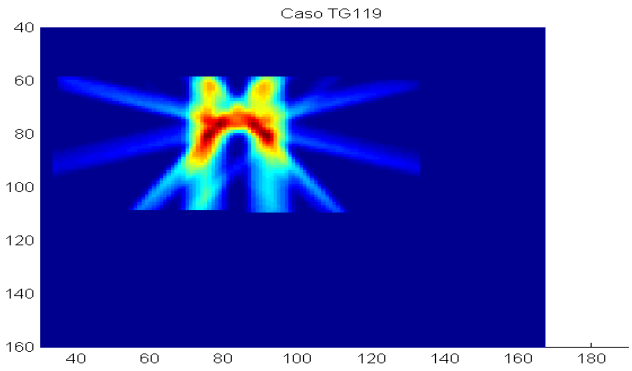


Figura 4: Região CT - Teste 2

5 Conclusões

Devido o grande avanço tecnológico, tornou-se possível obter imagens em 3D do interior do paciente com câncer e isto contribuiu significativamente para o avanço de pesquisas sobre o assunto incluindo a otimização matemática para o planejamento do tratamento do câncer.

O objetivo deste trabalho foi o estudo de otimização multiobjetivo, modelos e exemplos de aplicações em áreas promissoras de pesquisa, como em problemas de tratamento radioterápico.

Os testes computacionais realizados com dados apresentados na literatura (Craft, 2014), é um caso próximo da realidade, pois se trabalhou com dados de imagens tomográficas reais e permitiu observar a importância da teoria estudada. Devido a localização da região tumoral, percebemos a dificuldade de realização de planos de tratamento, principalmente planos otimizados. No caso TG119 observamos que a mudança do vetor de pesos modifica os resultados, deixando para o decisor a escolha da opção a ser utilizada.

Referências

CRAFT, D. *et al.* Shared data for intensity modulated radiation therapy (imrt) optimization research: the cort dataset. Giga Science, p. 12, 2014.

EHRGOTT M. ; HOLDER, A. Operation research methods for optimization in radiation oncology. Pesq. Oper., v. 2, 2009.

GOICOECHEA, A.; HANSEN, R.; DUCKSTEIN, L. Multiobjective Decision Analysis With Engineering And Business Applications. 2nd. ed. New York: John Wiley Sons, 1982.

GOLDBARG, M. C.; LUNA, H. P. L. Otimização Combinatória e Programação Linear. 2ed. ed. Rio de Janeiro: Elsevier, 2005.

HILLIER F. S. ; LIEBERMAN, G. J. Introduction to Operations Research. 8th. ed. New York: McGraw-Hill, 2005.

HOLDER, A. Designing radiotherapy plans with elastic constraints and interior point methods. Healt Care Management Science, n. 6, p. 516, 2003.

LIMA, A. A.; Miloca, S. A. Otimização Multiobjetivo : Conceitos e Aplicações. I EAICTI - Encontro Anual de Iniciação Científica Tecnológica e Inovação. UNIOESTE, 2015.

LOPES, H. S.; RODRIGUES, L.; STEINER, M. T. A. Meta-heurísticas em pesquisa operacional. 1ed. ed. Curitiba: Ominipax, 2013.

RIBEIRO, A. A.; KARAS, E. W. Um curso de otimização. São Paulo: Cengage Learning, 2013

TAHA, H. A. Pesquisa Operacional: uma visão geral. São Paulo: Pearson Prentice Hall, 2008.

ZIONTS, S. Linear and Integer Programming. Rio de Janeiro: Ed Prentice Hall, 2006.

O teorema fundamental do cálculo fracionário

Alexandre Carissimi[1]
Acadêmico da Universidade Estadual do Oeste do Paraná
alexandre.carissimi.96@gmail.com

Sandro Marcos Guzzo
Professor da Universidade Estadual do Oeste do Paraná
smguzzo@gmail.com

Resumo: Neste trabalho estudaremos definições para derivadas fracionárias, segundo Riemann-Liouville e Caputo; a definição de integral fracionária para Riemann-Liouville e as versões fracionárias do Teorema Fundamental do Cálculo, segundo Riemann-Liouville e Caputo. Para chegarmos ao Teorema Fundamental do Cálculo Fracionário, faremos um estudo preliminar sobre a definição de função gamma e do produto de convolução. Também estudaremos a Lei dos Expoentes para derivadas e integrais de ordem fracionária na versão de Riemann-Liouville.

Palavras-chave: Riemann-Liouville; Cálculo fracionário.

1 Função gamma

Definição 1 (Função gamma)**.** A integral imprópria abaixo é chamada de função gamma de x.

$$\Gamma(x) = \int_0^\infty t^{(x-1)} e^{-t} \mathrm{d}t.$$

Essa definição foi usada por Euler em seu texto *Institutiones calculi integralis*, em 1768. Algo interessante a se destacar sobre a função gamma é o seu comportamento fatorial quando restrita ao conjunto dos naturais, motivo pelo qual é chamada de fatorial generalizado.

Proposição 2 (Fatorial Generalizado)**.** *Considere a função gamma de x,então, $\forall n \in \mathbb{R}$ vale,*

$$\Gamma(n+1) = n\Gamma(n).$$

Se $n \in \mathbb{N}$, então

$$\Gamma(n) = (n-1)!.$$

Prova. Iniciaremos provando a primeira parte do teorema, em seguida usaremos o Princípio da Indução Finita para demonstrar a segunda. Usando integração por partes,

$$\begin{aligned}\Gamma(n+1) &= \int_0^\infty t^n e^{-t} \mathrm{d}t \\ &= -t^n e^{-t}\Big|_0^\infty + \int_0^\infty n t^{n-1} e^{-t} \mathrm{d}t\end{aligned}$$

[1]O autor agradece o apoio financeiro oferecido pelo CNPq através da bolsa de iniciação científica.

$$
\begin{aligned}
&= -\lim_{t\to\infty}\left(\frac{t^n}{e^t}\right) + \int_0^\infty nt^{n-1}e^{-t}\mathrm{d}t \\
&= n\cdot\int_0^\infty t^{n-1}e^{-t}\mathrm{d}t \\
\Gamma(n+1) &= n\Gamma(n).
\end{aligned}
$$

Agora, mostraremos que a segunda parte do teorema é válida, mostrando primeiramente, que vale para $n = 1$.

$$
\begin{aligned}
\Gamma(1) &= \int_0^\infty e^{-t}\mathrm{d}t \\
&= -e^{-t}\Big|_0^\infty \\
&= \lim_{t\to\infty}\Big(-e^{-t}\Big) + 1 \\
\Gamma(1) &= 1 = 0!.
\end{aligned}
$$

Agora, supomos que vale qualquer valor de n, ou seja, $\Gamma(n) = (n-1)!$ e mostraremos que vale para $n+1$, usando o fato provado na primeira parte do teorema.

$$
\begin{aligned}
\Gamma(n+1) &= n\Gamma(n) \\
&= n(n-1)! \\
&= n!
\end{aligned}
$$

□

2 Produto de convolução

Definição 3 (Convolução)**.** Sejam f e g duas funções, chamamos de produto de convolução de f e g, denotada por $(f * g)(x)$, a integral abaixo,

$$(f * g)(x) = \int_0^x f(u)\cdot g(x-u)\mathrm{d}u,$$

desde que ela exista.

3 Integral fracionária por Riemann-Liouville

Definição 4 (Integral fracionária por Riemann-Liouville)**.** Sejam ν um número real maior que zero, f uma função contínua por partes em $(0,\infty)$ e integrável em qualquer subintervalo de

$[0,\infty)$. Então para $t>0$ definimos a Integral de ordem ν de Riemann-Liouville da função f, denotada por $J^\nu f(t)$ como sendo:

$$J^\nu f(t)=\frac{1}{\Gamma(\nu)}\int_0^t (t-s)^{\nu-1}f(s)\mathrm{d}s.$$

Exemplo 1 Calcular $J^\nu t^n$, com $n>0$.

$$\begin{aligned}J^\nu t^n &= \frac{1}{\Gamma(\nu)}\int_0^t (t-s)^{\nu-1}s^n\mathrm{d}s\\ &= \frac{1}{\Gamma(\nu)}\int_0^t \Big[t\Big(1-\frac{s}{t}\Big)\Big]^{\nu-1}s^n\mathrm{d}s,\end{aligned}$$

fazendo $u=\dfrac{s}{t}$, temos $\dfrac{\mathrm{d}u}{\mathrm{d}s}=\dfrac{1}{t}\Rightarrow \mathrm{d}s=t\mathrm{d}u$,

$$\begin{aligned}J^\nu t^n &= \frac{1}{\Gamma(\nu)}\int_0^1 \Big[t\Big(1-u\Big)\Big]^{\nu-1}(ut)^n t\mathrm{d}u\\ &= \frac{1}{\Gamma(\nu)}\int_0^1 t^{\nu+n}(1-u)^{\nu-1}u^n\mathrm{d}u\\ &= \frac{t^{\nu+n}}{\Gamma(\nu)}\int_0^1 u^n(1-u)^{\nu-1}\mathrm{d}u,\end{aligned}$$

realizando outra mudança de variável, $u=\mathrm{sen}^2\theta$, obtemos $\dfrac{\mathrm{d}u}{\mathrm{d}\theta}=2\mathrm{sen}\theta\cos\theta$,

$$J^\nu t^n = 2\frac{t^{\nu+n}}{\Gamma(\nu)}\int_0^{\frac{\pi}{2}} \mathrm{sen}^{2n+1}\theta\cos^{2\nu+1}\theta\mathrm{d}\theta$$

Definiremos uma função $B(a,b)=2\int_0^{\frac{\pi}{2}}\mathrm{sen}^{2a-1}\theta\cos^{2b-1}\theta\mathrm{d}\theta$. Considere $\Gamma(p)\Gamma(q)$:

$$\begin{aligned}\Gamma(p)\Gamma(q) &= \int_0^\infty e^{-u}u^{p-1}\mathrm{d}u\int_0^\infty e^{-v}v^{q-1}\mathrm{d}v\\ &= \int_0^\infty\int_0^\infty e^{-(u+v)}u^{p-1}v^{q-1}\mathrm{d}u\mathrm{d}v,\end{aligned}$$

trocando novamente as variáveis, $u=x^2\Rightarrow\dfrac{\mathrm{d}u}{\mathrm{d}x}=2x$ e $v=y^2\Rightarrow\dfrac{\mathrm{d}v}{\mathrm{d}y}=2y$,

$$\begin{aligned}\Gamma(p)\Gamma(q) &= 4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}x^{2p-2}y^{2q-2}xy\mathrm{d}x\mathrm{d}y\\ &= 4\int_0^\infty\int_0^\infty e^{-(x^2+y^2)}x^{2p-1}y^{2q-1}\mathrm{d}x\mathrm{d}y,\end{aligned}$$

agora, introduzindo coordenadas polares $x=r\cos\theta$, $y=r\mathrm{sen}\theta$, temos:

$$\begin{aligned}\Gamma(p)\Gamma(q) &= 4\int_0^{\frac{\pi}{2}}\int_0^\infty e^{-r^2}(r\cos\theta)^{2p-1}(r\mathrm{sen}\theta)^{2q-1}r\mathrm{d}r\mathrm{d}\theta\\ &= 4\int_0^{\frac{\pi}{2}}\int_0^\infty e^{-r^2}r^{2p+2q-1}\cos^{2p-1}\theta\mathrm{sen}^{2q-1}\theta\mathrm{d}r\mathrm{d}\theta\\ &= 2\Big(2\int_0^{\frac{\pi}{2}}\mathrm{sen}^{2q-1}\theta\cos^{2p-1}\theta\mathrm{d}\theta\Big)\Big(\int_0^\infty e^{-r^2}r^{2p+2q-1}\mathrm{d}r\Big)\\ &= 2B(q,p)\int_0^\infty e^{-r^2}r^{2p+2q-1}\mathrm{d}r,\end{aligned}$$

sendo $t = r^2 \Rightarrow \dfrac{\mathrm{d}t}{\mathrm{d}r} = 2r$, temos

$$\Gamma(p)\Gamma(q) = B(q,p)\int_0^\infty e^{-t}t^{p+q-1}\mathrm{d}t,$$

donde segue que

$$\Gamma(p)\Gamma(q) = B(q,p)\Gamma(p+q) \Rightarrow B(q,p) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)},$$

voltando ao ponto no qual definimos $B(a,b)$

$$\begin{aligned} J^\nu t^n &= \frac{t^{\nu+n}}{\Gamma(\nu)}B(n+1,\nu) \\ &= \frac{t^{\nu+n}}{\Gamma(\nu)}\frac{\Gamma(\nu)\Gamma(n+1)}{\Gamma(\nu+n+1)} \\ J^\nu t^n &= \frac{\Gamma(n+1)}{\Gamma(\nu+n+1)}t^{\nu+n}. \end{aligned}$$

4 Lei dos expoentes para integrais fracionárias de Riemann-Liouville

No cálculo tradicional, onde as ordens de integração são inteiras, o operador integral possui uma propriedade importante, sabemos que

$$J^n J^m = J^{n+m},$$

onde $n, m \in \mathbb{N}$. Essa equação é chamada de *Lei dos Expoentes* para integrais, nesta seção iremos provar a Lei dos Expoentes para integrais fracionárias de Riemann-Liouville. Queremos mostrar que para $\alpha, \beta \geq 0$ vale

$$J^\alpha J^\beta = J^{\alpha+\beta}.$$

Proposição 5 (Lei dos Expoentes para integrais fracionárias de Riemann- Liouville)**.** *Sejam $J^\alpha f(t), J^\beta f(t)$, integrais de Riemann-Liouville de ordem α e β de $f(t)$, então vale:*

$$J^\alpha J^\beta f(t) = J^{\alpha+\beta} f(t).$$

Prova. Vimos anteriormente que:

$$J^\alpha f(t) = \frac{1}{\Gamma(\alpha)}t^{\alpha-1} * f(t),$$

então, vamos definir a função auxiliar:

$$\Phi_\alpha(t) = \frac{1}{\Gamma(\alpha)}t^{\alpha-1},$$

com $\alpha > 0$, onde Φ é nula para $t < 0$. Deste modo

$$J^{\alpha} f(t) = \Phi_{\alpha}(t) = \frac{1}{\Gamma(\alpha)} t^{\alpha-1} * f(t),$$

com $\alpha > 0$, vamos mostrar que $\Phi_{\alpha}(t) * \Phi_{\beta}(t) = \Phi_{\alpha+\beta}(t)$. Da definição de produto de convolução temos

$$\begin{aligned}
\Phi_{\alpha}(t) * \Phi_{\beta}(t) &= \int_0^t \frac{s^{\alpha-1}}{\Gamma(\alpha)} \frac{(t-s)^{\beta-1}}{\Gamma(\beta)} \mathrm{d}s \\
&= \frac{1}{\Gamma(\alpha)\Gamma(\beta)} \int_0^t s^{\alpha-1} t^{\beta-1} \Big(1 - \frac{s}{t}\Big)^{\beta-1} \mathrm{d}s \\
&= \frac{t^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^t s^{\alpha-1} \Big(1 - \frac{s}{t}\Big)^{\beta-1} \mathrm{d}s,
\end{aligned}$$

fazendo a mudança de variável $u = \dfrac{s}{t} \Rightarrow \dfrac{\mathrm{d}u}{\mathrm{d}s} = \dfrac{1}{t} \Rightarrow t\mathrm{d}u = \mathrm{d}s$,

$$\begin{aligned}
\Phi_{\alpha}(t) * \Phi_{\beta}(t) &= \frac{t^{\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 (ut)^{\alpha-1} (1-u)^{\beta-1} t\mathrm{d}u \\
&= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 u^{\alpha-1} (1-u)^{\beta-1} \mathrm{d}u,
\end{aligned}$$

notemos que a integral acima é justamente a definição da função $B(\alpha, \beta)$ usada no Exemplo 1, apenas realizando a mudança de variáveis $u = sen^2(t) \Rightarrow \dfrac{\mathrm{d}u}{\mathrm{d}t} = 2sen(t)cos(t)$, visto que

$$\int_0^1 u^{\alpha-1} (1-u)^{\beta-1} \mathrm{d}u = 2 \int_0^{\frac{\pi}{2}} sen^{2\alpha-1}\theta cos^{2\beta-1}\theta \mathrm{d}\theta$$

então usaremos o mesmo resultado obtido anteriormente,

$$\begin{aligned}
\Phi_{\alpha}(t) * \Phi_{\beta}(t) &= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} B(\alpha, \beta) \\
&= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \\
&= \frac{t^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} \\
&= \Phi_{\alpha+\beta}(t).
\end{aligned}$$

Escrevendo $J^{\alpha} J^{\beta} f(t)$ como um produto de convolução

$$\begin{aligned}
J^{\alpha} J^{\beta} f(t) &= \Phi_{\alpha} * J^{\beta} f(t) \\
&= \Phi_{\alpha}(t) * \Phi_{\beta}(t) \cdot f(t) \\
&= \Phi_{\alpha+\beta}(t) * f(t) \\
&= J^{\alpha+\beta} f(t).
\end{aligned}$$

□

5 Derivada fracionária por Riemann-Liouville

Definição 6 (Derivada fracionária por Riemann-Liouville)**.** Seja $\mu \in \mathbb{R}_+^*$, m o menor inteiro maior que μ e $\nu = m - \mu$, de modo que $0 < \nu \leq 1$. A derivada fracionária de ordem μ por Riemann-Liouville de $f(x)$, com $x > 0$ é

$$D_{RL}^{\mu} = D^m[J^{\nu} f(x)],$$

onde D^m é a derivada de ordem m inteira e J^{ν} é a integral fracionária de ordem ν de Riemann-Liouville.

Exemplo 2 Calcular $D_{RL}^{\frac{1}{2}} t$.

$$D_{RL}^{\frac{1}{2}} t = D^1 J^{\frac{1}{2}} t,$$

como vimos no Exemplo 1, a integral acima vale

$$\begin{aligned} D_{RL}^{\frac{1}{2}} t &= D^1 \frac{\Gamma(2)}{\Gamma(\frac{5}{2})} t^{\frac{3}{2}} \\ &= \frac{3\Gamma(2)}{2\Gamma(\frac{5}{2})} t^{\frac{1}{2}}, \end{aligned}$$

de acordo com a propriedade da função gamma provada na Proposição 2, temos

$$\begin{aligned} D_{RL}^{\frac{1}{2}} t &= \frac{3 \cdot 1}{2\frac{3}{4}\sqrt{\pi}} t^{\frac{1}{2}} \\ &= \frac{2}{\sqrt{\pi}} t^{\frac{1}{2}} \\ &= \frac{2\sqrt{\pi t}}{\pi}. \end{aligned}$$

6 Derivada fracionária por Caputo

Definição 7 (Derivada fracionária por Caputo)**.** Seja $\mu \in \mathbb{R}_+^*$, m o menor inteiro maior que μ e $\nu = m - \mu$, de modo que $0 < \nu \leq 1$. A derivada fracionária de ordem μ por Caputo de $f(x)$, com $x > 0$ é

$$D_C^{\mu} = J^{\nu}[D^m f(x)]$$

novamente, D^m é a derivada de ordem m inteira e J^{ν} é a integral fracionária de ordem ν de Riemann-Liouville.

Exemplo 3 Calcular $D_C^{\frac{1}{2}} t$.

$$D_C^{\frac{1}{2}} t = J^{\frac{1}{2}} D^1 t$$

$$= J^{\frac{1}{2}}1,$$

pelo Exemplo 1, temos

$$D_C^{\frac{1}{2}}t = \frac{\Gamma(1)}{\Gamma(\frac{3}{2})}t^{\frac{1}{2}},$$

pelo fato provado na Proposição 2,

$$\begin{aligned} D_C^{\frac{1}{2}}t &= \frac{1}{\frac{\sqrt{\pi}}{2}}t^{\frac{1}{2}} \\ &= \frac{2}{\sqrt{\pi}}t^{\frac{1}{2}} \\ &= \frac{2\sqrt{\pi t}}{\pi}. \end{aligned}$$

7 Lei dos Expoentes para derivadas fracionárias de Riemann-Liouville

Novamente do cálculo tradicional, com ordens de derivação inteiras, o operador derivada possui a seguinte propriedade chamada de *Lei dos expoentes* para derivadas:

$$D^n D^m = D^{n+m},$$

onde $n, m \in \mathbb{N}$. No entanto, para a derivada de ordem fracionário, esta igualdade não se cumpre em todos os casos. Podem ser dados alguns contra exemplos para mostrar isso, mas apenas citaremos os casos que podem ocorrer na definição dada por Riemann-Liouville:

$$\begin{aligned} D_{RL}^{\alpha} D_{RL}^{\beta} = D_{RL}^{\beta} D_{RL}^{\alpha} \neq D_{RL}^{\alpha+\beta} \\ D_{RL}^{\alpha} D_{RL}^{\beta} \neq D_{RL}^{\beta} D_{RL}^{\alpha} = D_{RL}^{\alpha+\beta} \end{aligned}$$

onde $\alpha, \beta \geq 0$.

8 Riemann-Liouville

Proposição 8 (Teorema fundamental do cálculo fracionário por Riemann-Liouville)**.** *Sejam $\alpha \in \mathbb{R}_+^*$ e $n \in \mathbb{N}^*$ tais que $n-1 \leq \alpha < n$. Suponha f uma função contínua em um intervalo $[a, b]$, então para $t \in [a, b]$, então*

$$D_{RL}^{\alpha}[J^{\alpha} f(t)] = f(t)$$

e se $D_{RL}^{\alpha} f(t)$ existir e for integrável, então

$$J^{\alpha}[D_{RL}^{\alpha} f(t)] = f(t) - \sum_{k=1}^{n} \frac{(t-\alpha)^{\alpha-k}}{\Gamma(\alpha-k+1)}(D^{n-k}(J^{n-\alpha} f))(a).$$

Prova. Para a primeira parte do teorema, pela Definição 6 e a Proposição 5, temos

$$\begin{aligned} D^{\alpha}_{RL}[J^{\alpha}f(t)] &= D^{n}[J^{n-\alpha}(J^{\alpha}f(t))] \\ &= D^{n}[J^{n-\alpha+\alpha}f(t)] \\ &= D^{n}[J^{n}f(t)] \\ &= f(t). \end{aligned}$$

Para a segunda igualdade, temos

$$\begin{aligned} J^{\alpha}[D^{\alpha}f(t)] &= \frac{1}{\Gamma(\alpha)}\int_a^t (t-s)^{\alpha-1} D^{\alpha}_{RL}f(s)\mathrm{d}s \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\Big[\int_a^t (t-s)^{\alpha} D^{\alpha}_{RL}f(s)\mathrm{d}s\Big]. \end{aligned}$$

Mas temos que

$$\frac{1}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha} D^{\alpha}_{RL}f(s)\mathrm{d}s = \frac{1}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha}\frac{\mathrm{d}^n}{\mathrm{d}s^n}\Big[J^{n-\alpha}f(s)\Big]\mathrm{d}s,$$

integrando por partes, fazendo $u(s) = (t-s)^{\alpha}$ e $\frac{\mathrm{d}v}{\mathrm{d}s} = \frac{\mathrm{d}^n}{\mathrm{d}s^n}[J^{n-\alpha}f(s)]$ e obtendo $\frac{\mathrm{d}u}{\mathrm{d}s} = -\alpha(t-s)^{\alpha-1}$ e $v(s) = \frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}[J^{n-\alpha}f(s)]$, usando também a Proposição 2 chegamos a

$$\begin{aligned} \frac{1}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha} D^{\alpha}_{RL}f(s)\mathrm{d}s &= \frac{1}{\Gamma(\alpha+1)}\Big[(t-s)^{\alpha}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}[J^{n-\alpha}f(s)\Big]_{s=\alpha}^{t} \\ &\quad + \frac{\alpha}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}[J^{n-\alpha}f(s)]\mathrm{d}s \\ &= -\frac{(t-a)^{\alpha}}{\Gamma(\alpha+1)}\Big(\frac{\mathrm{d}^{n-1}[J^{n-\alpha}f]}{\mathrm{d}t^{n-1}}\Big)(a) \\ &\quad + \frac{1}{\Gamma(\alpha)}\int_a^t (t-s)^{\alpha-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}[J^{n-\alpha}f(s)]\mathrm{d}s. \end{aligned}$$

Integrando por partes novamente, obtemos,

$$\begin{aligned} \frac{1}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha} D^{\alpha}_{RL}f(s)\mathrm{d}s &= -\frac{(t-a)^{\alpha}}{\Gamma(\alpha+1)}\Big(\frac{\mathrm{d}^{n-1}[J^{n-\alpha}f]}{\mathrm{d}t^{n-1}}\Big)(a) \\ &\quad + \frac{1}{\Gamma(\alpha)}\int_a^t (t-s)^{\alpha-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}[J^{n-\alpha}f(s)]\mathrm{d}s \\ &= -\frac{(t-a)^{\alpha}}{\Gamma(\alpha+1)}\Big(\frac{\mathrm{d}^{n-1}[J^{n-\alpha}f]}{\mathrm{d}t^{n-1}}\Big)(a) \\ &\quad - \frac{(t-a)^{\alpha-1}}{\Gamma(\alpha)}\Big(\frac{\mathrm{d}^{n-2}[J^{n-\alpha}f]}{\mathrm{d}t^{n-2}}\Big)(a) \\ &\quad + \frac{1}{\Gamma(\alpha-1)}\int_a^t (t-s)^{\alpha-2}\frac{\mathrm{d}^{n-2}}{\mathrm{d}s^{n-2}}J^{n-\alpha}f(s)\mathrm{d}s, \end{aligned}$$

integrando por partes repetidamente até que não hajam mais derivadas em s, temos

$$\begin{aligned} \frac{1}{\Gamma(\alpha+1)}\int_a^t (t-s)^{\alpha} D^{\alpha}_{RL}f(s)\mathrm{d}s &= -\frac{(t-a)^{\alpha}}{\Gamma(\alpha+1)}\Big(\frac{\mathrm{d}^{n-1}[J^{n-\alpha}f]}{\mathrm{d}t^{n-1}}\Big)(a) \\ &\quad - \frac{(t-a)^{\alpha-1}}{\Gamma(\alpha)}\Big(\frac{\mathrm{d}^{n-2}[J^{n-\alpha}f]}{\mathrm{d}t^{n-2}}\Big)(a) \end{aligned}$$

$$
\begin{aligned}
&- \cdots - \frac{(t-\alpha)^{\alpha-n+1}}{\Gamma(\alpha-n+2)}(J^{n-\alpha}f)(a) \\
&+ \frac{1}{\Gamma(\alpha-n+1)}\int_a^t (t-s)^{\alpha-n} J^{n-\alpha}f(s)\mathrm{d}s \\
&= -\sum_{k=1}^{n} \frac{(t-a)^{\alpha-k+1}}{\Gamma(\alpha-k+2)}\Big(\frac{\mathrm{d}^{n-k}J^{n-\alpha}f}{\mathrm{d}t^{n-k}}\Big)(a) + J^{\alpha-n+1}J^{n-\alpha}f(t) \\
&= J^1 f(t) - \sum_{k=1}^{n} \frac{(t-a)^{\alpha-k+1}}{\Gamma(\alpha-k+2)}\Big(\frac{\mathrm{d}^{n-k}J^{n-\alpha}f}{\mathrm{d}t^{n-k}}\Big)(a).
\end{aligned}
$$

Voltando ao início de nossa prova, temos

$$
\begin{aligned}
J^\alpha[D^\alpha f(t)] &= \frac{\mathrm{d}}{\mathrm{d}t}\Big[\int_a^t (t-s)^\alpha D^\alpha_{RL} f(s)\mathrm{d}s\Big] \\
&= \frac{\mathrm{d}}{\mathrm{d}t}J^1 f(t) - \frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=1}^{n} \frac{(t-a)^{\alpha-k+1}}{\Gamma(\alpha-k+2)}\Big(\frac{\mathrm{d}^{n-k}J^{n-\alpha}f}{\mathrm{d}t^{n-k}}\Big)(a) \\
&= f(t) - \sum_{k=1}^{n} \frac{(\alpha-k+1)(t-a)^{\alpha-k}}{\Gamma(\alpha-k+2)}\Big(\frac{\mathrm{d}^{n-k}[J^{n-\alpha}f]}{\mathrm{d}t^{n-k}}\Big)(a) \\
&= f(t) - \sum_{k=1}^{n} \frac{(t-a)^{\alpha-k}}{\Gamma(\alpha-k+1)}\Big(\frac{\mathrm{d}^{n-k}}{\mathrm{d}t^{n-k}}[J^{n-\alpha}f]\Big)(a).
\end{aligned}
$$

□

9 Caputo

Agora o teorema fundamental do cálculo na versão de Caputo. Como na versão de Riemann-Liouville também temos $D^\alpha_C[J^\alpha f(t)] = f(t)$ e portanto vamos nos concentrar na segunda expressão. Esta difere do resultado da seção anterior.

Proposição 9 (Teorema fundamental do cálculo fracionário por Caputo)**.** *Sejam $\alpha \in \mathbb{R}^*_+$ e $n \in \mathbb{N}^*$ tais que $n-1 \leq \alpha < n$. Se f possuir derivada de ordem n contínua em $[a,b]$, então*

$$
J^\alpha[D^\alpha_C f(t)] = f(t) - \sum_{k=0}^{n} \frac{(t-a)^k}{k!} f^{(k)}(a).
$$

Prova. Temos que

$$
\begin{aligned}
J^\alpha[D^\alpha_C f(t)] &= J^\alpha[J^{n-\alpha}\frac{\mathrm{d}^n}{\mathrm{d}t^n}f(t)] \\
&= J^n[\frac{\mathrm{d}^n}{\mathrm{d}t^n}f(t)] \\
&= \frac{1}{(n-1)!}\int_a^t (t-s)^{n-1}\frac{\mathrm{d}^n}{\mathrm{d}s^n}f(s)\mathrm{d}s.
\end{aligned}
$$

Integrando por partes chegamos a

$$
J^\alpha[D^\alpha_C f(t)] = \frac{1}{(n-1)!}\int_a^t (t-s)^{n-1}\frac{\mathrm{d}^n}{\mathrm{d}s^n}f(s)\mathrm{d}s
$$

$$= \frac{1}{(n-1)!}\Big[(t-s)^{n-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}f(s)\Big]_{s=a}^{t} + \frac{(n-1)}{(n-1)!}\int_a^t (t-s)^{n-2}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-s}}f(s)\mathrm{d}s$$
$$= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) + \frac{1}{(n-2)!}\int_a^t (t-s)^{n-2}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}f(s)\mathrm{d}s.$$

Integrando por partes novamente,

$$\begin{aligned}
J^{\alpha}[D_C^{\alpha}f(t)] &= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) + \frac{1}{(n-2)!}\int_a^t (t-s)^{n-2}\frac{\mathrm{d}^{n-1}}{\mathrm{d}s^{n-1}}f(s)\mathrm{d}s \\
&= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) + \frac{1}{(n-2)!}\Big[(t-s)^{n-2}\frac{\mathrm{d}^{n-2}}{\mathrm{d}s^{n-2}}f(s)\Big]_{s=a}^{t} \\
&\quad + \frac{(n-2)}{(n-2)!}\int_a^t (t-s)^{n-3}\frac{\mathrm{d}^{n-2}}{\mathrm{d}s^{n-2}}f(s)\mathrm{d}s \\
&= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) - \frac{(t-a)^{n-2}}{(n-2)!}\Big(\frac{\mathrm{d}^{n-2}f}{\mathrm{d}s^{n-2}}\Big)(a) \\
&\quad + \frac{1}{(n-3)!}\int_a^t (t-s)^{n-3}\frac{\mathrm{d}^{n-2}}{\mathrm{d}s^{n-2}}f(s)\mathrm{d}s.
\end{aligned}$$

Repetindo o processo até a última derivada em s, temos

$$\begin{aligned}
J^{\alpha}[D_C^{\alpha}f(t)] &= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) - \frac{(t-a)^{n-2}}{(n-2)!}\Big(\frac{\mathrm{d}^{n-2}f}{\mathrm{d}s^{n-2}}\Big)(a) \\
&\quad - \cdots - \frac{(t-a)}{1!}\Big(\frac{\mathrm{d}f}{\mathrm{d}s}\Big)(a) + \frac{1}{0!}\int_a^t \frac{\mathrm{d}}{\mathrm{d}s}f(s)\mathrm{d}s \\
&= -\frac{(t-a)^{n-1}}{(n-1)!}\Big(\frac{\mathrm{d}^{n-1}f}{\mathrm{d}s^{n-1}}\Big)(a) - \frac{(t-a)^{n-2}}{(n-2)!}\Big(\frac{\mathrm{d}^{n-2}f}{\mathrm{d}s^{n-2}}\Big)(a) \\
&\quad - \cdots - \frac{(t-a)}{1!}\Big(\frac{\mathrm{d}f}{\mathrm{d}s}\Big)(a) + f(t) - f(a) \\
&= f(t) - \sum_{k=0}^{n-1}\frac{(t-a)^k}{k!}f^{(k)}(a).
\end{aligned}$$

□

Referências

ZILL, D. G.. **Equações Diferenciais.** São Paulo: Makron Books, 2001.

CAMARGO, R. de F.. **Cálculo fracionário e aplicações.** 2009. 141 f. Tese (Doutorado) - Curso de Matemática, Departamento de Matemática, Univesidade Estadual de Campinas, Campinas, 2009.

DANNON, H. Vic.. The fundamental theorem of the fractional calculus, and the meaning of fractional derivatives. **Gauge Institute Journal,** Minneapolis, p.1-26, fev. 2009.

PODLUBNY, I.. **Fractional differential equations.** 198. ed. Kosice: Academic Press, 1999.

RAHIMY, M.. **Applications of fractional differential equations**. Applied Mathematical Sciences, Vol. 4, 2010, No. 50, 2453-2461.

As tendências em educação matemática e suas relações

Ana Cristina Dellabetta
Universidade Estadual do Oeste do Paraná
anacristinadellabetta@hotmail.com

Ana Maria Foss
Universidade Estadual do Oeste do Paraná
anafoss@bol.com.br

Daniele Donel
Universidade Estadual do Oeste do Paraná
danidonel@hotmail.com

Viviane Fátima Ribeiro
Universidade Estadual do Oeste do Paraná
www.vivianefribeiro@hotmail.com

Resumo: As Tendências em Educação Matemática (Resolução de Problemas, Modelagem Matemática, Tecnologias em Educação Matemática, Etnomatemática, Investigação Matemática e História da Matemática) estão previstas nas Diretrizes Curriculares para o ensino de Matemática do Estado do Paraná. Acredita-se que uma proposta metodológica, fundamentada nas Tendências em Educação Matemática faça a diferença na compreensão, no significado e aplicação do conhecimento matemático, e seja capaz de tornar a matemática uma disciplina agradável, mais fácil de aprender e de ser ensinada. Assim no presente trabalho argumentaremos sobre como essas Tendências se relacionam e consequentemente como podem vir a contribuir significativamente no processo ensino aprendizagem se utilizadas de forma adequada. Ressaltamos em nosso trabalho que todas elas podem estão interligadas, não havendo necessidade que o professor as utilize separadamente, mas sim há a possibilidade de interligar essas Tendências, pois se acredita que a educação inovadora pode ser um caminho para alcançar a aprendizagem, sendo a sala de aula o local estimulante para que aluno possa mostrar sua criatividade e seu potencial.

Palavras-chave: Tendências em Educação Matemática; processo ensino aprendizagem.

1 Introdução

Podemos observar que a Matemática é a disciplina que mais faz "inimigos", muitos alunos do ensino básico não gostam dela justamente por não terem aprendido conceitos básicos dessa ciência, necessários para a construção de novos conceitos ou até mesmo não entenderem alguns procedimentos práticos utilizados na resolução de problemas e exercícios. Na busca de proporcionar um ensino de matemática mais significativo, surgem as Tendências em Educação Matemática. Elas estão previstas nas Diretrizes Curriculares para o ensino de Matemática

do Estado do Paraná (DCE) (PARANÁ, 2008). Acredita-se que uma proposta metodológica, fundamentada nas Tendências em Educação Matemática possibilite uma melhor compreensão e torne a construção do conhecimento matemático mais significativo e seja capaz de tornar a matemática uma disciplina agradável, fácil de aprender e de ser ensinada.

Apresentaremos neste trabalho como as Tendências da Educação Matemática se relacionam e o que as mesmas apresentam em comum; os aspectos motivacionais, o papel do professor e do aluno.

2 As Tendências em Educação Matemática

Em geral as aulas de Matemática são trabalhadas da mesma forma como era no século passado, a chamada metodologia Tradicional, de modo que torna difícil despertar no aluno, o qual vivencia nessa sociedade tecnológica, o interesse pelos conteúdos programáticos desenvolvidos em sala de aula. Assim faz-se necessária uma atitude de mudança. Nesse sentido, as Tendências em Educação Matemática, orientadas pelas DCEs, (Resolução de Problemas, Modelagem Matemática, Tecnologias em Educação Matemática, Etnomatemática, Investigação Matemática e História da Matemática) propõem novas formas de ensinar e aprender Matemática.

Na utilização destas Tendências vê-se a importância de um bom planejamento por parte do professor, pois é preciso sair de sua "zona de conforto", o que gera uma "quebra de rotina", fugindo das metodologias tradicionais que são baseadas na reprodução do conteúdo presente no livro didático. Para isso faz-se necessário uma prévia preparação de sua aula refletindo sobre as diversas situações e dúvidas que podem vir a surgir, buscando evitar imprevistos para propiciar aos alunos uma aprendizagem significativa.

Vemos a importância desse planejamento também na fala de Carneiro e Passos (2014, p. 104) quando falam da utilização de Tecnologias em Educação Matemática:

> O professor deve estar preparado para enfrentar muitos imprevistos, questões e dúvidas às quais poderá não saber responder, muito mais que em aulas sem as tecnologias. [...]. Existe necessidade da formação contínua do professor, pois as TIC permitem novas formas de abordar os conteúdos, o que requer um maior domínio da matéria, assim como o conhecimento técnico.

Ainda segundo Fernando Carneiro e Cármen Lúcia Brancaglion Passos (2014) a utilização de tecnologias auxilia na formação do cidadão, ou seja, levando o estudante a refletir sobre aspectos sociais, políticos, culturais, entre outros. Segundo os Parâmetros Curriculares Nacionais (PCN) (1998, p. 140, apud CARNEIRO, PASSOS, 2014, p. 102) "a tecnologia deve servir para enriquecer o ambiente educacional, propiciando a construção de conhecimentos por meio de uma

atuação ativa, crítica e criativa por parte de alunos e professores".

Além disso, em todas as Tendências faz-se necessário um papel ativo do aluno, pois, por exemplo, em uma aula que se faz uso de Tecnologias de Informação e Comunicação (TIC), segundo Carneiro e Passos (2014) a dinâmica da sala é alterada, os alunos trabalham em equipes e "as possibilidades de elaboração de conhecimentos são muito diferentes das produzidas em aulas sem as TIC, porque o estudante é um participante ativo desse processo". Dessa forma, professor e aluno tornam-se atores cooperativos, desenvolvem-se e constroem novos conhecimentos. A relação professor-aluno toma uma dimensão diferente daquela que ocorre normalmente na sala de aula, em que o professor é a autoridade e o detentor do conhecimento, pois a construção do conhecimento se dá por meio do trabalho em conjunto, ambos são responsáveis pela aprendizagem. O professor precisa participar de forma ativa do processo de construção do conhecimento do aluno, sendo um mediador, motivador e orientador da aprendizagem.

Do ponto de vista da História da Matemática, um fator positivo acerca da abordagem histórica dos conteúdos matemáticos, segundo Silva e Ferreira (2011, apud LOPES, FERREIRA, 2013), é permitir ao docente a previsão dos possíveis erros dos alunos. De acordo com Berlinghoff e Gouvêa (2008, apud LOPES, FERREIRA, 2013), entender que muitas pessoas tinham dificuldades em lidar com certos assuntos matemáticos, mesmo depois de um tempo de divulgação de suas ideias básicas, "nos ajuda a compreender as dificuldades que os estudantes possam ter. Saber como foram superadas essas dificuldades historicamente também pode indicar um modo de ajudar os estudantes a superarem tais obstáculos" (p. 3). Assim, estratégias e questionamentos podem ser preparados antecipadamente pelo professor, promovendo sua postura como mediador entre o saber e o aluno.

O resgate da história dos saberes matemáticos ensinados no espaço escolar, proporcionado pela História da Matemática traz a construção de um olhar crítico sobre o assunto em questão, proporcionando reflexões acerca do conhecimento matemático e sua relação com a realidade já que podemos associar o surgimento de grande parte dos conhecimentos matemáticos às necessidades do ser humano. Como recurso em sala de aula, os PCN afirmam que a História da Matemática contribui para a construção de um olhar mais crítico aos objetos de conhecimento.

Além disso, conceitos abordados em conexão com sua história constituem-se veículos de informação cultural, sociológica e antropológica de grande valor formativo. "A História da Matemática é, nesse sentido, um instrumento de resgate da própria identidade cultural". (BRASIL, 1997, p.34 apud LOPES, FERREIRA, 2013). Lopes e Ferreira afirmam que a História da Matemática motiva os alunos no estudo, uma vez que eles podem comparar processos do passado com os do presente, apontando sobre a evolução de um determinado conteúdo, o que

contribui para quebrar o mito de uma matemática pronta e acabada, que não depende dos esforços humanos.

Também cabe ao professor a formação crítica dos seus alunos e as Tendências podem vir a auxiliar para que eles entendam seu papel na sociedade podendo exercer sua cidadania. Ubiratan D' Ambrósio enfatiza isso quando fala do papel do Programa Etnomatemática, pois ele cumpre o papel de formar jovens e adultos nos aspectos tanto críticos como de exercer a sua cidadania. Dessa forma, o autor salienta que deve ser pensado um novo currículo voltado para o desenvolvimento da criatividade e da curiosidade para que exista uma postura crítica e de questionamento constante. Dessa forma, a matemática ensinada nas escolas deixaria de ter um caráter perverso e alienante (D'AMBROSIO, 2008) e passaria a ser mais acessível, mostrando a sua importância nos aspectos social, político, econômico e ideológico.

A etnomatemática defende que cabe ao professor se preparar para agir com novas dinâmicas, buscando aproximar a Matemática à realidade dos alunos, com um olhar mais amplo e com práticas que acolham os saberes e fazeres em todo contexto sociocultural dos alunos. Pois assim poderá caminhar lado a lado com seu educando, ajudando-o na construção do conhecimento.

Referindo-se à Modelagem Matemática os seguintes autores afirmam:

> No ambiente de Modelagem, o aluno é incentivado a trabalhar em grupo, possibilitando o convívio social e o desenvolvimento do senso de cooperação, responsabilidade, criticidade e comunicação oral entre os membros do grupo. (SONEGO, 2009, p. 21 apud QUARTIERI; KNIJNIK, 2012).

Caldeira (2005, apud KLÜBER, BURAK, 2008) em sua proposta, diz que há, enquanto aprendizagem e uso de Modelagem, uma abrangência maior do que o simples ensino de conteúdos de matemática. Incitam-se decisões concernentes à participação dos alunos e professores como cidadãos e agentes de mudança da comunidade em que estão inseridos.

Biembengut (1999, p. 36 apud KLÜBER, BURAK, p. 24, 2008) defende que a modelagem é "um caminho para despertar no aluno o interesse por tópicos matemáticos que ainda desconhece, ao mesmo tempo que aprende a arte de modelar, matematicamente".

A Modelagem, proporciona ao aluno ser co-construtor do seu conhecimento, pois apresenta problemas diários dos estudantes onde estes são convidados a investigar por meio da Matemática a solução para os problemas, permitindo que o conhecimento matemático se torne mais interessante e estimulante.

Acerca da Resolução de problemas e da Investigação matemática, pode-se despertar a criticidade de alunos, partindo de problemas sociais, políticos e econômicos que envolvam a

Matemática.

A estratégia de Resolução de Problemas, no ensino da Matemática, deve voltar-se para o desenvolvimento do pensamento criador, visto atualmente a sociedade vem exigindo cada vez mais do ser humano sua capacidade criativa, buscando novos caminhos matemáticos para solucionar os problemas com que possam se deparar em seu cotidiano. É preciso tornar os alunos, pessoas capazes de enfrentar situações e contextos variáveis, que exijam deles a aprendizagem de novos conhecimentos e habilidades. Nesse sentido, um dos veículos mais acessíveis pode vir a ser a resolução de problemas.

O professor deve ser capaz de estabelecer um espaço de discussão oral, onde o educando seja levado a elaborar estratégias, apresentar hipóteses, fazendo o registro das soluções encontradas. Isto vem favorecer o pensamento matemático, passando a ser uma ação criativa. O aluno pode resolver o problema através da oralidade, do desenho até chegar a possibilidade de utilização dos critérios formais impostos pelas regras matemáticas.

Investigar é descobrir o que não se sabe. É esse o espírito de uma proposta pedagógica de Investigação Matemática. Nesse processo o aluno é conduzido a investigar, a fazer conjecturas, a realizar provas e refutações, discutindo os seus argumentos com colegas e professor. Nessa proposta, a investigação é encaminhada na resolução de problemas ou exercícios. Na investigação o aluno é conduzido ao prazer da descoberta, motivado por um problema em aberto que poderá apresentar diferentes formas de resolução.

Na Resolução de Problemas e na Investigação Matemática, o professor tem o papel de encaminhar as atividades e certificar-se de que os alunos entenderam o que foi proposto e acompanhar o desenvolvimento do trabalho, não dando respostas imediatas em relação às dúvidas que surgirem, mas colocar outras questões fazendo com que os alunos reflitam, também o professor precisa estar atento ao desenvolvimento das atividades para que os alunos possam evoluir e se necessário intervir direcionando-os para que o objetivo da aula seja alcançado.

3 Relações entre as Tendências em Educação Matemática

As seis Tendências da Educação Matemática analisadas estão todas interligadas. Percebemos que há um forte vínculo entre Resolução de Problemas, Modelagem e Investigação Matemática. Segundo Bassanezi (2002 apud QUARTIEI, KNIJNIK, 2012) a Modelagem apresenta como característica partir de uma situação-problema da realidade ou de um recorte dela e a partir destes desenvolver questionamentos que possibilitam ao aluno uma atitude de investigação.

Portanto está intimamente ligado à Resolução de Problemas e à Investigação Matemática. Ainda vale ressaltar que esta pode ter vínculo com a Etnomatemática, assim como destaca Caldeira (2009, p.38 apud QUARTIERI; KNIJNIK, 2012, p.12) a Modelagem proporciona que os conhecimentos possam ser "construídos de acordo com os interesses sociais, políticos culturais, estando profundamente relacionado à realidade do estudante". Esse entendimento sobre modelagem é pautado na indagação, que não é uma simples explicitação do problema, mas uma atitude que acompanha todo o processo de resolução. A indagação conduz à investigação, sendo essa a busca, seleção, organização e manipulação de informações.

A Modelagem Matemática é a arte de expressar por meio da linguagem Matemática situações-problema do dia-a-dia, sendo que tem estado presente desde os tempos mais primitivos. Ou seja, a Modelagem é tão antiga quanto à própria Matemática, surgindo de aplicações na rotina diária dos povos antigos. A Modelagem Matemática apresenta como característica partir de uma situação-problema da realidade ou de um recorte da realidade e, sobre este, desenvolver questionamentos que possibilitam ao aluno uma atitude de investigação e uso de seus conhecimentos matemáticos para resolver as questões propostas. Dessa forma, a Modelagem pode estar estritamente ligada com a História da Matemática, Investigação Matemática, Etnomatemática e Resolução de Problemas. A Modelagem Matemática, bem como a Resolução de Problemas, podem ser aliadas as Tecnologias em Educação Matemática utilizando-se ferramentas no contexto da informática que podem vir a proporcionar aos alunos novas formas de aprender.

Na verdade, assim como a Modelagem, a Resolução de Problemas também pode ser articulada à História da Matemática tendo visto que as regras utilizadas para resolver os problemas eram apenas ferramentas para a obtenção do resultado correto. Os problemas mais interessantes são aqueles que nasceram no seio das necessidades de grupos sociais e culturais (etnomatemática).

Assim, o professor pode trabalhar com Resolução de problemas buscando contribuições em outras Tendências como a História da Matemática e a Etnomatemática, pois:

> Resolver problemas é uma parte integrante de toda a aprendizagem matemática e, assim, ela não deveria ser uma parte isolada do programa de Matemática. [...] Os contextos dos problemas podem variar desde experiências familiares envolvendo as vidas dos estudantes ou seu dia-a-dia na escola, até aplicações envolvendo as ciências ou o mundo do trabalho. [...] Bons problemas dão aos estudantes a oportunidade de solidificar e estender sua compreensão e estimular nova aprendizagem. [...] Muitos conceitos matemáticos podem ser introduzidos através de problemas baseados nas experiências familiares vividas pelos estudantes ou de contextos matemáticos (STANDARDS, 2000, p. 52 apud Nunes, 2010, p. 82).

E ainda Carlos Vianna (2000, p. 3-4) também afirma que:

> A História da Matemática pode ser uma fonte relevante de problemas para serem trabalhados na Resolução de Problemas, o estudo da solução dada aos problemas reais que foram enfrentados em épocas diversas pode fornecer contribuições relevantes para o desenvolvimento de técnicas de modelagem e para o aprimoramento de modelos já elaborados, o conhecimento da História da Matemática dos diversos povos entrelaça-se inevitavelmente com os trabalhos de Etnomatemática...

A Etnomatemática também está intimamente ligada a outras Tendências entre elas a História da Matemática, pois procura investigar o fazer matemático de diversos povos, fazendo essa relação com a História, traz a matemática mais próxima do aluno e dessa maneira o educando entendendo como se deu a construção do conhecimento encontrará sentido no que esta aprendendo, compreendendo assim a evolução dos conceitos matemáticos ao longo dos tempos. Além disso, proporcionar que os estudantes conheçam diferentes matemáticas (ou etnomatemáticas), de povos desfavorecidos economicamente e politicamente, constitui-se como um caminho para a valorização do conhecimento que o próprio aluno traz consigo. Afinal, (re) conhecer as contribuições de diferentes povos, fugindo de uma visão única da (etno) matemática eurocentrista, possibilita atribuir valor à própria cultura ao perceber-se inserido no contexto do conhecimento escolar.

> É de extrema importância que em situações de ensino sejam consideradas as contribuições significativas de culturas que não tiveram hegemonia política e, também, que seja realizado um trabalho que busca explicar, entender e conviver com procedimentos, técnicas e habilidades matemáticas desenvolvidas no entorno sociocultural próprio a certos grupos culturais. (MIGUEL; MIORIM, 2011, p.54 apud Lopes e Ferreira, 2013).

A Etnomatemática ajuda a compreender a História da Matemática e seus diferentes caminhos de construção fazendo o educando entender o mundo em que vive, propondo um caráter interdisciplinar. Nesse sentido, observa-se que a Etnomatemática também está ligada à História da Matemática, assim como a Modelagem e a resolução de problemas, de forma que o educando possa realmente entender a aplicação dos conteúdos matemáticos e ainda questionar se os mesmos serão de fato úteis para a sociedade em que vive.

Ainda podemos buscar ajuda nas Tecnologias em Educação Matemática, pois esta amplia a possibilidade de incorporar a Etnomatemática na educação matemática, os recursos tecnológicos facilitam experimentações matemáticas, contribuindo para o raciocínio lógico e ainda permite a circulação e divulgação de conhecimentos de diversos locais e diferentes povos.

Pode parecer contraditório falarmos em uma matemática sofisticada quando falamos de Etnomatemática, mas justamente o essencial da Etnomatemática é incorporar a matemática cultural, contextualizada, na Educação Matemática e o uso de mídias amplia essa possibilidade, pois os recursos tecnológicos têm favorecido as experimentações matemáticas, visto que o raciocínio lógico com qualidade é essencial para situar-se no mundo moderno e chegar a uma

nova organização de sociedade, que permita exercer a crítica e análise do mundo moderno que vivemos.

Portanto, é possível ensinar Matemática utilizando as Tendências da Educação Matemática de forma articulada, por exemplo, a partir de um problema de uma situação real, pode-se buscar a sua solução construindo um modelo matemático, o qual permite que o educando entenda que a Matemática não é uma ciência pronta e acabada, mas que se desenvolve ao longo do tempo, conforme visto pela História da Matemática, que aproveita os conhecimentos que o educando tem de suas experiências fora do contexto escolar, mostrando características da Etnomatemática, e, ao solucionar o problema o educando pode realmente verificar se essa solução trouxe uma vantagem para a sociedade que ele vivencia, levando também nessa busca, aspectos de investigação, podendo-se também haver auxílio das tecnologias.

Enfim as Tendências em Educação Matemática podem estar todas interligadas, sendo que não há necessidade do professor seguir uma única tendência, mas sim, trabalhar de forma articulada entre elas.

4 Considerações Finais

A Educação Matemática precisa ser continuamente repensada de forma crítica e reflexiva, mediante aos avanços tecnológicos faz-se necessário a adequação, a melhoria e a inovação na forma de lecionar. Sendo assim, o profissional de Educação precisa conhecer as várias Tendências para reconstruir sua prática de ensino, transformando sua aula mais atrativa e interessante, vindo a proporcionar ao educando a oportunidade de analisar, investigar e participar ativamente na construção do conhecimento. Trabalhar com as Tendências pode contribuir para o ensino e aprendizagem, bem como abrir caminhos para uma aprendizagem significativa, onde o aluno consegue visualizar a Matemática em situações da sua vida diária.

Como ressaltamos em nosso trabalho que todas as Tendências estão interligadas, não havendo necessidade que o professor aborde cada uma delas separadamente, mas sim há a possibilidade de interligá-las, pois se acredita que a educação inovadora e interdisciplinar pode ser um caminho para alcançar a melhoria na aprendizagem, sendo a sala de aula o local estimulante para que aluno possa mostrar sua criatividade e seu potencial.

Ademais se tem observado que vários professores ao trabalharem com as Tendências em Educação Matemática destacam o fator motivacional que elas têm. Elas proporcionam ao aluno conhecer verdadeiramente a matemática, pois os conhecimentos matemáticos ensinados na escola

aparecem descontextualizados e sem funcionalidade. Conforme D'Ambrósio (2012) apud Lopes e Ferreira (2013) "do ponto de vista de motivação contextualizada, a matemática que se ensina hoje nas escolas é morta". Desta maneira, os alunos pensam que todos os assuntos tratados em sala de aula estão em sua forma mais acabada, mais pronta e, além disso, não é permitido questionar tal perfeição. As Tendências permitem que o aluno perceba que a Matemática não se reduz apenas a resolver problemas através de algoritmos memorizados como geralmente ela é apresentada aos alunos, porém ao abordar essas Tendências o professor faz com que o educando deixe de ser passivo, tenha curiosidade e perceba a beleza da Matemática.

Referências

CARNEIRO, R. F., PASSOS, C. L. B. **A utilização das Tecnologias da Informação e Comunicação nas aulas de Matemática: Limites e possibilidades**. Revista Eletrônica de Educação, v. 8, n. 2, p. 101-119, 2014.

D'AMBROSIO, U. **O Programa Etnomatemática: uma síntese**. Canoas: Acta Scientiae, v. 10, n.1, p.7-16, jan./jun. 2008.

KLÜBER, T. E., BURAK, D. **Concepções de modelagem matemática: contribuições teóricas**. São Paulo: Educação Matemática Pesquisa, 2008.

LOPES, L. S., FERREIRA, A. L. A. **Um olhar sobre a história nas aulas de matemática**. Belo Horizonte: Abakós, v. 2, n. 1, p. 75–88, nov. 2013.

NUNES, C. B. **O Processo Ensino-Aprendizagem-Avaliação de Geometria através da Resolução de Problemas: perspectivas didático-matemáticas na formação inicial de professores de matemática**. São Paulo: Universidade Estadual Paulista, 2010.

PARANÁ. Secretaria do Estado da Educação. **Diretrizes Curriculares da Rede Pública da Educação Básica do Estado do Paraná – Matemática**. Curitiba: SEED, 2008.

QUARTIERI, M. T., KNIJNIK, G. **Modelagem matemática na escola básica: surgimento e consolidação**. Lajeado: Caderno Pedagógico, 2012.

VIANNA, C. R. **História da matemática na educação matemática**. Anais VI Encontro Paranaense de Educação Matemática. Londrina: Editora da UEL, 2000. p. 15-19.

Teoria de pontos críticos de funcionais e aplicações

Paula Isabel Becker
Universidade Estadual do Oeste do Paraná - Cascavel
e-mail: paula_be10@hotmail.com

Sandro Marcos Guzzo
Universidade Estadual do Oeste do Paraná - Cascavel
e-mail: smguzzo@gmail.com

Resumo: O objetivo deste trabalho é mostrar a existência de uma solução para um Problema Variacional. Um Problema Variacional consiste em obter pontos críticos para uma função real, definida em um subconjunto de um espaço de funções, que esteja associada ao problema diferencial. Assim, procuramos uma formulação variacional de um problema inicial dado, cujo ponto crítico deste funcional remete a resolução do problema inicial.

Palavras-chave: Pontos críticos; métodos variacionais; equações diferenciais.

1 Introdução

A resolusão de uma Equação diferencial usando a Teoria dos Pontos Críticos, também conhecida como Método Variacional, será abordada neste trabalho. A importância de estudar este assunto se deve à necessidade de tratar um Problema Variacional de maneira elegante e simples. O Método Direto de Cálculo das Variações supre a necessidade de fazer apelo à equação de Euler-Lagrange, procurando apenas obter o ponto crítico de um funcional associado à Equação Diferencial proposta. Em meados do século XIX Hilbert foi o pioneiro a usar este método empregando-o para provar a existência de solução para o problema de Dirichlet. Hoje este método é eficaz e de grande importância para os estudos referentes à resolução de Equações Diferenciais Ordinárias.

Assim, procura-se explanar, de maneira sucinta, a resolução de um problema de valor de contorno linear através de métodos variacionais.

2 Teoria dos Pontos Críticos

O principal objetivo é mostrar a existência de uma solução para o seguinte problema de contorno linear utilizando a teoria dos pontos críticos. A base para os estudos feitos neste trabalho é encontrada no artigo do (Figueiredo, 1988). É importante destacar que as integrais deste trabalho são limitadas de a até b, porém serão denotadas apenas como $\int$ ao invés de $\int_a^b$.

Segue o problema de contorno

$$-(pu')' + qu = f, u(a) = u(b) = 0, \tag{1}$$

Tomaremos as seguintes condições nos coeficientes

$$p \in C^1[a,b], q \in C^0[a,b]; p(t) > 0, q(t) \geq 0, \tag{2}$$

para qualquer t no intervalo $[a,b]$.

Vamos supor que a função f é dada tal que $f \in C^0[a,b]$. Uma função $u \in C^2[a,b]$ que satisfaz a equação (1) e que se anula na extremidade do intervalo dado é chamada de solução clássica do problema 1. Assim, será mostrada a existência de uma solução u e será provado que essa solução pertence a $C^2[a,b]$.

Tomando então a equação (1), multiplica-se a equação por uma função $v \in C^1[a,b]$, de modo que $v(a) = v(b) = 0$, assim segue que

$$-v(p(t)u')' + q(t)uv = fv,$$

integrando por partes temos

$$\int p(t)u'v'dt + \int q(t)uvdt - \int fvdt = 0, \tag{3}$$

para qualquer $v \in C_0^1[a,b]$, de forma que a notação C_0^1 indica que $v(a) = v(b) = 0$ são satisfeitas.

Inicia-se de forma a procurar uma função limitada e mensurável que seja válida para todo conjunto de funções $\Phi \in C_0^1[a,b]$, este funcional será denominado como solução fraca para a equação (1). Nesse sentido consideremos o funcional a seguir

$$\Phi(v) = \frac{1}{2}\int p\left|v'\right|^2 dt + \frac{1}{2}\int qv^2dt - \int fvdt, \tag{4}$$

com $v \in C_0^1[a,b]$.

Dado $u_0 \in C_0^1[a,b]$ segue que a derivada direcional de Φ, em u_0, na direção de v é dada por

$$\begin{aligned}
\Phi'(u_0)v &= \lim_{t\to 0}\frac{1}{t}\left\{\Phi(u_0+tv) - \Phi(u_0)\right\} \\
&= \lim_{t\to 0}\frac{1}{t}\left\{\frac{1}{2}\int p\left|(u_0+tv)'\right|^2 dt + \frac{1}{2}\int q(u_0+tv)^2dt - \int f(u_0+tv)dt\right. \\
&\qquad \left. - \frac{1}{2}\int p\left|(u_0)'\right|^2 dt - \frac{1}{2}\int qu_0^2dt + \int fu_0dt\right\} \\
&= \lim_{t\to 0}\frac{1}{t}\left\{\frac{1}{2}\int p(u_0'+v+v't)^2dt + \frac{1}{2}\int qu_0^2dt + \int qu_0tvdt + \frac{1}{2}\int qt^2v^2dt\right.
\end{aligned}$$

$$- \int fu_0 dt - \int ftvdt - \frac{1}{2}\int pu_0^2 dt - \frac{1}{2}\int qu_0^2 dt + \int fu_0 dt\Big\}$$
$$= \int pu_0' v' dt + \int qu_0 v dt - \int fv dt$$

para $v \in C_0^1[a,b]$ qualquer. No Cálculo de Variações a expressão acima é conhecida como a primeira variação do funcional Φ.

Considerando os resultados acima e tomando u_0 como um mínimo do funcional Φ em $C_0^1[a,b]$ temos que

$$\Phi(u_0) \leq \Phi(u_0 + tv) \forall t \in \mathbb{R},$$

para qualquer v em $C_0^1[a,b]$ e também

$$\Phi'(u_0)v = 0$$

ou seja, u_0 é uma solução fraca para o problema em questão. Assim, basta que provemos agora que o funcional Φ tem um mínimo em $C_0^1[a,b]$. Se isto for provado, tem-se a existência de uma solução fraca de (1).

Para provarmos que o funcional Φ tem um mínimo teremos que provar que este é limitado inferiormente e que assume o ponto de mínimo. Inicia-se com a demonstração de que este funcional é limitado inferiormente em $C_0^1[a,b]$. Tomando $\Phi(v)$ podemos observar que

$$\Phi(v) \geq p_0 \int |v'|^2 dt - \left(\int f^2 dt\right)^{\frac{1}{2}} \cdot \left(\int v^2 dt\right)^{\frac{1}{2}}, \tag{5}$$

para todo $v \in C_0^1[a,b]$, de forma que $0 < p_0 = \min\{p(t) : t \in [a,b]\}$ e das condições citadas na equação (2) e usando a desigualdade de Cauchy-Schwarz, que é descrita em (Lima, 2003) como

$$|\langle x, y\rangle| \leq |x| \cdot |y|,$$

ou ainda neste caso específico

$$\int fv dt \leq \left(\int f^2 dt\right)^{\frac{1}{2}} \cdot \left(\int v^2 dt\right)^{\frac{1}{2}}.$$

Dando continuidade, considera-se também a desigualdade abaixo

$$\int |v'|^2 dt \geq c \int v^2 dt, \tag{6}$$

para qualquer $v \in C_0^1[a,b]$ e $c > 0$ uma constante independente de v. A desigualdade (6) é conhecida como a desigualdade de Wirtinger e pode ser demonstrada usando o Teorema Fundamental do Cálculo. De (5) e (6) segue que

$$\Phi(v) \geq \frac{p_0}{2} \int |v'|^2 dt - \left(\frac{1}{c}\int f^2 dt\right)^{\frac{1}{2}} \cdot \left(\int |v'|^2 dt\right)^{\frac{1}{2}}, \tag{7}$$

tomando neste caso

$$\left(\int |v'|^2\, dt\right)^{\frac{1}{2}} = x$$

segue que

$$\begin{aligned}\Phi(v) &\geq \frac{p_0}{2}x^2 - \left(\frac{1}{c}\int f^2 dt\right)^{\frac{1}{2}} x \\ &\geq -\frac{1}{2p_0}\cdot\frac{1}{c}\int f^2 dt \\ &= -\frac{1}{2cp_0}\int f^2 dt\end{aligned}$$

para todo v em $C_0^1[a,b]$.

Assim, observa-se que o funcional Φ é limitado inferiormente. Basta agora que mostremos que o ínfimo de Φ é assumido. Para este próximo passo, deve-se considerar o teorema de Bolzano-Weierstrass da Análise que afirma que toda função real contínua definida em um intervalo fechado e limitado da reta assume seu ínfimo nesse intervalo. Este teorema está relacionado com a parte topológica do funcional, logo, pode-se notar que é conveniente introduzir uma topologia em $C_0^1[a,b]$. Segue abaixo um resultado da Topologia que generaliza o teorema de Bolzano-Weierstrass e que servirá de ferramenta para a resolução do problema 1.

Teorema 1. *Seja X um espaço topológico compacto e seja $\Phi : X \to \mathbb{R}$ uma função real semi-contínua inferiormente definida em X. Então o ínfimo de Φ existe e é assumido em um ponto $u_0 \in X$.*

Sabe-se que o espaço $C_0^1[a,b]$, com suas definições usuais de soma e produto por um escalar, é um espaço vetorial sobre $\mathbb{R}$. Busca-se agora que este seja um espaço normado, e para isto vamos muni-lo com a seguinte norma

$$\|v\| = \left(\int |v'|^2\, dt\right)^{\frac{1}{2}}. \tag{8}$$

Para verificar que (8) define uma norma basta usarmos a desigualdade (6) e a desigualdade de Minkowski para integrais.

Entende-se que Φ é contínuo em X, de fato podemos observar que se $v_n \to v \in X$ então pela desigualdade (6) $v_n \to v \in L^2$, e assim concluimos que

$$\left|\int p\,|v_n'|^2\, dt - \int p\,|v'|^2\, dt\right| \leq \widehat{p}\int \left||v_n'|^2 - |v'|^2\right| dt \to 0$$

$$\left|\int q{v_n}^2 dt - \int qv^2 dt\right| \leq \widehat{q}\int |v_n^2 - v^2|\, dt \to 0$$

$$\left|\int f v_n dt - \int f v dt\right| \leq \left(\int f dt\right)^{\frac{1}{2}}\left(\int |v_n - v|^2\, dt\right)^{\frac{1}{2}} \to 0$$

onde $\widehat{p} = \max\{p(t) : t \in [a,b]\}$ e $\widehat{q} = \max\{q(t) : t \in [a,b]\}$. Assim, mostramos que o funcional Φ é contínuo e, portanto, é semicontínuo inferiormente.

Considerando agora um real $R > 0$ escolhido de forma conveniente, temos que

$$\inf_X \Phi = \inf\{\Phi(v) : v \in X, \|v\| \leq R\}.$$

Como consequência disso, podemos restringir nossas análises à bola $B_R = \{v \in X : \|v\| \leq R\}$, e assim tentar aplicar o teorema citado acima ao funcional Φ restrito à B_R. Porém, para isto, teremos que lidar com algumas dificuldades, como por exemplo: B_R não é compacta e um dos motivos para isso é o fato de X não ser completo. Com isso, o próximo passo na resolução deste problema é completar o espaço X.

Vamos completar o espaço X na norma (8). Considerando, inicialmente, o espaço $L^2[a,b]$ das funções reais mensuráveis a Lebesgue, definidas no intervalo $[a,b]$ e tais que $\int f^2 dt < \infty$. Tomando o produto interno $\langle f, g\rangle_{L^2} = \int fg dt$, teremos este espaço como sendo o espaço de Hilbert denotado por $H_0^1[a,b]$, de modo que este é um subespaço de $L^2[a,b]$ das funções u que possuem derivada fraca em $L^2[a,b]$ e de modo que $u(a) = u(b) = 0$.

Seguindo o conceito de derivada fraca temos que: u tem derivada fraca em $L^2[a,b]$ se existir $v \in L^2[a,b]$ tal que

$$\int v\varphi dt = -\int u\varphi' dt \tag{9}$$

para todo φ em $C_c^1[a,b]$, no qual o subíndice c tem por objetivo indicar que a função φ se anula fora de um intervalo fechado contido em (a,b).

Denotamos por u' a única derivada fraca de u, logo, pode ser ratificado que u é contínua em $[a,b]$. No espaço $H_0^1[a,b]$ define-se a norma

$$\|u\| = \left(\int u^2 dt + \int |u'|^2 dt\right)^{\frac{1}{2}} \tag{10}$$

e pode-se provar que $C_0^1[a,b]$ é denso em $H_0^1[a,b]$. Disto e por continuidade segue que a desigualdade de Wirtinger é válida $u \in H_0^1[a,b]$. Logo segue que a norma (10) é equivalente a

$$\|u\|_{H^1} = \left(\int |u'|^2 dt\right)^{\frac{1}{2}}. \tag{11}$$

Logo, a partir disso pode-se considerar $H_0^1[a,b]$ munido da norma (11), de modo que este é completo. Assim o complemento do espaço X na norma (8) pode ser identificado como o espaço de funções $H_0^1[a,b]$.

Agora, o funcional Φ está definido em $H_0^1[a,b]$, de modo que este é contínuo, porém $B_R = \{v \in H_0^1[a,b] \cdot \|v\|_{H^1} \leq R\}$ persiste não sendo compacta. Isto se deve ao fato de $H_0^1[a,b]$

ser um espaço de dimensão infinita e uma bola unitária em um espaço normado é compacta se, e somente se, o espaço for de dimensão finita.

Neste caso, sabe-se que o espaço $H_0^1[a,b]$ é um espaço de Hilbert definido com o seguinte produto interno

$$\langle u, v\rangle_{H^1} = \int u'v'dt,$$

ou seja, a norma desse produto interno é exatamente a norma (11). Dentre as características de um espaço de Hilbert, deve-se citar que as bolas deste espaço são fracamente compactas.

Para conseguirmos aplicar o teorema 1 com X sendo a bola B_R em $H_0^1[a,b]$ munido da topologia fraca, resta verificar que Φ seja semicontínua inferiormente na topologia fraca de $H_0^1[a,b]$. Para isto, necessitamos do teorema a seguir:

Teorema 2. *Seja E um espaço de Banach e $\Phi : E \to \mathbb{R}$ um funcional semicontínuo inferiormente (na norma) e convexo. Então Φ é semicontínuo inferiormente na topologia fraca de E.*

Lembremos que uma função convexa é uma função que satisfaz

$$\alpha f(u) + (1-\alpha)f(v) \geq f(\alpha u + (1-\alpha)y),$$

para quaisquer u e v no domínio da função e $\alpha \in [0,1]$. Desta forma Φ é claramente convexa pois é uma combinação de funções convexas (quadrática e modular) e operadores lineares (derivada e integral).

Observa-se então, que o teorema 2 só será válido se usarmos a semicontinuidade inferior. Assim, usando o teorema 1 pode–se concluir que Φ assume seu ínfimo em $u_0 \in B_R$. Tomando R tal que $\Phi(v) > \Phi(0) = 0$, para $\|v\|_{H^1} = R$, conclui-se que u_0 está no interior de B_R. Podemos observar que o funcional Φ em $H_0^1[a,b]$ é diferenciável a Fréchet. Assim seja

$$\nabla\Phi : H_0^1[a,b] \to H_0^1[a,b]$$

a diferencial do funcional, de forma que se usou o teorema de Riesz-Fréchet para identificar o espaço dos funcionais lineares contínuos em $H_0^1[a,b]$ com o próprio $H_0^1[a,b]$. Logo

$$\langle \nabla\Phi u_0, v\rangle_{H^1} = \Phi'(u_0)v,$$

em que $\Phi'(u_0)$ já foi definida anteriormente no início desta seção. Portanto, dado $u_0 \in H_0^1[a,b]$, este satisfaz a seguinte relação

$$\int pu_0'v'dt + \int qu_0vdt = \int fvdt, \tag{12}$$

para qualquer $v \in H_0^1[a,b]$.

Assim, pode-se concluir que um determinado u_0 é a solução fraca do problema 1. A primeira condição para a existência de uma solução para o problema está concluída, em que provou-se que este problema possui uma solução fraca. O próximo passo é provar a unicidade desta solução. Para isto, suponha que existam duas soluções $u_1, u_2 \in H_0^1[a,b]$. Com base na equação (12) segue que

$$\int p(u_1' - u_2')v'dt + \int q(u_1 - u_2)vdt = 0,$$

para qualquer v em $H_0^1[a,b]$.

Usando a hipótese citada na equação (2) e tomando $v = u_1 - u_2$ temos

$$\int p(u_1' - u_2')^2 dt = 0$$

implicando em $u_1' = u_2'$. Tomando a seguinte expressão

$$\int u_i'\varphi dt = -\int u_i\varphi' dt,$$

para todo $\varphi \in C_c^1[a,b]$. Utilizando a igualdade acima citada obtemos

$$\int (u_1 - u_2)\varphi' dt = 0\forall,$$

para todo $\varphi \in C_c^1[a,b]$.

Pela igualdade acima e como $(u_1 - u_2) \in L^2[a,b]$, segue que $u_1 - u_2 =$ constante. Portanto utilizando (2) segue que $u_1 = u_2$.

Assim, mostramos que um determinado u_0 é solução fraca do problema 1 e este é único.

Em seguida, vamos provar que $u_0 \in C^2$. Tomando então a equação (12) e reescrevendo-a, temos

$$\int pu_0'v'dt = -\int [qu_0 - f]vdt,$$

para todo $v \in H_0^1[a,b]$.

Esta expressão mostra que pu_0' possui derivada fraca em $L^2[a,b]$ e ainda

$$(pu_0')' = qu_0 - f. \tag{13}$$

Conclui-se então, que pu_0'é contínua, logo u_0' também é contínua. De (13) obtemos

$$pu_0'' = -p'u_0' + qu_0 - f$$

mostrando que u_0'' também é contínua.

Portanto, o problema 1 admite solução.

Referências

Figueiredo, Djairo G. **Métodos Variacionais em Equações Diferenciais**. Matemática Universitária N.7, Junho de 1988.

Lima, Elon Lages. **Espaços Métricos**. Rio de Janeiro: Associação Instituto Nacional de Matemática pura e aplicada, 2003.

A importância do Laboratório de Ensino de Matemática (LEM) no ensino e aprendizagem de Matemática e relatos de experiências quanto à implementação de um LEM e a utilização de jogos

Ana Maria Foss
Universidade Estadual do Oeste do Paraná
anafoss@bol.com.br

Daniele Donel
Universidade Estadual do Oeste do Paraná
danidonel@hotmail.com

Resumo: O Laboratório de Ensino de Matemática pode ser um instrumento eficaz tanto para propiciar ao aluno formas de conhecer, criar, manipular, levantar hipóteses, discutir afirmações, desenvolver e construir instrumentos matemáticos que possam ser utilizados como facilitadores de sua aprendizagem, como para proporcionar ao educador um local para pesquisa, reflexão e trabalho. De modo a sanar, em primeiro lugar, as dificuldades dos próprios professores no que diz respeito à sua formação. Seria fundamental, do ponto de vista da aprendizagem dos alunos e dos próprios professores, a implementação do Laboratório de Ensino de Matemática em qualquer instituição de ensino. Conscientes da necessidade da construção de um LEM, juntamente a direção do Colégio Estadual Ieda Baggio Mayer, do município de Cascavel, Paraná foi pedido um espaço para a implantação do LEM, sendo cedida uma sala, porém que necessita de reforma. Dessa forma impossibilitados de organizar o espaço físico, estamos empenhados em confeccionar jogos e materiais manipuláveis e preparando sugestões de atividades e orientações para o uso dos mesmos, que irão compor o futuro LEM do colégio. Também desenvolvemos atividades referentes ao uso de jogos com alunos da sala de apoio do 6º ano do Ensino Fundamental II na disciplina de Matemática do mesmo..

Palavras-chave: Laboratório de Ensino de Matemática; jogos matemáticos; materiais manipuláveis.

1 Introdução

Em sala de aula, os alunos nem sempre são colocados em situações em que tenham de agir para se tornarem sujeitos de sua aprendizagem, por exemplo, para encontrar diferentes soluções para a mesma questão ou para argumentar sobre a validade ou não de certa solução. Ensinar matemática exige do professor não só um conhecimento dos conteúdos, como também de métodos de ensino mais eficazes para promover a aprendizagem de seus alunos, métodos estes que não se reduzam somente a livros, quadro e giz. Um dos procedimentos que pode auxiliar o professor a tornar os conhecimentos matemáticos trabalhados na escola mais significativos e a tornar suas aulas mais interessantes é o uso de jogos e materiais manipuláveis.

Professores que querem realizar um trabalho com o uso de jogos e materiais manipuláveis encontram dificuldades em fazê-lo, pois como a maioria das escolas públicas não possui um espaço físico para organizar e guardar esses materiais, os docentes não têm a sua disposição um local apropriado para desenvolverem essas atividades pedagógicas, para pesquisa, reflexão e trabalho, que auxiliaria neste desafio sobre as melhores formas de ensinar e aprender Matemática. O espaço ao qual nos referimos e que defenderemos sua importância é o Laboratório de Ensino de Matemática (LEM). Embora sejam diversas as atividades pedagógicas que podem ser trabalhadas em um LEM, para efeito de estudo o enfoque será dado para as que vivenciamos no Programa Institucional de Bolsa de Iniciação à Docência (PIBID), assim como as demais experiências tidas no LEM da Universidade Estadual do Oeste do Paraná (Unioeste) e a implementação de um LEM no Colégio Estadual Ieda Baggio Mayer do município de Cascavel, Paraná.

2 Importância da criação de um LEM nas escolas

De acordo com Lorenzato (2006), no decorrer da história vários educadores destacaram a importância do apoio visual ou visual tátil como facilitador para a aprendizagem. Relata que, por volta de 1650 Comenius escreveu que o ensino deveria dar-se do concreto para o abstrato, justificando que o conhecimento começa pelos sentidos e que só se aprende fazendo. Locke, no século XVIII, dizia da necessidade da experiência sensível para alcançar o conhecimento. Atualmente, a tarefa dos educadores em geral não é mais a de transmitir, mas dar condições para que a aprendizagem realmente aconteça. O interesse na aprendizagem depende das situações estimuladoras criadas pelo educador para proporcionar ao educando o maior número possível de descobertas e desafios, estimulando, assim, a curiosidade dos alunos.

Partindo disso, o principal objetivo do Laboratório de Ensino de Matemática (LEM) é desenvolver e difundir atividades para o ensino de Matemática, de modo que os alunos aprendam a fazer fazendo. O Laboratório de Ensino de Matemática pode ser ao mesmo tempo, um instrumento eficaz tanto para propiciar ao aluno formas de conhecer, criar, manipular, levantar hipóteses, discutir afirmações, desenvolver e construir instrumentos matemáticos que possam ser utilizados como facilitadores de sua aprendizagem, como para proporcionar ao educador um local para pesquisa, reflexão e trabalho, auxiliando-o neste grande desafio sobre as melhores formas de ensinar e aprender Matemática.

Segundo Lorenzato (2006), o Laboratório de Ensino da Matemática na escola é uma sala-ambiente reservada para que as aulas de Matemática aí aconteçam de maneira a estruturar, organizar, planejar e construir o fazer matemático, facilitando tanto para o professor como para

o aluno o questionamento, a procura, a experimentação, a análise, a compreensão de conceitos e a conclusão de uma determinada aprendizagem, inclusive com a produção de materiais que possam facilitar o aprimoramento da prática pedagógica. Deve ser um local de referência para as atividades matemáticas, em que os professores possam se empenhar em tornar a matemática mais compreensível para seus alunos. Neste local, o professor poderá também planejar aulas e realizar outras atividades como exposições, olimpíadas, jogos, avaliações, entre outras.

Desta maneira, seria importante que os estabelecimentos de ensino tivessem um espaço no qual os educadores pudessem realizar pesquisas, construir materiais didáticos diversos em um processo de elaboração de conceitos, de aplicação dos mesmos em situações-problema, de modo a sanar, em primeiro lugar, as dificuldades dos próprios professores no que diz respeito à sua formação. Porque só assim eles terão condições para, depois, utilizar tais materiais com os alunos, suscitando neles o interesse e a participação desejada e auxiliando-os na elaboração e compreensão dos conhecimentos. Ou seja, seria fundamental, do ponto de vista da aprendizagem dos alunos e dos próprios professores, a implementação do Laboratório de Ensino de Matemática em qualquer instituição de ensino.

Porém diferentemente dos laboratórios de Física e Química, o Laboratório de Ensino de Matemática é pouco comum em escolas de Ensino Fundamental II e Ensino Médio em razão da inexistência de um espaço físico, recursos para investimento no mesmo ou até desconhecimento de sua importância. Durante a nossa primeira vida escolar ouvíamos dizer que havia na escola laboratório de Química e Física e adorávamos quando as aulas eram em tais espaços. Mas por que não tínhamos um Laboratório de Ensino de Matemática? Só fomos nos deparar com tal laboratório no Ensino Superior. Além do fascínio que tínhamos ao entrar naquelas salas, com todos aqueles aparelhos, mesas maiores, informativos pelas paredes, víamos por meio das experiências vivenciadas nos laboratórios o conteúdo e de maneira mais dinâmica o compreendia, obtendo uma aprendizagem mais significativa, relacionada à teoria trabalhada em sala de aula nessas disciplinas.

O laboratório, portanto, é um ambiente propício para estimular no aluno o gosto pela matemática, a perseverança na busca de soluções e a confiança em sua capacidade de aprender e fazer matemática. Além de contribuir para a construção de conceitos, procedimentos e habilidades matemáticas, pode propiciar também a busca de relações, propriedades e regularidades, estimulando o espírito investigativo. Por isso, deve ser neste local da escola onde se respire Matemática o tempo todo e possa ser também um ambiente permanente de busca e redescoberta.

De acordo com Mendes (2009, apud IRINEU, SANTOS, RODRIGUES, 2015) os ma-

teriais que são fornecidos pelo LEM, contribuem para que o professor desenvolva um trabalho melhor com seus alunos. As atividades que são feitas através desses materiais "[...] tem uma estrutura matemática a ser redescoberta pelo aluno que, assim, se torna um agente ativo na construção de seu próprio conhecimento matemático." (MENDES, 2009, p. 25).

Contudo segundo Antonio e Andrade (2011),

> O LEM não é uma poção mágica que resolverá todos os problemas da aprendizagem da matemática, nem se configura em uma estratégia para ser usada em todos os conteúdos. Além do que, deve-se tomar cuidado com "uso pelo uso" dele. Assim, o professor ao planejar sua aula, perceberá a necessidade ou não do uso dos materiais e jogos disponíveis no LEM, bem como se sua utilização deve ser individual, em grupos ou de observação apenas, cabendo ao professor a sua manipulação.

O uso do laboratório depende muito da atitude do professor e a sua busca pelo conhecimento, pois mesmo que o laboratório tenha uma infinidade de materiais, estes de nada adiantarão se não forem devidamente explorados ou se estes simplesmente forem mostrados ao aluno.

3 Primeiros passos da implementação do LEM no Colégio Estadual Ieda Baggio Mayer

Conscientes da necessidade da construção de um LEM, apoiados teoricamente por artigos que relatam os benefícios propiciados pelo Laboratório no Ensino da Matemática, pelas experiências tidas no LEM da Universidade Estadual do Oeste do Paraná (Unioeste) e por discussões feitas pelos bolsistas juntamente com os professores coordenadores e pelos professores supervisores das escolas beneficiadas pelo Programa Institucional de Bolsa de Iniciação á Docência (PIBID), juntamente a direção do Colégio Estadual Ieda Baggio Mayer do município de Cascavel, Paraná foi pedido um espaço no colégio para a implantação do LEM, sendo cedida uma sala do colégio, porém que necessita de reforma. Dessa forma impossibilitados de organizar o espaço físico, estamos empenhados em confeccionar os materiais didáticos que irão compô-lo.

De acordo com Lorenzato (2006, apud OTTESBACH, PAVANELLO, 2007, p.7-8) um Laboratório de Ensino de Matemática, de modo geral, pode ter:

- revistas, jornais e artigos;
- livros didáticos, paradidáticos e outros;
- jogos;

- quebra-cabeças;
- problemas desafiadores e de lógica;
- questões de olimpíadas, ENEM e vestibulares;
- textos sobre história da matemática;
- cds, transparências, fotos;
- figuras;
- sólidos;
- modelos estáticos ou dinâmicos;
- materiais didáticos industrializados;
- instrumentos de medidas;
- computadores, calculadoras;
- materiais didáticos construídos pelos alunos e professores;
- materiais e instrumentos necessários à produção de materiais didáticos e outros.

Faz-se importante, para a confecção de jogos e materiais didáticos manipuláveis, a utilização de materiais recicláveis, tanto pela fácil aquisição de material como pelo universo de criatividade que pode proporcionar e também é muito importante a participação dos alunos no processo de construção e que haja material didático suficiente para se trabalhar nas aulas. Os aspectos visuais do laboratório de Matemática devem ser atraentes, os materiais devem estar expostos e não fechados em armários, para que despertem a curiosidade dos alunos. O docente deve procurar meios para aumentar a motivação para a aprendizagem.

Assim estamos confeccionando jogos e materiais didáticos manipuláveis e ainda preparando sugestões de atividades e orientações para o uso dos jogos e materiais manipulativos. Dentre os materiais confeccionados podemos destacar os jogos Hex Multiplicativo, Caixinha de Sorteio, Divisão em Linha, Bingo de Tabuada e jogo do Positivo e Negativo e como material didático manipulável temos os tangrans.

De acordo com Lorenzato (2006, p. 6).

> Um LEM pode ser inicialmente um depósito/arquivo de instrumentos, livros, materiais manipuláveis, transparências, filmes, matérias-primas, entre outros e, posteriormente, se tornar um espaço organizado com a colaboração de educandos e educadores.

Segundo Lorenzato (2006, p. 11), a construção de um LEM não é objetivo para ser atingido a curto prazo, uma vez construído, ele demanda constante complementação, a qual por sua vez, exige que o professor se mantenha atualizado. Sabemos que estamos dando os primeiros passos para a implementação desse LEM, porém devemos agora, manter acesa a chama de manutenção deste, sempre alimentando-a e encorajando e preparando os professores.

4 Experiências com jogos no ensino da Matemática

Durante o primeiro semestre do ano de 2016, desenvolvemos atividades com alunos da sala de apoio[1] do 6º ano do Ensino Fundamental II na disciplina de Matemática do Colégio Estadual Ieda Baggio Mayer, do município de Cascavel, no estado do Paraná. Nesse período foi privilegiado o trabalho com as quatro operações fundamentais da aritmética (adição, subtração, multiplicação e divisão). A primeira atividade desenvolvida teve como objetivo as multiplicações, principalmente as tabuadas. Para trabalhar esse conteúdo confeccionamos o Bingo de Tabuada. Este jogo é similar ao bingo comum com o objetivo de fixação do conteúdo tabuada. O Número de jogadores é ilimitado, o professor faz o sorteio das fichas com as multiplicações e o jogador deverá marcar em sua cartela as respostas que correspondem ao número em seu cartão. O professor determina o tempo que aguardará para resolução do cálculo e ganhará o jogador que preencher primeiro toda a sua cartela. Além disso, o professor pode estabelecer ganhadores com o preenchimento de apenas uma linha ou o "azarão" (último a marcar).

Antes de iniciar a atividade, convém sempre esclarecer regras ou dificuldades com o conceito envolvido no jogo. Assim, foi trabalhado anteriormente o conceito da tabuada, fazendo os alunos entenderem que ela não foi inventada, mas sim construída. Muitos alunos não sabiam, por exemplo, que o resultado de 5×4 era obtido por meio da adição $4+4+4+4+4$. A partir daí eles foram estimulados a não olhar uma tabuada pronta quando não soubessem, mas sim, calcular o seu resultado. Porém posteriormente para agilizar outros cálculos é necessário que depois de compreendido o processo, o professor estimule o aluno para que este memorize-a.

Os alunos se envolveram bastante nesta atividade. Após sorteada a multiplicação cada aluno calculava-a em uma folha de registro para verificar se tinha em sua cartela. O trabalho com o jogo durou duas aulas geminadas, baseado no fato de que o mesmo exige tempo para cálculos e registros. Na folha de registro, todas as operações são escritas e esses registros nos permitem avaliar os avanços e as dificuldades dos alunos, mostrando se há necessidade de retomada pelo

[1]Sala de apoio – Programa da Secretaria do Estado da Educação do Paraná que visa atender alunos de 6º ano com defasagem de conteúdos de Língua Portuguesa e Matemática.

professor.

Constatou-se com essa atividade que o jogo prendeu a atenção dos alunos, o que não acontecia nas atividades de resolução de exercícios ou situações problemas, nas quais muitos se distraiam e não realizavam as atividades. Com o jogo, a maior parte deles participou e mostraram-se mais interessados em saber quando voltaríamos a trabalhar com eles. Esse primeiro contato foi válido e o ponto positivo que destacamos, é que alguns alunos no final da aula disseram que gostaram da possibilidade de não decorar a tabuada sem compreender o que estavam fazendo. Acreditamos que assim eles terão a oportunidade de decorar a tabuada pela prática constante, entendendo o processo e não simplesmente pela obrigação.

Um ponto negativo que destacamos é a de que os alunos que já sabiam o resultado de algumas multiplicações ou as faziam mais rapidamente que seus colegas queriam jogar rapidamente, por isso falavam os resultados para os colegas sem deixá-los pensar para jogar. Nesse momento, se fez necessário uma intervenção sobre os objetivos do jogo na aula e o tempo necessário a cada um para o raciocínio e registro dos resultados.

A segunda atividade foi desenvolvida para o encerramento do semestre. Esta teve como objetivo abordar todos os conteúdos trabalhados durante o semestre, ou seja, as quatro operações fundamentais da aritmética. O jogo utilizado foi Caixinha de Sorteio, o qual consiste em caixinhas que contém os números 1 a 99, em seu fundo, nos quais o jogador deve arremessar dois grãos de feijão, obtendo assim dois números. Em seguida o jogador deve arremessar um dado que contêm os símbolos das quatro operações em suas faces (figura 1). Assim ele deverá fazer a operação sorteada com os dois números obtidos na caixinha. Faz-se isso para todos os jogadores (sugere-se que seja jogado em dupla) pontua (um ponto) quem obtiver como resposta o maior número, o jogador que obtiver um número que não é divisível pelo outro número sorteado é desclassificado da rodada. O jogo termina quando alguém obtiver dez pontos.

Figura 1: Jogo Caixinha de Sorteio (Fonte: as autoras).

Por ainda não terem alguns conhecimentos matemáticos acerca dos números negativos e de como utilizar o algoritmo da divisão para quando o dividendo era menor que o divisor, estipulamos mais algumas regras ao jogo, por exemplo, quando eles sorteassem a operação de divisão (ou subtração) eles deviam operar dividindo o maior número pelo menor (ou subtraindo o menor do maior). Também pedimos a eles que registrassem seus cálculos (a operação sorteada por seu oponente e a sua) em uma folha de registro.

Destacamos alguns pontos positivos dessa atividade. Os pares comparavam seus resultados para as mesmas operações como uma forma de correção e quando havia diferença entre os resultados revisavam suas contas a fim de encontrar um erro. Além disso, com a folha de registro podemos detectar que apenas um aluno apresentou dúvidas quanto ao algoritmo da multiplicação e assim pudemos planejar estratégias para ajudar esse aluno.

Os jogos nessas atividades foram usados como treinamento. Lara (2003, p. 25 apud Strapason, 2011, p.38) trata os jogos de treinamento na seguinte perspectiva:

> [...] é necessário que o/a aluno/a utilize várias vezes o mesmo tipo de pensamento e conhecimento matemático, não para memorizá-lo, mas, sim, para abstraí-lo, estendê-lo, ou generalizá-lo, como também, para aumentar sua autoconfiança e sua familiarização com o mesmo. O treinamento pode auxiliar no desenvolvimento de um pensamento dedutivo ou lógico mais rápido. Muitas vezes, é através de exercícios repetitivos que o/a aluno/a percebe a existência de outro caminho de resolução que poderia ser seguido aumentando, assim, suas possibilidades de ação e intervenção. Além disso, o jogo de treinamento

> pode ser utilizado para verificar se o/a aluno/a construiu ou não determinado conhecimento, servindo como um "termômetro" que medirá o real entendimento que o/a aluno/a obteve. Entretanto, com a participação do/a aluno/a nos jogos e sua necessária participação ativa, o/a professor/a poderá perceber as suas reais dificuldades, auxiliando-o a saná-las.

Foi observado que ao trabalhar com jogos, é possível encontrar motivação para transpor os paradigmas que envolvem a matemática. O uso dos jogos aliado aos materiais manipuláveis mostraram ser alternativas para despertar o interesse dos alunos potencializando a aprendizagem. A organização de um espaço para esses materiais traz praticidade às aulas, uma vez que após a confecção dos mesmos, ele estará sempre à mão para ser usado. Ao professor cabe um embasamento e uma preparação prévia definindo os objetivos que almeja atingir.

5 Considerações finais

Para melhorar seu desempenho em sala de aula o professor necessita aprimorar seus métodos e a prática de ensino. No trabalho pedagógico, os desafios surgem a todo o momento levando o professor à reflexão. Uma alternativa de trabalho é o Laboratório de Ensino da Matemática que, embora não seja a solução para os problemas relacionados ao ensino da Matemática, é certamente um caminho que pode auxiliar os professores com novas possibilidades de ação.

O professor pode também utilizar desse espaço para explorar atividades interdisciplinares, construir e elaborar materiais didáticos. Faz-se importante, para a confecção destes, a utilização de materiais recicláveis, tanto pela fácil aquisição de material como pelo universo de criatividade que pode proporcionar e também é muito importante a participação dos alunos no processo de construção. Para uma aprendizagem significativa, a teoria e a prática não podem estar separadas, a união delas se faz necessária. Sem o experimento sempre fica uma parte da teoria em que o aluno acredita que se aceita, sem se convencer ou sem confirmar a verdade dos fatos.

Foi observado que ao trabalhar com jogos, era possível encontrar motivação para transpor os modelos que envolvem a matemática. O uso dos jogos aliado aos materiais manipuláveis mostraram ser alternativas para despertar o interesse dos alunos fortalecendo a aprendizagem. A organização de um espaço para esses materiais traz praticidade às aulas, uma vez que após a confecção dos mesmos, ele estará sempre à mão para ser usado.

O laboratório, portanto, é um ambiente propício para estimular no aluno o gosto pela matemática, a perseverança na busca de soluções e a confiança em sua capacidade de aprender e fazer matemática. Pois o aluno contará com materiais, que se bem escolhidos pelo professor e

bem trabalhados lhe propiciarão um aprendizado significativo. Além de que, os alunos estarão compartilhando suas descobertas e se comunicando com os colegas e com o professor, que é seu companheiro na busca pelo conhecimento.

Segundo Lorenzato (2006, p. 11), a construção de um LEM não é objetivo para ser atingido a curto prazo, uma vez construído, ele demanda constante complementação, a qual por sua vez, exige que o professor se mantenha atualizado. Sabemos que estamos dando os primeiros passos para a implementação do LEM do Colégio Estadual Ieda Baggio Mayer, porém devemos agora, manter acesa a chama de manutenção deste, sempre alimentando-a e encorajando e preparando os professores.

Referências

ANTONIO, Fátima de Carvalho. ANDRADE, Susimeire Vivien Rosotti de. **A importância do Laboratório de Ensino Aprendizagem de Matemática**. XIII CIAEM-IACME, Recife, Brasil, 2011.

IRINEU, Jailson Fernandes. SANTOS, Priscila Geralda Claudino. RODRIGUES, Raiane. **Laboratório de ensino e suas implicações na formação inicial de professores de Matemática**. Instituto Federal de Minas Gerais/Campus São João Evangelista. 2015.

LORENZATO, Sérgio (Org.). **O Laboratório de Ensino de Matemática na formação de professores**. Campinas, SP: Autores Associados, 2006. Coleção Formação de Professores

OTTESBACH, Rosângela Cristina. PAVANELLO, Regina Maria. **Laboratório de ensino e aprendizagem da Matemática na apreciação de professores**. Universidade Estadual de Maringá. Maringá (PR). 2007.

STRAPASON, Lísie Pippi Reis. **O uso de jogos como estratégia de ensino e aprendizagem da Matemática no 1^o ano do Ensino Médio**. Centro Universitário Franciscano (UNIFRA). Santa Maria (RS), 2011.

A Transformada de Laplace para a resolução de problemas de valor inicial para EDOs

Paula Alessandra Fabricio[1]
Universidade Estadual do Oeste do Paraná
Discente do Curso de Matemática e bolsista de
Iniciação Científica do Conselho Nacional de
Desenvolvimento Científico e Tecnológico (CNPq)
ale_paulinha@hotmail.com

Daniela Maria Grande Vicente
Universidade Estadual do Oeste do Paraná
Docente do Curso de Matemática
daniela.grande@unioeste.br

Resumo: Neste trabalho apresentamos um breve estudo que tem como objetivo principal descrever a utilidade da Transformada de Laplace para encontrar soluções de equações diferenciais ordinárias (EDOs). Este método é extremamente útil para resolver problemas de valor inicial para EDOs com coeficientes constantes, desta forma nos atentaremos a estes tipos de equações.

Palavras-chave: Transformada de Laplace; Equações Diferenciais Ordinárias.

1 Introdução

As equações diferenciais modelam problemas que descrevem o comportamento do mundo físico. Problemas envolvendo o movimento de fluidos, o fluxo de corrente elétrica em circuitos, a dissipação de calor em objetos sólidos, o aumento ou a diminuição de populações, entre muitos outros, muitas vezes envolvem equações diferenciais, que são equações que contém derivadas.

Neste trabalho, trataremos de um método muito eficaz para resolver equações diferenciais usando a transformada de Laplace. Começaremos definido a transformada de Laplace e estabelecendo as condições para a sua existência. Apresentaremos alguns resultados que nos permite usar a transformada de Laplace para resolver problemas de valor inicial para equações diferenciais lineares com coeficientes constantes. O método da Transformada de Laplace permite levar a resolução de equações diferenciais à resolução de equações algébricas, que são muito mais simples de resolver, e por este fato muitas vezes é mais conveniente optar por este método.

[1]Agradecimento ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) pelo auxílio financeiro.

2 Transformada de Laplace

Sabemos que a diferenciação e a integração são transformações que basicamente transformam uma função em outra. Essas tranformações possuem a propriedade de lineridade, ou seja, a transformada de uma combinação linear de funções é uma combinação linear das transformadas. Para α e β constantes $\frac{d}{dx}[\alpha f(x) + \beta g(x)] = \alpha f'(x) + \beta g'(x)$ e $\int [\alpha f(x) + \beta g(x)]\, dx = \alpha \int f(x)\, dx + \beta \int g(x)\, dx$ desde que as derivadas e as integrais existam.

Seja uma integral definida como $\int_a^b K(s,t)f(t)\, dt$ que transforma uma função f da variável t na função F da variável s, nosso interesse está em uma transformada integral cujo intervalo de integração é $[0, \infty)$.

Se $f(t)$ estiver definida para $t \geq 0$, então a relação da forma

$$F(s) = \int_\alpha^\beta K(s,t)f(t)\, dt, \tag{1}$$

é chamada de transformada integral, tal que $K(s,t)$ é uma função dada, chamada de **núcleo** da tranformação e os limites de integração também são dados, sendo possível que $\alpha = -\infty$ ou $\beta = \infty$, ou ambos. A ideia é que a função f é tranformada em outra função F, a qual a entendemos como a transformada de f. Especificamente, vamos estudar uma determinada transformada, a qual é definida da seguinte maneira.

Definição 1. Seja f uma função definida para todo $t \geq 0$. A integral

$$\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\, dt, \tag{2}$$

será chamada **Transformada de Laplace** de f, desde que a integral convirja. Quando a integral (2) convergir temos como resultado uma função de s.

O parâmetro s é complexo, no entanto, neste texto, consideraremos somente valores para os quais s seja real. Detonaremos por $\mathcal{L}\{f(t)\}$ ou por $F(s)$ a transformada de Laplace de f, ou seja, usaremos letras minúsculas para denotar a função que está sendo transformada e a letra maiúscula correspondente para denotar sua transformada de Laplace, por exemplo,

$$\mathcal{L}\{f(t)\} = F(s) \;\; e \;\; \mathcal{L}\{y(t)\} = Y(s).$$

A transformada de Laplace usa como núcleo a função $k(s,t) = e^{-st}$. Uma vez que, as soluções das equações diferenciais lineares com coeficientes constantes estão baseadas na função exponencial, a transformada de Laplace é extremamente útil para a resolução dessas equações.

O método de resolução de uma equação diferencial usando a transformada de Laplace, consiste basicamente em três etapas: na primeira, a equação diferencial dada é transformada em

uma equação algébrica, em seguida, esta equação é resolvida por pura manipulação algébrica. Por fim, a solução da equação algébrica é transformada em sentido contrário de tal maneira que fornece a solução da equação diferencial original. Nessa última etapa "invertemos"a transformada.

A transformada de Laplace de uma determinada função f vai existir se forem satisfeitas algumas condições, ou seja, para que exista $\mathcal{L}\{f(t)\}$ é suficiente que f seja contínua por partes e de ordem exponencial para $t \geq M$. Uma função é contínua por partes em um intervalo $a \leq t \leq b$ se o intervalo puder ser dividido por um número finito de pontos $a = t_0 < t_1 < ... < t_n = b$ tal que f é contínua em cada subintervalo aberto $t_{i-1} < t < t_i$ e f tenda a um limite finito nos extremos de cada subintervalo por pontos no interior do intervalo. Ou seja, f é contínua por partes em $a \leq t \leq b$ se for contínua aí, exceto por um número finito de descontinuidades do tipo salto. O conceito de ordem exponencial está definido a seguir.

Definição 2. Dizemos que uma função f é de **ordem exponencial** a se existem constantes a, $M > 0$ e $K > 0$ tal que $|f(t)| \leq Ke^{at}$ para todo $t \geq M$.

A fim de demonstrar o teorema que estabelece as condições para a existência da transformada de Laplace, precisamos enunciar um resultado sobre a convergência de integrais impróprias que é o teorema da comparação.

Teorema 3. *Se f for contínua por partes para $t \geq a$, se $|f(t)| \leq g(t)$ quando $t \geq M$ para alguma constante positiva M e se $\int_M^\infty g(t)\,dt$ converge, então $\int_a^\infty f(t)\,dt$ também converge. Por outro lado, se $f(t) \geq g(t) \geq 0$ para $t \geq M$ e se $\int_M^\infty g(t)\,dt$ diverge, então $\int_a^\infty f(t)\,dt$ também diverge.*

Teorema 4. *Se $f(t)$ é contínua por partes no intervalo $[0, \infty)$ e de ordem exponencial a, então $\mathcal{L}\{f(t)\}$ existe para $s > a$.*

Prova. Pela propriedade aditiva de intervalos de integrais definidas, podemos escrever

$$\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)dt = \int_0^M e^{-st} f(t)dt + \int_M^\infty e^{-st} f(t)dt = I_1 + I_2.$$

A integral I_1 existe, pois pode ser escrita como uma soma de integrais sobre intervalos nos quais $e^{-st}f(t)$ é contínua. Agora, uma vez que f é de ordem exponencial, existem constantes $a, M > 0, K > 0$ de modo que $|f(t)| \leq Ke^{at}$ para $t \geq M$. Podemos então escrever

$$|e^{-st} f(t)| \leq Ke^{-st}e^{at} = Ke^{(a-s)t},$$

e assim se $\int_M^\infty e^{(a-s)t}dt$ convergir, pelo teste da comparação para integrais impróprias, Teorema (3), então I_2 converge. De fato,

$$\int_M^\infty e^{(a-s)t}\,dt = \left.\frac{e^{(a-s)t}}{a-s}\right|_M^\infty = -\frac{e^{(a-s)M}}{a-s}.$$

Portanto, I_2 existe para $s > a$. A existência de I_1 e I_2 implica que $\mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)\, dt$ existe para $s > a$. □

Vejamos alguns exemplos de transformadas de Laplace de algumas funções conhecidas.

Exemplo 5. Seja $f(t) = a$, então

$$\begin{aligned}\mathcal{L}\{a\} = \int_0^\infty e^{-st}(a)\, dt &= a \lim_{b\to\infty} \int_0^b e^{-st}\, dt \\ &= a \lim_{b\to\infty} \frac{-e^{-st}}{s}\Big|_0^b \\ &= a \lim_{b\to\infty} \frac{-e^{-sb}+1}{s} = \frac{a}{s}\end{aligned}$$

desde que $s > 0$. Em outras palavras, quando $s > 0$, então $-sb < 0$ o que garante que $e^{-sb} \to 0$ quando $b \to \infty$. A integral diverge para $s < 0$.

Vamos adotar a notação $|_0^\infty$ como abreviação de $\lim_{b\to\infty}(\ )|_0^b$, ficando subentendido que, no limite superior, $e^{-st} \to 0$ quando $t \to \infty$ para $s > 0$.

Exemplo 6. Seja $f(t) = t$, temos que $\int_0^\infty e^{-st} t\, dt$. Integrando por partes e usando $\lim_{b\to\infty} \frac{-be^{-sb}}{s} = 0$, $s > 0$, com o resultado do exemplo 5 para $a = 1$, obtemos

$$\mathcal{L}\{t\} = \frac{-te^{-st}}{s}\Big|_0^\infty + \frac{1}{s}\int_0^\infty e^{-st}\, dt = \frac{1}{s}\mathcal{L}\{1\} = \frac{1}{s}\left(\frac{1}{s}\right) = \frac{1}{s^2}.$$

Exemplo 7. Seja $f(t) = e^{at}$, com $a \in \mathbb{R}$, para todo $s > a$ temos

$$\begin{aligned}\mathcal{L}\{e^{at}\} &= \int_0^\infty e^{-st}e^{at}\, dt = \int_0^\infty e^{(a-s)t}\, dt \\ &= \frac{1}{a-s}e^{(a-s)t}\Big|_0^\infty = \frac{1}{s-a},\end{aligned}$$

desde que $a - s < 0$.

Exemplo 8. Calcule $\mathcal{L}\{sen\, 2t\}$. Temos que

$$\begin{aligned}\mathcal{L}\{sen\, 2t\} &= \int_0^\infty e^{-st}\, sen\, 2t\, dt = \frac{-e^{-st}\, sen\, 2t}{s}\Big|_0^\infty + \frac{2}{s}\int_0^\infty e^{-st}\, cos\, 2t\, dt \\ &= \frac{2}{s}\int_0^\infty e^{-st}\, cos\, 2t\, dt, \quad s > 0 \\ &= \frac{2}{s}\left[\frac{-e^{-st}\, cos\, 2t}{s}\Big|_0^\infty - \frac{2}{s}\int_0^\infty e^{-st}\, sen\, 2t\, dt\right] \\ &= \frac{2}{s^2} - \frac{4}{s^2}\,\mathcal{L}\{sen\, 2t\}.\end{aligned}$$

Temos uma equação em $\mathcal{L}\{sen\, 2t\}$ que aparece nos dois lados da igualdade. Resolvendo essa equação obtemos

$$\mathcal{L}\{sen\, 2t\} = \frac{2}{s^2+4}, \quad s > 0.$$

Teorema 9. *($\mathcal{L}$ é uma transformação linear) Suponha que as funções f e g cujas transformadas de Laplace existem para $s > a_1$ e $s > a_2$, respectivamente. Então, para s maior do que o máximo de a_1 e a_2, e α e β números reais quaisquer,*

$$\mathcal{L}\{\alpha f(t) + \beta g(t)\} = \alpha\mathcal{L}\{f(t)\} + \beta\mathcal{L}\{g(t)\}$$

Prova. Pela definição,

$$\begin{aligned}\mathcal{L}\{\alpha f(t) + \beta g(t)\} &= \int_0^\infty e^{-st}[\alpha f(t) + \beta g(t)]\,dt \\ &= \alpha\int_0^\infty e^{-st}f(t)\,dt + \beta\int_0^\infty e^{-st}g(t)\,dt \\ &= \alpha\mathcal{L}\{f(t)\} + \beta\mathcal{L}\{g(t)\}.\end{aligned}$$

□

O seguinte teorema, retirado do Boyce (2012), explicita a relação entre a transformada de f' e a transformada da f, e estabelece uma maneira de usar a transformada de Laplace para resolver problemas de valor inicial para equações diferenciais lineares com coeficientes constantes.

Teorema 10. (Transformada de uma derivada). *Suponha que f é contínua e que f' é contínua por partes em qualquer intervalo $0 \le t \le A$. Suponha, além disso, que f é de ordem exponencial, ou seja, que existem constantes K, a e M tais que $|f(t)| \le Ke^{at}$ para $t \ge M$. Então $\mathcal{L}\{f(t)\}$ existe para $s > a$ e, além disso,*

$$\mathcal{L}\{f'(t)\} = s\,\mathcal{L}\{f(t)\} - f(0).$$

Prova. Consideremos a integral

$$\int_0^A e^{-st}\,f'(t)\,dt, \tag{3}$$

caso f' tenha pontos de descontinuidade no intervalo $0 \le t \le A$, vamos denotá-la por $t_1, t_2, ..., t_n$, então de (3) temos

$$\int_0^A e^{-st}\,f'(t)\,dt = \int_0^{t_1} e^{-st}\,f'(t)\,dt + \int_{t_1}^{t_2} e^{-st}\,f'(t)\,dt + ... + \int_{t_n}^{A} e^{-st}\,f'(t)\,dt.$$

Integrando por partes cada parcela à direita do sinal de igualdade, obtemos

$$\begin{aligned}\int_0^A e^{-st}\,f'(t)\,dt &= e^{-st}f(t)\,|_0^{t_1} + e^{-st}f(t)\,|_{t_1}^{t_2} + ... + e^{-st}f(t)\,|_{t_n}^{A} \\ &+ s\left[\int_0^{t_1} e^{-st}\,f(t)\,dt + \int_{t_1}^{t_2} e^{-st}\,f(t)\,dt + ... + \int_{t_n}^{A} e^{-st}\,f(t)\,dt\right] \\ &= e^{-st_1}\,f(t_1) - f(0) + e^{-st_2}f(t_2) - e^{-st_1}f(t_1) + ... \\ &+ e^{-sA}f(A) - e^{-st_n}f(t_n) + s\int_0^A e^{-st}\,f(t)\,dt\,.\end{aligned}$$

Como f é contínua, as parcelas em $t_1, t_2, ..., t_n$ se cancelam,

$$\int_0^A e^{-st} f'(t)\, dt \,=\, e^{-sA} f(A) - f(0) \,+\, s \int_0^A e^{-st} f(t)\, dt\,.$$

Para $A \geq M$, temos $|f(A)| \leq Ke^{aA}$, consequentemente, $|e^{-sA} f(A)| \leq Ke^{-(s-a)A}$. Portanto, $e^{-sA} f(A) \to 0$ quando $A \to \infty$, sempre que $s > a$. Logo para $s > a$,

$$\mathcal{L}\{f'(t)\} \,=\, s\mathcal{L}\{f(t)\} \,-\, f(0)$$

e segue a prova do teorema. □

Se f' e f'' satisfazem as mesmas condições f e f' no teorema (10), então para $s > a$

$$\begin{aligned} \mathcal{L}\{f''(t)\} &= s\mathcal{L}\{f'(t)\} - f'(0) \\ &= s[s\mathcal{L}\{f(t)\} - f(0)] - f'(0), \end{aligned}$$

isto é,

$$\mathcal{L}\{f''(t)\} = s^2\mathcal{L}\{f(t)\} - sf(0) - f'(0).$$

Semelhantemente,

$$\mathcal{L}\{f'''(t)\} = s^3\mathcal{L}\{f(t)\} - s^2 f(0) - sf'(0) - f''(0).$$

Desde que f e suas derivadas satisfaçam condições adequadas, podemos obter a *n-ésima* derivada $f^{(n)}$ e o resultado é dado pelo corolário a seguir.

Corolário 11. *Seja $n \in \mathbb{N}$. Se $f, f', ..., f^{(n-1)}$ forem contínuas em $[0, \infty)$ e de ordem exponencial, e se $f^{(n)}(t)$ for contínua por partes em $[0, \infty)$, então $\mathcal{L}\{f^{(n)}(t)\}$ existe para $s > a$ e é dado por*

$$\mathcal{L}\{f^{(n)}(t)\} \,=\, s^n F(s) \,-\, s^{(n-1)} f(0) \,-\, s^{(n-2)} f'(0) \,-\, ... \,-\, f^{(n-1)}(0),$$

onde $F(s) = \mathcal{L}\{f(t)\}$.

Vamos ilustrar com os exemplos seguintes como a transformada de Laplace pode ser usada para resolver problemas de valor inicial. Nestes exemplos as equações envolvidas são equações diferenciais homogêneas com coeficientes constantes, porém a transformada de Laplace também é extremamente útil para resolver problemas que envolvem equações diferencias não homogêneas. Também vamos supor que a solução $y = \phi(t)$ satisfaz as condições do corolário 11.

Exemplo 12. Vamos considerar a equação diferencial

$$y'' - y' - 6y = 0 \tag{4}$$

com condições iniciais

$$y(0) = 1 \quad e \quad y'(0) = -1 \tag{5}$$

Calculando a transformada de Laplace da equação diferencial, obtemos

$$\mathcal{L}\{y''\} - \mathcal{L}\{y'\} - 6\mathcal{L}\{y\} = 0, \tag{6}$$

onde usamos a linearidade da transformada de Laplace para escrever a soma das transformadas separadamente. Usando o Corolário 11 para expressar $\mathcal{L}\{y''\}$ e $\mathcal{L}\{y'\}$ em função de $\mathcal{L}\{y\}$ a eq.(6) fica

$$s^2\mathcal{L}\{y\} - sy(0) - y'(0) - [s\mathcal{L}\{y\} - y(0)] - 6\mathcal{L}\{y\} = 0.$$

Tomando $\mathcal{L}\{y\} = Y(s)$,

$$s^2Y(s) - sy(0) - y'(0) - [sY(s) - y(0)] - 6Y(s) = 0$$

$$Y(s)(s^2 - s - 6) + (1 - s)y(0) - y'(0) = 0 \tag{7}$$

Substituindo os valores de $y(0)$ e $y'(0)$ dados pelas condições iniciais (5) na eq.(7) e depois resolvendo para $Y(s)$,

$$Y(s)(s^2 - s - 6) + (1 - s) + 1 = 0$$

$$Y(s) = \frac{s-2}{(s^2 - s - 6)} = \frac{s-2}{(s-3)(s+2)}. \tag{8}$$

A eq.(8) expressa a transformada de Laplace $Y(s)$ da solução $y = \phi(t)$ do problema de valor inicial, e para determinar a função ϕ precisamos encontrar a função cuja transformada de Laplace é $Y(s)$.

Expandindo a expressão a direita do sinal de igualdade na eq.(8) em frações parciais obtemos:

$$Y(s) = \frac{s-2}{(s^2 - s - 6)} = \frac{a}{(s-3)} + \frac{b}{(s+2)} = \frac{a(s+2) + b(s-3)}{(s-3)(s+2)}. \tag{9}$$

Igualando os numeradores em (9), temos

$$s - 2 = a(s+2) + b(s-3),$$

que deve satisfazer para todos os valores de s. Em particular se $s = 3$, temos $a = 1/5$, analogamente se $s = -2$, então $b = 4/5$.

Substituindo os valores para a e b em (9), obtemos

$$Y(s) = \frac{1/5}{(s-3)} + \frac{4/5}{(s+2)}. \tag{10}$$

Usando o resultado do exemplo 7, temos que $\frac{1}{5}e^{3t}$ tem transformada $\frac{1}{5}(s-3)^{-1}$, analogamente a transformada de Laplace de $\frac{4}{5}e^{-2t}$ é $\frac{4}{5}(s+2)^{-1}$ portanto pela linearidade da transformada de Laplace,

$$y = \phi(t) = \frac{1}{5}e^{3t} + \frac{4}{5}e^{-2t}$$

tem transformada (10) e é a solução do problema de valor inicial.

2.1 Transformada de Laplace Inversa

Definição 13. Se $F(s)$ representa a transformada de Laplace de uma função $f(t)$, isto é, $\mathcal{L}\{f(t)\} = F(s)$, dizemos então que $f(t)$ é a **transformada inversa de Laplace** de $F(s)$ e escrevemos $f(t) = \mathcal{L}^{-1}\{F(s)\}$.

Notemos que $\mathcal{L}^{-1}$ também é uma transformação linear, ou seja, para as constantes α e β,

$$\mathcal{L}^{-1}\{\alpha F(t) + \beta G(t)\} = \alpha\mathcal{L}^{-1}\{F(s)\} + \beta\mathcal{L}^{-1}\{G(s)\},$$

onde F e G são as transformadas das funções f e g.

Exemplo 14. Considere a equação diferencial

$$y" + 5y' + 6y = 0,$$

com condições iniciais

$$y(0) = 1 \quad e \quad y'(0) = 0.$$

Calculando a sua transformada, obtemos

$$\mathcal{L}\{y"\} + 5\mathcal{L}\{y'\} + 6\mathcal{L}\{y\} = 0$$
$$s^2Y(s) - sy(0) - y'(0) - 5[sY(s) - y(0)] + 6Y(s) = 0.$$

Utilizando as condições iniciais e resolvendo para $Y(s)$ temos

$$Y(s) = \frac{s-5}{s^2-5s+6} = \frac{s-5}{(s-2)(s-3)}$$

e usando frações parciais para resolver a equação acima encontramos

$$\begin{aligned}\frac{s-5}{(s-2)(s-3)} &= \frac{A}{(s-2)}+\frac{B}{(s-3)}\\ &= \frac{A(s-3)+B(s-2)}{(s-2)(s-3)}.\end{aligned}$$

Fazendo $s=3$, obtemos que $B=-2$ e com $s=2$ temos $A=3$, assim

$$\frac{3}{(s-2)}-\frac{2}{(s-3)}=\frac{s-5}{(s-2)(s-3)}.$$

Da linearidade da transformada inversa e utilizando o resultado do exemplo (7)

$$\begin{aligned}\mathcal{L}^{-1}\left\{\frac{s-5}{(s-2)(s-3)}\right\} &= 3\mathcal{L}^{-1}\left\{\frac{1}{(s-2)}\right\}-2\mathcal{L}^{-1}\left\{\frac{1}{(s-3)}\right\}\\ y &= \phi(t)=3\mathcal{L}^{-1}\left\{\frac{1}{(s-2)}\right\}-2\mathcal{L}^{-1}\left\{\frac{1}{(s-3)}\right\}\\ y &= \phi(t)=3e^{2t}-2e^{3t}\end{aligned}$$

que é a solução do problema de valor inicial.

Referências

BOYCE, E. W. ; DIPRIMA, C. R. Equações Diferenciais Elementares e Problemas de Valores de Contorno. 9ª ed. Rio de Janeiro: LTC, 2012.

G.ZILL, Dennis; CULLEN, Michael R. Equações diferenciais. 3ª ed. Sao Paulo: Eugenia Pessotti, 2001. 473 p. Tradução de Antonio Zumpano, revisão tecnica: Antonio Pertence Jr.

Um modelo de programação linear mista aplicada a um problema de transportes

Amarildo de Vicente
Universidade Estadual do Oeste do Paraná - UNIOSTE
amarildo.vicente@gmail.com

André Wilson de Vicente
Universidade Estadual do Oeste do Paraná - UNIOSTE
madarasenju63@gmail.com

Rogério Luiz Rizzi
Universidade Estadual do Oeste do Paraná - UNIOSTE
rogeriorizzi@hotmail.com

Cláudia Brandelero Rizzi
Universidade Estadual do Oeste do Paraná - UNIOSTE
claudia_rizzi@hotmail.com

Resumo: Este trabalho apresenta um problema relacionado a um sistema de transportes, que é objeto de estudo de uma empresa de informatização. O problema consiste no planejamento de entrega de dinheiro a uma rede de agências bancárias, por meio de uma empresa de transporte de valores, e tem por objetivos diminuir custos com esta tarefa de entrega bem como reduzir riscos provenientes da estocagem do dinheiro. A fim de apresentar uma solução inicial para o sistema pretendido foi criado um modelo de programação linear mista e sua implementação foi feita por meio da linguagem utilizada pelo software livre Glpk. Testes diversos foram realizados através de simulações com dados fictícios e as soluções obtidas em cada caso satisfizeram plenamente as expectativas geradas.

Palavras-chave: Minimização; programação linear mista; problema de transporte.

1 Introdução

O problema aqui tratado surgiu da necessidade de uma empresa de informatização ao implementar um sistema de transporte de dinheiro da sede de uma transportadora de valores para diversas agências de uma rede bancária. A empresa quer viabilizar uma possível redução com estes transportes e também assegurar uma certa cautela com a guarda de dinheiro nos cofres, o que traz sérios riscos ao patrimônio do seu cliente. Na modelagem do problema, que será descrita na seção 3, houve a necessidade de trabalhar com variáveis binárias a fim de fazer manipulações em algumas restrições. Esta técnica será descrita brevemente na seção 2. O modelo apresentado neste trabalho é um protótipo inicial que precisa ser aperfeiçoado com dados e informações ainda não disponibilizadas. Todavia, para as simulações realizadas os resultados obtidos foram bastante satisfatórios em relação às expectativas iniciais.

2 Embasamento Teórico e Técnicas de Modelagem

Como se sabe da literatura de otimização, ver por exemplo Bradley et al. (1977), Goldbarg e Luna (2005), todo problema de programação linear (PPL) pode ser escrito na forma

Maximizar $z = \sum_{j=1}^{n} c_j x_j$ (01)

sujeita a:

$g_1(x_1, x_2, \ldots, x_n) \leq b_1$,

$g_2(x_1, x_2, \ldots, x_n) \leq b_2$,

...

$g_m(x_1, x_2, \ldots, x_n) \leq b_m$,

$x_j \geq 0,\ j = 1, 2, \ldots, n$,

sendo que $g_1, g_2, \ldots, g_m$ são funções lineares de $x_1, x_2, \ldots, x_n$.

Quando as variáveis do problema são todas inteiras ele é chamado de problema de programação linear inteira (PPLI) e quando o problema contém tanto variáveis contínuas quanto inteiras ele é chamado de problema de programação linear mista (PPLM). Uma variedade muito grande de problemas se enquadram nesta última categoria e muitos deles são aplicações à área de transportes, como pode ser visto por exemplo em Goldbarg e Luna (2005). É nesta área que o problema aqui proposto se aplica.

Dentre as variáveis inteiras destacam-se aquelas que assumem apenas os valores 0 ou 1, chamadas de variáveis binárias. Elas são empregadas na composição de modelos clássicos como o Problema do Caixeiro Viajante, o Problema do Carteiro Chinês, problemas de fluxo em rede, etc., e também são bastante úteis na manipulação de restrições, que será descrita na sequência.

Na resolução de um PPL, a menos que algo seja dito em contrário, pressupõe-se que todas as restrições devam ser atendidas. De um ponto de vista da lógica matemática o conjunto de restrições em (01) forma uma proposição composta conjuntiva da forma " $g_1(x_1, x_2, \ldots, x_n) \leq b_1$ e $g_2(x_1, x_2, \ldots, x_n) \leq b_2$ e ... ", É este fato que faz com que as restrições determinem uma região viável convexa. Porém, há situações em que a satisfação de uma restrição é condicionada à satisfação de outra. Em outras palavras, às vezes existe a necessidade de que, sendo uma restrição atendida, outra precisa deixar de ser atuante. No modelo (01), poderia ser de interesse do programador que apenas uma das duas primeiras restrições fosse atendida, o que poderia ser escrito do ponto de vista da lógica como $g_1(x_1, x_2, \ldots, x_n) \leq b_1$ ou $g_2(x_1, x_2, \ldots, x_n) \leq b_2$, onde o conectivo "ou" é exclusivo. A fim de realizar esta tarefa torna-se necessário fazer uma manipulação com as restrições, e é neste ponto que as variáveis binárias desempenham um papel fundamental. Como ilustração consideremos o PPL a seguir

Maximizar $Z = x + y$ (02)

sujeita a

$x + y \leq 5$ (a)

$2x + 3y \leq 12$ (b)

$x \geq 0$, $y \geq 0$.

A região viável neste caso, denominada R, está apresentada na Figura 1.

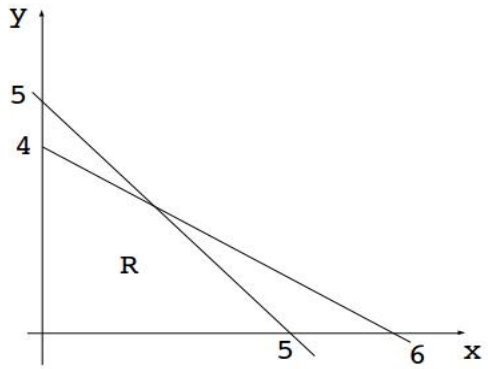


Figura 1: Região viável para o modelo (02)

Vamos supor que, por algum motivo, deseja-se que apenas uma das duas restrições seja atendida, ou seja, que $x + y \leq 5$ ou $2x + 3y \leq 12$, onde o conectivo "ou" é do tipo exclusivo. A seguir o modelo anterior será reproduzido com um novo formato, no qual figura uma constante M cujo papel será explicado na sequência.

Maximizar $Z = x + y$ (03)

sujeita a

$x + y \leq 5 + Mt$, (c)

$2x + 3y \leq 12 + M(1 - t)$, (d)

$x \geq 0$, $y \geq 0$, $t \in \{0, 1\}$.

No último modelo, quando $t = 1$ a restrição (d) se torna $2x + 3y \leq 12$, que é a restrição (b) do modelo (02) já que, neste caso, $1 - t = 0$. Já a restrição (c) terá no seu lado direito o acréscimo do valor M, que deverá ser positivo e grande o suficiente para que esta restrição seja sempre atendida não importando os valores de x e y que atendem à restrição (d). Por outro lado, quando $t = 0$ a restrição (c) se torna $x + y \leq 5$, ao passo que (d) passa a ter o acréscimo do valor M do seu lado direito. Novamente, este valor de M deve ser positivo e grande o suficiente para que esta restrição seja sempre atendida não importando os valores de x e de y que atendem à

restrição (c). Este tipo de artifício converte um conjunto de duas restrições da forma conjuntiva, quando as duas precisam ser atendidas, para a forma disjuntiva exclusiva, quando apenas uma deve ser atendida. Por exemplo, sendo $M = 3$ e $t = 1$ a região viável de (03) é representada pelo triângulo da Figura 2. Neste caso a restrição (c) fica desabilitada, ou seja, não tem mais utilidade.

Outro recurso que pode ser utilizado diz respeito ao controle de custos fixos, que será útil no modelo apresentado neste trabalho. Consideremos o PPLM a seguir, onde c_1 e c_2 representam custos fixos, sendo que c_1 só ocorre se $x > 0$ e c_2 só ocorre se $y > 0$.

Minimizar $Z = c_1t_1 + c_2t_2$

Sujeita a

$x \leq 10t_1$,

$y \leq 12t_2$,

$2x + y \geq 4$,

$x \geq 0$, $y \geq 0$, $t_1, t_2, \in \{0, 1\}$.

Neste modelo, ocorrendo por exemplo $t_1 = 0$ e $t_2 = 1$, obrigatoriamente deverá ocorrer $x = 0$ enquanto y poderá assumir qualquer valor de 4 a 12. Nesta situação apenas o custo c_2 será computado em Z. Há muito outros artifícios que permitem manipular restrições, como restrições de múltiplas escolhas, restrições condicionais, entre outras. Maiores detalhes podem ser vistos em Vicente (2012).

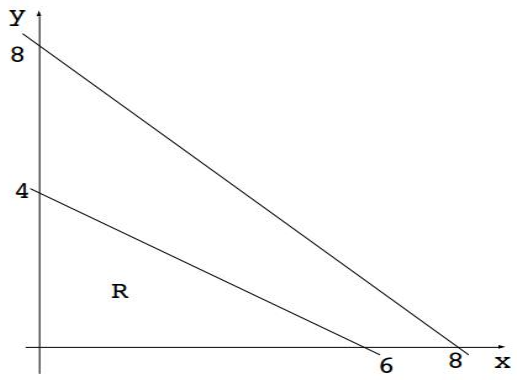


Figura 2: Região viável para o modelo (03)

3 Descrição e Modelagem do Problema

O problema apresentado neste trabalho originou-se na necessidade de aperfeiçoar o sistema de abastecimento feito por uma empresa de informatização, destinado ao atendimento de uma rede bancária. Um veículo de uma transportadora de valores deve deslocar periodicamente de sua sede até às agências para fazer a reposição de dinheiro. Estas entregas possuem um custo que o banco deseja minimizar. Os custos podem ser divididos em duas categorias: a primeira é um custo fixo ocorrido em cada viagem feita pela transportadora; a segunda é uma espécie de custo de translado feito de uma agência para outra e que só ocorre se houver mais de uma entrega na mesma viagem. Não está prevista aqui a escolha da melhor rota em cada situação pois isto já é conhecido pela transportadora. Por questões de segurança o banco estipulou um limite máximo de dinheiro que pode permanecer em cada agência durante a noite. Este montante varia de uma agência para outra conforme o grau de risco. Além disso, a fim de garantir a funcionalidade do banco o responsável quer que um limite mínimo de dinheiro seja mantido em cada agência, limite que também varia de uma agência para outra conforme a movimentação. Esta movimentação se refere à retirada média de dinheiro feita diariamente por clientes. O controle de estoque é feito diariamente e sempre que o limite mínimo for atingido ao final do dia uma nova remessa deverá envida no dia seguinte. Não é viável para o banco que uma remessa inferior a uma quantia K estipulada seja feita. A ideia é que o maior número possível de agências sejam abastecidas em um mesmo dia, aproveitando uma mesma viagem, o que corresponde a diminuir o total de viagens necessárias.

O modelo para atender ao que foi exposto foi formulado como segue.

Constantes

m: Número de agências a serem atendidas pela transportadora.

Cv: Custo fixo por viagem realizada.

Ct: Custo de translado de uma agência para outra (entregas secundárias).

K: Quantidade mínima transportada em uma viagem.

D: Total de dias estipulado para o planejamento.

$Lmin_i$: Limite mínimo em estoque na agência i.

$Lmax_i$: Limite máximo em estoque na agência i.

R_i: Retiradas diárias, em média, ocorridas na agência i.

M_1, M_2, M_3: Constantes para manipulação de restrições.

Variáveis

x_{ij}: Quantidade transportada para a empresa i no j-ésimo dia.

E_{ij}: Estoque existente na empresa i no j-ésimo dia.

y_{ij}: Variável binária que assume valor 1 se for transportado dinheiro para a agência i no j-ésimo dia e 0 em caso contrário.

z_j: Variável binária que assume valor 1 se ocorrer alguma entrega no j-ésimo dia e 0 em caso contrário.

t_j: Variável binária que assume valor 1 se ocorrer mais de uma entrega no j-ésimo dia e 0 em caso contrário.

w_j: Total de translados, ou entregas secundárias, ocorridas no j-ésimo dia.

C: Custo total no período estipulado.

Em todas as constates e variáveis indexadas o índice i varia de 1 até m e j varia de 1 até D.

Modelo

Minimizar $C = \sum_{j=1}^{D} Cv * z_j + \sum_{j=1}^{D} Ct * wj$

sujeita a

1. $Lmin_i \leq E_{ij} \leq Lmax_i$,
2. $E_{i1} = x_{i1} - R_i$,
3. $E_{ij} = x_{ij} - R_i + E_{ij-1}$,
4. $x_{ij} \leq M_1 y_{ij}$
5. $y_{ij} \leq z_j$,
6. $x_{ij} \geq K y_{ij}$,
7. $\sum_{i=1}^{m} y_{i,j} \leq 1 + M_2 t_j$,
8. $w_j \geq \sum_{i=1}^{m} y_{ij} - 1 - M_3(1 - t_j)$,

$w_j \geq 0$, $x_{ij} \geq 0$, $E_{ij} \geq 0$, y_{ij}, t_j e $z_j \in \{0, 1\}$, $i = 1, ..., m$, $j = 1, ..., D$.

As restrições que compõem o modelo têm as finalidades descritas a seguir.

1, 2 e 3: Dizem respeito às limitações e ao controle de estoque.

4 e 5: São restrições lógicas e servem para controlar os dias em que houve entregas e quais agências receberam dinheiro nestes dias.

6: Restrição lógica que serve para controlar uma quantidade mínima de dinheiro transportado para uma determinada agência.

7 e 8: São restrições lógicas de implicação e sevem para controlar se houve entregas secundárias em um mesmo dia, ou seja, se houve mais de uma entrega em uma mesma viagem.

As constantes M_1, M_2 e M_3 que aparecem no modelo devem ser positivas e "grandes", conforme exposto da Seção 2.

4 Validação do Modelo

O modelo proposto foi implementado por meio do software livre Glpk. A fim de avaliar seu desempenho foram feitas simulações com dados fictícios, já que dados reais não estão à disposição dos autores deste trabalho. Como ilustração, consideremos uma situação em que o planejamento deve ser feito para um período de dez dias, $D = 10$, entre envolvendo quatro agências, $m = 4$. Vamos supor que $Cv = 2000$, $Ct = 500$ e que $K = 5000$. Os demais dados se encontram na Tabela 1 .

Tabela 1: Dados para simulação do modelo

Agência	$Lmin_i$	$Lmax_i$	R_i
1	20000	80000	18000
2	15000	60000	14000
3	12000	50000	10000
4	30000	80000	28000

De acordo com a execução do programa as remessas devem ser efetuadas conforme os dados apresentados na Tabela 2. O custo total, nesta simulação, deve ser $C = 12500$.

Tabela 2: Resultado obtidos para os dados simulados

Agência	1	2	3	4	5	6	7	8	9	10
1	980000	0	0	48000	0	0	0	54000	0	0
2	740000	0	0	42000	0	0	39000	0	0	0
3	60000	0	0	30000	0	0	22000	0	0	0
4	118000	0	0	84000	0	0	24000	84000	0	0

5 Análise dos Resultados e Conclusões

Apesar de o modelo proposto apresentar variáveis binárias, que normalmente gera dificuldades no processamento, para as simulações feitas as respostas foram atingidas em tempos muito pequenos, plenamente viáveis. Tendo em vista a indisponibilidade de dados reais para testes diversas simulações com dados fictícios foram feitas e em todos os casos os resultados alcançados atingiram as expectativas. Apesar de ter apresentado um bom funcionamento o modelo carece de aperfeiçoamentos, o que só será possível após a realização de testes com dados reais e com a inserção de novas informações que o aproxime melhor do problema real.

Referências

BRADLEY, Stephen, P.; HAX, Arnoldo C.; MAGNANTI, Thomas L. **Applied Mathematical Programming**. Addison-Wesley Publishing Company, 1977. 735 p.

GOLDBARG, Marco Cesar; LUNA, Henrique Pacca, L. **Otimização Combinatória e Programação Linear**: modelos e algoritmos. 2a. Ed., Rio de Janeiro: Elsevier, 2005. 518 p.

VICENTE, A. **Utilização de Variáveis Binárias para Manipulação de Restrições em Programação Linear Mista**. In: SANTOS, R. F., SIQUEIRA, J. A. C. Fontes Renováveis: Agroenergia, v.1., Cascavel, Edunioeste, 2012, p. 153-181.

Métodos de busca linear com direções de Newton e BFGS para problemas de otimização irrestrita

Jesus Marcos Camargo
Universidade Estadual do Oeste do Paraná - UNIOESTE.
marco.s.camargo@hotmail.com

Resumo: Este trabalho é o resultado de uma revisão bibliográfica a respeito do método de Newton e do método BFGS com busca linear para resolução de problemas de otimização irrestrita.

Palavras-chave: Otimização irrestrita; Método de Newton; Método BFGS.

1 Introdução

As ideias apresentadas nesse trabalho são baseadas nas obras de Martínez e Santos (1998), Nocedal e Wright (2006), Luenberger e Ye (2008) e Bertsekas (1999) e foram utilizadas como suporte teórico para a dissertação de mestrado Camargo (2015).

São apresentados aqui alguns resultados a cerca da otimização irrestrita, o método de Newton e o método de quase-Newton conhecido como BFGS, ambos com busca linear. O trabalho traz os algoritmos de ambos os métodos e a demonstração da convergência global dos algoritmos de busca linear com *backtracking*.

2 Otimização Irrestrita

O estudo de otimização é constituído pelo desenvolvimento de técnicas para localizar em um determinado conjunto Ω um ponto x^* que torne o valor funcional mínimo ou máximo. Usualmente tratamos apenas da minimização de funções, uma vez que maximizar uma função f é equivalente a minimizar $-f$. Neste trabalho será apresentada a teoria de otimização de funções em que $\Omega = \mathbb{R}^n$, conhecida como otimização irrestrita. O conteúdo aqui apresentado é resultado do estudo de Martínez e Santos (1998) e Nocedal e Wright (2006), incluindo definições e demonstrações.

Para o desenvolvimento de tais teorias consideramos que sempre é possível determinar um mínimo local para a função objetivo e que tal função é de classe C^1.

3 Busca Linear

Ao estabelecer um algoritmo para minimização irrestrita de uma função, estamos, geralmente, desenvolvendo um processo iterativo de modo que dado um certo ponto x_n, onde $\nabla f(x_n) \neq 0$, seja possível determinar um ponto x_{n+1}, tal que $f(x_{n+1}) < f(x_n)$.

Na busca linear, dado x_k procuramos o próximo ponto em uma direção d_k, que seja *direção de descida*, isto é, existe um certo $\varepsilon > 0$ tal que para todo $t \in (0, \varepsilon]$ temos $f(x_k + td_k) < f(x_k)$. Podemos ainda caracterizar tais direções utilizando o gradiente da função, conforme o lema a seguir.

Lema 1. *Se $\nabla f(x)^T d < 0$ então para todo $t > 0$ suficientemente pequeno, $f(x + td) < f(x)$.*

Prova. Observe que $\nabla f(x)^T d = \lim_{t\to 0} \dfrac{f(x+td) - f(x)}{t}$. Como na hipótese temos que $\nabla f(x)^T d < 0$ então segue que para todo $t > 0$ suficientemente pequeno $f(x + td) < f(x)$. □

Observe que, usando a caracterização dada pelo lema anterior temos que $-\nabla f(x)$ é uma direção de descida (chamada *direção de máxima descida*). Deste modo é assegurado que dado um ponto x_k que não anula o gradiente, sempre é possível obter uma direção de descida, lembrando que estamos assumindo que a função é de classe C^1.

Infelizmente pedir somente que $f(x+td) < f(x)$ não é suficiente para garantir que depois de um determinado número de iterações atinja-se um ponto estacionário, isso porque não estamos impondo um controle sobre o tamanho do passo definido por t. Perceba que se tomarmos para t escolhas ruins isso pode gerar uma convergência muito lenta, ou ainda, o método pode ficar preso em um ponto não estacionário. Assim precisamos fazer escolhas inteligentes, de modo que t forneça um decréscimo suficiente. Para realizar essa escolha temos a chamada condição de Armijo que estabelece uma relação entre o tamanho do passo e a norma do vetor gradiente, esta condição é apresentada pelo teorema a seguir.

Teorema 2 (Condição de Armijo)**.** *Sejam $x, d \in \mathbb{R}^n$ tais que $\nabla f(x) \neq 0$, $\nabla f(x)^T d < 0$ e $\alpha \in (0, 1)$. Existe um $\varepsilon(\alpha) > 0$ tal que*

$$f(x + td) \leq f(x) + \alpha t \nabla f(x)^T d \qquad \forall t \in (0, \varepsilon]\,. \tag{1}$$

Prova. Como $\nabla f(x) \neq 0$ então temos que

$$\lim_{t\to 0} \frac{f(x+td) - f(x)}{t} = \nabla f(x)^T d \neq 0.$$

Deste modo podemos escrever

$$\lim_{t\to 0} \frac{f(x+td) - f(x)}{t\nabla f(x)^T d} = 1.$$

Assim existe um certo $\varepsilon > 0$ tal que $\forall t \in (0, \varepsilon]$ tem-se

$$\frac{f(x+td) - f(x)}{t\nabla f(x)^T d} \geq \alpha,$$

e portanto $f(x+td) \leq f(x) + \alpha t \nabla f(x)^T d$. □

Ainda que tenhamos agora determinado um controle sobre o tamanho do passo, esse não é completamente efetivo uma vez que a direção d também pode se tornar demasiadamente pequena, fazendo com que a sequência gerada na busca linear convirja novamente para um ponto não estacionário, este tipo de situação é contornada por meio da *condição* β que estabelece um controle sobre a norma de d_k a cada passo.

Dado um parâmetro $\beta > 0$ estabelecemos que a cada iteração a direção d_k satisfaça

$$\|d_k\| \geq \beta \|\nabla f(x_k)\|. \tag{2}$$

Assim, controlamos a norma de d_k em função do gradiente da função no ponto x_k.

Essas duas condições são suficientes para garantir o controle dos passos dados em cada iteração. Como $f(x+td) < f(x)$ não era uma condição suficiente, também não é suficiente pedir somente que $\nabla f(x)^T d < 0$. De fato, em algumas situação as direções d_k escolhidas em cada iteração podem gerar uma sequência convergindo para uma certa direção d ortogonal a $\nabla f(x)$. Como não desejamos que isso aconteça, vamos impor a *condição do ângulo*.

Sabemos da geometria analítica que dados dois vetores $\nabla f(x_k)$ e d_k vale a expressão

$$\frac{\nabla f(x_k)^T d_k}{\|\nabla f(x_k)\| \|d_k\|} = \cos\phi$$

onde ϕ representa o ângulo formado entre tais vetores. Nesse sentido como desejamos afastar as direções d_k de direções ortogonais ao gradiente no ponto x_k, impomos por meio de um parâmetro $\theta \in (0,1)$ que a cada iteração, a direção usada satisfaça

$$f(x_k)^T d_k \leq -\theta \|\nabla f(x_k)\| \|d_k\|. \tag{3}$$

Antes de estabelecermos um algoritmo conceitual, vamos destacar que embora tenhamos definido o decréscimo suficiente para a função f pela equação 1, não foi estabelecido como fazer a escolha de um t apropriado. Entre os métodos que podem ser escolhidos, optamos por usar *backtracking*, que consiste em, a cada passo, escolher o maior $t \in \{\frac{1}{2}, \frac{1}{4}, \cdots\}$ satisfazendo a condição de Armijo.

Algoritmo de busca linear com *backtracking*: Defina $x_0 \in \mathbb{R}^n, \alpha \in (0,1), \beta > 0$ e $\theta \in (0,1)$. Dado x_k pela k-ésima iteração, determine x_{k+1} da seguinte forma:

1. Se $\nabla f(x_k) = 0$, pare!

2. Escolha $d_k \in \mathbb{R}^n$ tal que

$$\|d_k\| \geq \beta \|\nabla f(x_k)\|$$

$$\nabla f(x_k)^T d_k \leq -\theta \|\nabla f(x_k)\| \|d_k\|;$$

3. Defina $t = 1$;

4. Enquanto $f(x_k + td_k) > f(x_k) + \alpha t \nabla f(x_k)^T d_k$ tome $t \leftarrow \bar{t} \in [0, 1t, 0.9t]$;

5. Faça $x_{k+1} = x_k + td_k$.

Na prática escolhemos valores bem pequenos para α, em geral $\alpha = 10^{-4}$ (NOCEDAL; WRIGHT, 2006), e $\theta = 10^{-6}$ (MARTÍNEZ; SANTOS, 1998), entretanto a escolha de β é um tanto mais complexa e deve ser avaliada conforme o problema empregado, é sugerido que β assuma o inverso de uma cota superior para a norma da matriz hessiana, pois isso não inibe a aceitação das direções de Newton, que serão esclarecidas posteriormente.

Com base no desenvolvimento dado pelo texto acima, é fácil verificar que o algoritmo está bem definido, além disso é possível provar sua convergência global.

Teorema 3 (Convergência global)**.** *Se x^* é ponto limite de uma sequência gerado pelo algoritmo dado, então $\nabla f(x^*) = 0$.*

Prova. Seja $s_k = x_{k+1} - x_k = td_k \ \forall k \in \mathbb{N}$. Seja K_1 um subconjunto infinito de $\mathbb{N}$ tal que $\lim\limits_{k \in K_1} x_k = x^*$, vamos dividir a demonstração em dois casos:

1. $\lim\limits_{k \in K_1} \|s_k\| = 0$;
2. Existe K_2 subconjunto de K_1 e $\varepsilon > 0$ tal que para todo $k \in K_2$ tem-se $\|s_k\| \geq \varepsilon$.

Supondo que vale o item (1) temos ainda duas outras possibilidades:

1.1 Se existe uma subsequência de s_k, com índices em $K_3 \subset K_1$ tal que $s_k = d_k$, então

$$\|\nabla f(x^*)\| = \lim_{k \in K_3} \|\nabla f(x_k)\| \leq \lim_{k \in K_3} \frac{\|d_k\|}{\beta} = \lim_{k \in K_3} \frac{\|s_k\|}{\beta} = 0;$$

1.2 Se para todo $k \in K_1$ com $k > k_0$ temos $t < 1$, então existe um $\overline{s_k}$ múltiplo de s_k tal que $\|\overline{s}_k\| \leq 10\|s_k\|$ e $f(x_k + \overline{s}_k) > f(x_k) + \alpha \nabla f(x_k)^T \overline{s}_k$.

Sabemos que $\lim_{k\in K_1} \|\overline{s}_k\| = 0$ e para todo $k \in K_1, k > k_0$ temos

$$\nabla f(x_k)^T \overline{s}_k \leq -\theta \|\nabla f(x_k)\| \|\overline{s}_k\|. \tag{4}$$

Seja v um ponto de acumulação da sequência $\dfrac{\overline{s}_k}{\|\overline{s}_k\|}$, então temos que $\|v\| = 1$ e existe uma subsequência com índice em $K_4 \subset K_1$ tal que $\lim \dfrac{\overline{s}_k}{\|\overline{s}_k\|} = v$.

Portanto,

$$\nabla f(x^*)^T v = \lim \nabla f(x_k)^T v = \lim \nabla f(x_k) \frac{\overline{s}_k}{\|\overline{s}_k\|}.$$

De (4) segue que

$$\nabla f(x_*)^T v \leq -\theta \lim \|f(x_k)\|. \tag{5}$$

Agora para todo $k \in K_4$

$$f(x_k + \overline{s}_k) - f(x_k) = \nabla f(x_k + \xi_k \overline{s}_k)^T \overline{s}_k, \qquad \xi_k \in (0,1).$$

Veja que a condição de Armijo não foi satisfeita para $\overline{s}_k$, assim

$$\nabla f(x_k + \xi_k \overline{s}_k)^T \frac{\overline{s}_k}{\|\overline{s}_k\|} > \alpha \nabla f(x_k)^T \overline{s}_k \qquad \forall k \in K_4,$$

ou seja,

$$\nabla f(x_k + \xi_k \overline{s}_k)^T \frac{\overline{s}_k}{\|\overline{s}_k\|} > \alpha \nabla f(x_k)^T \frac{\overline{s}_k}{\|\overline{s}_k\|}.$$

Passando o limite temos

$$\nabla f(x^*)^T v \geq \alpha \nabla f(x^*)^T v$$

ou

$$(1-\alpha) \nabla f(x^*)^T v \geq 0$$

Logo $\nabla f(x^*)^T v \geq 0$ e por (5) segue que $\nabla f(x^*)^T v = 0$

Se $\nabla f(x_*) \neq 0$, novamente por (5), para k suficientemente grande

$$0 = \nabla f(x^*)^T v \leq -\theta \|\nabla f(x_k)\| < 0.$$

Contradição, portanto temos que $\nabla f(x^*) = 0$

Vamos agora considerar que é válido o item (2).

Como $\|s_k\| \geq \varepsilon \ \forall k \in K_2$, pela condição de Armijo temos que

$$f(x_k + s_k) \quad \leq \quad f(x_k) + \alpha \nabla f(x_k)^T s_k$$

$$\begin{aligned} &\leq f(x_k) - \alpha\theta\|\nabla f(x_k)\|\|s_k\| \\ &\leq f(x_k) - \alpha\theta\varepsilon\|\nabla f(x_k)\| \quad \forall k \in K_2. \end{aligned}$$

Portanto,

$$\frac{f(x_k) - f(x_{k+1})}{\alpha\theta\varepsilon} \geq \|\nabla f(x_k)\| \geq 0.$$

Passando ao limite concluímos que $\nabla f(x_k) = 0$. □

O método de busca linear fornece um algoritmo com convergência global. No entanto ele não define uma direção de descida a ser usada, por este motivo precisamos recorrer a outros métodos que forneçam uma escolha inteligente para tal direção. Existem uma série de métodos que podem ser utilizados nesse sentido. Aqui, conforme mencionado, apresentaremos dois métodos, Newton e BFGS, ambos dentro do contexto de busca linear.

4 O método de Newton

Umas das direções mais importantes, possivelmente, são as chamadas direções de Newton. Tais direções são obtidas da expansão da série de Taylor de segunda ordem (NOCEDAL; WRIGHT, 2006) e são expressas pela seguinte equação $d_k = -(\nabla^2 f(x_k))^{-1}\nabla f(x_k)$, onde estamos assumindo que a hessiana $\nabla^2 f(x_k)$ é definida positiva. O método de Newton é uma variação do método usado para encontrar raízes de funções. Para melhor compreendê-lo vamos fazer algumas considerações acerca de funções quadráticas.

Dada uma função quadrática da forma $g(x) = x^T Gx + b^T x + c$ onde G é uma matriz $n \times n$ definida positiva e $x \in \mathbb{R}^n$, sabemos do cálculo diferencial que esta função possui um mínimo global, mais que isso, partindo de qualquer ponto de $\mathbb{R}^n$ usando a direção de Newton em um passo atingimos esse mínimo. De fato $\nabla g(x) = Gx + b$ é o gradiente da função $g(x)$ assim, é fácil ver que tomando a direção d_k zeramos o gradiente. Deste modo, temos direções que funcionam muito bem para quadráticas e assim, para utilizar essa vantagem, tomamos uma aproximação local da função objetivo para uma forma quadrática, usando a fórmula de Taylor de segunda ordem.

É importante ressaltar que as direções fornecidas por Newton só podem ser assumidas como direção de descida se temos que a hessiana da função f é definida positiva (NOCEDAL; WRIGHT, 2006), entretanto para isso seria necessário pedir que a função f fosse convexa (ver convexidade em Martínez e Santos (1998)), mas desejamos estabelecer um método de convergência global, mantendo assim a principal característica da busca linear. Assim usamos uma

atualização da diagonal da matriz hessiana quando esta não for definida positiva. Salvo essas alterações temos o seguinte algoritmo.

Algoritmo de Newton com Busca Linear: Defina $x_0 \in \mathbb{R}^n, \alpha \in (0,1), \beta > 0$ e $\theta \in (0,1)$. Dado x_k pela k-ésima iteração determine x_{k+1} da seguinte forma:

1. Se $\nabla f(x_k) = 0$, pare!
2. Obter a fatoração de Cholesky: $\nabla^2 f(x_k) = LDL^T$;
3. Se (2) falhar defina $B_k = \nabla^2 f(x_k) + \mu I$ para $\mu > 0$ de modo que $B_k > 0$, então obtenha a fatoração Cholesky: $B_k = LDL^T$;
4. Definir d_k resolvendo
$$Ly = -\nabla f(x_k) \text{ e } DL^T d_k = y;$$
5. Se $\nabla f(x_k)^T d_k > -\theta \|\nabla f(x_k)\| \|d_k\|$ faça $\mu \leftarrow \max\{2\mu, 10\}$ e volte em (3);
6. Se $\|d_k\| < \beta \|\nabla f(x_k)\|$, faça
$$d_k \leftarrow \beta \frac{\|\nabla f(x_k)\|}{\|d_k\|} d_k;$$
7. Obter t por *backtracking* que satisfaça
$$f(x_k + t d_k) \leq f(x_k) + \alpha t \nabla f(x_k)^T d_k;$$
8. Defina $x_{k+1} = x_k + t d_k$ e volte a (1).

Observe que a escolha das direções de Newton não alteram a convergência global do algoritmo.

5 A fórmula BFGS

O método de Newton possui boas propriedades de convergência (BERTSEKAS, 1999). Contudo, pode não ser apropriado em alguns casos por fazer-se necessário o cálculo da hessiana da função. Este cálculo pode ser um tanto quanto inconveniente por diversos motivos: custo computacional, complexidade da função e mesmo a propensão a erros humanos. Assim, faz sentido desenvolver métodos que conservam, de certo modo, as boas propriedades de convergência e dispense o cálculo da hessiana. Essa é a ideia primordial dos chamados métodos quase-Newton, que buscam direções aproximadas das direções de Newton utilizando uma matriz B_k próxima

de $\nabla^2 f(x_k)$ a cada iteração. Para fazer essas atualizações busca-se uma matriz B_k que satisfaça a equação secante

$$B_{k+1}s_k = y_k \text{ onde } s_k = x_{k+1} - x_k \text{ e } y_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

pois esta, em geral, fornece boas aproximações da hessiana (MARTÍNEZ; SANTOS, 1998).

Existem muitas fórmulas para a atualização de B_k. A fórmula que vamos tratar aqui impõe que B_k seja simétrica, e é dada pela seguinte expressão:

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}.$$

Conhecida como fórmula BFGS, desenvolvida independentemente por Broyden, Fletcher, Goldfarb e Shanno, em 1970 (LUENBERGER; YE, 2008), esta é uma atualização de posto 2 que possui uma propriedade conveniente acerca da positividade da matriz B_{k+1} apresentada no teorema a seguir.

Teorema 4. *Na fórmula BFGS, se B_k é simétrica definida positiva e $s_k^T y_k > 0$, então B_{k+1} também é simétrica e definida positiva.*

Não apresentaremos aqui a demonstração deste teorema devido a sua simplicidade. Este teorema é fundamental, pois um dos requisitos para se obter direções de descida apresentado na seção anterior, era que a matriz hessiana fosse definida positiva. Deste modo temos controle sobre a positividade em cada atualização. Uma vez que possa ser determinada a aproximação B_k da hessiana de $f(x_k)$ simétrica e definida positiva, sabemos por meio de um cálculo simples de $s_k^T y_k$ se a matriz B_{k+1} será também definida positiva. Deste modo para obter uma direção de descida d_k basta resolver o sistema linear

$$B_k d_k = -\nabla f(x_k).$$

Uma vez que temos B_k definida positiva, o sistema acima possui solução única, determinada por $d_k = -B_k^{-1}\nabla f(x_k)$. Assim, seria conveniente que em vez de resolver o sistema linear, encontremos uma atualização para B_k^{-1}. Como a fórmula BFGS é uma atualização de posto 2, a inversa de B_{k+1} já pode ser calculada como uma atualização da inversa de B_k. Isso é feito por meio da seguinte formulação

$$B_{k+1}^{-1} = B_k^{-1} + \frac{(s_k - B_k^{-1}y_k)s_k^T + s_k(s_k - B^{-1}y_k)^T}{s_k^T y_k} - \frac{(s_k - B_k^{-1}y_k)^T y_k s_k s_k^T}{(s_k^T y_k)^2}.$$

Por uma questão de simplicidade de notação usaremos H_k para representar B_k^{-1}.

Agora que definimos uma forma adequada de encontrar direções de descida, podemos definir um algoritmo de busca linear usando as direções fornecidas com o cálculo da fórmula

BFGS. Apenas faz-se necessário uma correção para o caso em que d_k não satisfaça a condição 3. Neste caso, simplesmente descartaremos a direção fornecida por BFGS e usaremos a direção de máxima descida $-\nabla f(x_k)$, embora na prática tal situação seja rara.

Algoritmo de busca linear com direções BFGS: Defina $x_0 \in \mathbb{R}^n, \alpha \in (0,1), \beta > 0$ e $\theta \in (0,1)$ e uma matriz $H_0 \in \mathbb{R}^{n\times n}$ simétrica e definida positiva. Dado x_k pela k-ésima iteração determine x_{k+1} da seguinte forma:

1. Se $\nabla f(x_k) = 0$, pare!
2. Defina $d_k = -H_k \nabla f(x_k)$;
3. Se $\nabla f(x_k)^T d_k > -\theta \|\nabla f(x_k)\| \|d_k\|$ então

$$H_k \leftarrow I$$

$$d_k \leftarrow -\nabla f(x_k);$$

4. Se $\|d_k\| < \beta \dfrac{\|\nabla f(x_k)\|}{\|d_k\|}$ então corrija usando

$$d_k \leftarrow \beta \frac{\|\nabla f(x_k)\|}{\|d_k\|} d_k;$$

5. Obter t por *backtracking* que satisfaça

$$f(x_k + td_k) \leq f(x_k) + \alpha t \nabla f(x_k)^T d_k;$$

6. Defina:

$$x_{k+1} = x_k + td_k$$

$$s_k = x_{k+1} - x_k$$

$$y_k = \nabla f(x_{k+1}) - \nabla f(x_k);$$

7. Se $s_k^T y_k > 0$ defina

$$H_{k+1} = H_k + \frac{(s_k - H_k y_k)s_k^T + s_k(s_k - H_K y_k)^T}{s_k^T y_k} - \frac{(s_k - H_k y_k)^T y_k s_k s_k^T}{(s_k^T y_k)^2}$$

   caso contrário defina $H_{k+1} = H_k$;

8. Volte ao passo (1).

Referências

BERTSEKAS, Dimitri P.. **Nonlinear programming.** 2 ed. Belmont: Athena Scientific, 1999.

CAMARGO, Jesus Marcos. **Detecção de curvas em imagens com auxilio de funções ordenadas.** 2015. 50 f. Dissertação (Mestrado) - Curso de Matemática, Universidade Estadual de Maringá, Maringá, 2015.

LUENBERGER, David G.; YE, Yinyu. **Linear and nonlinear programming.** 3 ed. New York: Springer, 2008.

MARTÍNEZ, José Maria; SANTOS, Sandra Augusta. **Métodos computacionais de otimização.** 1 ed. Campinas: IMECC-UNICAMP, 1998.

NOCEDAL, Jorge; WRIGHT, Stephen J.. **Numerical optimization.** 2 ed. New York: Springer, 2006.

Agrupamento de dados baseado em colônia de formigas

Carina Moreira Costa[1]
Universidade Estadual do Oeste do Paraná
carina.costa@unioeste.br

Rosangela Villwock
Universidade Estadual do Oeste do Paraná
rosangela.villwock@unioeste.br

Resumo: Este trabalho tem o objetivo de apresentar uma revisão sobre o Agrupamento de dados baseado em Colônia de Formigas, bem como das principais modificações já realizadas no algoritmo. Para isso, foi realizada uma pesquisa na literatura e análise das modificações propostas por alguns autores, observando quais melhorias estas proporcionaram ao algoritmo. Este estudo proporcionou maior entendimento sobre a dinâmica do algoritmo, sobre as motivações envolvidas nas modificações e quais aspectos deve-se levar em consideração para obter um bom agrupamento.

Palavras-chave: Mineração de Dados; Agrupamento de Dados; Metaheurística.

1 Agrupamento de dados baseado em colônia de formigas

O agrupamento de dados ou *clusterização* é uma das principais tarefas de mineração de dados e tem sido amplamente utilizado em muitas aplicações reais para a obtenção de informações novas e úteis implícitas nos dados. Segundo Hair Jr *et al.* (2005) a análise de agrupamento é um conjunto de técnicas multivariadas que tem como objetivo agrupar objetos (indivíduos ou variáveis) com base em suas características. Na análise de agrupamento o objetivo é dividir o conjunto de objetos em dois ou mais grupos (*clusters*) de acordo com a similaridade dos objetos, sendo que os grupos formados são mutuamente excludentes.

Há vários métodos de análise de agrupamento na área de mineração de dados. Alguns dos métodos requerem a definição de muitos parâmetros de entrada, sendo um destes parâmetros o número de grupos existente na base de dados. Porém, esta informação nem sempre é conhecida pelo usuário.

O algoritmo de agrupamento baseado em colônia de formigas – foco deste trabalho - possui uma vantagem com relação à maioria dos outros algoritmos de agrupamento, este não solicita que o usuário informe a quantidade de grupos (LAURO, 2008; ESPENCHITT, 2008).

[1]Os autores agradecem à Fundação Araucária por bolsa de iniciação científica concedida à primeira autora e pelo apoio financeiro ao projeto de pesquisa "Identificação do Perfil dos Municípios do Estado do Paraná por meio do Processo de Descoberta de Conhecimento em Bases de Dados".

O agrupamento baseado em Formigas foi proposto inicialmente por Deneubourg *et al.* (1991). Segundo Lauro (2008, p. 22), "o agrupamento de corpos mortos pela espécie de formiga *Pheidole pallidula* inspirou estes autores sobre o agrupamento por um grupo de agentes homogêneos". A formação de cemitérios e organização da ninhada são dois exemplos notáveis de comportamento coletivo de insetos sociais. A formação de montes, como por exemplo, o cemitério (agrupamento de cadáveres), tem sido observado na espécie *Pheidole pallidula*. Conforme as formigas morrem, elas são carregadas para fora dos ninhos pelas formigas operárias. O processo de formar montes (agrupamento) surge em função da atração entre corpos já depositados e as operárias que estão transportando outros corpos. Os pequenos montes formados atraem as operárias para depositarem novos corpos e assim estes pequenos montes vão aumentando de tamanho (HARTMANN, 2005).

No agrupamento baseado em formigas proposto por Deneubourg *et al.* (1991) as formigas foram representadas como agentes simples que se moviam aleatoriamente em uma grade quadrada. Os padrões foram dispersos nesta grade e poderiam ser carregados, transportados e descarregados pelos agentes (formigas). Estas operações são baseadas na similaridade e na densidade dos padrões distribuídos dentro da vizinhança local dos agentes, padrões isolados ou cercados por dissimilares são mais prováveis de serem carregados e então descarregados numa vizinhança de similares.

Após um número estipulado de iterações, se obtém uma separação espacial dos padrões na grade, os grupos formados tendem a formar regiões densas na grade.

2 Funções e parâmetros envolvidos no algoritmo

As decisões de carregar e descarregar padrões são tomadas pelas probabilidades P_{pick} e P_{drop}, descritas por Deneubourg *et al.*, dadas a seguir pelas equações 1 e 2, respectivamente.

$$P_{pick} = \left(\frac{k_p}{k_p + f(i)}\right)^2. \tag{1}$$

$$P_{drop} = \left(\frac{f(i)}{k_d + f(i)}\right)^2. \tag{2}$$

Nestas equações, $f(i)$ é uma estimativa da fração de padrões na vizinhança que são semelhantes ao padrão atual i da formiga e k_p e k_d são constantes reais. Os autores usaram k_p= 0,1 e k_d= 0,3.

Os parâmetros k_p e k_d demonstram a influência que a função densidade $f(i)$ exerce no cálculo da probabilidade de um agente carregar ou descarregar um objeto em determinada posição. A probabilidade de carregar decresce com o aumento de $f(i)$, de 1 quando $f(i) = 0$ para $\frac{1}{4}$ quando $f(i) = k_p = 0,1$ e possui o menor valor quando $f(i)$ tende para 1. A probabilidade de descarregar aumenta com $f(i)$, de 0 quando $f(i) = 0$ para $\frac{1}{4}$ quando $f(i) = k_d = 0,3$ e possui maior valor quando $f(i)$ tende para 1. Observa-se este comportamento na Figura 1 a seguir.

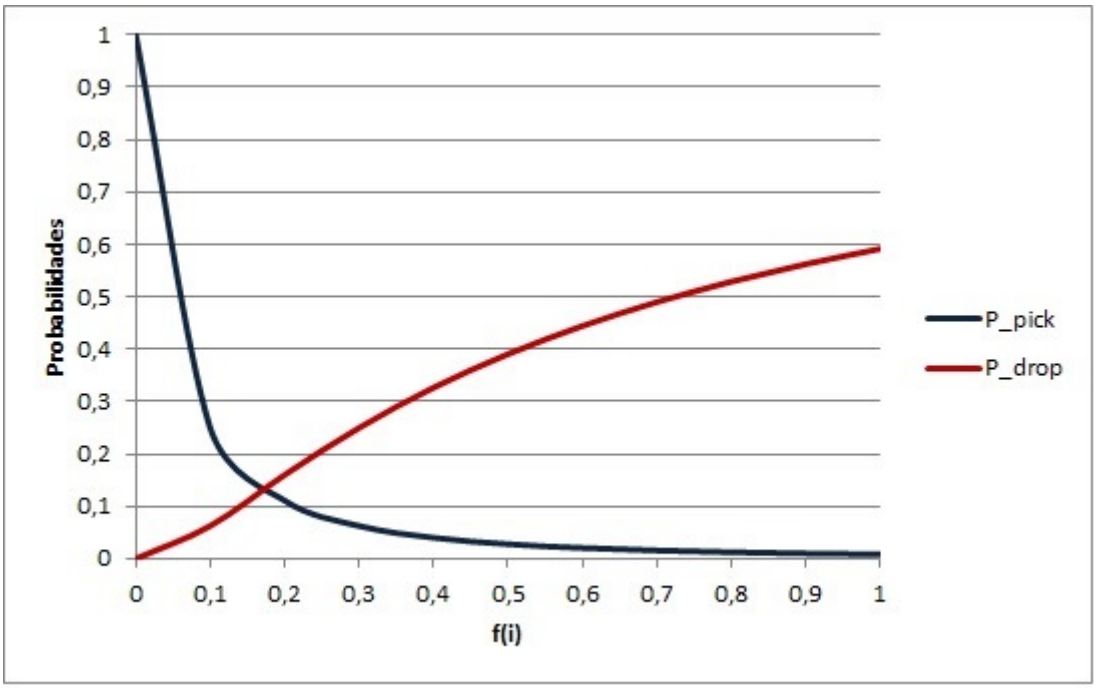


Figura 1: Probabilidade de carregar e descarregar um objeto em função de $f(i)$.

Desta forma, a função de vizinhança $f(i)$ tem grande influência nas decisões de carregar e descarregar padrões. Em Deneubourg *et al.* (1991) a estimativa f foi obtida por meio de uma memória de curto prazo de cada formiga, em que o conteúdo das últimas células analisadas é armazenado, sendo que o número de células depende do tamanho de memória escolhido. A escolha da função de vizinhança $f(i)$ foi essencialmente motivada pela sua simplicidade e facilidade de implementação.

2.1 Principais modificações já realizadas no algoritmo

De acordo com Lumer e Faieta (1994), Deneubourg *et al.* restringiram seus estudos para ambientes com objetos idênticos ou com dois tipos de objetos e assim a tarefa dos agentes era bastante trivial, somente para transportar objetos. Desta forma, Lumer e Faieta (1994) incluíram modificações ao modelo que permitiram a manipulação de dados numéricos com o uso de diferentes tipos de objetos e melhoraram a qualidade da solução e o tempo de convergência

do algoritmo.

A decisão de carregar padrões é baseada na probabilidade P_{pick} dada pela equação 1 anterior e, a decisão de descarregar padrões é baseada na probabilidade P_{drop} que foi adaptada, dada pela equação 3 a seguir, em que k_d é uma constante e $f(i)$ é dada pela equação 4.

$$P_{drop} = \begin{cases} 2f(i), \ se \ f(i) < k_d \\ 1, \ se \ f(i) \geq k_d \end{cases} \tag{3}$$

$$f(i) = max \left\{ 0, \ \frac{1}{\sigma^2} \sum_{j \in L} \left[1 - \frac{d(i,j)}{\alpha} \right] \right\}. \tag{4}$$

Na equação 4 $d(i,j)$ é uma função de dissimilaridade entre padrões i e j pertencente ao intervalo [0,1]. A métrica utilizada, neste caso, foi a distância Euclidiana, dada por:

$$d(i,j) = \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + ... + |x_{in} - x_{jn}|^2} \tag{5}$$

em que $i = (x_{i1}, x_{i2}, \ldots, x_{in})$ e $j = (x_{j1}, x_{j2}, \ldots, x_{jn})$ representam pontos (ou objetos) n-dimensionais. O parâmetro α vai escalar a dissimilaridade, ele é dependente dos dados e pertencente ao intervalo $[0, 1]$. L é a vizinhança local de tamanho igual a σ^2.

O termo σ^2 corresponde ao tamanho da vizinhança da célula em que o agente está e, desta forma, $f(i)$ é dependente da densidade. O agente se localiza no centro desta vizinhança e este percebe a vizinhança pelo raio de percepção, dado por $r = \frac{\sigma - 1}{2}$. A Figura 2 a seguir ilustra o agente no centro e sua vizinhança, com raio de percepção igual a 1 e tamanho da vizinhança igual a 9.

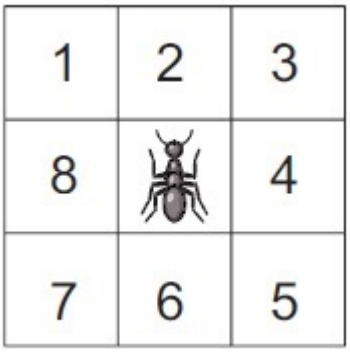


Figura 2: Representação das células da vizinhança de uma formiga (HARTMANN, 2005).

A função $f(i)$ atinge valor máximo se, e somente se, todos os itens na vizinhança são similares, ou seja, $d(i,j) = 0$, qualquer que seja o item j pertencente à vizinhança. Note que

$f(i) \in [0,1]$. Lumer e Faieta utilizaram em seu trabalho $k_p = 0,1$, $k_d = 0,15$ e $\alpha = 0,5$.

O objetivo de Lumer e Faieta (1994) era definir uma medida de similaridade ou dissimilaridade entre os padrões pois no algoritmo proposto inicialmente os objetos eram similaridades se fossem idênticos e dissimilares se os objetos não fossem idênticos.

Além destas modificações, Lumer e Faieta (1994) introduziram uma memória de curto prazo que acelerou significativamente o processo de agrupamento. A formiga poderia lembrar-se dos últimos itens que foram descarregados, bem como os locais onde foram descarregados. Quando a formiga carrega um novo item, ela faz uma comparação com os itens anteriormente descarregados, que estão na memória, e transporta este item para o local onde está o item mais similar a esse.

Os algoritmos de agrupamento baseados em formigas estão principalmente apoiados nas versões propostas por Deneubourg *et al.* (1991) e Lumer e Faieta (1994). Várias modificações foram introduzidas para melhorar a qualidade do agrupamento e, em particular, a separação espacial entre os grupos na grade (BORYCZKA, 2009).

Alterações que melhoram a separação espacial dos grupos e permitem que o algoritmo seja mais robusto foram introduzidas por Handl, Knowles e Dorigo (2005), numa versão melhorada do algoritmo chamada *ATTA*. Uma das alterações é a inclusão de uma restrição na função de vizinhança $f^*(i)$, dada pela equação 6 a seguir.

$$f^*(i) = \begin{cases} \frac{1}{\sigma^2} \sum_{j \in L} \left[1 - \frac{d(i,j)}{\alpha}\right], \ se \ \ \forall j \left(1 - \frac{d(i,j)}{\alpha}\right) > 0 \\ 0 \quad , \ caso \ contrário. \end{cases} \tag{6}$$

em que: a restrição $\forall j \left(1 - \frac{d(i,j)}{\alpha}\right) > 0$ serve para penalizar dissimilaridades elevadas. Esta restrição melhora a separação espacial entre os grupos.

Os autores também modificaram o cálculo das probabilidades de carregar e descarregar padrões, obtidas experimentalmente, dadas pelas equações 7 e 8, respectivamente:

$$P^*_{pick}(i) = \begin{cases} 1, \ se \ f^*(i) \leq 1 \\ \frac{1}{f^*(i)^2}, \ caso \ contrário. \end{cases} \tag{7}$$

$$P^*_{drop}(i) = \begin{cases} 1, \ se \ f^*(i) \geq 1 \\ f^*(i)^4, \ caso \ contrário. \end{cases} \tag{8}$$

2.2 Parâmetros da Função de Vizinhança

A definição dos parâmetros da função de vizinhança é um fator decisivo na qualidade do agrupamento. Para o raio de percepção r é mais atrativo empregar vizinhanças maiores para melhorar a qualidade do agrupamento e da distribuição na grade. Entretanto, este procedimento é mais caro computacionalmente (pois o número de células a serem analisadas em cada ação cresce quadraticamente com o aumento do raio de percepção) e ainda inibe a formação rápida dos grupos na fase inicial (HANDL; KNOWLES; DORIGO, 2005).

Um raio de percepção que aumenta gradualmente com o tempo acelera a dissolução de grupos pequenos preliminares (HANDL; KNOWLES; DORIGO, 2005). No trabalho dos autores, eles utilizaram um raio de percepção inicial igual a 1 ($\sigma = 3$) com um incremento linear até chegar em 5 ($\sigma = 11$). Enquanto ocorre este incremento, os autores mantiveram o parâmetro escalar $\frac{1}{\sigma^2}$ inalterado na função de vizinhança $f^*(i)$, de forma que, na fase inicial $f^*(i)$ está limitada ao intervalo [0,1], depois, o limite superior aumenta com cada aumento do raio de percepção.

A separação espacial dos grupos na grade é crucial para que grupos individuais sejam bem definidos, possibilitando a sua recuperação automática. A proximidade espacial, quando ocorrer, pode indicar a formação prematura do agrupamento (HANDL; KNOWLES; DORIGO, 2005).

Desta forma, após a fase inicial de agrupamento durante um curto intervalo de iterações, entre os tempos $t_{inicial} = 0,45.N$ e $t_{final} = 0,55.N$ em que N é o número total de iterações, os autores substituíram o parâmetro escalar $\frac{1}{\sigma^2}$ por $\frac{1}{N_{occ}}$ na equação 6, em que N_{occ} é o número de células da grade ocupadas, observadas dentro da vizinhança local. Assim, somente a semelhança e não a densidade será levada em consideração. Isto acarreta no espalhamento dos dados na grade, mas de uma forma ordenada, na qual cada grupo ocupa seu próprio lugar na grade. Após esse período, a função vizinhança volta à sua forma anterior. Então novos grupos serão formados, mas agora com maior probabilidade de serem gerados próximo aos centros destas regiões, pelo fato das fronteiras de suas vizinhanças serem de baixa qualidade.

Este efeito ocorre pois nas equações 7 e 8 propostas as decisões de carregar e descarregar padrões se tornam mais cautelosas, sendo determinadas exclusivamente pelo valor da função de vizinhança $f^*(i)$.

O parâmetro α é um limitante que vai "escalar" a dissimilaridade na função de vizinhança $f(i)$, sendo que o funcionamento do algoritmo é crucialmente dependente deste parâmetro. A escolha de um valor muito pequeno para α impede a formação de grupos na grade, por outro

lado, a escolha de um valor muito grande para α resulta na fusão de grupos. No caso limite, $\alpha = 1$, todos os itens formariam um único grupo.

Fixar parâmetro α não é simples e a sua escolha é altamente dependente da estrutura do conjunto de dados. Um valor inadequado resulta em uma excessiva ou extremamente baixa atividade na grade. Por conta disso, Handl, Knowles e Dorigo (2005) propuseram uma adaptação automática de α por meio da quantidade de atividade bem sucedida do agente, ou seja, se faz o registro do sucesso e insucesso da operação de descarregar os itens.

3 O algoritmo básico

Numa fase inicial, todos os padrões são aleatoriamente espalhados na grade, na qual cada célula deve conter apenas um padrão, assim, haverá células da grade ocupadas e outras que estarão livres. Depois, cada formiga escolhe aleatoriamente um único padrão para carregar e é colocada em uma posição aleatória na grade.

Na próxima fase, chamada de fase de distribuição, em um laço (*loop*) simples, cada formiga é selecionada aleatoriamente. Esta formiga se desloca na grade executando um passo de comprimento l numa direção aleatória. A formiga então decide, probabilisticamente, se descarrega seu padrão nesta posição.

Se a decisão for de não descarregar o padrão, escolhe-se aleatoriamente outra formiga e recomeça-se o processo. Se a decisão for de descarregar, a formiga descarrega o padrão em sua posição atual na grade, se esta estiver livre. Se esta célula da grade estiver ocupada por outro padrão, o mesmo deve ser descarregado numa célula imediatamente vizinha desta, que esteja livre, por meio de uma procura aleatória.

A formiga procura, então, por um novo padrão para carregar. Dentre os padrões livres na grade, ou seja, dentre os padrões que não estão sendo carregados por nenhuma formiga, a formiga seleciona aleatoriamente um, vai para a sua posição na grade (posição em que o padrão se encontra), faz a avaliação da função de vizinhança e decide probabilisticamente se carrega este padrão. Este processo de escolha de um padrão livre na grade é executado até que a formiga encontre um padrão que deva ser carregado.

Só então esta fase é reiniciada, escolhendo-se outra formiga até que um critério de parada seja satisfeito. Uma iteração ocorre quando todas as formigas já tenham sido selecionadas. Ao final, um processo para recuperação dos grupos é aplicado.

Considerações Finais

A realização do estudo bibliográfico foi de extrema importância para a compreensão da dinâmica do algoritmo e das propostas de melhoria, visto que esta meta-heurística ainda exige muita investigação para melhorar seu desempenho, estabilidade e outras características, consideradas "chaves", que fariam de tal algoritmo uma ferramenta madura para mineração de dados.

Referências

BORYCZKA, U.. Finding groups in data: Cluster analysis with ants. Applied Soft Computing 9, 61-70, 2009.

DENEUBOURG, J. L.; GOSS, S.; FRANKS, N.; SENDOVA-FRANKS, A.; DETRAIN, C.; CHRÉTIEN, L. The dynamics of collective sorting: Robot-like ants and ant-like robots. Proceedings of the First International Conference on Simulation of Adaptive Behaviour: From Animals to Animats 1, 356–365, 1991.

ESPENCHITT, D. G.. Segmentação de Dados em um Número Desconhecido de Grupos Utilizando Algoritmo de Colônia de Formigas. Tese (Doutorado) - Rio de Janeiro: COPPE/UFRJ, 2008.

HAIR JR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. Análise Multivariada de Dados. Tradução de: SANTANNA, A. S.; CHAVES NETO, A. Porto Alegre: Bookman, 2005.

HANDL, J.; KNOWLES, J.; DORIGO, M. Ant-Based Clustering and Topographic Mapping. Artificial Life v. 12, 35-61, 2005.

HARTMANN, V.. Evolving agents warms for clustering and sorting. Proceedings of the 2005 Conference on Genetic and Evolutionary Computation, 217-224, Washington DC, 2005.

LAURO, A. L.. Agrupamento de Dados Utilizando Algoritmo de Colônia de Formigas. Dissertação (Mestrado) - Curso de Programa de Pós-graduação em Engenharia Civil, Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2008.

LUMER, E.; FAIETA, B. Diversity and adaptation in populations of clustering ants. Proceedings of the Third International Conference on Simulation of Adaptive Behaviour: From Animals to Animats 3, 501–508, 1994.

O corpo ordenado e completo dos números reais

Alexandre Batista de Souza
Universidade Estadual do Oeste do Paraná
aledaron@gmail.com

Flavio Roberto Dias Silva
Universidade Estadual do Oeste do Paraná
frdsilva@yahoo.com.br

Resumo: O presente trabalho é resultado da iniciação científica voluntária do discente Alexandre Batista de Souza. Aqui objetivamos apresentar a construção do corpo ordenado e completo dos números reais, de um ponto de vista da álgebra. Para isso, utilizamos a abordagem clássica conhecida como cortes de Dedekind.

Palavras-chave: cortes de Dedekind; números; construção.

1 Introdução

Nosso trabalho busca apresentar a construção do corpo ordenado e completo dos números reais, por cortes de Dedekind. A motivação e/ou relevância desse estudo, dá-se entre outras coisas, pela importância do conjunto dos números reais dentro da matemática e das ciências exatas, nas quais é elemento fundamental para a formalização, por exemplo, do cálculo diferencial e integral. O conceito de completude, entre outros, é o que permite falar em cálculo infinitesimal.

Com tal propósito, e tendo como elemento guia o caminho traçado por Deparis, (2009) faz-se necessário inicialmente formalizar o conceito intuitivo de número natural, por meio de axiomas que garantam a existência e o comportamento usual do conjunto dos números naturais $\mathbb{N}$. Em seguida, utiliza-se os elementos deste conjunto associado ao conceito de classes de equivalência, para construir ao anel de integridade dos números inteiros $\mathbb{Z}$, que contém uma cópia algébrica $\mathbb{N}$. Analogamente, utiliza este anel de integridade para construir um corpo ordenado $\mathbb{Q}$, no qual $\mathbb{Z}$ está imerso, pela existência de um isomorfismo entre os anéis, $\mathbb{Z}$ e $\mathbb{Q}$. Destacamos que a construção de $\mathbb{N}$, $\mathbb{Z}$ e $\mathbb{Q}$, não será feita aqui e pode ser encontrada no trabalho da autora supracitada bem como em Ferreira (2013). Neste artigo nos dedicaremos a apresentar a construção do corpo ordenado e completo dos números reais $\mathbb{R}$, assumindo a existência do corpo ordenado dos números racionais $\mathbb{Q}$. Por último, é importante lembrar que este trabalho se utiliza de alguns conceitos clássicos da álgebra como; relações de ordem, aplicações, operações, anéis, corpos, isomorfismos que podem ser encontrados, por exemplo em Domingues (2003).

2 O corpo ordenado e completo dos números reais

Considerando o corpo das frações do anel de integridade dos números inteiros $\mathbb{Q}$ e a relação $\prec$, definida sob o mesmo, construiremos subconjuntos deste que serão elementos de um novo conjunto que também será totalmente ordenado e terá estrutura de corpo. Ainda, além de possuir essas propriedades, esse conjunto possuirá também a completude.

2.1 O corte de Dedekind

Definição 1. O corte de Dedekind é denotado por α, tal que $\alpha \subset \mathbb{Q}$, e satisfaz as seguintes propriedades,

a) $\alpha \neq \emptyset$ e $\gamma \neq \mathbb{Q}$;

b) $\forall p, q \in \mathbb{Q}$, se p $\in \alpha$ *e* $q \prec p$, então $q \in \alpha$;

c) $\forall p \in \alpha, \exists\ q \in \alpha,$ *tal que* $p \prec q$.

Definição 2. Para todo $r \in \mathbb{Q}$, o conjunto α é definido por

$$\alpha = \{p \in \mathbb{Q}; p \prec r\}.$$

Teorema 3. *O conjunto α, dado pela definição 2 satisfaz a definição 1.*

Prova. Temos que mostrar que α satifaz as condições a, b e c da definição 1.

Para tanto, note que para todo $r \in \mathbb{Q}$, existe $s \in \mathbb{Q}$, tal que $s \prec r$. Disso decorre que, $\alpha \neq \emptyset$. É fato também, que para todo $r \in \mathbb{Q}$, existe $t \in \mathbb{Q}$, tal que $r \prec t$, implicando que, $\alpha \neq \mathbb{Q}$. Assim, α satisfaz a condição (a).

Sejam $p, q \in \mathbb{Q}$, com $p \in \alpha$ e $q \prec p$. Sendo assim, temos da definição de α, que $q \prec r$ e, portanto $q \in \alpha$. Logo, α satisfaz a condição (b).

Finalmente, para todo $p \in \alpha$, existe $q = \frac{p+r}{2} \in \alpha$. Pela construção $q \prec r$ e $p \prec q$. Analogamente, existe $q_1 = \frac{q+r}{2}$, com $p \prec q_1$. Repetindo este processo encontramos $q_n = \frac{p+(\ 2^{n+1}-1)\cdot r}{2^{n+1}}$, para todo $n \in\ \mathbb{N}^*$. Decorre então que α não contém o supremo e a condição (c) é satisfeita.

□

Definição 4. O conjunto dos cortes de Dedekind α é denotado por $\mathbb{R}$ e definido por

$$\mathbb{R} = \{\alpha \subset \mathbb{Q}; \alpha \text{ é dado pela definição 2 }\}$$

2.2 Operações em $\mathbb{R}$

Com o intuito de que $\mathbb{R}$ tenha estrutura de corpo, vamos definir duas operações neste conjunto, que irão satisfazer as propriedades desejadas.

Definição 5. Sejam $\alpha, \beta \in \mathbb{R}$. O conjunto denotado por γ é definido por

$$\gamma = \{a + b; a \in \alpha \text{ e } b \in \beta\}$$

Lema 6. *Sejam $\alpha \in \mathbb{R}$ e $x \notin \alpha$. Então $p \prec x, \forall p \in \alpha$.*

Prova. Sejam $\alpha \in \mathbb{R}$ e $x \notin \alpha$. Suponha que exista $p \in \alpha$, com $x \preceq p$. Disso, pela propriedade (b), da definição 1, temos que $x \in \alpha$, contradizendo a hipótese. □

Teorema 7. *O conjunto γ, dado pela definição 5 é um corte de Dedekind.*

Prova. Novamente, temos que verificar as três condições da definição 1.

Note que, sendo α, $\beta \in \mathbb{R}$, existem $a \in \alpha$ e $b \in \beta$, tal que $a + b \in \gamma$, implicando que $\gamma \neq \emptyset$. Pelo mesmo motivo, existem $c \in \alpha$ e $d \in \beta$, tal que $c \notin \alpha$ e $d \notin \beta$. Pelo lema 6, para todo $a \in \alpha$ e $b \in \beta$, $a \prec c$ e $b \prec d$ implicando que $a + b \neq c + d$. Logo, $c + d \notin \gamma$ e $\gamma \neq \mathbb{Q}$. Temos portanto, que γ atende a condição (a).

Agora sejam $x \in \gamma$ e $y \prec x$. Assim $x = a + b$, para algum $a \in \alpha$ e $b \in \beta$. Disso decorre que $y - a \prec b$ e sendo $\beta \in \mathbb{Q}$, temos que $y - a \in \beta$. Portanto, tomando $y = a + (y - a)$, para algum $a \in \alpha$ e $y - a \in \beta$, $y \in \gamma$, e a condição (b) é satisfeita.

Para comprovar a condição (c) da referida definição, tome $x \in \gamma$. Pela definição do conjunto γ, $x = a + b$, para algum $a \in \alpha$ e $b \in \beta$. Sendo $\alpha, \beta \in \mathbb{R}$, existem $c \in \alpha$ e $d \in \beta$, tais que $a \prec c$ e $b \prec d$, para quaisquer $a \in \alpha$ e $b \in \beta$. Em decorrência disso temos que $a + b \prec c + d$ para todo $a \in \alpha$ e $b \in \beta$, e portanto para qualquer $x \in \gamma$ existe $y \in \gamma$ tal que $x \prec y$.

□

Definição 8. A adição em $\mathbb{R}$ é denotada por $+$ e definida por

$$\begin{array}{rccc} + : & \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ & (\alpha, \beta) & \mapsto & +((\alpha, \beta)) = \gamma. \end{array}$$

Definição 9. O zero é denotado por 0^* e definido por

$$0^* = \{p \in \mathbb{Q}; p \prec 0, \text{ com } 0 \in \mathbb{Q}\}.$$

Este conjunto é o elemento neutro para a operação de adição, fato que será demonstrado no teorema 17.

Definição 10. A relação menor que em $\mathbb{R}$ é denotada por $\prec$ e definida por

$$\alpha \prec \beta \Leftrightarrow \ \alpha \subset \beta.$$

Definição 11. Sejam $\alpha, \beta \in \mathbb{R}$, com $0^* \prec \alpha$ e $0^* \prec \beta$. O conjunto denotado por γ é definido por

$$\gamma = \mathbb{Q}_- \bigcup \{a \cdot b; a \in \alpha \ e \ b \in \beta \ com \ 0 \prec a \ e \ 0 \prec b\}.$$

Teorema 12. *O conjunto γ, dado pela definição 11 é um corte de Dedekind.*

Prova. É um argumento análogo ao caso da operação de adição e pode ser encontrado em Deparis, (2009), ou mesmo em Ferreira, (2013). □

Definição 13. A multiplicação em $\mathbb{R}$ é denotada por $\cdot$ e definida por

$$\begin{array}{rccl} \cdot : & \mathbb{R} \times \mathbb{R} & \to & \mathbb{R} \\ & (\alpha, \beta) & \mapsto & \cdot((\alpha, \beta)) = \gamma \end{array}$$

tais que,

$\alpha \cdot \beta = 0^*$, se $\alpha = 0^*$ ou $\beta = 0^*$;

$\alpha \cdot \beta = \gamma$, se $\alpha \prec 0^*$ e $\beta \prec 0^*$ ou se $0^* \prec \alpha$ e $0^* \prec \beta$;

$\alpha \cdot \beta = -\gamma$, se $0^* \prec \alpha$ e $\beta \prec 0^*$ ou se $\alpha \prec 0^*$ e $0^* \prec \beta$.

Observamos que o conjunto $-\gamma$ é o elemento simétrico de γ com respeito a operação de adição, como será provado no teorema 17.

2.3 Relação de ordem em $\mathbb{R}$

Objetivamos agora, definir uma relação de ordem total no conjunto dos cortes de Dedekind, para que este possa ter estrutura de corpo ordenado, munido das operações que foram definidas na subseção 2.2.

Definição 14. A relação menor ou igual em $\mathbb{R}$ é denotada por $\preceq$ e definida por

$$\alpha \preceq \beta \Leftrightarrow \ \alpha \subseteq \beta.$$

Proposição 15. *A relação $\preceq$ é relação de ordem em $\mathbb{R}$.*

Prova. Precisamos mostrar que a relação $\preceq$ é reflexiva, anti-simétrica e transitiva.

De fato,como α é um conjunto temos que $\alpha \subset \alpha$, o que implica em $\preceq$ ser reflexiva.

Tome quaisquer α e $\beta \in \mathbb{R}$ com $\alpha \preceq \beta$ e $\beta \preceq \alpha$. Assim, pela definição desta relação, $\alpha \subseteq \beta$ e $\beta \subseteq \alpha$ e pela igualdade de conjuntos $\alpha = \beta$. Logo, $\preceq$ é anti-simétrica.

Também, para todo α, β e $\gamma \in \mathbb{R}$, tais que $\alpha \preceq \beta$ e $\beta \preceq \gamma$, observe que $\alpha \subseteq \beta$ e $\beta \subseteq \gamma$ temos as respectivas inclusões que satisfazem a propriedade transitiva, e portanto a relação menor ou igual é transitiva. □

Proposição 16 (Tricotomia)**.** $\forall \alpha, \beta \in \mathbb{R}$, *temos que*

$$\alpha \preceq \beta \text{ ou } \beta \preceq \alpha$$

Prova. Suponhamos sem perda de generalidade que, $\alpha \nsubseteq \beta$. Assim existe $x \in \mathbb{Q}$, tal que $x \in \alpha$ e $x \notin \beta$. Pelo lema 6, $p \prec x$ para qualquer $p \in \beta$. Como $x \in \alpha$ e $p \in \alpha$, para todo $p \in \beta$. Decorre então que, $\beta \subseteq \alpha$ e logo $\beta \preceq \alpha$. □

As proposições 15 e 16 mostram que relação menor ou igual é uma relação de ordem total sob o conjunto $\mathbb{R}$ (Domingues, 2003).

2.4 O corpo ordenado e completo dos números reais

Teorema 17. *O conjunto $\mathbb{R}$ munido das operações $\cdot$ e $+$ possui estrutura de corpo.*

Prova. Ver Deparis, (2009). □

Proposição 18. *Para todo α, β e $\gamma \in \mathbb{R}$, se $\alpha \preceq \beta \Rightarrow \alpha + \gamma \preceq \beta + \gamma$.*

Prova. Seja $a + c \in \alpha + \gamma$ para algum $a \in \alpha$ e $c \in \gamma$. Temos da hipótese que $\alpha \preceq \beta$, então, $\alpha \subseteq \beta$. Logo $a + c \in \beta + \gamma$ para algum $a \in \alpha$ e $c \in \gamma$. Logo $\alpha + \gamma \subseteq \beta + \gamma$ e portanto, pela definição da relação menor ou igual, $\alpha + \gamma \preceq \beta + \gamma$. □

Proposição 19. *Para todo α, β e $\gamma \in \mathbb{R}$, com $0^* \prec \gamma$,se $\alpha \preceq \beta$ então $\alpha \cdot \gamma \preceq \beta \cdot \gamma$*

Prova. Seja $a \cdot c \in \alpha \cdot \gamma$, para algum $a \in \alpha$ e $b \in \beta$. Como supõe-se que $\alpha \preceq \beta$ temos que $\alpha \subseteq \beta$, e disso decorre que $a \cdot c \in \beta \cdot \gamma$, para algum $a \in \beta$ e $c \in \gamma$. Logo $\alpha \cdot \gamma \subseteq \beta \cdot \gamma$ e portanto, pela definição da relação menor ou igual $\alpha \cdot \gamma \preceq \beta \cdot \gamma$. □

As proposições 15, 16, 18 e 19, juntamente com o teorema 17, garantem, segundo Domingues (2003) que $\mathbb{R}$ possui estrutura de corpo ordenado. Sendo assim, a partir de agora caminharemos para mostrar que $\mathbb{R}$ é um corpo ordenado e completo.

Definição 20. O corte racional é denotado por r^* e definido por

$$r^* = \{p \in \mathbb{Q}; p \prec r, \forall r \in \mathbb{Q}\}.$$

Definição 21. O conjunto cortes racionais é denotado por $\mathbb{Q}'$ e definido por

$$\mathbb{Q}' = \{r^*; r \in \mathbb{Q}\}.$$

Teorema 22. *O conjunto $\mathbb{Q}'$ dotado das operações $+$ e $\cdot$ é um subcorpo em $\mathbb{R}$.*

Prova. Teremos que verificar que 0^* e $1^* \in \mathbb{Q}'$, e além disso se este conjunto é fechado para as operações $+$ e $\cdot$ definidas sobre $\mathbb{R}$.

Como $0, 1 \in \mathbb{Q}$, pela definição do conjunto dos cortes racionais, 0^* e $1^* \in \mathbb{Q}'$.

Tome $r^*, s^* \in \mathbb{Q}'$. Pela definição de adição temos que $r^* + s^* = (r + s)^*$ e pela definição do conjunto $\mathbb{Q}'$ ocorre que $(r + s)^* \in \mathbb{Q}'$.

Analogamente, pela definição de multiplicação temos que $r^* \cdot s^* = (r + s)^*$ e logo pela definição do conjunto $\mathbb{Q}'$ este é fechado para multiplicação. Portanto, o conjunto dos cortes racionais possuem estrutura de subcorpo em $\mathbb{R}$. □

Teorema 23. *Os anéis $\mathbb{Q}$ e $\mathbb{Q}'$ são isomorfos.*

Prova. Defina a função

$$\begin{array}{rccc} f: & \mathbb{Q} & \to & \mathbb{Q}' \\ & r & \mapsto & f(r) = r^*. \end{array}$$

Tome $f(r) = f(s)$ e note que pela definição da função f que $r^* = s^*$. Logo pela definição de corte $r = s$ e a função f é injetora.

Tome $r^* \in \mathbb{Q}'$ e note que sempre existe r tal que $f(r) = r^*$, indicando que f é sobrejetora.

Observe que pela definição da função f, $f(r + s) = (r + s)^*$ e pela definição de adição $(r + s)^* = r^* + s^* = f(r) + f(s)$.

Também pela definição da função f, $f(r \cdot s) = (r \cdot s)^*$. Da definição de multiplicação temos que$(r \cdot s)^* = r^* \cdot s^* = f(r) \cdot f(s)$.

Portanto concluímos que f é um isomorfismo entre os anéis $\mathbb{Q}$ e $\mathbb{Q}'$. □

O teorema 23 nos diz que $\mathbb{Q}$ está imerso em $\mathbb{R}$, ou seja, que o conjunto dos números racionais está contido no conjunto dos cortes de Dedekind, e juntamente com o teorema 22, sugere que $\mathbb{Q}' \neq \mathbb{R}$ ou seja, que existe $\alpha \in \mathbb{R}$ e $\alpha \notin \mathbb{Q}'$. Isso é o que comprova o próximo resultado.

Teorema 24. *Existe $\alpha \in \mathbb{R}$, tal que, $\alpha \notin \mathbb{Q}'$.*

Prova. Defina $\alpha = \mathbb{Q}_- \bigcup \{x \in \mathbb{Q}_+; x \cdot x \prec 2, \text{ com } 2 \in \mathbb{Q}\}$. Vamos mostrar apenas que este conjunto é um corte de Dedekind. O argumento para o fato de que $\alpha \notin \mathbb{Q}'$ pode ser observado em Ferreira (2013).

É imediato que $\alpha \neq \emptyset$, pois $\mathbb{Q}_- \subset \alpha$. Note também que, $4 \in \mathbb{Q}$, mas $4 \notin \alpha$, pois $2 \prec 4$. Logo, $\alpha \neq \mathbb{Q}$ e consequentemente α satisfaz a condição (a).

Agora sejam $p, q \in \mathbb{Q}$, tal que $p \in \alpha$ e $q \prec p$. Se $p \in \mathbb{Q}_-$, então $q \in \mathbb{Q}_-$, então $q \in \alpha$. Se $0 \prec p$ e $q \preceq 0$, então $q \in \mathbb{Q}_-$. Logo $q \in \alpha$. Se $0 \prec pe0 \prec q$, temos disso e da hipótese que $q \cdot q \prec p \cdot p$. Assim como $p \in \alpha$ temos que $p \cdot p \prec 2$. Decorre portanto que, $q \cdot q \prec 2$ e a condição (b) está atendida.

Seja $p \in \alpha$. Temos que para todo $n \in \mathbb{N}^*$,

$$(p + \tfrac{1}{n}) \cdot (p + \tfrac{1}{n}) = p \cdot p + 2p \cdot \tfrac{1}{n} + \tfrac{1}{n} \cdot \tfrac{1}{n} \preceq p \cdot p + 2p \cdot \tfrac{1}{n} + \tfrac{1}{n}$$

Note que para $\frac{2p+1}{2-p \cdot p} \prec n$ temos,

$$p \cdot p + 2p \cdot \tfrac{1}{n} + \tfrac{1}{n} \prec 2 \text{ e disso, } (p + \tfrac{1}{n}) \cdot (p + \tfrac{1}{n}) \in \alpha$$

Como $p \prec (p + \frac{1}{n}) \cdot (p + \frac{1}{n})$, a condição (c) está verificada. □

Definição 25. O elemento α dado pelo teorema 24 é denominado corte irracional.

Definição 26. Seja $\mathbb{K}$ um corpo ordenado e $A \subset \mathbb{K}$. O subconjunto A é limitado superiormente se existe $k \in \mathbb{K}$ tal que $a \preceq k$, para qualquer $a \in A$. Assim k é denominado cota superior do conjunto A. A menor das cotas superiores de A - quando existe - é denominado supremo do conjunto A.

Definição 27. Seja $A \subseteq \mathbb{K}$ com A um conjunto limitado superiormente e $\mathbb{K}$ um corpo ordenado. Definimos que $\mathbb{K}$ é um corpo ordenado e completo se e somente se,

$$\forall A \subset \mathbb{K}, \exists k \in \mathbb{K}; a \preceq k, \forall a \in A.$$

Proposição 28. *Se $\alpha, \beta \in \mathbb{R}$, com $\alpha \prec \beta$, existe um corte racional r^* tal que $\alpha \prec r^* \prec \beta$.*

Prova. Sendo $\alpha \prec \beta$, existe $p \in \beta$ tal que $p \notin \alpha$. Segue pela condição (c) da definição 1 que existe $r \in \beta$ tal que $p \prec r$. Como temos que $r \in \beta$, mas pela definição de $r^*, r \notin r^*$, segue disso que $r^* \prec \beta$. Pela condição (b) da definição 1, $p \in r^*$ mas da maneira que foi tomado, $p \notin \alpha$, o que implica em $\alpha \prec r^*$. □

O próximo resultado mostra que $\mathbb{R}$ possui a completude (Dedekind, 1901), e muitas vezes é conhecido como o teorema de Dedekind ou teorema do completamento (Rudin, 1971).

Teorema 29. *Sejam A e B subconjuntos de $\mathbb{R}$ tais que,*

i) $\mathbb{R} = A \cup B$;

ii) $A \cap B = \emptyset$;

iii) $A \neq \emptyset$ *e* $B \neq \emptyset$;

iv) se $\alpha \in A$ *e* $\beta \in B$, *então* $\alpha \prec \beta$.

Então,

$$\exists!\ \gamma \in \mathbb{R}; \alpha \preceq \gamma \preceq \beta, \forall \alpha \in A \text{ e } \beta \in B.$$

Prova. Primeiro provaremos a existência γ.

Considere o conjunto $\gamma = \{r \in \mathbb{Q} | r \in \alpha$ para algum $\alpha \in A\ \}$. Vamos mostrar que γ satisfaz a definição 1, ou seja, é um corte de Dedekind.

Com efeito, como $A \neq \emptyset$ temos que $\alpha \neq \emptyset$, implicando imediatamente que $\gamma \neq \emptyset$. Por outro lado, $\alpha \subset \beta$, para todo $\alpha \in A$ e para todo $\beta \in B$, existe $p \in \beta$ talque $p \notin \alpha$. Logo, $p \notin \gamma$, pela definição deste conjunto e decorre que $\gamma \neq \mathbb{Q}$. Portanto a condição (a) esta satisfeita.

Sejam $r \in \gamma$ e $s \prec r$. Disso, temos que $r \in \alpha$ para algum $\alpha \in A$. Sendo $\alpha \in \mathbb{R}$, temos que $s \in \alpha$ e segue disso que $s \in \gamma$. Logo γ satisfaz a condição (b).

Seja $r \in \gamma$, o que implica que $r \in \alpha$, para algum $\alpha \in A$. Como $\alpha \in \mathbb{R}$, existe $s \in \alpha$ tal que $r \prec s$ e portanto $s \in \gamma$. mostrando que γ cumpre a condição (c).

Agora tendo que $\gamma \in \mathbb{R}$ devemos mostrar ainda que $\alpha \preceq \delta \preceq \beta$ para todo $\alpha \in A$ e para todo $\beta \in B$. De fato, pela definição do conjunto γ temos que $\alpha \preceq \gamma$, para todo $\alpha \in A$. Suponha por absurdo que exista $\beta \in B$ com $\beta \prec \gamma$. Disso temos que existe $r \in \gamma$ e $r \notin \beta$, e portanto $r \in \alpha$ para algum $\alpha \in A$. Logo $\beta \prec \alpha$, contrariando a hipótese.

Por fim suponha que existam γ_1 e γ_2, com $\gamma_1 \prec \gamma_2$ tais que $\alpha \preceq \gamma_1 \prec \gamma_2 \preceq \beta$ para todo $\alpha \in A$ e para todo $\beta \in B$. Pela proposição 28, existe γ_3 tal que $\gamma_1 \prec \gamma_3 \prec \gamma_2$. Disso temos que, $\gamma_3 \prec \beta$ para todo $\beta \in B$ e como $A \bigcup B = \mathbb{R}$, então $\gamma_3 \in A$. Analogamente temos que, $\alpha \prec \gamma_3$ implicando que $\gamma_3 \in B$. Mas $\gamma_3 \in A \bigcap B$ contradiz a hipótese e portanto, γ é único.

□

Corolário 30. *O corpo ordenado $\mathbb{R}$ é completo (satisfaz a definição 27).*

Prova. Seja $M \subset \mathbb{R}$ um conjunto limitado superiormente e considere o conjunto das cotas superiores de M denotado por S, bem como seu complementar denotado por S^c. Note que para S^c e S as hipóteses do teorema anterior se verificam e temos que,

$$\exists! \gamma \in \mathbb{R}; \alpha \preceq \gamma \preceq \beta, \forall \alpha \in S^c \text{ e } \beta \in S$$

Com o objetivo de mostrar que γ é o supremo do conjunto M,devemos considerar dois casos;

1º caso: Se ocorre que,

$$\exists! \gamma \in \mathbb{R}; \alpha \prec \gamma \preceq \beta, \forall \alpha \in S^c \text{ e } \beta \in S$$

obtemos que $\gamma \preceq \beta, \forall \beta \in S$, ou seja, γ é a menor das cotas superiores de M, e sendo assim seu o supremo.

2º caso: Por outro lado, aceitando como verdadeiro que,

$$\exists! \gamma \in \mathbb{R}; \alpha \preceq \gamma \prec \beta, \forall \alpha \in S^c \text{ e } \beta \in S.$$

e supondo que γ não é o supremo de M - ou seja, existe $\mu \in M$ tal que $\gamma \prec \mu$ - obtemos que, $\mu \in M$. Pela proposição 28, ocorre que existe $\eta \in M$ tal que $\gamma \prec \eta \prec \mu$, e isso implica que $\eta \notin S$. Logo $\eta \in S^c$,resultando que η é o supremo de S^c, contradizendo então, o fato de γ ser o supremo. Diante disso não podemos admitir como verdadeiro o segundo caso e temos que só o primeiro pode ocorrer, o que prova o corolário. □

Definição 31. A estrutura denotada por $(\mathbb{R}, +, \cdot)$ ou simplesmente por $\mathbb{R}$, é o corpo ordenado e completo dos números reais.

3 Considerações finais

Este trabalho buscou apresentar a formalização para o conceito de número real, baseada na abordagem dos cortes de Dedekind. Assumindo a existência do corpo ordenado dos números racionais, construímos um conjunto dotado de duas operações, $\cdot$ e $+$, e de uma relação de ordem total $\preceq$, que lhe conferem estrutura de corpo ordenado que é também dotado de completude.

Referências

DEDEKIND, Richard. Essay on theory of Numbers. Publishing Company, Chicago, n. 3, p. 1-6,1901. Disponível em: ¡www.personal.psu.edu/t20/courses/math140/dedekind-book.pdf¿. Acesso em: 29 ago. 2016.

DEPARIS, Dafne de Morais. Construção do corpo ordenado dos números reais.2009. Monografia (Licenciatura em Matemática)-Centro de Ciências Exatas e Tecnológicas, Universidade Estadual do Oeste do Paraná, Cascavel, Pr.

DOMINGUES, Hygino; IEZZI, Gelson. Álgebra Moderna. 4 ed. São Paulo: Atual, 2003.386 p.

FERREIRA, Jamil. A Construção dos Números. 3 ed. Rio de Janeiro: SBM, 2013. 133 p.

RUDIN, Walter. Princípios de Ánalise Matemática. 3 ed. Brasília: Universidade de Brasília, 1971. 296 p.

Estudo dos métodos geoestatísticos envolvidos na determinação da estrutura de dependência espacial em uma variável com tendência direcional

Rodrigo Lorbieski
Universidade Estadual do Oeste do Paraná
rodrigo.lorbieski@hotmail.com

Luciana P. C. Guedes
Universidade Estadual do Oeste do Paraná
luciana_pagliosa@hotmail.com

Miguel A. Uribe-Opazo
Universidade Estadual do Oeste do Paraná
mopazo@unioeste.br

Resumo: Na geoestatística existe um método de interpolação chamado krigagem, que possibilita estimar valores de uma variável georreferenciada em qualquer localização não amostrada dentro da área em estudo, sem tendência e com variância mínima. Essa estimação espacial de localizações não amostradas é feita considerando a estrutura de dependência espacial da variável, expressa pela função semivariância. Além disso, espera-se que os valores dentro da área de interesse não apresentem tendência que possam afetar os resultados quanto a suposição de estacionaridade. Entretanto, isso nem sempre ocorre, pois há situações em que a variável exibe uma variação sistemática, sendo necessário então sua incorporação no estudo de dependência espacial. Objetiva-se nesse trabalho testar a influência da tendência direcional na predição de valores em um conjunto de dados e com isso testar se essa tendência pode afetar na análise de uma determinada variável. Para isso foram realizadas simulações onde foram criados diferentes ambientes nas quais havia a presença ou não desta tendência e avaliar a sua influência na predição dos dados. Os resultados mostraram diferenças significativas nos valores preditos nos casos estudados o que mostra que a presença de tendência direcional no conjunto de dados pode influenciar na predição de valores em locais não amostrados.

Palavras-chave: Covariável; estacionaridade; simulação de dados..

1 Introdução

Na estatística clássica geralmente se supõe que realizações vizinhas de variáveis aleatórias são independentes entre si, ou seja, uma não exerce influência sobre a outra. Normalmente fenômenos naturais apresentam alguma relação nas variações entre valores de uma variável observadas em localizações vizinhas o que nos permite dizer que existe algum grau de dependência espacial. A análise espacial de dados surge como alternativa e/ou complemento à análise clássica da estatística, pois este tipo de análise considera as correlações entre os dados observados em

distintas localizações. Dentre os procedimentos que a literatura nos apresenta para se trabalhar com esses tipos de dados a que ganhou mais ênfase nos últimos tempos foi a Geoestatística (GUIMARÃES, 2004).

Para Landim e Sturaro (2002), a geoestatística calcula estimativas dentro de um contexto regido por um fenômeno natural com distribuição no espaço. Sendo assim, ela supõe que os valores das variáveis regionalizadas são espacialmente correlacionadas e, portanto, tendo grande aplicação principalmente para realizar estimativas e/ou simulações em locais não amostrados. De maneira geral esta técnica procura extrair de uma aparente aleatoriedade de dados as características estruturais probabilísticas do fenômeno regionalizado.

Segundo Garcia (2015), uma variável regionalizada $Z(s)$ é expressa por um modelo linear espacial gaussiano, que consiste na soma dos seguintes termos: $\mu(s)$ que representa a componente determinística estrutural de $Z(s)$, ou seja, um valor médio ou uma tendência; $\varepsilon'(s)$ que representa um termo estocástico correlacionado, que varia localmente, em que s identifica uma posição em duas dimensões representadas pelos pares coordenados (x, y); e ε'' que representa um ruído aleatório não correlacionado, com distribuição normal com média zero e variância σ^2.

Segundo Uribe-Opazo, et al. (2012) a função $\mu(s)$ pode ser decomposta numa combinação linear de covariáveis $x_j(s_i)$ (funções conhecidas associadas às coordenadas) e β_j (parâmetro desconhecido a ser estimado), em que os coeficientes associados a essas covariáveis são considerados parâmetros do modelo linear espacial gaussiano, que devem ser estimados.

Dessa forma quando essa média ou tendência é constante, na região do estudo, e a variância é constante, dentro dos limites de continuidade espacial (estacionariedade), podemos utilizar as técnicas geoestatísticas de Krigagem Ordinária e Simples. No entanto, existem situações nas quais essa tendência não é constante, ou seja, diz-se que os dados são tendenciosos. Nesse caso, segundo Andriotti (1988), as semivariâncias, que são medidas de variabilidade condicionada pela distância, crescerão mais rapidamente que a distância h - o que indica presença de deriva - e, assim, o semivariograma experimental irá superestimar o semivariograma real. Para Bohling (2005), os valores da semivariância apresentados no semivariograma continuam aumentando, mesmo além da variância global, o que é um forte indicativo de tendência direcional espacial na variável. A presença de tendência direcional em um conjunto de dados de uma determinada variável pode afetar consideravelmente a construção do semivariograma, podendo induzir, dependendo da extensão e da intensidade de amostragem, a conclusões totalmente equivocadas (STARKS e FANG, 1982). Assim existem na literatura ao menos duas opções recomendadas para trabalhar em casos nos quais os dados são tendenciosos: a) Ignorar o problema e seguir as análises utilizando uma modelagem linear ou potencial; b) Modelar a superfície de

tendência no termo $\mu(s)$, sendo que esse termo não será explicado por uma constante. Dessa forma os valores do resíduo obtidos pela diferença entre $Z(s)$ e $\mu(s)$ será estacionário, ou seja, sem tendência direcional (BOHLING, 2005).

Objetiva-se, portanto, nesse trabalho realizar um estudo comparativo da análise geoestatística de variáveis simuladas com diferentes valores de tendência direcional, quanto à incorporação ou não dessa tendência. A importância desse trabalho se justifica ao se supor que a tendência nos dados pode alterar os resultados da análise geoestatística e interferir na confecção do mapa temático.

2 Material e Métodos

Nesse trabalho realizou-se a simulação de dados por meio software livre R (R DEVELOPMENT CORE TEAM, 2008) e o pacote geoR (RIBEIRO JUNIOR e DIGGLE, 2001). Para a simulação uma variável georreferenciada com estrutura de dependência espacial e com tendência direcional foi considerada. Para essa variável foi fixado o modelo exponencial para a função semivariância. Também foram fixados os parâmetros do modelo exponencial: efeito pepita igual a 0, patamar igual a 10, e alcance igual a 20 equivalente a $\frac{1}{3}$ do alcance prático. Considerou-se uma área quadrada, sendo a amplitude dos eixos x e y de 0 a 100. Nessa área, os dados foram simulados considerando uma amostragem aleatória. Foi pré-definida também a direção de 90° (eixo Y) para a ocorrência de tendência direcional em $Z(s)$. Para a ocorrência dessa tendência direcional, foram pré determinados os valores dos termos referentes a função determinística $\mu(s)$, β_0 e β_1. Para β_0 o valor fixado é igual a 10 e para β_1, os valores foram fixados em 0,03, 0,06 e 0,6. Dessa forma classificaram-se cada um desses casos como Caso 1 ($\beta_0 = 10$ e $\beta_1 = 0,03$), Caso 2 ($\beta_0 = 10$ e $\beta_1 = 0,06$) e Caso 3 ($\beta_0 = 10$ e $\beta_1 = 0,6$) respectivamente. Os valores de β_0 e β_1 foram escolhidos de forma que a correlação entre a variável $Z(s)$ e as coordenadas do eixo Y fossem, fraca, média e forte respectivamente, para que assim se pudesse verificar o comportamento da variável simulada em cada um desses cenários. Foram realizadas 100 simulações para cada caso estudado.

Para a criação dos mapas temáticos primeiro se estimou os parâmetros do modelo geoestatístico, inclusive os valores de β_0 e β_1, que definem a tendência direcional, por meio do método de máxima verossimilhança. Em seguida realizou-se a krigagem universal (que considera a média como não sendo constante) e por fim elaborou-se os mapas temáticos.

Para a comparação entre os valores preditos com os dados simulados originais, e os valores preditos nos casos nas quais a tendência direcional foi incorporada, utilizou-se os métodos de

Exatidão Global (EG), Kappa (K) e Tau (T). A exatidão global (EG), é uma medida de avaliação utilizada para mensurar a similaridade entre o mapa de referência e o mapa modelo e segundo Anderson et al. (1976), o nível mínimo de precisão é de 0,85. O índice Kappa (K) por sua vez, fornece uma medida de similaridade entre os valores do mapa de referência e o mapa modelo classificado com baixa exatidão se $K < 0,67$, média exatidão se $0,67 < K < 0,80$ e alta exatidão se $K > 0,80$ (CONGALTON e GREEN, 1999). O índice de concordância Tau (T), também conhecido como Kappa modificado se mostra menos subjetivo e de mais fácil compreensão e utilização, esse índice segue a mesma classificação do índice Kappa (MA e REDMOND, 1995). Além dessa metodologia os valores também foram comparados utilizando a soma do quadrado das diferenças entre os valores preditos nas duas situações criadas para cada caso estudado.

3 Resultados e Discussões

De acordo com os resultados obtidos pode-se observar que ao aumentar o valor de β_1 (o efeito da tendência dessa direção), aumentou-se também o coeficiente de correlação, de acordo com a Tabela 1. Ao ajustar o modelo pelo método de máxima verossimilhança considerando a média como sendo constante, dessa forma desconsiderando a existência da tendência notou-se que as estimativas do alcance e do patamar se distanciaram do seu valor nominal (real), assim como a estimativa da média. Ainda conforme tabela 1 é possível verificar que o efeito pepita em todos os casos é igual a 0, conforme foi estabelecido nos parâmetros de criação das simulações, o que nos garante forte dependência espacial para os dados analisados. Ao observar o valor do desvio padrão das médias em cada caso, nota-se que há um aumento desse parâmetro

Tabela 1: Média das estimativas dos parâmetros encontrados nos dados originais simulados

Caso	**Correlação**	μ	σ_μ	a	ϕ_1	$\phi_1 + \phi_2$
1	0,2723	11,480	1,152	67,86	0	9,789
2	0,4975	12,950	1,197	102,50	0	13,680
3	0,9876	39,05	2,173	7205	0	1658

μ: média; σ_μ: desvio padrão da média; a: alcance; ϕ_1: efeito pepita; $\phi_1 + \phi_2$: patamar.

De acordo com esses resultados nota-se a influência da tendência direcional na estimativa dos parâmetros do modelo que expressa a função semivariância. É possível então afirmar que a tendência direcional afetou o comportamento da estimação dos parâmetros analisados. Essa afirmação pode ser confirmada quando observa-se o comportamento desses mesmos parâmetros

no caso onde a tendência direcional foi incorporada, ou seja, quando a média não foi considerada constante, e, e portanto, sem a influência dessa tendência no comportamento desses parâmetros. De acordo com a Tabela 2 observa-se que quando a tendência direcional foi incluída na estimação dos parâmetros da função semivariância e do termo determinístico os valores estimados do alcance e do patamar se mantiveram próximos do valor nominal real. Os valores para β_0 e β_1 que foram estimados estão de acordo com os valores pré-determinados em cada caso simulado indicando uma boa aproximação do método de ajuste.

Tabela 2: Média das estimativas dos parâmetros encontrados nos dados no qual houve a incorporação da tendência direcional.

Caso	β_0	σ_{β_0}	β_1	σ_{β_1}	**a**	β_1	$\phi_1 + \phi_2$
1	10,010	1,864	0,0303	0,026	50,12	0	7,928
2	10,010	1,864	0,0603	0,026	50,12	0	7,928
3	9,984	1,799	0,599	0,028	48,81	0	7,743

β_0, β_1: média, σ_β: desvio padrão da média; a: alcance; ϕ_1: efeito pepita; $\phi_1 + \phi_2$: patamar

Também comparou-se os valores obtidos da predição espacial dos valores da variável georreferenciada simulado em localizações não amostradas, entre os dados com nas quais a média foi considerada constante durante o ajuste e os dados nas quais a média não foi considerada constante durante o ajuste (na qual a tendência foi incorporada) . Para isso utilizou-se as medidas de similaridade Exatidão Global (EG), Kappa (K) e Tau (T) (Tabela 3 e Figuras 1 a 3). Também se observou os valores da soma do quadrado das diferenças entre os valores preditos pelo método da krigagem (Tabela 3 e Figura 4).

De acordo com a Tabela 3 e Figura 1 verifica-se que de acordo com a medida exatidão global no Caso 1, na qual há pouca influência da tendência direcional, os valores preditos são altamente similares, conforme pode-se ver na Tabela 2, na qual 98% dos valores estão acima da classificação mínima (0,85) para ser considerada altamente similar. Observa-se ainda que no Caso 2 ocorre uma diminuição dessa similaridade, ainda assim sendo considerado alto, conforme Tabela 3 94% dos valores são considerados altamente similares. No Caso 3, porém, percebe-se relevante falta de similaridade dos valores preditos obtidos, sendo que 53% dos valores obtidos estão abaixo de 0,85, o que nesta metodologia indica significativa diferença entre os dados obtidos. Esses resultados também são observados quando olhamos para os gráficos bloxplot em cada caso simulado.

Tabela 3: Porcentagem dos valores que estão abaixo e acima dos considerados como altamente similares para o método de exatidão global.

	$< \mathbf{85}\%$	$> \mathbf{85}\%$
CASO 1	2%	98%
CASO 2	6%	94%
CASO 3	53%	47%

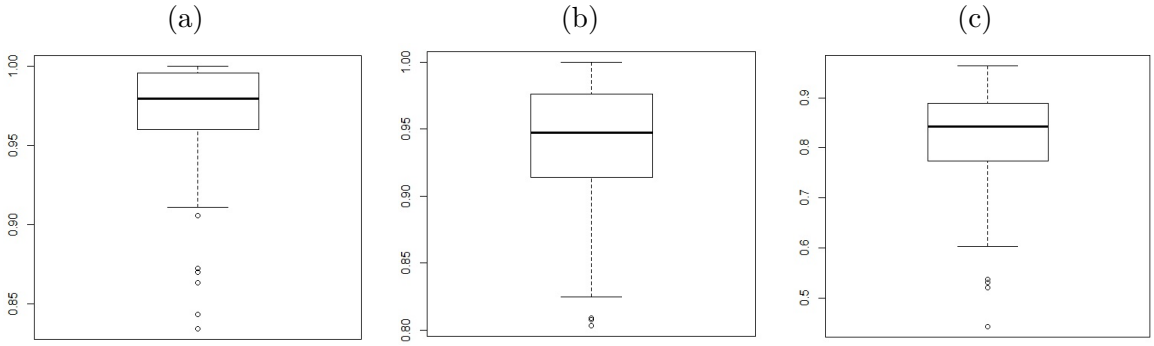


Figura 1: Comparação do método de exatidão global nos três casos de simulação: (a): Caso1; (b): Caso2; (c): Caso3.

De forma semelhante a Exatidão Global o método de comparação TAU indica alta similaridade dos dados no Caso 1, com 100% dos valores acima de 0,80 que nesta metodologia indica significativa similaridade entre os dados obtidos, conforme Tabela 4 e uma pequena diminuição dessa similaridade no Caso 2, com 97% dos valores acima de 0,80. No Caso 3 a média dos valores encontrados após as 100 simulações está abaixo de 0,80, conforme tabela 6, entretanto somente 40% desses valores estão abaixo desse valor, como pode ser visto na tabela 4, o que indica uma alta similaridade entre os dados obtidos. No entanto observa-se uma significativa diminuição dessa similaridade se comparado com os outros dois casos, e conforme figura 2.

Tabela 4: Porcentagem dos valores que estão abaixo e acima dos considerados como altamente similares para o método Tau.

	$< \mathbf{80}\%$	$> \mathbf{80}\%$
CASO 1	0%	100%
CASO 2	3%	97%
CASO 3	40%	60%

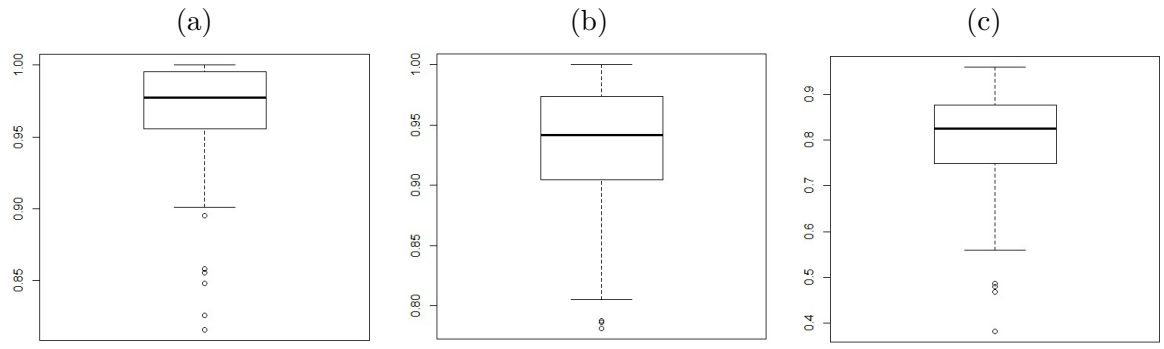


Figura 2: Comparação do método TAU nos três casos de simulação: (a): Caso1; (b): Caso2; (c): Caso3.

O método KAPPA de medida de acurácia também concorda com os outros métodos citados. De acordo com a tabela 5 e figura 3, nota-se alta similaridade dos dados preditos no Caso 1, com 100% dos valores acima de 0,80, que nesta metodologia indica significativa similaridade entre os dados obtidos e também no Caso 2, com 95% dos valores acima de 0,80. No Caso 3, na qual ocorre uma forte influência da tendência direcional, observa-se que 59% dos valores estão acima de 0,80 o que indica alta similaridade nos dados obtidos, conforme tabela 5, entretanto a média dos valores simulados fica abaixo de 0,80, conforme tabela 6. Assim como nos outros métodos de avaliação de similaridade, houve um aumento na diferença dos valores obtidos conforme se aumenta a influência da tendência direcional no conjunto de dados, ou seja, os valores encontrados no Caso 1 são mais parecidos com os valores reais do que no Caso 3, na qual há grande influência da tendência direcional.

Tabela 5: Porcentagem dos valores que estão abaixo e acima dos considerados como altamente similares para o método Kappa.

	$< \mathbf{80}\%$	$> \mathbf{80}\%$
CASO 1	0%	100%
CASO 2	5%	95%
CASO 3	41%	59%

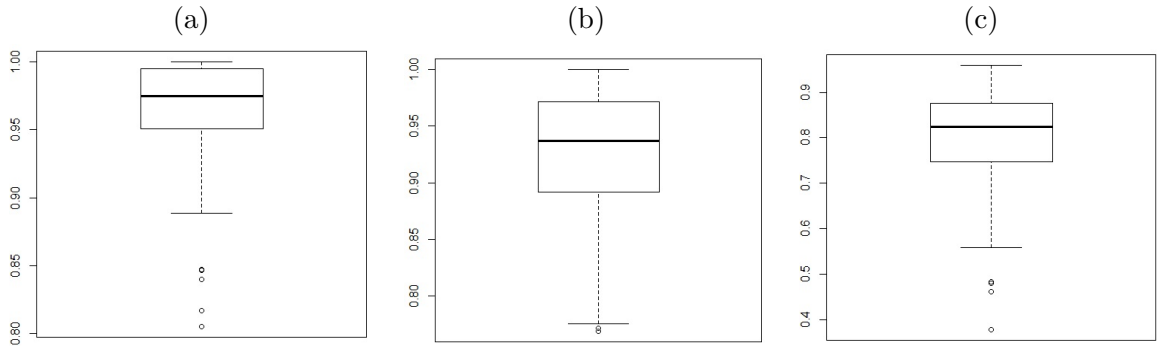


Figura 3: Comparação do método KAPPA nos três casos de simulação: (a): Caso1; (b): Caso2; (c): Caso3.

Quando observamos a soma do quadrado das diferenças, observamos que a maior diferença entre os valores preditos ocorre no Caso 3 e a menor diferença no Caso 1, havendo, portanto um aumento gradativo dessa diferença o que acompanha o aumento da correlação linear e da tendência direcional, conforme tabela 6 e figura 4.

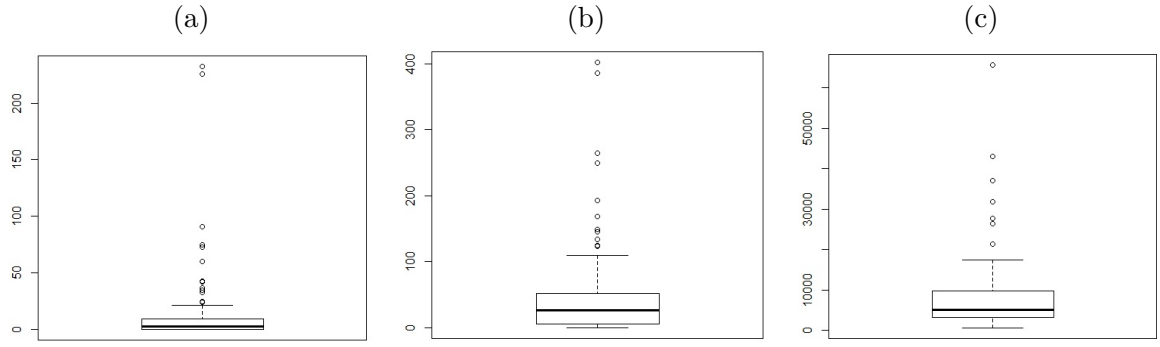


Figura 4: Comparação dos valores da soma dos quadrados da diferença nos três casos de simulação: (a): Caso1; (b): Caso2; (c): Caso3.

Tabela 6: Comparação das medidas dos índices de similaridade entre os mapas temáticos nos diferentes casos simulados.

Caso	**EG**	**TAU**	**KAPPA**	**SQD**
1	0,969	0,966	0,963	13,350
2	0,937	0,929	0,924	47,620
3	0,819	0,799	0,798	8236,000

EG: exatidão Global; SQD: soma do quadrado das diferenças.

Ao observar os gráficos boxplots nota-se alta similaridade entre os valores simulados mesmo com o decréscimo dessas medidas quando se aumentou os valores de β_1. A maior diferença ocorre quando se observa a soma quadrada das diferenças. Essa diferença pode ter sido gerada

devido a que as medidas de similaridade (EG, T, K) comparam os valores das variáveis em classes ou intervalos, já a soma quadrada das diferenças compara os exatamente os valores obtidos. Esses resultados confirmam o efeito da tendência direcional na predição de valores no conjunto de dados, pois conforme se aumenta a influência dessa tendência aumenta-se também a dissimilaridade entre os valores encontrados.

4 Considerações Finais

A geoestatística é uma importante ferramenta quando o que se deseja é predizer o comportamento de certas variáveis em uma determinada área por meio de coleta de somente alguns pontos dessa área. Isso é possível graças ao estudo da dependência espacial da variável desejada por meio da análise da correlação entre os dados observados em diferentes localizações.

Essa ferramenta terá uma melhor eficiência se na análise da variabilidade espacial de dados não estacionários, for incorporada a presença da tendência direcional por meio de uma covariável no termo determinístico.

Os resultados obtidos nesse trabalho mostram que a tendência direcional afetou os resultados quanto a eficiência da estimação do modelo geoestatístico e na predição espacial dos valores de uma variável georreferenciada em localizações não amostradas. Dessa forma, se não incorporada essa tendência pode-se ocorrer erros na predição de valores para locais não amostrados e consequentemente erros de avaliação da variável em estudo.

Referências

ANDRIOTTI, J.L.S. **Introdução à geoestatística linear**. Porto Alegre: Companhia de Pesquisas Minerais, 1988. 99p.

BOHLING, G. **Introduction to Geostatistics and Semi-Variogram Analysis**, 2005. Disponível em: <http://people.ku.edu/~gbohling/cpe940>. Acesso em: 20 Jul. 2016

CONGALTON, R. G.; GREEN, K. **Assesing the accuracy of remotely sensed data: principles and practices**. New York: Lewis Publisher, 1999, 130p.

GARCIA, T. J. F. **Geoestatística Aplicada Às Normais Climatológicas de Temperaturas Médias Compensadas no Brasil**. Dissertação de Mestrado. Universidade Federal de Lavras - UFLA. Lavras, MG. 2015. Disponível em: <http://repositorio.ufla.br/bitstream/1/9783/1/DISSERTACAO_Geoestat%C3%ADstica%20aplicada%20%C3%A0s%20normais%20climatol%C3%B3gicas%20de.pdf> Acesso em 05 Nov 2015.

GUIMARÃES, E. C. **Geoestatística Básica e Aplicada**. UFU - Universidade Federal de Uberlândia. Uberlândia, Minas Gerais, 2004.

LANDIM, P.M.B.; STURARO, J.R. **Krigagem Indicativa Aplicada à Elaboração de Mapas Probabilístico de Riscos**. DGA, IGCE, UNESP/ Rio Claro, Lab. Geomatemática, Texto Didático 06, 19 pp. 2002. Disponível em <http://www.rc.unesp.br/igce/aplicada/DIDATICOS/LANDIM/kindicativa.pdf> Acesso em 07 Abr 2015.

MA, Z. ; REDMOND, R. L. **Tau coefficients for accuracy assessment of classification of remote sensing 312 data**. Photogrammetric Engineering and Remote Sensing. Bethesda, 61(4), 453-439. 1995.

R Development Core Team. 2008. R: a language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Áustria. ISBN 3-900051 07-0, URL http://www.R-project.org.

RIBEIRO JUNIOR, P. J.; DIGGLE, P. J. **geoR: A package for geostatistic alanalysis**. RNEWS, New York, v. 1,n. 2, p. 15-18, 2001.

STARKS, T.H.; FANG, J.H. **The effect of drift on the experimental semivariogram**. Mathematical Geology, v. 14, n. 4, p. 309-319, 1982.

URIBE-OPAZO, M.A.; BORSSOI, J.A.; GALEA, M. **Influence diagnostics in Gaussian spatial linear models**. Journal of Applied Statistics, 39:3, 615-630. 2012. doi: 10.1080/02664763.2011.607802

Perfil do emprego na construção civil

Gustavo Rosa[1]
Universidade Estadual do Oeste do Paraná
gugo_rosa@hotmail.com

Rosangela Villwock
Universidade Estadual do Oeste do Paraná
rosangela.villwock@unioeste.br

Resumo: Diante da busca constante pelo desenvolvimento econômico nacional, o governo brasileiro tem incentivado nos últimos anos a construção civil por meio de diversos investimentos. Do fato do aumento de investimentos, o número de obras tende a crescer e por consequência também o contingente de trabalhadores. Diante dessa realidade, este trabalho tem como objetivo delinear o perfil do emprego na construção civil brasileira, por meio da Análise de Agrupamentos. Os grupos encontrados foram caracterizados conforme seus centroides. A partir da avaliação dos resultados obtidos, observou-se que no grupo onde o índice de primeiro emprego e reemprego são maiores, o contingente de trabalhadores com escolaridade inferior é maior. Além disso, no grupo onde a taxa de construção de rodovias, ferrovias, obras urbanas e obras de arte especiais é maior, o índice de trabalhadores com grau de escolaridade elevado também é maior.

Palavras-chave: mineração de dados; K-Médias; emprego na construção civil.

1 Introdução

A construção civil no Brasil é uma área de grande importância para a economia, dada a sua capacidade de gerar as bases para o desenvolvimento do país. Tal importância pode ser notada pelo setor da construção civil produzir a infraestrutura econômica necessária para o funcionamento de outros setores da economia, ou seja, a partir da construção de portos, aeroportos, rodovias, ferrovias, sistemas de fornecimento de energia, permite-se o funcionamento correto de serviços primários, secundários e terciários. Destarte, a capacidade e qualidade do setor construtivo, influencia diretamente o desenvolvimento dos mais diversos setores da economia nacional (TEIXEIRA; CARVALHO, 2005).

Nos últimos anos quando se trata do setor da construção civil, o seu aumento considerável tem sido impulsionado através de programas habitacionais de cunho governamental e o Programa de Aceleração do Crescimento (PAC), tendo como consequência o aumento da procura de mão de obra envolvida diretamente no setor citado. Além de questões habitacionais e o PAC, investimentos feitos sobre o país como realização da Copa do Mundo de 2014 e dos Jogos

[1]O autor agradece à Fundação Araucária por bolsa de iniciação científica

Olímpicos de 2016 também estão envolvidos com o crescimento e destaque da construção civil (CARDOSO, 2013).

Conforme dados levantados pelo CAGED (Cadastro Geral de Empregados e Desempregados) em 2014, as admissões envolvidas no setor em pauta apontam cerca de 704 mil profissionais contratados. Porém, quando em confronto com os dados obtidos em 2013 pela mesma fonte, apresentou uma decadência de aproximadamente 2,52%. Ainda, o CAGED (2014) destaca elementos em relação a demissões, indicando uma leve e considerável expansão de 0,13%. Contudo, a mesma resposta é considerada de certa forma como uma consequência da taxa avaliada de emprego, com uma redução de 23,84%. Já quando comparamos índices de desemprego na construção civil, entre 2003 que apresentava porcentagem de 8,9%, e 2014 que apontava proporção de 2,5%, podemos afirmar que o mesmo sofreu uma perda expressiva.

Nesse sentido, frente à grande capacidade de geração de emprego e influência na economia causada pela construção civil, o projeto de pesquisa de iniciação científica buscará delinear o perfil do emprego neste setor.

Um meio para atingir tal objetivo é a análise minuciosa de dados, a qual combina diferentes áreas, tais como a Estatística, Matemática e nesse caso a Construção Civil. A análise será efetuada a partir de técnicas de agrupamento de dados, ou seja, por meio de uma das áreas mais promissoras e úteis da atualidade: a Descoberta de Conhecimento em Bases de Dados, a qual busca a organização e ordenação de tais dados, com o objetivo de prover os mesmos de um sentido dentro de um contexto, gerando assim a informação e, em um estágio final, a partir dos dados e informações, a geração do conhecimento. Uma definição mais formal para a Descoberta do Conhecimento em Base de Dados ou simplesmente KDD (sigla que corresponde ao termo em inglês, Knowledge Discovery in Databases) foi dada por Fayyad et al. (1996): é um processo não trivial, que tem como objetivo a identificação de padrões compreensíveis, válidos, novos e úteis a partir de grandes conjuntos de dados, de forma interativa e iterativa. A Descoberta de Conhecimento em Bases de Dados pode se dividir em: seleção, pré-processamento, formatação, mineração de dados e interpretação.

Segundo Goldschimidt, Passos e Bezerra (2015) o processo de KDD está inserido em um cenário de complexidade, dada à dificuldade de percepção frente aos diversos elementos que surgem durante o processo e como forma de orientação pode ser utilizado o modelo CRISP-DM, o qual tem como objetivo fornecer as orientações básicas para o processo de Descoberta de Conhecimento em Bases de Dados.

2 Materiais e Métodos

Os dados foram obtidos na CBIC – Câmara Brasileira da Indústria da Construção (disponível em http://www.cbicdados.com.br/menu/emprego/rais-ministerio-do-trabalho-e-emprego), cujos dados são retirados da Relação Anual de Informações Sociais (RAIS) do Ministério do Trabalho –, as quais se referem ao ano de 2014, sendo elas: mulheres empregadas; primeiro emprego; reemprego; analfabetos empregados; trabalhadores com: primeiro ciclo completo, fundamental completo, e ensino superior completo; construção de rodovias, ferrovias, obras urbanas e obras de arte especiais; construção de outras obras de infraestrutura. Referentes aos 26 estados da federação e o Distrito Federal.

Dentre os diversos algoritmos de Agrupamento de Dados, para Tan, Steinbach e Kumar (2006), os métodos de Agrupamento fundamentais podem ser classificados em Método de Partição, Método Hierárquico, Método Baseado em Densidade e Método Baseado em Malhas. Neste trabalho um Método de Partição (o K-Médias) foi utilizado.

Os métodos de partição ou particionamento são explicados por Tan, Steinbach e Kumar (2006) da seguinte forma: dado um número n de objetos, um método de particionamento constroi k partições, onde cada partição representa um grupo e $k \leq n$, ou seja, o método divide o conjunto de dados em k grupos, devendo cada grupo conter pelo menos um objeto.

Segundo Johnson e Wichern (1998), o método de particionamento mais conhecido é o K-Médias, sendo que os k grupos produzidos normalmente são de melhor qualidade em comparação com os produzidos por métodos hierárquicos. O funcionamento básico do K-Médias é baseado na definição de k centroides iniciais (pontos do conjunto de dados, por exemplo) e em seguida cada ponto do conjunto de dados é atribuído ao grupo representado pelo centroide de menor distância. Após a alocação inicial, o método segue iterativamente atualizando os centroides de cada grupo (calculando o ponto médio dos pontos alocados ao grupo) e realocando cada ponto do conjunto de dados ao centroide mais próximo, até que um critério de parada seja satisfeito.

3 Resultados e Discussão

A área da construção civil empregou 3.019.427 de trabalhadores, sendo notadamente a contribuição da região sudeste, para o total de empregados superior as demais regiões. Desse total de empregados, pode-se ainda segmentar os trabalhadores conforme o sexo. Em 2014, observou-se um número de 286.317 mulheres empregadas e 2.733.110 de homens. O contexto mostra que o ramo da construção civil ainda demostra uma grande disparidade entre o total de

homens e mulheres empregados. De forma percentual com referência ao total de empregados, ao que tange às mulheres empregadas, o estado que mostrou menor valor foi o de Piauí com 5,20% e o maior foi o de Amazonas com 12,97%. O desvio padrão, tanto do total quanto dos homens e mulheres empregados, é maior que a média, o que leva a um coeficiente de variação acentuado. Sendo assim há uma grande variabilidade dos dados com relação à média e por consequência pode-se inferir que o conjunto de dados apresenta uma heterogeneidade.

Dos empregados no setor, alguns foram contratados em 2014, sendo que parte teve o primeiro emprego e parte foi reempregada no setor da construção civil, o primeiro representa 3,52% e o segundo 47,63% do total de empregados no ano. O total de carteiras assinadas foi de 1.544.520, representando 51,15% do total, 3.019.427. A partir desta observação, de um grande número de admitidos, é possível notar o quão o mercado da construção civil estava aquecido e necessitando de profissionais no ano em questão. Proporcionalmente ao total de empregados os estados do Rio de Janeiro com 2,29% e Acre com 7,59% figuraram os extremos da amostra quanto ao primeiro emprego. E quanto ao reemprego, Pernambuco com 40,18% e Tocantins com 59,10% apresentaram o menor e o maior valor proporcional respectivamente. O conjunto de dados referente às variáveis é heterogêneo, variando muito de estado para estado.

No que afeta o nível de escolaridade foram observadas as variáveis analfabetos, primeiro ciclo do fundamental completo, fundamental completo, ensino médio completo e ensino superior completo, não sendo computados os trabalhadores que deixaram a escola em algum ponto que não fosse a completude seja do ensino fundamental primeiro ciclo, fundamental segundo ciclo, ensino médio ou ensino superior. Proporcionalmente ao total de empregados, os índices extremos de analfabetos se encontraram nos estados do Rio de Janeiro com 0,50% e Pernambuco 2,63%; do primeiro ciclo do ensino fundamental completo no Tocantins com 3,90% e Piauí com 10,13%; para o ensino fundamental completo no Acre com 11,24% e Santa Catarina com 21,57%; com relação ao ensino médio completo no Piauí com 21,68% e Roraima com 52,99%; e por fim para o ensino superior completo em Rondônia com 2,67% e São Paulo com 7,43%. Todas as variáveis apresentaram uma alta variação entre cada unidade federação, sendo a variação referente ao ensino superior completo a mais marcante, representando um coeficiente de variação, em porcentagem, de aproximadamente 184%.

Ao que compete às obras na construção civil selecionou-se construção de rodovias, ferrovias, obras urbanas e obras de arte especiais (pontes, viadutos, etc) e construção de outras obras de infraestrutura. Proporcionalmente ao total de empregados quanto à construção de rodovias, ferrovias, obras urbanas e obras de arte especiais os estados de Roraima com 0,52% e Rondônia com 18,57%, se encontraram nos extremos. Já para construção de outras obras de

infraestrutura Roraima com 1,52% e Amapá com 24,13% representaram as duas extremidades da amostra. A variação de cada variável entre estes estados foi de alta a extremamente alta, sendo clara a heterogeneidade de dados. Tal heterogeneidade pode ser devido a diversos fatores, dentre os quais pode-se citar a situação econômica de cada estado, o que gera a demanda para o desenvolvimento da construção civil.

Os resultados apresentados abaixo foram obtidos pela aplicação do método K-Médias à base de dados criada, sendo tais analisados por dois procedimentos: análise de cada variável e, posteriormente, de cada grupo.

Ao que se refere à distribuição de gênero, foi anteposto a variável mulheres empregadas (% Feminino), como mostra a Tabela 1. É possível observar que o valor para variável percentual de emprego feminino no centroide variou de 9,16% a 9,80%. E, por consequência, os percentuais de emprego masculino tiveram uma variação de 90,20% no grupo 2 a 90,84% no grupo 1. Nota-se que não houve uma variação determinante entre os valores desta variável nos centroides.

No que tange às admissões no período, foram selecionados apenas as variáveis primeiro emprego e reemprego, a porcentagem faltante se refere principalmente aos trabalhadores já empregados em anos anteriores. Verifica-se na Tabela 1, que os valores da variável primeiro emprego no centroide variaram de 3,49% a 5,44%. Já o reemprego apresentou uma variação de 46,89% a 51,06%. A variação referente ao primeiro emprego foi acentuada, no grupo 1 é 55,87% maior que no grupo 2. Ambas variáveis representam um relativo crescimento na construção civil, pois o aumento no número de empregados é um indicativo de mais obras em andamento.

Tabela 1: Valores para as variáveis emprego feminino, primeiro emprego e reemprego no centroide.

Grupo	% Feminino	% Primeiro emprego	% Reemprego
Grupo 1	9,16	5,44	51,06
Grupo 2	9,80	3,49	46,89

No que concerne o nível de escolaridade foram analisados: analfabetos, fundamental primeiro ciclo completo, fundamental completo, ensino médio completo, ensino Superior completo. Os desistentes em algum nível escolar não foram computados nas variáveis. A Tabela 2 mostra os valores das variáveis no centroide.

Tabela 2: Valor das variáveis analfabetos, fundamental primeiro ciclo completo, fundamental completo, ensino médio completo, ensino superior completo no centroide.

Grupo	% Analfabetos	% Fundamental: primeiro ciclo completo	% Fundamental completo	% Ensino médio completo	% Ensino superior completo
Grupo 1	1,46	6,06	14,71	39,26	3,44
Grupo 2	0,75	6,28	17,05	38,21	4,80

Ao que compete aos analfabetos a variação foi de 0,75% a 1,46%, representando 94,67% de variação entre os grupos. A variável fundamental primeiro ciclo completo, a qual se refere a completude de 1º ao 5º ano, variou de 6,06% a 6,28%, sendo maior no grupo 2 em 3,63%. Para o fundamental completo o valor da variável no centroide variou de 14,71% a 17,05%, com variação entre o grupo 1 e o grupo 2 de 15,91%. Já a variável ensino médio completo teve uma variação de 38,21% a 39,26%, estando o grupo 1 maior em 2,75%. Por fim, a variável ensino superior completo variou entre 3,44% a 4,80%, no qual o grupo 2 é 40% maior que o grupo 1.

Dentre estas variáveis, pode-se verificar uma grande variação na variável analfabetos, sendo que no grupo 1 é 94,67% maior que no grupo 2. Esta foi a variação mais notável seguida da variável ensino superior completo, sendo que no grupo 2 é 40% maior que no grupo 1. A grande variação no atributo analfabetos pode ser interpretada como uma demanda de profissionais para atribuições que não exigem conhecimento específico, sendo uma opção na falta de trabalhadores com um nível de escolaridade mais alto.

Para a modalidade de construção, as variáveis utilizadas foram a construção de rodovias, ferrovias, obras urbanas e obras de arte especiais e construção de outras obras de infraestrutura. Como mostra a Tabela 3, é possível verificar que para a variável construção de rodovias, ferrovias, obras urbanas e obras de arte o valor no centroide variou de 7,68% a 10,94%, sendo que entre grupos a variação atingiu 42,45%. Já para a variável de construção de outras obras de infraestrutura ocorreu uma variação de 7,90% a 9,72%, com o grupo 2 sendo 23,04% maior que o grupo 1. A variável construção de rodovias, ferrovias, obras urbanas e obras de especiais apresentou uma discrepância relevante, a terceira maior no conjunto de variáveis.

Tabela 3: Valor para as variáveis na construção de rodovias, ferrovias, obras urbanas e obras de arte especiais e construção de outras obras de infraestrutura no centroide.

Grupo	% Construção de rodovias, ferrovias, obras urbanas e obras de arte especiais	% Construção de outras obras de infraestrutura
Grupo 1	7,68	7,90
Grupo 2	10,94	9,72

Após a análise de cada variável os grupos podem ser caracterizados, interpretando-se os aspectos mais relevantes.

O grupo 1 ficou caracterizado por apresentar as maiores taxas de primeiro emprego e reemprego. Este grupo também apresentou um índice de analfabetismo muito elevado com relação ao grupo 2. Além de a construção de rodovias, ferrovias, obras urbanas e obras de arte especiais e construção de outras obras de infraestrutura serem diminutas. O grupo em questão é formado por 11 estados, caracterizado essencialmente por estados do Nordeste e do Norte e apenas um do Centro Oeste, possuindo 914.397 trabalhadores no setor, aproximadamente 30,28% do total.

O grupo 2 apresentou um maior número de estados, 16 no total, com um contingente de trabalhadores de 2.105.030, por volta de 69,72%. Constituído por todos os estados do Sul e Sudeste, além de 4 estados do Nordeste, 2 estados do Norte e 3 estados do Centro Oeste. Este grupo ficou caracterizado por possuir um percentual de construção de rodovias, ferrovias, obras urbanas e obras de arte especiais e construção de outras obras de infraestruturas consideravelmente mais alto em comparação com o grupo 1. Este grupo comporta o maior contingente de graduados.

4 Conclusão

A partir da análise dos resultados obtidos foi possível observar que no grupo onde o índice de primeiro emprego e reemprego são maiores, o contingente de trabalhadores com escolaridade inferior é maior, podendo tal fato decorrer da necessidade imediata de profissionais e o estoque de trabalhadores com maior grau de instrução não ter sido capaz de absorver toda a demanda.

Além disso, no grupo onde a taxa de construção de rodovias, ferrovias, obras urbanas

e obras de arte especiais é maior, o índice de trabalhadores com grau de escolaridade elevado também é maior, podendo tal fato ser interpretado como, onde há uma maior demanda de obras, há um maior número de trabalhadores e, por consequência, uma elevada procura de trabalhadores com o ensino superior completo. Observa-se aqui a importância na qualificação do trabalhador quando programadas novas obras.

Referências

CARDOSO, F. H. Incentivo do estado e desenvolvimento: uma análise sobre o crescimento da área da construção civil. 2013. 9 f. Dissertação (Mestrado) - Curso de Ciências Sociais, Universidade Estadual de Londrina, Londrina, 2013.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHRUSAMY, R. **Advances in knowledge Discovery & Data Mining**. California: AAAI/MIT, 1996.

GOLDSCHIMIDT, R; PASSOS, E; BEZERRA, E. **Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações**. 2 ed. Rio de Janeiro: Elsevier Editora Ltda, 2015.

HAIR JR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. **Análise Multivariada de Dados**. Tradução de: SANTANNA, A. S.; CHAVES NETO. 5 ed. A. Porto Alegre: Bookman, 2005.

Johnson, R.A.; Wichern, D.W.. **Applied Multivariate Statistical Analysis**. New Jersey: Prentice Hall, 2007.

SINDICATO DE ARQUITETURA E ENGENHARIA. **Movimentação do emprego no setor da arquitetura e engenharia consultiva**. 2014.

TAN, P. N; STEINBACH, M; KUMAR, V. **Introdução ao DATA MINING Mineração de Dados**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2009.

TEIXEIRA, L. P.; CARVALHO, F. M. A. de. A construção civil como instrumento do desenvolvimento da economia brasileira. Revista Paranaense de Desenvolvimento, Curitiba, v. 109, p.9-26, jul. 2005.

Teorema da Função Implícita

Bruno Belorte[1]
Acadêmico do Curso de Matemática - CCET
Universidade Estadual do Oeste do Paraná
Caixa Postal 711 - 85819-110 - Cascavel - PR - Brasil
bruno.belorte@gmail.com

André Vicente
Docente do Curso de Matemática - CCET
Universidade Estadual do Oeste do Paraná
Caixa Postal 711 - 85819-110 - Cascavel - PR - Brasil
andre.vicente@unioeste.br

Resumo: Neste trabalho apresentamos uma demonstração do Teorema da Função Implícita. A essência da demonstração consiste em criar uma função auxiliar a qual permite utilizar o Teorema da Função Inversa.

Palavras-chave: Teorema da Função Implícita; Análise matemática; Análise no Rn.

1 Introdução

A ideia central do Teorema da Função Implícita é estabelecer condições para que: dada uma equação

$$F(x, y) = 0$$

então seja possível dizer que há funções definidas implicitamente pela equação acima, ou seja, existe uma função φ tal que $y = \varphi(x)$ e

$$F(x, \varphi(x)) = 0.$$

Na seção 3 apresentaremos um exemplo mais detalhado.

Para a demonstração do Teorema da Função Implícita usaremos o clássico Teorema da Função Inversa, o qual garante, mediante hipóteses apropriadas, que uma dada função é um difeomorfismo local. Assim, dedicamos uma seção à uma breve discussão do Teorema da Função Inversa.

Este texto é resultado de um trabalho de iniciação científica no qual estudamos o Teorema da Função Implícita. O livro texto para a realização dos trabalhos foi (CIPOLATTI, 2002), no qual baseamos estas notas e onde podem ser consultados maiores detalhes.

[1]Agradecemos ao PIBIC/UNIOESTE pelo apoio financeiro e pelo incentivo à pesquisa.

Este trabalho está organizado da seguinte forma: na seção 2 discutimos o Teorema da Função Inversa; na seção 3 apresentamos o resultado principal, ou seja, Teorema da Função Implícita e, finalmente, na seção 4 apresentamos uma aplicação do Teorema da Função Implícita.

2 Teorema da Função Inversa

Se $\Omega \subset \mathbb{R}^n$ é um conjunto aberto e $f : \Omega \to \mathbb{R}^n$ é uma função diferenciável em $x_0 \in \Omega$, dizemos que $f'(x_0) : \mathbb{R}^n \to \mathbb{R}^n$ é função linear que melhor aproxima f nas proximidades de x_0. Para uma melhor explicação, relembremos a definição de uma função diferenciável ou Frechét-derivável em $x_0 \in \Omega$ se existem funções $L, \varepsilon : \mathbb{R}^n \to \mathbb{R}^m$ tais que

$$f(x_0 + h) = f(x_0) + L(h) + \varepsilon(h), \tag{1}$$

com L linear e ε função de $o(\|h\|)$, isto é $\lim_{h\to 0} \frac{\|\varepsilon(h)\|}{\|h\|} = 0$. Denotamos L por $f'(x_0)$. Prova-se que a matriz de $f'(x_0)$, na base canônica do $\mathbb{R}^n$, é dada por

$$[f'(x_0)] = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \frac{\partial f_n}{\partial x_2}(x_0) & \dots & \frac{\partial f_n}{\partial x_n}(x_0) \end{bmatrix}$$

onde $f_1, \dots, f_n$ são as funções coordenadas de f. A matriz acima é chamada de matriz Jacobiana e seu determinante é chamado de Jacobiano.

Da expressão (1) temos que

$$f(x_0 + h) - f(x_0) = L(h) + \varepsilon(h). \tag{2}$$

Podemos interpretar a equação (2) dizendo que para um x_0 fixo e um h pequeno, o lado esquerdo da equação é aproximadamente igual a $f'(x_0) = L(h)$, ou seja, o valor da transformação linear aplicada a h.

Como motivação para o estudo do Teorema da Função Inversa, vamos considerar uma função linear $g : \mathbb{R}^n \to \mathbb{R}^n$ definida por $g(x) = Ax$ onde A é uma matriz quadrada $n \times n$. Sabemos que se $\det A \neq 0$ então a matriz é inversível, logo g é inversível e sua inversa é dada por $g^{-1}(x) = A^{-1}x$. Prova-se que g e g^{-1} são diferenciáveis em $\mathbb{R}^n$, $g'(x_0) = A$ e $(g^{-1})(x_0) = A^{-1}$, para qualquer $x_0 \in \mathbb{R}^n$.

Assim, parece plausível que, se $f : \Omega \to \mathbb{R}^n$ é diferenciável em $x_0 \in \Omega$ e $J_f(x_0) := \det[f'(x_0)] \neq 0$ então f é invertível nas proximidades de x_0. Porém não é isto o que ocorre mesmo quando $n = 1$, vamos a um exemplo.

Exemplo 1. Seja $f : \mathbb{R} \to \mathbb{R}$ tal que

$$f(x) = \begin{cases} \frac{x}{2} + x^2 \operatorname{sen} \frac{1}{x} & \text{se } x \neq 0, \\ 0 & \text{se } x = 0. \end{cases}$$

Para o caso onde $x \neq 0$ calculamos a derivada usando as regras de derivação, assim

$$\begin{aligned} f'(x) &= \frac{d}{dx}\frac{x}{2} + \frac{d}{dx}\left(x^2 \operatorname{sen} \frac{1}{x}\right) \\ &= \frac{1}{2} + \left[\left(\frac{d}{dx}x^2\right) \operatorname{sen} \frac{1}{x} + x^2 \left(\frac{d}{dx} \operatorname{sen} \frac{1}{x}\right)\right] \\ &= \frac{1}{2} + 2x \operatorname{sen} \frac{1}{x} + x^2 \cos \frac{1}{x} \left(\frac{-1}{x^2}\right) \\ &= \frac{1}{2} + 2x \operatorname{sen} \frac{1}{x} - \cos \frac{1}{x}. \end{aligned}$$

Para o caso $x = 0$ temos que calcular a derivada pela definição, logo

$$\begin{aligned} f'(0) = \lim_{h \to 0} \frac{f(0+h) - f(0)}{h} &= \lim_{h \to 0} \frac{\frac{h}{2} + h^2 \operatorname{sen} \frac{1}{h} - 0}{h} \\ &= \lim_{h \to 0} \left(\frac{1}{2} + h \operatorname{sen} \frac{1}{h}\right) = \frac{1}{2}. \end{aligned}$$

Deste modo temos a função

$$f'(x) = \begin{cases} \frac{1}{2} + 2x \operatorname{sen} \frac{1}{x} - \cos \frac{1}{x} & \text{se } x \neq 0, \\ \frac{1}{2} & \text{se } x = 0. \end{cases}$$

que é diferenciável em todos os pontos de $\mathbb{R}$. Se f fosse invertível na vizinhança $x_0 = 0$ então seria necessariamente injetora nessa vizinhança. E como $f'(0) = \frac{1}{2} > 0$, seria necessariamente crescente nessa vizinhança. Mas isto não é possível pois $f'(x)$ muda de sinal infinitas vezes, como mostra a figura 1, em qualquer vizinhança que contenha $x_0 = 0$.

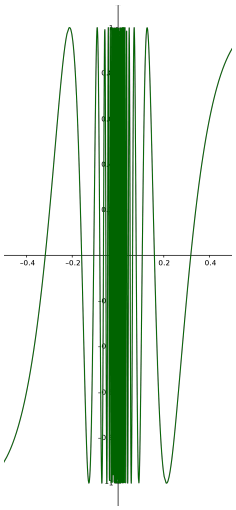


Figura 1: Ideia do gráfico da função $\frac{1}{2} + 2x \operatorname{sen} \frac{1}{x} - \cos \frac{1}{x}$

Então, só teríamos o resultado desejado se f' fosse contínua em $x_0 = 0$, logo $f'(x) > 0$ para todo x suficientemente próximo de $x_0 = 0$.

Os espaços vetoriais normados para os quais todas as sequências de Cauchy são covergentes são denominados Espaços de Banach. Estes são fundamentais para Análise, pois neles ficam assegurados os processos de limite. O Teorema da Função Inversa ocorre para funções $f : V \to V$ onde V é um espaço de Banach. No nosso caso, consideramos $\mathbb{R}^n$ que é um espaço de Banach.

Teorema 2 (Teorema da Função Inversa)**.** *Seja $\Omega \subset \mathbb{R}^n$ aberto e $f : \Omega \to \mathbb{R}^n$ função de classe C^1 tal que $J_f(x_0) \neq 0$. Então existe $\delta_0 > 0$ tal que*

1. *A função f é injetora em $U = B_{\delta_0}(x_0)$;*
2. *O conjunto $V = f(U)$ é aberto;*
3. *A inversa $f^{-1} : V \to U$ é de classe C^1 e $[(f^{-1})'(f(x_0))] = [f'(x_0)]^{-1}$.*

Vejamos um exemplo.

Exemplo 3. Sejam $f_1(x, y) = e^x \cos y$ e $f_2(x, y) = e^x \operatorname{sen} y$. Para a transformação $f : \mathbb{R}^2 \to \mathbb{R}^2$ com $(x, y) \to f((x, y)) = (f_1(x, y), f_2(x, y))$, temos

$$f'(x, y) = \begin{bmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{bmatrix} = \begin{bmatrix} e^x \cos y & -e^x \operatorname{sen} y \\ e^x \operatorname{sen} y & e^x \cos y \end{bmatrix}$$

e assim

$$\det[f'(x, y)] = e^{2x}(\cos^2 y + \operatorname{sen}^2 y) = e^2 x \neq 0.$$

Pelo Teorema da Função Inversa, temos que a aplicação f é localmente inversível.

3 O Teorema da Função Implícita

Nesta seção apresentaremos o resultado principal deste trabalho. Inicialmente faremos uma descrição informal dos objetivos relacionados à conclusão do Teorema.

Em $\mathbb{R}^2$ temos que o gráfico de uma função $f(x)$ é o conjunto de todos os pontos da forma (x, y) tais que $f(x) = y$, sendo que podemos ver o gráfico da função como uma curva a qual toda linha vertical cruza a curva somente uma vez. Porém há outros tipos de curvas no plano

as quais não passam no teste da linha vertical. Uma destas é a circunferência unitária, a qual é descrita como $x^2+y^2=1$. Ou seja, o círculo de raio 1 centrado na origem é o conjunto de todos os pontos (x,y) tais que $x^2+y^2=1$. Sabemos que se removermos uma metade do círculo, como a metade inferior, teremos o gráfico de uma função. Esta função obtida é o que chamamos de função implícita.

O Teorema da Função Implícita está relacionado com a solução do seguinte problema:

Dada $f:\mathbb{R}^{k+m}\to\mathbb{R}^m$ e $(x_0,y_0)\in\mathbb{R}^{k+m}$ tal que $f(x_0,y_0)=0$, deseja-se saber se existe $\Omega\subset\mathbb{R}^k$ aberto e uma função $\varphi:\Omega\to\mathbb{R}^m$ tal que

1. O elemento $x_0\in\Omega$ cumpre $\varphi(x_0)=y_0$;
2. A função $f(x,\varphi(x))=0$, para qualquer $x\in\Omega$.

No caso em que $m=k=1$ podemos obter uma resposta para o problema acima por meio da Teoria das Equações Diferenciais Ordinárias. De fato, supondo f e φ diferenciáveis, temos pela Regra da Cadeia

$$\frac{\partial f}{\partial x}+\varphi'(x)\frac{\partial f}{\partial y}=0.$$

Se $f\in C^1$ então $\frac{\partial f}{\partial y}\neq 0$, podemos obter φ como solução do problema de valor inicial

$$\begin{cases}\frac{\partial\varphi}{\partial x}=\Phi(x,\varphi)\\ \varphi(x_0)=y_0,\end{cases}\tag{3}$$

onde estamos denotando

$$\Phi(x,y)=-\left(\frac{\partial f}{\partial y}(x,y)\right)^{-1}\frac{\partial f}{\partial x}(x,y).$$

Com as hipóteses estabelecidas sobre f temos que φ enquadra-se nas hipóteses do Teorema de Picard, o qual garante a existência de solução para (3).

Podemos tratar a equação acima com o Teorema da Função Inversa. Consideremos o caso particular onde $f:\mathbb{R}^n\to\mathbb{R}^m$, a função linear definida por $f(z)=Az$, onde A é a matriz $m\times n$ com $n=k+m$. Denotando $z=(x,y)=(x_1,\ldots,x_k,y_1,\ldots,y_m)$, como f é linear, podemos escrever $f(x,y)=Az=Bx+Cy$, onde B e C são submatrizes de ordem $m\times k$ e $m\times m$, isto é, $A=[B\quad C]$ é composta por dois blocos B e C.

Observemos que neste caso particular em que $n=k=1$, os blocos B e C são as derivadas parciais de f, assim

$$B=\tfrac{\partial f}{\partial x}(x_0,y_0)\quad,\quad C=\tfrac{\partial f}{\partial y}(x_0,y_0)$$

e

$$\varphi = -\left[\frac{\partial f}{\partial y}(x_0, y_0)\right]^{-1}\left[\frac{\partial f}{\partial x}(x_0, y_0)\right]. \tag{4}$$

O modo para resolver este problema por meio do Teorema da Função Inversa pode ser observado se reescrevermos a equação $f(x, y) = 0$. Seja $F : \mathbb{R}^n \to \mathbb{R}^n$ com $n = k + m$, a função linear definida por $F(x, y) = (x, f(x, y))$. Então $F(z) = \mathcal{A}z$ onde $\mathcal{A}$ é a matriz

$$\begin{bmatrix} I_k & 0 \\ B & C \end{bmatrix},$$

onde I_k é a matriz identidade de ordem $k \times k$ e 0 é a matriz nula de ordem $k \times m$. Temos que $\det \mathcal{A} = \det C$. Assim se C é inversível, a matriz $\mathcal{A}$ também o é, sendo que

$$\mathcal{A}^{-1} = \begin{bmatrix} I_k & 0 \\ -C^{-1}B & C^{-1} \end{bmatrix}.$$

Portanto

$$F(x, y) = (x, 0) \Leftrightarrow F^{-1}F(x, y) = F^{-1}(x, 0) \Leftrightarrow (x, y) = F^{-1}(x, 0) = (x, -C^{-1}Bx)$$

daqui

$$Y = -C^{-1}Bx,$$

que é equivalente a (4).

Teorema 4 (Teorema da Função Implícita)**.** *Seja $f : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$ uma função de classe C^1. Suponha $f(x_0, y_0) = 0$ e*

$$\det\left[\frac{\partial f}{\partial y}(x_0, y_0)\right] \neq 0$$

Então existe um conjunto aberto $\Omega \subset \mathbb{R}^k$ e $\varphi : \Omega \to \mathbb{R}^m$ função de classe C^1 tais que

1. *O elemento $x_0 \in \Omega$ cumpre $\varphi(x_0) = y_0$;*
2. *A função $f(x, \varphi(x)) = 0$, para qualquer $x \in \Omega$.*

Prova. Seja a função definida

$$\begin{aligned} F : \mathbb{R}^k \times \mathbb{R}^m &\to \mathbb{R}^k \times \mathbb{R}^m \\ (x, y) \mapsto F(x, y) &= (x, f(x, y)) = (x_1, \ldots, x_k, f_1(x, y), \ldots, f_m(x, y)) \\ &= (x_1, \ldots, x_k, z_1, \ldots, z_m). \end{aligned}$$

Como $x, f(x, y) \in C^1$ então $F \in C^1$ e a matriz jacobiana de F em $z_0 = (x_0, y_0)$, $[F'(z_0)]$ é igual a

$$\left(\begin{array}{cccc:cccc}
\frac{\partial F_1}{\partial x_1}(z) & \frac{\partial F_1}{\partial x_2}(z) & \dots & \frac{\partial F_1}{\partial x_k}(z) & \frac{\partial F_1}{\partial y_{k+1}}(z) & \frac{\partial F_1}{\partial y_{k+1}}(z) & \dots & \frac{\partial F_1}{\partial y_{k+m}}(z) \\
\frac{\partial F_2}{\partial x_1}(z) & \frac{\partial F_2}{\partial x_2}(z) & \dots & \frac{\partial F_2}{\partial x_k}(z) & \frac{\partial F_2}{\partial y_{k+1}}(z) & \frac{\partial F_2}{\partial y_{k+1}}(z) & \dots & \frac{\partial F_2}{\partial y_{k+m}}(z) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial F_k}{\partial x_1}(z) & \frac{\partial F_k}{\partial x_2}(z) & \dots & \frac{\partial F_k}{\partial x_k}(z) & \frac{\partial F_k}{\partial y_{k+1}}(z) & \frac{\partial F_k}{\partial y_{k+1}}(z) & \dots & \frac{\partial F_k}{\partial y_{k+m}}(z) \\
\hdashline
\frac{\partial F_{k+1}}{\partial x_1}(z) & \frac{\partial F_{k+1}}{\partial x_2}(z) & \dots & \frac{\partial F_{k+1}}{\partial x_k}(z) & \frac{\partial F_{k+1}}{\partial y_1}(z) & \dots & \frac{\partial F_{k+1}}{\partial y_{m-1}}(z) & \frac{\partial F_{k+1}}{\partial y_m}(z) \\
\frac{\partial F_{k+2}}{\partial x_1}(z) & \frac{\partial F_{k+2}}{\partial x_2}(z) & \dots & \frac{\partial F_{k+2}}{\partial x_k}(z) & \frac{\partial F_{k+2}}{\partial y_1}(z) & \dots & \frac{\partial F_{k+2}}{\partial y_{m-1}}(z) & \frac{\partial F_{k+2}}{\partial y_m}(z) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial F_{k+m}}{\partial x_1}(z) & \frac{\partial F_{k+m}}{\partial x_2}(z) & \dots & \frac{\partial F_{k+m}}{\partial x_k}(z) & \frac{\partial F_{k+m}}{\partial y_1}(z) & \dots & \frac{\partial F_{k+m}}{\partial y_{m-1}}(z) & \frac{\partial F_{k+m}}{\partial y_m}(z)
\end{array}\right)$$

e assim

$$[F'(z_0)] = \begin{array}{c} \\ k \\ \\ \\ \\ m \\ \\ \\ \end{array}\!\!\left\{\begin{array}{c} \overbrace{}^{k} \quad \overbrace{}^{m} \\ \left(\begin{array}{cccc:cccc}
1 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\
0 & 1 & \dots & 0 & 0 & \dots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & 0 & 0 \\
0 & 0 & \dots & 1 & 0 & \dots & 0 & 0 \\
\hdashline
\frac{\partial z_1}{\partial x_1}(z) & \frac{\partial z_1}{\partial x_2}(z) & \dots & \frac{\partial z_1}{\partial x_k}(z) & \frac{\partial z_1}{\partial y_1}(z) & \dots & \frac{\partial z_1}{\partial y_{m-1}}(z) & \frac{\partial z_1}{\partial y_m}(z) \\
\frac{\partial z_2}{\partial x_1}(z) & \frac{\partial z_2}{\partial x_2}(z) & \dots & \frac{\partial z_2}{\partial x_k}(z) & \frac{\partial z_2}{\partial y_1}(z) & \dots & \frac{\partial z_2}{\partial y_{m-1}}(z) & \frac{\partial z_2}{\partial y_m}(z) \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\
\frac{\partial z_m}{\partial x_1}(z) & \frac{\partial z_m}{\partial x_2}(z) & \dots & \frac{\partial z_m}{\partial x_k}(z) & \frac{\partial z_m}{\partial y_1}(z) & \dots & \frac{\partial z_m}{\partial y_{m-1}}(z) & \frac{\partial z_m}{\partial y_m}(z)
\end{array}\right)
\end{array}$$

logo, podemos agrupar os elementos em quatro submatrizes como abaixo

$$[F'(z_0)] = \begin{bmatrix} I_k & 0 \\ \left[\frac{\partial f}{\partial x}(z_0)\right] & \left[\frac{\partial f}{\partial y}(z_0)\right] \end{bmatrix}.$$

Como

$$J_F(z_0) = \det[F'(z_0)] = \det\left[\frac{\partial f}{\partial y}(z_0)\right] \neq 0,$$

pelo Teorema da Função Inversa, existe $U \in \mathbb{R}^k \times \mathbb{R}^m$ vizinhança aberta de z_0 tal que $V = F(U)$ é aberto e $F : U \to V$ é difeomorfismo de classe C^1.

Denotamos por $(\tilde{x}, \tilde{y}) = F(x, y)$ para $(x, y) \in U$, então $F^{-1}(\tilde{x}, \tilde{y}) = F^{-1}(F(x, y))$ se e somente se $F^{-1}(\tilde{x}, \tilde{y}) = (x, y)$, com $(\tilde{x}, \tilde{y}) \in V$ devido a F ser homeomorfismo. Como $x = \tilde{x}$, decorre da definição que F^{-1} tem a forma

$$F^{-1}(\tilde{x}, \tilde{y}) = (\tilde{x}, g(\tilde{x}, \tilde{y})), \quad (\tilde{x}, \tilde{y}) \in V,$$

onde $g : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}^m$. Devido ao difeomorfismo temos que g é injetora e de classe C^1, logo $\tilde{y} = f(x, y)$ se e somente se $y = g(x, \tilde{y})$. Em particular,

$$f(x, y) = 0 \Leftrightarrow y = g(x, 0).$$

Pois se $f(x, y) = 0$, então $\tilde{y} = 0$ e assim $y = g(\tilde{x}, 0) = g(x, 0)$. Reciprocamente, se $y = g(x, 0)$, então $g(x, \tilde{y}) = g(x, 0)$ então $\tilde{y} = y = 0$. Assim como $\tilde{y} = f(x, y)$ se e somente se $y = g(x, \tilde{y})$ então $f(x, y) = 0$.

Concluímos a demonstração definindo $\varphi = g(x, 0)$, $\forall x \in \Omega = U \cap \mathbb{R}^k$. Notemos também que $\varphi(x_0) = g(x_0, 0) = y_0$ pois $y = g(x_0)$ se e somente se $0 = f(x_0, y)$, então $y = y_0$ por hipótese. □

4 Multiplicadores de Lagrange

Nesta seção, vamos demonstrar uma importante aplicação do Teorema da Função Implícita, que é o Método dos Multiplicadores de Lagrange para o cálculo de extremos de funções sujeitas a restrições. Como um exemplo geral, imaginemos uma função $f : \mathbb{R} \to \mathbb{R}^n$ contínua e um problema de otimização em que queremos determinar o mínimo global da função f sobre uma bola fechada $B = \overline{B_R(x_0)}$, isto é, determinar $\bar{x}$ de B tal que $f(\bar{x}) < f(x)$ para todo elemento x da bola B.

Como B é um conjunto compacto e a função f é contínua, sabemos que a solução para este problema existe. Se f é diferenciável e $\bar{x}$ pertence ao interior de B, então a solução pode ser determinada dentre os pontos críticos de f. Porém, determinar o ponto mínimo $\bar{x}$ quando este está na fronteira pode não ser válido. Assim o resultado a seguir fornece um método caso a função f seja suficiente regular.

Teorema 5. *Sejam $f, g : \mathbb{R}^n \to \mathbb{R}$ funções de classe C^1 e um conjunto $S = \{x \in \mathbb{R}^n : g(x) = 0\}$. Suponhamos $x_0 \in S$ tal que*

$$g'(x_0) \neq 0 \quad e \quad f(x_0) = \min\{f(x) : x \in S\}.$$

Então $f'(x_0)$ e $g'(x_0)$ são linearmente dependentes (combinação linear uma da outra), ou seja, existe um multiplicador de Lagrange $\lambda \in \mathbb{R}$ tal que $\nabla f(x_0) = \lambda \nabla g(x_0)$.

Prova. Se $g'(x_0) \neq 0$ para $x_0 \in S$, então há uma derivada parcial não nula para pelo menos uma coordenada. Vamos então supor sem perda de generalidade que $\frac{\partial g}{\partial x_n}(x_0) \neq 0$. Seja $\lambda \in \mathbb{R}$ tal que

$$\frac{\partial f}{\partial x_n}(x_0) = \lambda \frac{\partial g}{\partial x_n}(x_0).$$

Assim, para concluirmos a prova basta demonstrarmos o resultado acima para as outras $n-1$ derivadas parciais

$$\frac{\partial f}{\partial x_i}(x_0) = \lambda \frac{\partial g}{\partial x_i}(x_0),$$

onde $i = 1, \ldots, n-1$.

Se denotarmos $x = (\tilde{x}, y) \in \mathbb{R}^{n-1} \times \mathbb{R}$, $x_0 = (\tilde{x}_0, y_0)$, então g é de classe C^1, $g(\tilde{x}_0, y_0) = 0$ e

$$\frac{\partial g}{\partial x_n}(x_0) = \frac{\partial g}{\partial y}(\tilde{x}_0, y) \neq 0,$$

segue do Teorema da Função Implícita que existe uma vizinhança aberta $\Omega \subset \mathbb{R}^{n-1}$ de $\tilde{x}_0$ e uma função $\varphi : \Omega \to \mathbb{R}$ de classe C^1 tais que $\varphi(\tilde{x_0}) = y_0$ e

$$g(\tilde{x}, \varphi(\tilde{x})) = 0, \text{ para todo } \tilde{x} \in \Omega. \tag{5}$$

Além disso, como da hipótese

$$f(\tilde{x}_0, \varphi(\tilde{x}_0)) \leq f(\tilde{x}, \varphi(\tilde{x})), \text{ para todo } \tilde{x} \in \Omega,$$

verificamos que $\tilde{x}_0 \in \Omega$ é ponto mínimo para a função diferenciável $\tilde{x} \mapsto \psi(\tilde{x}) = f(\tilde{x}, \varphi(\tilde{x}))$. Portanto, $\psi'(\tilde{x}_0) = 0$ e temos da Regra da Cadeia,

$$[\psi'(\tilde{x}_0)] = \left[\frac{\partial f}{\partial \tilde{x}}(x_0)\right] + \frac{\partial f}{\partial y}(x_0)[\psi'(\tilde{x}_0)] = 0. \tag{6}$$

Derivando da equação 5 em relação a $\tilde{x}$, obtemos

$$\left[\frac{\partial g}{\partial \tilde{x}}(x_0)\right] + \frac{\partial g}{\partial y}(x_0)[\psi'(\tilde{x}_0)] = 0. \tag{7}$$

Multiplicando a equação 7 por λ e subtraindo esta da equação 6 temos que

$$\left[\frac{\partial f}{\partial \tilde{x}}(x_0)\right] = \lambda \left[\frac{\partial g}{\partial \tilde{x}}(x_0)\right].$$

□

Exemplo 6. Vamos encontrar o máximo e o mínimo global da função $f : \mathbb{R}^2 \to \mathbb{R}$ com $(x, y) \mapsto f(x, y) = xy$ sobre a circunferência $x^2 + y^2 = 1$. Podemos reescrever a circunferência como $g(x, y) = 0$ para $g(x, y) = x^2 + y^2 - 1$. A condição de Lagrange nos dá

$$\nabla f = \lambda \nabla g \Rightarrow (y, x) = \lambda(2x, 2y)$$

a qual é equivalente ao sistema de equações

$$y = 2\lambda x, \quad x = \lambda 2y.$$

Combinando com a circunferência $x^2 + y^2 = 1$, temos então três equações e três incógnitas (x, y e λ). Como $x^2 + y^2 = 1$ nós obtemos

$$1 = x^2 + y^2 = (2\lambda y)^2 + (2\lambda x)^2 = 4\lambda^2$$

então temos que $\lambda = \pm\frac{1}{2}$. Então $y = \pm x$, o que nos dá quatro possíveis soluções da equação $x^2 + y^2 = 1$ com a substituição anterior, ou seja

$$(x, y) = \left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right).$$

Temos o valor máximo em $\frac{1}{2}$ quando substituímos os dois pontos $(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ e $(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$ em f e valores mínimos em $-\frac{1}{2}$ o qual gera mais dois pontos (ROSENBAUM, 2014).

Referências

CIPOLATTI, R. **Cálculo Avançado**. 2. ed. Rio de Janeiro: UFRJ/IM, 2002. v. 1.

ROSENBAUM, W. **Inverse and Implicit Function Theorems**. Los Angeles: [s.n.], 2014. Notas de Aula. Disponível em: http://www.willrosenbaum.com/wp-content/teaching/14F32AH/IFT.pdfs.

Um teorema de existência de solução para uma classe de equações diferenciais ordinárias não lineares

Valdecir de Oliveira Teixeira[1]
Acadêmico do Curso de Matemática - CCET da Universidade Estadual do Oeste do Paraná
valdecirdeoliveirateixeira@hotmail.com

André Vicente
Professor do Curso de Matemática - CCET da Universidade Estadual do Oeste do Paraná
andre.vicente@unioeste.br

Resumo: Neste trabalho estabelecemos condições que garantam a existência de solução para uma classe de equações diferenciais ordinárias de primeira ordem. Mais precisamente, sob condição de f, uma função definida em um subconjunto do $\mathbb{R}^2$ e com valores reais, ser contínua e Lipschitziana provamos, utilizando o teorema do ponto fixo de Banach, que o problema $x'(t) = f(t, x(t))$ admite solução $x = x(t)$ em um aberto I contido em $\mathbb{R}$. Este resultado é conhecido na literatura como Teorema de Picard.

Palavras-chave: Teorema do Ponto Fixo de Banach; Teorema de Picard; Equações Diferenciais Ordinárias.

1 Introdução

Em matemática, no estudo de equações diferenciais, o Teorema de Picard é um importante teorema de existência e unicidade de soluções para as equações de primeira ordem com determinadas condições iniciais. A ideia da prova se baseia em transformar a equação diferencial em uma equação integral, e após isso, construir uma aplicação entre espaços métricos e aplicar o Teorema do Ponto Fixo de Banach.

A condição de continuidade de Lipschitz exclui as equações diferenciais que não tenham essa propriedade. Para poder aplicar o Teorema do Ponto Fixo de Banach, é necessário definir um operador F entre dois espaços funcionais de funções contínuas, conhecido como "operador de Picard". Mediante a hipótese de Lipschitz nós estabelecemos que o operador de Picard é uma contração sobre espaços de Banach com a métrica induzida pela norma usual. Isso nos permite aplicar o Teorema do Ponto Fixo de Banach e concluir que o operador tem um único ponto fixo.

Este artigo é resultado de um trabalho de Iniciação Científica qur foi baseado na referência [White, A. J.], onde pode ser encontrado maiores detalhes.

Este trabalho está organizado da seguinte forma: na seção 2 nós apresentamos algumas notações, definições e resultados que serão usados posteriormente, e na seção 3 enunciamos e demonstramos o resultado principal.

[1]Agradeço à Fundação Araucária pelo apoio financeiro.

2 Notações e Resultados Preliminares

Nesta seção apresentamos algumas notações e resultados que serão usados posteriormente. O desenvolvimento do conteúdo dessa seção pode ser visto com mais detalhe nas referências [LIMA, 2013/2014], [WHITE, 1973] e [KREYSZIG, 1978].

2.1 Notações e definições

Definição 1. (**espaço métrico, métrica**) Seja X um conjunto não-vazio. Uma função $d : X \longrightarrow \mathbb{R}$ será uma *métrica* de X se

i) $d(x, y) \geq 0 \qquad x, y \in X,$

ii) $d(x, y) = 0 \qquad$ sse $x = y,$

iii) $d(x, y) = d(y, x) \qquad x, y \in X,$

iv) $d(x, y) \leq d(x, z) + d(z, y) \qquad x, y, z \in X.$

Se X for um conjunto não-vazio e d uma métrica de X, então o par (X, d) será chamado *espaço métrico.*

Definição 2. (**métrica usual de** $\mathbb{R}^n$) Se $n \in \mathbb{N}$ e, se $\mathbb{R}^n$ denota o produto cartesiano de $\mathbb{R}$ por ele próprio n vezes, então os elementos de $\mathbb{R}^n$ podem ser tomados como *n-uplas* ordenadas $(x_1, x_2, ..., x_n)$ de números reais. Se definirmos

$$d((x_1, ..., x_n), (y_1, ..., y_n)) = \sqrt{(x_1 - y_1)^2 + ... + (x_n - y_n)^2},$$

então d é uma métrica, chamada *métrica usual* de $\mathbb{R}^n$.

Definição 3. (**sequência real**) Uma *sequência* em um conjunto X é uma função de $\mathbb{N}^*$ em X; uma *sequência real* é uma sequência em $\mathbb{R}$. Uma sequência real $\{x_n\}$ é *limitada* se existe um número real K tal que, para $n \in \mathbb{N}^*$, $|x_n| \leq K$.

Definição 4. (**bola, vizinhança**) Se $p \in X$ e ε é um número real positivo, o conjunto

$$S_p(\varepsilon) = \{x : x \in X, d(p, x) < \varepsilon\}$$

é chamado a *ε-bola aberta de centro p*, a *ε-vizinhança* de p, ou, ainda, se não for exigida a precisão, simplesmente uma *vizinhança* de p ou, uma *bola.*

Definição 5. (**sequência convergente**) Seja $\{x_n\}$ uma sequência em X e $x \in X$. A sequência $\{x_n\}$ é dita *convergente* para x se, para todo ε real e positivo, existir um inteiro positivo N_ε tal que $x_n \in S_x(\varepsilon)$ sempre que $n > N_\varepsilon$.

Se $\{x_n\}$ converge para x, escrevemos $x_n \longrightarrow x$ quando $n \longrightarrow \infty$, ou $\lim_{n\to\infty} x_n = x$, e x é chamado de *limite* da sequência $\{x_n\}$. Se, para algum $x \in X$, $\{x_n\}$ converge para x, dizemos que $\{x_n\}$ é uma *sequência convergente*. Dizemos que uma sequência é *divergente*, se não é convergente.

Definição 6. (**conjunto aberto**) Um subconjunto A de X será *aberto* se, para cada $a \in A$ existir uma vizinhança de a contida em A; isto é, A será aberto se, para $a \in A$, existir um $\varepsilon > 0$ tal que $S_a(\varepsilon) \subset A$.

Definição 7. (**função contínua**) Suponha que (X, d) e (Y, ρ) sejam espaços métricos e que $x \in X$. Diz-se que uma função $f : X \longrightarrow Y$ é *contínua* em x (mais precisamente, (d, ρ)-*contínua* em x) se, para qualquer $\varepsilon > 0$, existe um $\delta > 0$ tal que $\rho(f(t), f(x)) < \varepsilon$ sempre que $d(t, x) < \delta$. Equivalentemente, f é contínua em x se, dada uma vizinhança qualquer $S_{f(x)}(\varepsilon)$ de $f(x)$, existe uma vizinhança $S_x(\delta)$ de x tal que $f[S_x(\delta)] \subset S_{f(x)}(\varepsilon)$. Se f não é contínua no ponto x, diz-se que ela é *descontínua* em x, ou que tem uma descontinuidade em x.

Definição 8. (**função uniformemente contínua**) Se (X, d) e (Y, ρ) forem espaços métricos, $f : X \longrightarrow Y$ será *uniformemente contínua* se, dado $\varepsilon > 0$, existir um $\delta > 0$ tal que

$$f[S_x(\delta)] \subset S_{f(x)}(\varepsilon) \quad \text{para todo} \quad x \in X.$$

Definição 9. (**conjunto compacto**) Suponhamos (X, d) um espaço métrico. Um subconjunto A de X é chamado de *compacto* se todo subconjunto infinito de A tem um ponto de acumulação em A. Se X for um subconjunto compacto de X, diremos que (X, d) é um espaço métrico compacto.

Definição 10. (**sequência de Cauchy**) Suponha (X, d) um espaço métrico e $\{x_n\}$ uma sequência em X, convergente para um ponto $x \in X$. Então, se $\varepsilon > 0$, existe um inteiro N tal que $d(x, x_n) < \varepsilon/2$, se $n > N$. Desde que $d(x_m, x_n) < d(x, x_n) + d(x, x_m)$, segue-se que uma sequência convergente tem a seguinte propriedade: para qualquer $\varepsilon > 0$, há um inteiro N tal que

$$d(x_m, x_n) < \varepsilon \quad \text{se} \quad m > N \quad \text{e} \quad n > N.$$

Uma sequência que tem a propriedade acima é chamada de *sequência de Cauchy*.

Definição 11. (**espaço métrico completo**) Se um espaço métrico tem a propriedade: "toda sequência de Cauchy converge", então diz-se que ele é *completo*.

Definição 12. (**o espaço** $C(X)$) Suponhamos que (X, d) seja um espaço métrico compacto. Denotamos por $C(X)$ a família de todas as funções contínuas a valores reais definidas sobre

X. Podemos mostrar que se $f \in C(X)$, então f é limitada, isto é, que $C(X) \subset B(X)$. Se denotarmos por σ a métrica para $B(X)$ restringida a $C(X)$, então, para f e g em $C(X)$,

$$\sigma(f,g) = \sup\{|f(x) - g(x)| : x \in X\}.$$

Definição 13. (**convergência simples (ponto a ponto)**) Suponhamos que S seja um conjunto, que (Y,ρ) seja um espaço métrico, e que $\{f_n\}$ seja uma sequência de funções de S em Y; suponhamos também que f seja uma função de S em Y. Diz-se que a sequência $\{f_n\}$ converge *ponto a ponto* para f sobre S se, para cada $s \in S$, a sequência $\{f_n(s)\}$ converge para $f(s)$, e, então, escrevemos $f_n \longrightarrow f$ (ponto a ponto sobre S). Em detalhes, temos $f_n \longrightarrow f$ (ponto a ponto sobre S) se, para cada $s \in S$ e para cada $\varepsilon > 0$ existe um $n_0 \in \mathbb{N}^*$ tal que $\rho(f_n(s), f(s)) < \varepsilon$ sempre que $n \geq n_0$.

Observe que a escolha de n_0 na proposição acima é feita depois de s e ε terem sido escolhidos, de modo que n_0 depende de s e ε.

Definição 14. (**convergência (limite) uniforme**) Suponhamos que S seja um conjunto, que (Y,ρ) seja um espaço métrico, e que $\{f_n\}$ seja uma sequência de funções de S em Y; suponhamos também que f seja uma função de S em Y. Dizemos que a sequência $\{f_n\}$ *converge uniformemente* para f sobre S se, para todo $\varepsilon > 0$, existe um $n_0 \in \mathbb{N}^*$ tal que

$$\rho(f_n(s), f(s)) < \varepsilon, \quad \text{se} \quad n > n_0 \quad \text{e} \quad s \in S.$$

Algumas vezes dizemos que f é o *limite uniforme* de $\{f_n\}$ e escrevemos $f_n \longrightarrow f$ (uniformemente sobre S).

Definição 15. (**equação diferencial ordinária, solução**) Sejam Ω um subconjunto do espaço $\mathbb{R}\times\mathbb{E}$ onde $\mathbb{R}$ é a reta real e $\mathbb{E} = \mathbb{R}^n$ um espaço euclidiano n-dimensional. Um ponto de $\mathbb{R}\times\mathbb{E}$ será denotado por (t,x), $t \in \mathbb{R}$ e $x = (x_1, x_2, ..., x_n)$ em $\mathbb{E}$; salvo menção em contrário, adotaremos em $\mathbb{R} \times \mathbb{E}$ a métrica:

$$|(t,x)| = \max\{|t|, |x|\}$$

onde $|x|$ denota a métrica usual de $\mathbb{R}^n$.

Seja $f : \Omega \longrightarrow \mathbb{E}$ uma aplicação contínua e seja I um intervalo não degenerado da reta, isto é, um subconjunto conexo de $\mathbb{R}$ não reduzido a um ponto. O intervalo I pode ser aberto, fechado ou semi-fechado, finito ou infinito.

Uma função diferenciável $\varphi : I \longrightarrow \mathbb{E}$ chama-se *solução da equação*

$$\frac{dx}{dt} = f(t,x) \tag{1}$$

no intervalo I se:

i) o gráfico de φ em I, isto é, $\{(t, \varphi(t)); t \in I\}$ está contido em Ω e

ii) $\dfrac{d\varphi}{dt} = f(t, \varphi(t))$ para todo $t \in I$. Se t é um ponto extremo do intervalo, a derivada é a derivada lateral respectiva.

A equação (1) chama-se *equação diferencial ordinária de primeira ordem* e é denotada abreviadamente por

$$x' = f(t, x). \tag{1}$$

Sejam $f_i : \Omega \longrightarrow \mathbb{R}, i = 1, ..., n$ as componentes de f; $\varphi = (\varphi_1, ..., \varphi_n)$ com $\varphi_i : I \longrightarrow \mathbb{R}$ é uma solução de (1) sse cada φ_i é diferenciável em I, $(t, \varphi_1(t), ..., \varphi_n(t)) \in \Omega$ para todo $t \in I$ e

$$\begin{aligned} \frac{d\varphi_1}{dt}(t) &= f_1(t, \varphi_1(t), ..., \varphi_n(t)) \\ \frac{d\varphi_2}{dt}(t) &= f_2(t, \varphi_1(t), ..., \varphi_n(t)) \\ &\vdots \\ \frac{d\varphi_n}{dt}(t) &= f_n(t, \varphi_1(t), ..., \varphi_n(t)) \end{aligned} \tag{2}$$

para todo $t \in I$. Por esta razão diz-se que a equação diferencial "vetorial" (1) é equivalente ao sistema de equações diferenciais escalares

$$\frac{dx_i}{dt} = f_i(t, x_1, ..., x_n) \quad i = 1, ..., n. \tag{2}$$

Definição 16. (**ponto fixo**) Se X for um conjunto não-vazio e $F : X \longrightarrow X$ uma função, um ponto $x_0 \in X$ é chamado de *ponto fixo* de F se $F(x_0) = x_0$, isto é , se x_0 é deixado fixo pela aplicação de F.

Definição 17. (**contração**) Suponha (X, d) um espaço métrico. Então a função $F : X \longrightarrow X$ é uma *contração* se existe um número real k, tal que $0 < k < 1$ e

$$d(F(x), F(y)) \leq kd(x, y) \qquad (x \in X, y \in X).$$

É fácil ver que toda contração é contínua. Geometricamente isso significa que quaisquer pontos x e y têm imagens que estão mais próximas do que os pontos x e y; mais precisamente, a razão $d(F(x), F(y))/d(x, y)$ não pode exceder uma constante k que é estritamente menor do que 1.

2.2 Resultados Preliminares

O resultado apresentado nessa seção servirá como pré-requisito para a demonstração do Teorema de Picard.

Teorema 18 (**Teorema do Ponto Fixo de Banach ou Teorema da Contração**). *Se* (X, d) *é um espaço métrico completo e* $F : X \longrightarrow X$ *é uma contração, então* F *tem um, e um só, ponto fixo.*

Prova. Suponhamos que $0 < k < 1$ e que F satisfaça a

$$d(F(x), F(y)) \leq kd(x, y) \qquad (x \in X, y \in X).$$

Seja x_0 e definimos $\{x_n\}$ indutivamente pela equação

$$x_n = F(x_{n-1}) \qquad\qquad (n \in \mathbb{N}^*).$$

Se $n \in \mathbb{N}^*$,

$$\begin{aligned} d(x_n, x_{n+1}) &= d(F(x_{n-1}), F(x_n)) \\ &\leq kd(x_{n-1}, x_n) \qquad \text{(definição de contração)} \\ &= kd(F(x_{n-2}), F(x_{n-1})) \qquad \text{(definição de } \{x_n\}\text{)} \\ &\leq k^2 d(x_{n-2}, x_{n-1}) \qquad \text{(definição de contração)} \\ &\vdots \\ &\leq k^n d(x_0, x_1). \end{aligned}$$

Assim, se $m > n$,

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{r=n}^{m-1} d(x_r, x_{r+1}) \\ &\leq d(x_0, x_1) \sum_{r=n}^{m-1} k^r \\ &\leq \frac{d(x_0, x_1)}{1-k}(k^n - k^m), \end{aligned}$$

e, assim, como $\{k^n\}$ é uma sequência de Cauchy (porque $0 < k < 1$), então também $\{x^n\}$ será. Assim, como (X, d) é completo, existe um $x^* \in X$ tal que

$$\lim_{n\to\infty} x_n = x^*,$$

e, sendo F contínua, temos

$$\begin{aligned} F(x^*) &= \lim_{n\to\infty} F(x_n) \\ &= \lim_{n\to\infty} x_{n+1} \\ &= x^*. \end{aligned}$$

Consequentemente x^* é um ponto fixo de F.

Suponhamos y^* um ponto fixo de F e $x^* \neq y^*$. Então $d(x^*, y^*) > 0$ e (desde que $F(x^*) = x^*$, $F(y^*) = y^*$)

$$\begin{aligned} d(x^*, y^*) &= d(F(x^*), F(y^*)) \\ &\leq kd(x^*, y^*) \end{aligned}$$

o que é impossível, uma vez que $0 < k < 1$. □

3 Resultado Principal

Vamos usar o Teorema de Banach para provar o Teorema de Picard que, embora não seja o mais forte do seu tipo que conhecemos, tem um papel vital na teoria das equações diferenciais ordinárias. A ideia de abordagem é bastante simples: a equação diferencial ordinária de primeira ordem será convertida em uma equação integral, a qual nos permite definir uma aplicação F, e as condições do teorema implica que F é uma contração de modo que o seu ponto fixo torna-se a solução de nosso problema. Ver detalhes em [WHITE, 1973] e [KREYSZIG, 1978].

Teorema 19 (**Teorema de Picard**). *Suponhamos que $U \subset \mathbb{R}^2$ seja um conjunto aberto e que $F : U \longrightarrow \mathbb{R}$ seja uma função contínua que satisfaz à condição (Lipschitz) de que, para algum $K > 0$,*

$$|f(x, y_1) - f(x, y_2)| \leq K|y_1 - y_2|$$

para $(x, y_1) \in U$, e $(x, y_2) \in U$. Então, para qualquer $(x_0, y_0) \in U$, existe um número real positivo a e uma e uma só função $g : [x_0 - a, x_0 + a] \longrightarrow \mathbb{R}$ tal que

$$g'(x) = f(x, g(x)) \quad (x \in (x_0 - a, x_0 + a)) \tag{2}$$

e $g(x_0) = y_0$.

Prova. Se $a(> 0)$ e uma função $g : [x_0 - a, x_0 + a] \longrightarrow \mathbb{R}$ existem satisfazendo a (2) e à condição $g(x_0) = y_0$, então está implícito que $(x, g(x)) \in U$ para $x \in (x_0 - a, x_0 + a)$ e, como g é contínua, segue-se que a função $x \longrightarrow f(x, g(x))$ é contínua. Consequentemente, g' é contínua em $(x_0 - a, x_0 + a)$, e segue que

$$\begin{aligned} g'(x) &= f(x, g(x)) \\ g(x) &= \int_{x_0}^{x} f(t, g(t))dt + y_0 \quad (x \in (x_0 - a, x_0 + a)). \end{aligned} \tag{2}$$

Por outro lado, se $a(> 0)$ e

$$g : [x_0 - a, x_0 + a] \longrightarrow \mathbb{R}$$

existem satisfazendo (2), então, $g'(x) = f(x, g(x))$ $(x \in (x_0 - a, x_0 + a))$ e $g(x_0) = y_0$. Assim, o problema de encontrar um a e uma g satisfazendo a (1) é equivalente ao de encontrar um a e uma g satisfazendo a (2).

A ideia do final da prova é selecionar um número real $a(> 0)$ e considerar a aplicação

$$F : C[x_0 - a, x_0 + a] \longrightarrow C[x_0 - a, x_0 + a]$$

definida por

$$F(g)(x) = \int_{x_0}^{x} f(t, g(t))dt + y_0. \tag{3}$$

Então, encontrar uma g satisfazendo a (2) é equivalente a encontrar um ponto fixo para F e, com esse fim, temos de nos assegurar, por conveniente escolha de a, que F é uma contração. Além disso, temos que garantir que F está bem definida. Precisamos estar certos de que as funções com as quais operamos são tais que

$$t \in [x_0 - a, x_0 + a]$$

e, então, o integrando em (3) está definido. Essas considerações motivam a escolha de a e da bola B na prova que se segue.

Tomemos inicialmente $N > |f(x_0, y_0)|$. Escolhamos então a tal que as três condições seguintes sejam satisfeitas:

1) $0 < a < 1/K$,

2) $Q = \{(x, y) : |x - x_0| \leq a, |y - y_0| \leq aN\} \subset U$,

3) $(x, y) \in Q$ implica $|f(x, y)| < N$.

Essa escolha é possível porque $|f|$ é contínua em (x_0, y_0). Seja $I = [x_0 - a, x_0 + a]$, e seja $Y_0 : I \longrightarrow \mathbb{R}$, definida por $Y_0(x) = y_0 \quad (x \in I)$.

Finalmente, seja d a métrica usual para $C(I)$, e tomemos

$$B = \{g : g \in C(I), d(g, Y_0) \leq aN\},$$

isto é, B é a bola fechada de centro Y_0 e raio aN.

Observamos primeiramente que, se $t \in I$ e $g \in B$, então $(t, g(t)) \in Q$, pois, se $t \in I$, então $|t - x| \leq a$ e

$$\begin{aligned} |g(t) - y_0| &\leq \sup\{|g(x) - y_0| : x \in I\} \\ &= d(g, Y_0) \end{aligned}$$

$$\leq \quad aN$$

de modo que $(t, g(t)) \in Q$. Segue-se (pois $Q \subset U$) que a aplicação $t \longrightarrow f(t, g(t))$ de I em $\mathbb{R}$ é bem definida e define uma função contínua em I. Consequentemente, podemos definir $F : B \longrightarrow C(I)$ por

$$F(g)(x) = \int_{x_0}^{x} f(t, g(t))dt + y_0. \qquad (x \in I, g \in B).$$

Vamos agora observar que $F[B] \subset B$ (isto é, se $g \in B$, então $F(g) \in B$), pois, se $g \in B$, então

$$\begin{aligned}
d(F(g), Y_0) &= \sup\{|F(g)(x) - Y_0(x)| : x \in I\} \\
&= \sup\{|F(g)(x) - y_0| : x \in I\} \\
&= \sup\left\{\left|\int_{x_0}^{x} f(t, g(t))\right| dt : x \in I\right\} \\
&= \sup\left\{\int_{x_0}^{x} |f(t, g(t))|\, dt : x \in I\right\} \\
&\leq \sup\left\{N \int_{x_0}^{x} dt : x \in I\right\} \\
&= \sup\{N\,|x - x_0| : x \in I\} \\
&\leq \sup\{Na\} \\
&= aN
\end{aligned}$$

desde que $x \in I$ e $|x_0 - t| \leq |x_0 - x|$, então $t \in I$ e $(t, g(t)) \in Q$ e, portanto (por 3)), $|f(t, g(t))| < N$. Segue-se que

$$d(F(g), Y_0) \leq aN$$

e, então, $F(g) \in B$.

Assim, F aplica B em B e, como B é um espaço métrico completo (é um subconjunto fechado do espaço métrico completo $C[I]$), para mostrarmos que F tem um único ponto fixo, será suficiente mostrar que F é uma contração.

Para esse fim, suponhamos g_1 e g_2 pontos de B. Então, se $t \in I$, $(t, g_1(t))$ e $(t, g_2(t))$ pertencem a Q e, portanto, a U, logo

$$|f(t, g_1(t)) - f(t, g_2(t))| \leq K|g_1(t) - g_2(t)|$$

teremos

$$\begin{aligned}
d(F(g_1), F(g_2)) &= \sup\{|F(g_1)(x) - F(g_2)(x)| : x \in I\} \\
&= \sup\left\{\left|\int_{x_0}^{x} [f(t, g_1(t)) - f(t, g_2(t))]dt\right| : x \in I\right\}
\end{aligned}$$

$$\begin{aligned} &\leq \sup\left\{K\int_{x_0}^{x}|g_1(t)-g_2(t))|\,dt : x\in I\right\} \\ &\leq \sup\left\{\int_{x_0}^{x}dt\,[K\sup\{|g_1(t)-g_2(t)|\}] : x\in I\right\} \\ &= \sup\{|x-x_0|\,Kd(g_1,g_2)\} \\ &\leq aKd(g_1,g_2). \end{aligned}$$

Segue-se que, como $aK<1$ (por 1)), F é uma contração. O Teorema do Ponto Fixo de Banach implica que F tem um único ponto fixo $g\in C(I)$, isto é, uma função contínua $g\in C(I)$ satisfazendo $F(g)=g$. Assim, escrevendo $g=F(g)$, nós temos por (3)

$$g(x)=y_0+\int_{x_0}^{x}f(t,g(t))dt. \tag{4}$$

Desde que $(t,g(t))\in Q$ onde f é contínua e diferenciável. Por isso g é diferenciável e satisfaz (1). Por outro lado, cada solução de (1) deve satisfazer (4). Isto completa a prova. □

Referências

Kreyszig, Erwuin. *Introductory Funcional Analysis with Applications*, by John Wiley & Sons. Inc., 1978.

Lima, Elon L. *Curso de Análise*, Vol.1, 14a. edição Projeto Euclides, 2013.

Vicente, André. *Notas de Aula de Análise*, Unioeste - Cascavel: 2015-2016.

White, A. J. *Análise Real: uma Introdução*, tradução, Elza F. Gomide. São Paulo, Edgard Blücher, Ed. da Universidade de São Paulo, 1973.

Teixeira, Valdecir de O. *Teorema de Aproximação de Weierstrass*, Relatório de Iniciação Científica, 2015.

A torre de Hanói no ensino de função exponencial

Eliandra de Oliveira[1]
Universidade Estadual do Oeste do Paraná
eliandra.oliveiraaa@gmail.com

Roselaine Maria Gonçalves[2]
Universidade Estadual do Oeste do Paraná
roselainemgoncalves@gmail.com

Resumo: Este trabalho relata nossa experiência ao usar Jogos Matemáticos em sala de aula para introduzir o conteúdo de função exponencial. Utilizamos o jogo da torre de Hanói, com o intuito de motivar os alunos a aprender algo novo. Essa foi uma das atividades desenvolvidas no Programa Institucional de Bolsas de Iniciação à Docência – PIBID. No subprojeto de Matemática fomos incentivados a utilizar metodologias diferenciadas em sala, oportunizando o aprendizado por meio do manuseio das diferentes Metodologias de Ensino da Matemática e pela vivência em sala de aula, promovendo assim a prática docente e a reflexão sobre a mesma.

Palavras-chave: Jogos; torre de Hanói; função exponencial.

1 Introdução

O Programa Institucional de Bolsas de Iniciação à Docência – PIBID, é um projeto do governo federal que tem como principal objetivo incentivar a formação de professores em nível superior para a atuação na educação básica e contribuir para a valorização do magistério. Este programa concede bolsas aos acadêmicos de cursos de licenciatura que participam de projetos de iniciação à docência, desenvolvidos por instituições de ensino superior em parceria com escolas de educação básica da rede pública de ensino. Na Universidade Estadual do Oeste do paraná – Unioeste, bem como no Curso de Matemática do *campus* de Cascavel, esse programa é desenvolvido a partir de 2010.

Desde 2015, trabalhando no Colégio Estadual Pacaembu percebemos o quanto este programa tem feito a diferença na vida dos acadêmicos que querem seguir a carreira de docente, pois o trabalho do PIBID os tem ajudado a conviver com diversas situações e dificuldades enfrentadas em sala de aula, com as quais não teriam o privilégio de trabalhar apenas nos estágios supervisionados que a universidade oferece. Sem dúvida alguma, podemos afirmar que o PIBID possui muitos pontos positivos e é uma grande oportunidade de aprender a ensinar, de aprender a aprender e de integrar a teoria estudada na universidade com a prática docente cotidiana.

[1]Bolsista Pibid desde 2015
[2]Bolsista Pibid desde 2014

Uma das integrações que nos é proporcionada vem da utilização das diversas metodologias de ensino que podem ser aplicadas em sala, como os jogos matemáticos. Os professores coordenadores nos incentivaram a pensar em uma maneira de trabalhar com esta metodologia e ensinar por meio dela. Como já estávamos trabalhando havia um tempo com os primeiros anos do Ensino Médio do Colégio Estadual Pacaembu, pensamos em aplicar uma atividade para introduzir algum conteúdo usando jogos de uma forma recreativa e educativa. Com o PIBID temos a oportunidade de observar determinada turma durante várias semanas, por isso já conhecíamos a turma que iriamos aplicar os jogos, e percebemos que seria importante trazer algo relativo a jogos para os alunos, por serem bem agitados, porém esforçados.

A palavra jogo deriva do termo em latim "*jocus*" que significa gracejo, brincadeira, divertimento. Caracterizado como uma atividade física ou intelectual que integra um sistema de regras e define um indivíduo (ou um grupo) vencedor e outro perdedor. Segundo Strapason (2011, p.14) "o jogo deve ter um significado para quem joga, seja de entretenimento ou finalidade educativa, conforme o jogo escolhido. Em ambos os casos sempre propicia situações de prazer, de desprazer e de busca de estratégias para a melhor jogada".

Strapason (2011, p. 20) diz também que:

> O papel dos jogos como estratégia de ensino e aprendizagem da Matemática tem sido salientada em inúmeras pesquisas. Os jogos propiciam aprendizagens mais motivadoras e interessantes, tanto para o aluno quanto para o professor. Inúmeras habilidades matemáticas podem ser desenvolvidas através dos jogos, entre elas, o raciocínio reflexivo, pois é necessário sempre pensar muito bem antes de realizar qualquer jogada e a cada nova jogada, um novo raciocínio pode surgir. Os raciocínios lógicos utilizados pelos alunos durante o jogo sempre se assemelham à resolução de um problema matemático, mesmo que o jogo não seja em relação a um conteúdo matemático específico.

Todos nós sabemos que a utilização de jogos na escola não é algo novo e que muitos professores os utilizam a sala de aula, pois a aprendizagem que se dá através do jogo possibilita formas de aprendizagem diferenciadas das quais os alunos estão acostumados. Além de tornar a aula mais atrativa, faz com que os alunos possam investigar, resolver e descobrir durante o jogo a melhor jogada estabelecer relações entre os elementos do jogo e, consequentemente, os conceitos da matemática que estão no mesmo.

Segundo Grando (2000, p. 27),

> O jogo, pelo seu caráter propriamente competitivo, apresenta-se como uma atividade capaz de gerar situações-problema "provocadoras", onde o sujeito necessita coordenar diferentes pontos de vista, estabelecer várias relações, resolver conflitos e estabelecer uma ordem.

No Ensino Médio o jogo não é tão utilizado, pois para muitos não é visto como uma ferramenta de ensino, mas como uma "diversão" aos alunos para a aula não ficar tão cansativa.

Por mais que o jogo traga barulho, movimentos e alegria não podemos deixar que isto seja apenas uma diversão aos alunos, pois ele permite ao aluno corrigir seus erros, assim como rever suas respostas, que descubra em que está falhando e faz com que pense em uma nova estratégia. Algo que nos chama a atenção no jogo também, é que quando o aluno erra, ele não é criticado pelos demais e nem por si mesmo, os colegas se ajudam e dão dicas ao longo de sua execução. Além disso, o aluno não tem medo de perguntar, nem de se expor.

De acordo com Borin (*apud* Groenwald e Timm, p. 9. 2002),

> Outro motivo para a introdução de jogos nas aulas de matemática é a possibilidade de diminuir bloqueios apresentados por muitos de nossos alunos que temem a Matemática e sentem-se incapacitados para aprendê-la. Dentro da situação de jogo, onde é impossível uma atitude passiva e a motivação é grande, notamos que, ao mesmo tempo em que estes alunos falam Matemática, apresentam também um melhor desempenho e atitudes positivas frente a seus processos de aprendizagem.

Segundo Smole e Diniz, o jogador não aprende a pensar sobre jogo quando o joga apenas uma vez. Assim podemos afirmar que realmente o aluno precise jogar algumas vezes para aprender as regras e ter a aprendizagem desejada pelos professores, o que com certeza toma algum tempo da aula, por isso precisamos sempre estar preparados para todos os tipos de ocorrências.

Outro fato importante que o professor precisa ter em mente, qual jogo pode levar e utilizar em sala de aula. Ele deve conhecer bem o material que quer utilizar e saber auxiliar os alunos para que haja aprendizagem. Além disso é necessária uma socialização para que as dúvidas que haviam durante a realização do jogo sejam tiradas e também que os alunos coloquem as suas opiniões e sugestões, mostrando o que aprenderam o que poderia mudar para tornar o jogo mais atrativo e proveitoso.

O jogo por sua vez também pode frustrar os alunos, ser incompreensível, obrigatório, um passatempo, ou algo em que simplesmente, quem tem sorte vence. Uma forma de proporcionar uma aprendizagem significativa é proporcionar após a sua realização uma discussão sobre o jogo, indagando os alunos com relação às estratégias utilizadas e aos novos conhecimentos adquiridos.

No nosso caso, com a Torre de Hanói, pretendíamos motivar os alunos a aprender algo novo, como por exemplo, o jogo em si, o qual era novidade para todos os alunos e também aprender sobre função exponencial, que era o nosso objetivo principal. Por meio do raciocínio lógico-matemático, incentivamos os alunos de maneira a descrever uma lei de formação para a função exponencial, observando as jogadas feitas por eles nos fornecesse uma fórmula para que pudéssemos prever a movimentação necessária caso possuíssem n peças.

2 Torre de Hanói

A Torre de Hanói é um jogo muito antigo e conhecido pela maioria dos professores de Matemática. Ele nos possibilita construir uma fórmula matemática que relaciona o número de discos com o número de seus movimentos para serem transportados entre três pinos.

Tradicionalmente a Torre de Hanói é feita de madeira e é composta por três pinos e sete discos, com diâmetros diferentes. Mas, para fins didáticos é comum encontrarmos variações com menos ou mais discos, ou estudos com mais pinos. Além disso, elas podem ser confeccionadas com diversos materiais como emborrachados, isopor e materiais recicláveis.

Sobre o jogo, conta-se a seguinte lenda:

"No começo dos tempos, Deus criou a Torre de Brahma com três pinos de diamante alinhados e colocou, no primeiro, 64 discos de ouro maciço em ordem decrescente de tamanho. Deus, então, chamou seus Sacerdotes e ordenou a eles que transferissem todos os discos para o terceiro pino, movendo apenas um disco de cada vez e nunca colocando um disco maior sobre um menor. Os sacerdotes, então, obedeceram e começaram o seu trabalho, dia e noite. Quando eles terminarem, a Torre de Brahma irá ruir e o mundo acabará".

3 Atividade desenvolvida na escola

Elaboramos e desenvolvemos juntamente com o professor supervisor um jogo que tinha como objetivo introduzir o conceito de função exponencial. A princípio, apresentaremos a descrição da atividade desenvolvida e na sequência apresentamos os resultados e reflexões.

A atividade que desenvolvemos na sala de aula foi realizada em 2015 e para a sua execução foi necessário que os alunos ficassem em duplas ou trios. Após formarem as duplas ou trios, apresentamos o jogo a eles, contamos a lenda do mesmo e entregamos as torres às duplas, juntamente com uma tabela impressa, para que eles pudessem registrar os números de discos e jogadas feitas.

Explicamos as regras do jogo, em que o jogador deve transferir os discos do pino A para o pino C e contabilizar o número de jogadas realizadas, lembrando que o disco de maior diâmetro nunca deve ficar sobre o de menor diâmetro e que só é permitido mover um disco por vez. Pedimos que os alunos tentassem resolver o problema da torre com 2, 3, 4 e 5 discos, respectivamente, registrando o número de movimentos feitos para cada número de discos na tabela.

Após a manipulação, perguntamos aos alunos se eles poderiam predizer qual seria o menor número de movimentos necessários para mover n discos do primeiro para o último pino.

Fizemos outras perguntas a eles como, qual será o próximo número mínimo? Como determinar uma lei que relacione os termos n e o número mínimo $J(n)$ de movimentos, para qualquer número natural n?

Posteriormente pedimos que preenchessem uma tabela para que fizéssemos a construção do gráfico, o qual é diferente do que estão acostumados, pois até o momento haviam estudado apenas as funções afins e quadráticas. Portanto, mostramos uma nova função chamada função exponencial.

4 O desenvolvimento

No primeiro momento contamos aos alunos que nós iriamos trabalhar com eles aquele dia e que havíamos trazido um material diferente, chamado "Torre de Hanói", neste momento os alunos ficaram agitados, pois queriam pegar e começar a mexer. Com a explicação do que faríamos os alunos ficaram impressionados com o tamanho do número de movimentos que teria que acontecer para o "fim do mundo".

Após entregamos as torres aos alunos se mostraram muito interessados em manuseá-las, até porque os mesmos não conheciam o material. De início tentaram mover todos os seis discos, porém não obtiveram êxito, por ser algo complexo. Então pedimos que fizessem os movimentos com um número menor de discos e registrassem a quantidade de movimentos feitos. Eles deveriam repetir o processo três vezes para cada quantidade de discos, por exemplo, com três discos deveriam jogar três vezes e então marcar o número de movimentos a cada vez, para conseguirem diferenciar qual era o número mínimo de movimentos feitos.

Ao observar os grupos, percebemos que algumas alunas fizeram apenas uma movimentação e replicaram para as próximas colunas, então falamos a elas que cada um tem uma forma de jogar diferente e, elas falaram que não. Por isso pedimos para cada uma jogar. A primeira jogou e executou quatorze movimentos, a segunda jogou e também obteve o mesmo número de movimentos, porém, quando a terceira jogou deu a metade de movimentos feitos (sete). Por conta disso começaram a jogar as três vezes solicitadas. Muito dos grupos conseguiram chegar ao número mínimo de jogadas possíveis.

Como possuíamos aulas geminadas, optou-se em deixar que os alunos jogassem na primeira aula várias vezes. Ao término desta, fomos ao quadro e reproduzimos a tabela que eles

possuíam em mãos, perguntando quais eram os movimentos mínimos que haviam encontrado para cada quantidade de discos. Em seguida, mostramos os movimentos mínimos que poderiam ser feitos, no caso para um disco é feito um movimento; dois discos são três movimentos; três discos são sete movimentos; quatro discos são quinze movimentos; e assim por diante até n discos. O objetivo aqui era discutir sobre o jogo em si, se eles perceberam alguma estratégia para fazer o mínimo de movimentos possíveis, entre outras coisas.

Quando chegassem a n discos os alunos deveriam enunciar a lei de formação para saberem quantos movimentos deveriam ser feitos para esta quantidade de discos, generalizando a situação. Neste momento notamos que os alunos estavam dispersos com a possibilidade de movimentar os n discos, portanto se fez necessário chamar a atenção dos mesmos.

Com a atenção deles, questionamos se haveria uma relação entre a quantidade de movimentos e o número de discos. Um dos alunos falou que podíamos usar a seguinte expressão $2 \cdot 3 + 1 = 7$, na qual ele se baseou no resultado do movimento anterior para calcular o próximo movimento. Assim, no caso em que tivéssemos 50 discos seria necessário saber quantos movimentos foram utilizados para movimentar 49 discos, portanto não seria possível "prever" os n movimentos sem saber qual foi a movimentação anterior.

Todos os casos expressos pelos alunos foram generalizados, deste modo o caso acima citado tem a forma $f(x) = 2x + 1$. Os alunos concordaram que esta função não "funcionaria" quando não soubéssemos o que havia acontecido anteriormente. Questionamos se possuíam outra ideia de como fazer esta função de modo que fosse válida para qualquer quantidade de discos.

Os alunos não perceberam a lei de formação. Dessa maneira indicamos que deveriam utilizar a potência no meio da lei de formação. Mostramos também a forma geral da função exponencial $f(x) = a^x$, onde a variável é o x e o a é um número real fixo. A partir disto os alunos começaram a discutir a função exponencial e observando a função anterior eles disseram que $x = 2$ e que a iria variar de acordo com a quantidade de discos que possuíamos e somaríamos 1, ficando da seguinte forma: $f(x) = a^2 + 1$.

Assim, substituímos o a pelos possíveis números de discos, mas a função não "funcionou" para todos os casos, logo essa função também estava "descartada". Falamos para eles que se olhássemos para a função geral sabíamos que a incógnita variável sempre é o x, com isso eles concluíram que x dependia da quantidade de discos. Mas, e o valor do a?

Os alunos pensaram um pouco, analisando as duas funções feitas anteriormente. Uma das alunas disse que podíamos usar a função $f(x) = 2^x - 1$, perguntamos se todos concordavam

com esta função e alguns alunos concordaram. Fizemos então a substituição de x por valores reais e os alunos perceberam que esta lei de formação era válida.

Notamos que todos os alunos estavam pensando em grupo, pois para encontrarem a função muitos iam falando o que deveria ser feito e ao final um deles juntou tudo. A cada pergunta feita e respondida corretamente o aluno ganhava um chocolate, o que o incentivava a fazer as relações.

Em seguida fizemos a construção do gráfico no quadro, para que analisássemos o mesmo. Perguntamos-lhes o que podiam perceber com relação a este e algumas alunas disseram que ele era crescente. Nos questionaram, devido ao formato, se era uma reta ou uma parábola. Lhes devolvemos a pergunta, mas não souberam a resposta. Por isso utilizamos os números negativos para calcular os valores menores e iguais a zero, isso nos deu uma reta assíntota ao gráfico, ou seja, quando x tende a 0 (zero). Explicamos que este "pedaço" da reta iria se aproximar o máximo possível de 0, mas nunca seria este valor. Com isso perceberam que não era nem uma parábola nem uma reta, mas uma curva, que tende ao infinito.

Ao traçar a parte negativa da reta, destacamos aos alunos que neste problema isso não "seria correto", pois não temos um número negativo de discos. Outra condição deste exercício é do gráfico ser feito de pontos, ou seja, não desenharíamos uma reta, pois não nos é possível ter um número fracionário de discos, como por exemplo, 4,7 discos.

Destacamos aos alunos que a partir daquele momento, eles estudariam a função exponencial e deixamos para que o Professor da turma pudesse continuar com a aula, pois o nosso objetivo era que eles entendessem um pouco sobre o que iriam estudar a partir deste momento.

5 Conclusão

Durante o período que estávamos na escola, tivemos a oportunidade de observar os alunos e de realizar algumas atividades com eles. A turma que trabalhamos foi o primeiro ano do Ensino Médio, a qual tinha pouco mais de trinta alunos. Dessa maneira, pensamos em várias atividades diferentes para contribuir no conhecimento dos alunos e, consequentemente, ajudar o professor em suas aulas e também para nosso aprendizado enquanto acadêmicas de licenciatura em Matemática.

Nós percebemos que os alunos aprendem mais quando aquilo que está sendo ensinado está ao mesmo tempo sendo experimentado por eles, seja por materiais manipulativos, experimentos ou até mesmo no dia a dia, e foi com esse propósito que levamos esta atividade a eles. E

enquanto pibidianas, pretendemos sempre estar experimentando e vivenciando as práticas que nos permitam entender melhor como se dá a relação de ensino-aprendizagem junto aos alunos.

A atividade em si, foi significativa para nós e também para os alunos, pois durante as outras observações que fizemos, percebemos a facilidade com que os alunos estavam entendendo a nova função estudada, a função exponencial. Eles mostraram interesse pelo conteúdo e interagiam durante as aulas. Entendemos que alcançamos nosso objetivo de motivar a aprender e sempre buscar conhecer algo novo.

Referências

GRANDO, Regina Célia. **O conhecimento matemático e o uso de jogos na sala de aula**. Campinas: Universidade Estadual de Campinas Faculdade de Educação, 2000.

GROENWALD, Claudia Lisete Oliveira; TIMM, Ursula Tatiana. **Utilizando curiosidades e jogos matemáticos em sala de aula**. Disponível em: <http://www.somatematematica.com.br/> Acesso em: 02 set. 2016

Significado de jogos. Disponível em: <http://www.significados.com.br/jogo/>. Acesso em: 31 ago. 2016.

SMOLE, Kátia Stocco; DINIZ, Maria Ignez; PESSOA, Neide; ISHIHARA, Cristiane. Os jogos nas Aulas de Matemática do Ensino Médio. In: SMOLE, Kátia Stocco; DINIZ, Maria Ignez; PESSOA, Neide. Cadernos do Mathema: Jogos de Matemática. Porto Alegre: Artmed Editora S.A., 2008. p. 9-27.

STRAPASON, Lísie Pippi Reis. **O uso de jogos com estratégia de ensino e aprendizagem da matemática no 1º ano do Ensino Médio**. Santa Maria, RS – 2011

Torres de Hanói. Disponível em: <http://m3.ime.unicamp.br/recursos/1361>. Acesso em: 05 jan. 2016.

As funções de Mittag-Leffler e equações diferenciais de ordem arbitrária

Sarah Elusa de Melo Menoncin
Universidade Estadual do Oeste do Paraná
sarah_elusa@hotmail.com

Sandro Marcos Guzzo
Universidade Estadual do Oeste do Paraná
smguzzo@gmail.com

Resumo: Estudos recentes sobre mecânica de fluidos, finanças, biologia molecular e em muitos outros campos, detectaram que a influência de fatores aleatórios pode trazer muitos aspectos interessantes para o modelo, os quais fazem necessária a inclusão de operadores de ordem arbitrária na construção dos modelos matemáticos que descrevem tais fenômenos. Esses problemas envolvendo derivadas fracionárias tornam-se mais complexos, entretanto, com alguns ajustes, as técnicas desenvolvidas para modelos clássicos de problemas de valor inicial envolvendo equações diferenciais de ordem inteira, como por exemplo, a Transformada de Laplace podem ser aplicadas diretamente a eles. Desta forma, este trabalho de pesquisa tem como principal interesse estudar as funções de Mittag-Leffler de um, dois e três parâmetros e sua aplicação na determinação da solução de equações diferenciais de ordem arbitrária.

Palavras-chave: Equações diferenciais de ordem arbitrária; Função de Mittag-Leffler; Transformada de Laplace.

1 A Transformada de Laplace

A Transformada de Laplace mostra-se eficaz na resolução de problemas de valor inicial envolvendo equações gerais de segunda ordem com coeficientes constantes. Uma das vantagens deste método é que ele pode ser utilizado na resolução de problemas envolvendo agentes descontínuos ou impulsivos, para os quais outros métodos não seriam adequados. Caso seja de interesse do leitor, os resultados dessa seção podem ser verificados mais detalhadamente em Boyce (2009).

Definição 1. Dada uma função $f(t)$, definimos sua Transformada de Laplace pela equação

$$F(s) = \mathcal{L}\{f(t)\} = \int_0^\infty e^{-st} f(t)dt,$$

desde que a integral imprópria exista e seja convergente.

É importante destacar que a Transformada de Laplace da função f, denotada por $\mathcal{L}\{f(t)\}$ existe se as condições a seguir forem cumpridas:

1. f for seccionalmente contínua no intervalo $0 \leq t \leq$ A para todo $A > 0$;

2. $|f(t)| \leq Ke^{at}$ quando $t \geq M$, com $K, M, a \in \mathbb{R}$ e $K, M > 0$.

Adota-se a notação $\mathcal{L}^{-1}\{F(s)\}$ para indicar a Transformada Inversa de $F(s)$, ou seja, a função $f(t)$ cuja Transformada de Laplace é dada por $F(s)$. Segue abaixo uma tabela de Transformadas de Laplace e suas respectivas transformadas inversas.

Tabela 1: Tabela Transformada de Laplace

$F(s)$	$F(s)^{-1}$
$\frac{1}{s}$	1
$\frac{1}{s^2}$	t
$\frac{1}{s-a}$	e^{at}
$\frac{1}{(s+a)^2}$	te^{-at}
$\frac{n!}{s^{n+1}}$	t^n
$\frac{s}{s^2+a^2}$	$\cos at$
$\frac{a}{s^2+a^2}$	$\sin at$
$\frac{s}{s^2-a^2}$	$\cosh at$
$\frac{a}{s^2-a^2}$	$\sinh at$
$\frac{s^2-a^2}{(s^2+a^2)^2}$	$t\cos t$
$\frac{2as}{(s^2+a^2)^2}$	$t\sin t$

Uma propriedade importante desta transformada é o fato de ela ser um operador linear. Sendo assim, temos que,

$$\mathcal{L}\{c_1 f_1(t) + c_2 f_2(t)\} = c_1\mathcal{L}\{f_1(t)\} + c_2\mathcal{L}\{f_2(t)\},$$

com c_1 e c_2 constantes.

Além disto, se f é uma função contínua e, suponha que f' seja seccionalmente contínua em um determinado intervalo, então temos que

$$\mathcal{L}\{f'\} = s\mathcal{L}\{f\} - f(0).$$

2 A Função Gama

Definição 2. A função Gama é a função que a cada número real positivo $x > 0$ associa o número real representado por $\Gamma(x)$ determinado pela integral imprópria

$$\Gamma(x) = \int_0^\infty e^{-t}t^{(x-1)}dt.$$

Vejamos uma propriedade muito interessante da função Gama.

Proposição 3. *Se $x > 0$, então*

$$\Gamma(x+1) = x\Gamma(x).$$

Prova. Esta propriedade pode ser verificada facilmente por meio de integração por partes. De fato, tomando $u = t^x$ e $\frac{dv}{dt} = e^{-t}$, temos que $\frac{du}{dt} = xt^{x-1}$ e $v = -e^{-t}$. Assim, temos que

$$\begin{aligned}\Gamma(x+1) &= \int_0^\infty e^{-t}t^x dt \\ &= \left[-e^{-t}t^x\Big|_{t=0}^{\infty}\right] + x\int_0^\infty e^{-t}t^{x-1}dt.\end{aligned}$$

Sabemos que $\lim_{t\to\infty} e^{-t}a^t = 0$, donde

$$\Gamma(x+1) = (0+0) + x\int_0^\infty e^{-t}t^{x-1}dt = x\Gamma(x).$$

□

Outro resultado importante acerca da Função Gama é que, em particular, se $x = n$ com $n \in \mathbb{N}^*$, interpretamos a função gama como uma generalização do conceito de fatorial. Deste modo, temos que,

$$\Gamma(n+1) = n!$$

ou ainda,

$$\Gamma(n) = (n-1)!.$$

De fato, para $n = 1$ temos,

$$\Gamma(1) = \int_0^\infty e^{-t}dt = -e^{-t}\Big|_{t=0}^{\infty} = 1 = 0!.$$

Vamos agora supor, por um momento, que seja válido para $k \in \mathbb{N}$, ou seja, $\Gamma(k) = (k-1)!$. Assim, para $k+1$, da proposição anterior, segue que

$$\Gamma(k+1) = k\Gamma(k) = k(k-1)! = k! \tag{1}$$

e desta forma, segue por indução que $\Gamma(n+1) = n!$ para todo $n \in \mathbb{N}^*$. Mais detalhes sobre a função Gama podem ser encontrados em Grigoletto (2014).

3 As Funções de Mittag-Leffler

Nesta seção introduzimos a clássica função de Mittag-Leffler, denotada por $E_\alpha(x)$ que é a função de Mittag-Leffler de um parâmetro α, dada por

$$E_\alpha(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\alpha k + 1)}.$$

É relevante notar que no caso em que $\alpha = 1$ e da equação (1) temos que

$$E_1(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(k+1)} = \sum_{k=0}^{\infty} \frac{x^k}{x!} = e^x,$$

o que nos permite dizer que a função de Mittag-Leffler pode ser interpretada como uma generalização da função exponencial.

A função de Mittag-Leffler de dois parâmetros tem papel importante no desenvolvimento do cálculo fracionário. Assim, a mesma consiste na função que depende de dois parâmetros reais, α e β com $\alpha, \beta > 0$ e dada por

$$E_{\alpha,\beta}(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\alpha k + \beta)}.$$

Observe que quando $\beta = 1$, temos que

$$E_{\alpha,1}(x) = \sum_{k=0}^{\infty} \frac{x^k}{\Gamma(\alpha k + 1)} = E_\alpha(x)$$

e se reduz portanto à função de Mittag-Leffler de um parâmetro.

4 A Transformada de Laplace da Função de Mittag-Leffler

Nesta seção, faremos uso da generalização da função de Mittag-Leffler em termos da função exponencial para obter sua Transformada de Laplace. Deste modo, nos preocuparemos primeiramente em obter a Transformada de Laplace da função $t^k e^{at}$. Estes resultados podem ser verificados mais detalhadamente em Camargo (2006).

Utilizando a representação em série para e^x, temos que

$$\int_0^\infty e^{-t} e^{xt} dt = \sum_{k=0}^{\infty} \frac{x^k}{k!} \int_0^\infty e^{-t} t^k dt.$$

Pela definição da Função Gama, da existência da integral acima e da convergência da série obtida, resultados que podem ser verificados em livros de Cálculo I, segue que

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} \int_0^\infty e^{-t} t^k dt = \sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

donde

$$\int_0^\infty e^{-t}e^{xt}dt = \frac{1}{1-x}, \text{ se } |x| < 1.$$

Derivando a expressão anterior k vezes, com relação a x, temos

$$\int_0^\infty e^{-t}t^k e^{kt}dt = \frac{k!}{(1-x)^{k+1}}.$$

Lembre-se de que nosso objetivo inicial é encontrar uma expressão para a Transformada de Laplace da função $t^k e^{at}$. Desta forma, convenientemente faremos uma mudança de variáveis na equação anterior, tomando $a, s \in \mathbb{R}$ tais que $1 - x = s - a$. Assim, substituindo na equação, temos

$$\int_0^\infty e^{-st}t^k e^{at}dt = \frac{k!}{(s-a)^{k+1}}$$

e portanto, segue que

$$\mathcal{L}(t^k e^{at}) = \int_0^\infty e^{-st}t^k e^{at}dt = \frac{k!}{(s-a)^{k+1}}.$$

Procedendo da mesma forma para encontrar a Transformada de Laplace da função $t^{\beta-1}E_{\alpha,\beta}(at^\alpha)$ obtemos

$$\begin{aligned}\mathcal{L}(t^{\beta-1}E_{\alpha,\beta}(at^\alpha)) &= \int_0^\infty e^{-st}t^{\beta-1}E_{\alpha,\beta}(at^\alpha)dt \\ &= \sum_{k=0}^{\infty}\frac{a^k}{\Gamma(\alpha k+\beta)}\int_0^\infty e^{-st}t^{\alpha k+\beta-1}dt \\ &= \frac{k!s^{\alpha-\beta}}{(s^\alpha-a)^{k+1}}.\end{aligned}$$

5 Aplicação

Apresenta-se aqui a solução de um problema de valor inicial envolvendo uma equação de ordem arbitrária. Assim, o problema em questão é delimitado pelas seguintes condições e equação,

$$\begin{aligned}D_*^\alpha y(t) - \lambda y(t) = 0, t > 0, n-1 < \alpha < n \\ y^{(k)}(0) = b_k, b_k \in \mathbb{R}, k = 0, \cdots, n-1,\end{aligned}$$

sendo $D_*^\alpha y(t)$ a derivada fracionária segundo a definição de Caputo, a qual é descrita por Oliveira (2014).

Para resolver o problema, tomaremos a Transformada de Laplace da equação acima, obtendo

$$\mathcal{L}\left\{D_*^\alpha y(t) - \lambda y(t)\right\} = \mathcal{L}\left\{0\right\}$$

$$\Leftrightarrow \mathcal{L}\left\{D_*^{\alpha} y(t)\right\} - \lambda \mathcal{L}\left\{y(t)\right\} = 0.$$

Precisamos agora da Transformada de Laplace de $D_*^{\alpha} y(t)$, ou seja, da Transformada de Laplace da derivada de ordem arbitrária da função $y(t)$. Assim, utilizaremos a Transformada de Laplace para derivada fracionária segundo Caputo, a qual também é descrita com mais detalhes por Oliveira (2014), temos que

$$\mathcal{L}(D_*^{\alpha} f(t)) = s^{\alpha}\mathcal{L}[f(t)] - \sum_{k=0}^{n-1} s^{n-1-k} f^{(k)}(0).$$

Aplicando o resultado anterior na equação e escrevendo $\mathcal{L}(y(t)) = Y(s)$, temos

$$s^{\alpha} Y(s) - \sum_{k=0}^{n-1} s^{n-1-k} y^{(k)}(0) - \lambda Y(s) = 0,$$

e, resolvendo a equação com respetio a $Y(s)$, segue que

$$Y(s) = \sum_{k=0}^{n-1} \frac{s^{\alpha-k-1}}{s^{\alpha} - \lambda} y^{(k)}(0),$$

que, com pequenos ajustes, pode ser vista como a Transformada de Laplace da função de Mittag-Leffler, como visto na seção 4. Assim temos

$$\begin{aligned} Y(s) &= \sum_{k=0}^{n-1} \frac{s^{\alpha-k-1}}{s^{\alpha} - \lambda} y^{(k)}(0) \\ &= \sum_{k=0}^{n-1} \mathcal{L}(t^k E_{\alpha,k+1}(\lambda t^{\alpha}) b_k) \\ &= \mathcal{L}\left\{ \sum_{k=0}^{n-1} t^k E_{\alpha,k+1}(\lambda t^{\alpha}) \right\}. \end{aligned}$$

Usando a transformada inversa, obtemos a solução da equação pretendida, dada por

$$y(t) = y(t,\alpha) = \sum_{k=0}^{n-1} t^k E_{\alpha,k+1}(\lambda t^{\alpha}).$$

Referências

BOYCE, W. E.; DIPRIMA, R. C. **Equações Diferenciais Elementares e Problemas de Valores de Contorno,** 9.ed. Rio de Janeiro: LTC, 2009.

CAMARGO, R. F.; OLIVEIRA, E. C.; CHIACCHIO, A.O. **Sobre a Função de Mittag-Leffler**, R. P., 15/2006, UNICAMP, Campinas, SP, 2006.

GRIGOLETTO, E. C. **Equação Diferencial Fracionária e a Função de Mittag-Leffler.** 2014. 138 f. Tese (Doutorado em Matemática Aplicada), Instituto de Matemática, Estatística e Computação Científica, UNICAMP, Campinas, SP, 2014.

OLIVEIRA, D. S. **Derivada Fracionária e as Funções de Mittag-Leffler.** 2014. 106 f. Dissertação (Mestrado em Matemática Aplicada), Instituto de Matemática, Estatística e Computação Científica, UNICAMP, Campinas, SP, 2014.

Criptografia RSA

Guilherme de Loreno
Universidade Estadual do Oeste do Paraná-Unioeste
guilherme.loreno@unioeste.br

Daniela Maria Grande Vicente
Universidade Estadual do Oeste do Paraná-Unioeste
daniela.grande@unioeste.br

Resumo: Nesse trabalho apresentamos brevemente um dos métodos mais utilizados de criptografia de chave pública, o RSA, bem como um pouco da matemática requerida para a compreensão do mesmo. Descrevemos aqui, sucintamente, o funcionamento do método da Criptografia RSA, que nada mais é que um método para codificar mensagens, nas quais apenas o destinatário legítimo conseguirá decodificar. Neste artigo apresentamos também uma breve explicação de como funciona o método de codificação, assim como o de decodificação e uma discussão, em linhas gerais, do porquê que tal método sempre funciona.

Palavras-chave: Criptografia; Codificação; Decodificação.

1 Introdução

Atualmente, as pessoas cada vez mais vem usando a internet para realizar transações bancárias ou comerciais. Para que essas transações sejam seguras é preciso bons métodos de criptografia, pois é relativamente fácil interceptar mensagens enviadas por linha telefônica ou de internet, tornando-se necessário codificar essas mensagens de forma que só o seu destinatário possa decodifica-la.

Um dos métodos de criptografia de chave pública mais utilizados comercialmente é o RSA, e foi inventado em 1978 por R. L. Rivest, A. Shamir e L. Andleman. O método consiste basicamente em dois processos, o primeiro de codificar a mensagem e posteriormente decodificar. Mais especificamente, o método RSA consiste na escolha de dois números primos grandes (de 150 algarismos ou mais).

Para decodificar uma mensagem precisamos conhecer esses dois números primos, que chamaremos de p e q. A chave de codificação do RSA é a multiplicação deles $n = p.q$ e esse número é conhecido como a 'chave pública', pois n pode ser tornado público. Já a chave de decodificação é constituída pelos números p e q que são mantidos em segredo pelo usuário. Então, para decifrar o RSA basta fatorar o número n e encontrar p e q. Porém isso se torna muito trabalhoso e até mesmo praticamente impossível, pois no método RSA os números que devem ser fatorados são muito grandes, com mais de 150 algarismos. Podemos dizer então que

é disso que a segurança do RSA depende, da ineficiência dos métodos de fatoração atualmente conhecidos.

1.1 Aritmética Modular

Nesta seção apresentamos algumas definições e propriedades que nos serão úteis para a compreensão do método RSA. A Aritmética dos Restos, nada mais é que o estudo dos restos na divisão de números de inteiros por um inteiro n fixado. Em vista disso, podemos dizer que os restos das divisões desses inteiros repetem-se com período exatamente igual a n. As definição e propriedades de *Aritmética Modular* foram extraídas de COUTINHO (2013).

Definição 1. Diremos que dois inteiros a e b são *congruentes módulo* n se, e somente se, $n \mid b - a$ e escrevemos,

$$a \equiv b \, mod \, n.$$

Um outro modo de definir congruência modular é a seguinte: seja $n \in \mathbb{Z}$ e $n > 1$. Diremos que dois inteiros a e b são *congruentes módulo* n se, e somente se, a e b possuírem mesmo resto quando divididos por n. Simbolicamente,

$$a \equiv b \, mod \, n.$$

Proposição 2. *Tem-se que* $a \equiv b \, mod \, n$ *se, e somente se,* $n \mid b - a$.

Proposição 3. *Se* $a \equiv b \, mod \, n$ *e* $x \equiv y \, mod \, n$, *então:*

(i) $a + x \equiv b + y \, mod \, n$.

(ii) $a \cdot x \equiv b \cdot y \, mod \, n$.

Em particular,

$$a^k \equiv (b)^k \quad mod \, n, \quad \text{para qualquer } k \geq 0.$$

1.2 Potências

Uma outra aplicação das congruências é no cálculo de restos da divisão de uma potência por um número qualquer. Esse método é muito importante na implementação do RSA, pois reduz o problema de calcular o resto da divisão de um número muito grande, para valores menores, pois números menores são mais fáceis de serem trabalhados. Iremos ilustrar esse método através de dois exemplos.

Queremos calcular o resto da divisão de 10^{135} por 7. Observemos que $135 = 6 \cdot 22 + 3$ e $10^6 \equiv 1 \pmod 7$. Temos então as seguintes congruências módulo 7

$$10^{135} \equiv (10^6)^{22} \cdot 10^3 \equiv (1)^{22} \cdot 10^3 \equiv 6 \mod 7.$$

Logo, como $0 \leq 6 < 7$, o resto da divisão de 10^{135} por 7 é 6.

Agora queremos encontrar o resto da divisão de $2^{124\,512}$ por 31. A verdade é que o maior problema é encontrar o quociente dessa divisão, pelo motivo do dividendo ser um número muito grande. Já que, para o resto podemos utilizar congruências. Calculemos as potências de 2 módulo 31,

$$2^2 \equiv 4 \mod 31$$

$$2^3 \equiv 8 \mod 31$$

$$2^4 \equiv 16 \mod 31$$

$$2^5 \equiv 32 \equiv 1 \mod 31.$$

Dividindo 124 512 por 5 obtemos, $124\,512 = 5 \cdot 24\,902 + 2$. Portanto,

$$2^{124\,512} \equiv (2^5)^{24\,902} \cdot 2^2 \mod 31$$

Como $2^5 \equiv 1 \mod 31$, temos

$$2^{124\,512} \equiv (1)^{24\,902} \cdot 2^2 \quad \Rightarrow \quad 2^{124\,512} \equiv 4 \mod 31.$$

Portanto, como $0 \leq 4 < 31$, podemos concluir que o resto da divisão de $2^{124\,512}$ por 31 é 4.

Definição 4. Sejam $a, a', n \in \mathbb{Z}$ e $n > 1$. Diremos que a e a' são inversos módulo n se, se somente se,

$$a \cdot a' \equiv 1 \mod n.$$

Também dizemos que a' é o *inverso de a módulo n*, e vice-versa. Contudo nem todos os inteiros admitem inverso módulo n. Temos um teorema que nos garante que se dado um inteiro a quando que ele possuirá inverso módulo n.

Teorema 5. *Sejam $a < n$ inteiros positivos, a possui inverso módulo n se, e somente se, a e n não têm fatores primos em comum.*

Então temos que se, a admite inverso módulo n, em outras palavras, $mdc(a,n) = 1$ e $b, c \in \mathbb{Z}$ tais que

$$a \cdot b \equiv a \cdot c \ \ mod\, n$$

então a pode ser cancelado e podemos concluir que $b \equiv c \ \ mod\, n$.

Podemos utilizar o resultado desse teorema para resolver *congruências lineares*, que é uma equação do tipo

$$ax \equiv b \ \ mod\, n, \quad \text{para} \quad b \in \mathbb{Z} \tag{1}$$

Se $mdc(a,n) = 1$ então existe um $a' \in \mathbb{Z}$ tal que $a \cdot a' \equiv 1 \ \ mod\ n$. Multiplicando a equação (1) por a', obtemos

$$a' \cdot a \cdot x \equiv a' \cdot b \ \ mod\, n \quad \Rightarrow \quad x \equiv a' \cdot b \ \ mod\, n.$$

1.3 O Pequeno Teorema de Fermat

Teorema 6 (Teorema de Fermat)**.** *Se p é primo e $a \in \mathbb{Z}$ que não é divisível por p, então*

$$a^{p-1} \equiv 1 \ \ mod\, p.$$

Esse teorema, facilita muito a resolução de problemas que envolvem congruências, é considerado a base para a criação dos Testes de Primalidade atuais. Uma prova desse teorema pode ser encontrado em Coutinho (2008).

2 Criptografia

Nesta seção, trataremos da descrição do método RSA. O texto é baseado na referência Coutinho (2013), bem como os exemplos citados foram retidados desta obra.

2.1 Pré-codificação

Para usar o método RSA para codificar uma mensagem, primeiramente, devemos transformar o texto em uma sequência de números, essa etapa é chamada de *pré-codificação*. Suponha que a mensagem a ser codificada seja um texto composto somente por letras, a primeira etapa do processo é converter o texto em números, usando a seguinte ideia

$$A = 10, B = 11, C = 12, ..., Z = 35.$$

Onde o espaço entre duas palavras é representado pelo número 99. Após essa etapa, o que vamos obter é um grande número inteiro, e para continuar o método devemos determinar os parâmetros fundamentais do RSA, que são os dois números primos, que denotamos por p e q e, chamando n=$p.q$. A última fase de pré-codificação consiste em quebrar o número inicialmente encontrado em blocos menores que o tal número n e a escolha desses blocos não é única, devemos apenas tomar alguns cuidados, como exemplo, para que não iniciem em zero.

Para ilustrar tal ideia, digamos que queremos pré-codificar a frase *Paraty é linda* que é convertida no número

$$2510271029349914992118231310.$$

Precisamos determinar os parâmetros do sistema RSA que vamos usar, que são os dois primos distintos. A próxima etapa do processo é quebrar em blocos o longo número produzido anteriormente, os quais devem ser menores que n. Por exemplo, tomando $p = 11$ e $q = 13$, então $n = 143$. Para esse caso, a mensagem, pode ser quebrada nos seguintes blocos:

$$25 - 102 - 7 - 93 - 49 - 91 - 49 - 92 - 118 - 23 - 13 - 10.$$

Essa é a primeira etapa do método, em que apenas é feito a conversão da mensagem, neste caso um texto, para uma mensagem formada por blocos de números. A próxima etapa é a codificação.

2.2 Codificando e Decodificando

Para codificar uma mensagem precisamos do número n, que é o produto dos primos p e q e de um inteiro positivo e, de modo que

$$mdc(e, \varphi(n)) = 1$$

donde $\varphi(n) = (p-1)(q-1)$. Chamamos o par ordenado (n, e) de *chave de codificação* do sistema RSA que estamos usando. Então, é necessário codificar cada bloco, obtido da *pré-codificação*, e a mensagem codificada será dada pela sequência de blocos codificados. Para codificar um bloco $b < n$ e $b \in \mathbb{Z}_+$, vamos denotar por $C(b)$, devemos fazer

$$C(b) = \text{resto da divisão de } b^e \text{ por n}$$

que em termos de aritmética modular, é equivalente a

$$b^e \equiv C(b)\ mod\ n.$$

Levando em consideração o exemplo dado anteriormente, temos $p = 11$ e $q = 13$, e $n = 143$ e $\varphi(n) = 120$, precisamos escolher e, para esse caso, o menor valor possível para e é 7, o qual é o menor primo que não divide 120. Assim o bloco 102 da mensagem anterior é codificado como

$$C(102) = \text{o resto da divisão de } 102^7 \text{ por } 143\ ,$$

desenvolvendo os cálculos, obtemos que

$$C(102) = 119$$

pois,

$$102^7 \equiv (-41)^7 \equiv -41^7 \equiv -81 \cdot 138 \equiv -24 \equiv 119 \ \ mod(143).$$

Fazendo o mesmo para toda a mensagem, obtemos a seguinte sequência de blocos:

$$64 - 119 - 6 - 119 - 102 - 36 - 130 - 36 - 27 - 79 - 23 - 117 - 10.$$

Agora para decodificar um bloco de mensagem codificada, precisamos ter em mãos dois números que são, n e o inverso de e em $\varphi(n)$, que denotamos por d,

$$d = \text{inverso de } e \text{ módulo } \varphi(n).$$

É muito fácil calcular d, desde que $\varphi(n)$ e e sejam conhecidos, basta aplicar o algoritmo euclidiano estendido, que nada mais é que dados $a, b \in \mathbb{Z}_+$ e digamos que $d = mdc(a, b)$, então existem α e β tais que

$$\alpha \cdot a + \beta \cdot b = d \tag{2}$$

sendo que α e β podem não ser únicos.

Podemos afirmar que (2) sempre admite solução, pois ela é um tipo de equação bem particular, que nada mais é que um tipo de *equação diofantina*, pois $a, b, d \in \mathbb{Z}$ e α, β também são inteiros a serem determinados e temos um resultado acerca desse tipo de equação, que nos garante que, toda equação diofantina do tipo $a \cdot \alpha + b \cdot \beta = d$ admite solução se, e somente se, $mdc(a, b)$ divide d. Mas, como em (2), $d = mdc(a, b)$ e $d \mid d$, portanto a equação admite uma família de soluções, por isso dizemos que α e β não são únicos.

No exemplo que estamos seguindo, temos que $n = 143$ e $e = 7$. Aplicando o algoritmo euclidiano estendido para calcular d, dessa forma, dividindo $\varphi(143) = 120$ por 7 obtemos

$$120 = 7 \cdot 17 + 1 \Rightarrow 1 = 120 + (-17) \cdot 7.$$

logo, o inverso de 7 módulo 120 é -17. Como precisamos que d seja positivo, temos que $d = 120 - 17 = 103$ que é o menor inteiro positivo congruente a -17 módulo 120. Assim para decodificar o bloco 119 da mensagem codificada, calculamos a forma reduzida de

$$119^{103} \equiv 102 \ mod\, 143.$$

Vamos chamar o par (n, d) de *chave de decodificação*. Logo, seja a um bloco de mensagem codificada, no processo anterior, então iremos chamar de $D(a)$ como o resultado da decodificação, onde $D(a)$ é obtido da seguinte maneira,

$$D(a) = \text{resto da divisão de } a^d \text{ por } n$$

o que em termos de aritmética modular é equivalente a

$$a^d \equiv D(a) \ mod\, n.$$

Observemos que chamamos de a um bloco de mensagem codificada, ou seja,

$$a = C(b) \text{ para algum } b, \text{ que é um bloco de } n$$

Assim, para que o código seja útil deve ocorrer que

$$D(a) = D(C(b)) = b.$$

Portanto, para decodificar uma mensagem precisamos, além do próprio n, conhecer o inverso d de e módulo $\varphi(n)$.

2.3 Explicando o funcinamento do RSA

Dissemos anteriormente que se b é um bloco da mensagem original, só será legítimo chamar o processo acima de decodificação se

$$D(a) = D(C(b)) = b,$$

com $1 \leq b \leq n - 1$.

Para isso devemos nos convencer que isso, de fato, sempre acontece. Sendo assim, basta apenas provar que $D(C(b)) \equiv b \ (mod\ n)$, isso é suficiente pois tanto b quanto $D(C(b))$ estão no intervalo que vai de 1 a $n - 1$, logo como os dois estão em um mesmo intervalo e deixam o mesmo resto na divisão por n isso implica que tais números são iguais. Isso justifica o fato que

precisamos escolher $b < n$, e também manter os blocos separados, mesmo depois da codificação. Assim, temos por definição que

$$D(C(b)) \equiv (b^e)^d \equiv b^{ed} \pmod{n} \tag{3}$$

mas como d é o inverso de e módulo $\varphi(n)$, logo

$$ed = 1 + k \cdot \varphi(n)$$

para algum $k \in \mathbb{Z}$. Notemos que tanto e quanto d são inteiros maiores de 2 e $\varphi(n) > 0$,então $k > 0$. Substituindo em (3)

$$b^{ed} \equiv b^{1+k\cdot\varphi(n)} \equiv \left(b^{\varphi(n)}\right)^k \cdot b \mod n. \tag{4}$$

Lembremos que $n = p.q$, onde p e q são primos distintos. Calculando a forma reduzida de b^{ed} módulo p e módulo q, que é análogo para ambos, assim faremos apenas para um deles. Digamos que queremos calcular a forma reduzida de b^{ed} módulo p, assim como temos que

$$ed = 1 + k \cdot \varphi(n) = 1 + k(p-1) \cdot (q-1)$$

logo,

$$b^{ed} \equiv b \cdot \left(b^{p-1}\right)^{k(q-1)} \pmod{p}$$

Para que possamos usar o Teorema de Fermat, precisamos ter que *p não divide b*. Primeiramente, considerando o caso em que *p divide b*, nesses termos temos

$$b \equiv 0 \pmod{p}$$

e assim a congruência é imediatamente verificada e consequentemente

$$b^{e\cdot d} \equiv b \ (mod\, p),$$

para qualquer que seja b.

Caso contrário, *p não divide b* então $b^{p-1} \equiv 1 \ (modp)$ pelo Teorema de Fermat, e obtemos $b^{ed} \equiv b \ (mod\, p)$ e anteriormente vimos que $b^{e\cdot d} \equiv b \ (mod\, p)$, para quando $p \mid b$ e analagomente podemos mostrar que $b^{ed} \equiv b \ (mod\ q)$, ou seja, $b^{ed} - b$ é divisível por p e q que são primos distintos e como temos que $mdc(p, q) = 1$, donde , $pq \mid b^{ed} - b$ e como $n = pq$, concluímos que

$$b^{ed} \equiv b \ (mod\ n).$$

Mostrados os dois casos, encerra-se nossa demonstração.

2.4 Segurança do RSA

Um ponto crucial do método RSA é a escolha dos primos p e q, pois se tais números forem pequenos, o sistema seria fácil de quebrar, porém não basta escolhê-los grandes. Digamos que queremos implementar o RSA com chave pública (n, e), de modo que n seja um inteiro com aproximadamente r algarismos, sendo assim para construirmos n, devemos escolher um primo p que satisfaça

$$\frac{4r}{10} < p < \frac{45r}{100}$$

no qual esse intervalo, representa o número de algarismos de p. Em seguida, deve-se escolher q próximo de

$$\frac{10^r}{p}.$$

Para construir um tal número n precisamos de dois primos, suponha que tenham 104 e 127 algarimos, respectivamente, podemos perceber que esses primos estão longe bastante um do outro, o que acaba se tornando impraticável fatorar n pelo teorema de Fermat. Outro fator que deve-se tomar cuidado é que os números $p-1$, $q-1$, $p+1$ e $q+1$ não têm fatores primos pequenos, pois dessa forma seria de facilmente fatorado por alguns dos algoritmos de fatoração conhecidos.

Lembrando que a chave n é igual ao produto de dois primos p e q, que foram escolhidos conforme os critérios acima descritos, assim somos levados a pensar que tendo posse de n, bastaria fatorá-lo, descobrir p e q e usá-los para encontrar e, assim teríamos a chave de decodificação (n, d) e assim poderíamos reconstituir a mensagem original. Entretanto, tudo isso pode parecer muito simples, em princípio, mas na prática é totalmente inviável pois, não existem computadores rápidos o suficiente, nem algoritmos eficientes, que nos permitam fatorar um número inteiro muito grande que não tenham fatores relativamente pequenos, em suma podemos afirmar que não existe algoritmo conhecido capaz de fatorar inteiros grandes de modo realmente eficientes.

3 Considerações Finais

Podemos concluir que apesar do RSA ser um método de criptografia de chave-pública, ou seja, qualquer usuário pode ter acesso a mensagem criptografada, apenas o destinatário legítimo consegue ler a mensagem, isso por que ele é o único que possui a chave para decodificar a mensagem. A segurança do método é pautada na impossibilidade de se fatorar números inteiros, em outras palavras, não existem métodos de fatoração eficientes a fim de tornar viável a tentativa de decifrar mensagens criptografadas usando o RSA.

Referências

COUTINHO S. C. **Números Inteiros e Criptografia RSA**. Coleção Matemática e Aplicações. Rio de Janeiro: IMPA, 2013.

COUTINHO S. C. **Criptografia**. Rio de Janeiro: IMPA, 2008.

Um algoritmo de programação linear sequencial

Renato Massamitsu Zama Inomata
UNIOESTE - Universidade Estadual do Oeste do Paraná
Centro de Ciências Exatas e Tecnológicas - Curso de Engenharia Civil
renato.inomata@gmail.com

Paulo Domingos Conejo
UNIOESTE - Universidade Estadual do Oeste do Paraná
pconejo33@gmail.com

Resumo: A Programação Linear Sequencial (PLS) é destinada à solução de problemas de otimização não lineares. O método PLS é baseado em teoria de aproximações lineares, e consiste em resolver uma sequência de problemas de otimização lineares. Neste trabalho apresentamos um algoritmo de PLS e resolvemos um exemplo numérico. Foi introduzida a teoria de convergência deste algoritmo PLS, em que, a função de mérito, fundamental no estudo de convergência, é uma combinação convexa entre a função objetivo e uma medida de viabilidade.

Palavras-chave: Convergência; Otimização não Linear; PLS.

1 Introdução

A Otimização é a área da Programação Matemática que consiste em minimizar ou maximizar uma determinada função, f, denominada função objetivo, sujeita a determinadas condições, Ω, denominadas restrições (RIBEIRO e KARAS, 2013). Um problema de otimização é usualmente escrito na forma

$$\begin{aligned} &\text{minimizar} \quad f(x) \\ &\text{sujeito a} \quad x \in \Omega, \end{aligned} \tag{1}$$

com $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$, $h : \mathbb{R}^n \to \mathbb{R}^p$, $g : \mathbb{R}^n \to \mathbb{R}^q$ e $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ funções diferenciáveis.

Programação Linear Sequencial consiste em uma técnica de aproximação linear, em que, em cada iteração, é resolvido um problema de otimização linear. A linearização é feita a partir da fórmula de Taylor. Várias vertentes para a PLS estão disponíveis na literatura (SENNE, 2009) e os citados neste. Estas vertentes alteram a forma de como a função de mérito (fundamental no estudo de convergência) é definida para garantir a convergência. Este trabalho foi embasado na dissertação de mestrado de Senne (2009), sobre *Otimização Topológica de Mecanismos Flexíveis* e no artigo resultado desta dissertação (GOMES e SENNE, 2011).

2 Um algoritmo de Programação Linear Sequencial

Com a introdução de variáveis de folga, um problema de otimização não linear geral (com desigualdades) como em (1) pode ser transformado em um problema com somente restrições de igualdade. Neste trabalho temos interesse no problema não linear

$$\begin{aligned} \min \quad & f(x) \\ \text{s.a.} \quad & h_i(x) = 0 \qquad\qquad i = 1, \ \ldots \ , m \\ & x^{\min} \leq x \leq x^{\max}, \end{aligned} \tag{2}$$

onde $x = (x_1 \ \ldots \ x_n)^T \in \mathbb{R}^n$ é o vetor para as n variáveis do problema; $f : \mathbb{R}^n \to \mathbb{R}$ e $h_i : \mathbb{R}^n \to \mathbb{R}$ $(i = 1, \ \ldots \ , m)$ são funções com derivadas parciais primeiras Lipschitz contínua, de modo que f é a função objetivo e h_i com $i = 1, \ \ldots \ , m$ são as m restrições de igualdade do problema; $x^{\min} = (x_1^{\min} \ \ldots \ x_n^{\min})^T$ e $x^{\max} = (x_1^{\max} \ \ldots \ x_n^{\max})^T$ são os vetores que contêm as restrições que limitam, respectivamente, inferiormente e superiormente as variáveis de x.

Um ponto $\bar{x} \in \mathbb{R}^n$ é um ponto dito *factível* ou *viável* se todas as suas restrições forem satisfeitas.

As aproximações lineares de f e h_i, que são feitas a partir da fórmula de Taylor, são dadas por

$$f(x + s) \approx f(x) + \nabla f(x)^T s \quad e \quad h_i(x + s) \approx h_i(x) + \nabla h_i(x)^T s, \qquad i = 1, \ldots, m. \tag{3}$$

Denotamos

$$A(x) = [\nabla h_1(x) \ \ldots \ \nabla h_m(x)]^T \tag{4}$$

a matriz Jacobiana das restrições, e

$$C(x) = [h_1(x) \ \ldots \ h_m(x)]^T \tag{5}$$

o vetor que possui os valores de cada uma das restrições no ponto x. Definimos também o conjunto X que representa as restrições canalizadas

$$X = \{x \in \mathbb{R}^n \mid x^{\min} \leq x \leq x^{\max}\}.$$

Usando os itens (4) e (5), podemos reescrever (3). Assim, para a função objetivo

$$f(x + s) \approx f(x) + \nabla f(x)^T s \equiv L(x, s)$$

e para as funções de restrição

$$C(x + s) \approx C(x) + A(x)s.$$

É possível obter a solução aproximada de um problema não linear como (2) com o auxílio da *Programação Linear Sequencial* (PLS), que consiste num método iterativo que se utiliza de uma sequência de problemas de programação linear, nos quais tanto a função objetivo quanto as restrições são aproximações lineares das funções envolvidas no problema (2). Assim, o subproblema linearizado em cada iteração k fica

$$\begin{aligned} \min \quad & \nabla f(x^k)^T s \\ \text{s.a.} \quad & A(x^k)s + C(x^k) = 0 \\ & \max\{-\delta_k, x^{\min} - x^k\} \leq s \leq \min\{\delta_k, x^{\max} - x^k\}, \end{aligned} \tag{6}$$

onde δ_k é a *região de confiança* do subproblema na iteração k, cuja função é evitar que este problema seja ilimitado e também ajudar a garantir a convergência global do algoritmo. Essa região indica que a linearização (6) do problema (2) o representa suficientemente bem.

Como outros métodos iterativos, partindo de x^k buscamos obter x^{k+1}, que neste caso é calculado com s_c, que é a solução ótima para o subproblema da iteração k. Assim

$$x^{k+1} = x^k + s_c.$$

É possível que as restrições para o problema principal (2) não sejam satisfeitas, tendo assim uma *inviabilidade* no ponto. Para medir o quão inviável é um ponto, utilizamos a função $\varphi : \mathbb{R}^m \to \mathbb{R}$, dada por

$$\varphi(x) = \frac{1}{2}\|C(x)\|_2^2.$$

Para o problema linearizado utilizamos a função $M : \mathbb{R}^m \to \mathbb{R}$ para medir a *inviabilidade*, sendo

$$M(x, s) = \frac{1}{2}\|A(x)s + C(x)\|_2^2.$$

Ainda sobre a inviabilidade, um ponto é dito φ-estacionário se x é um ponto que atende as condições de Karush-Kuhn-Tucker (KKT), descritas por Kuhn e Tucker (1951), do problema

$$\begin{aligned} \min \quad & \varphi(x) \\ \text{s.a.} \quad & x \in X. \end{aligned} \tag{7}$$

Quando se resolve o problema original (2) de maneira aproximada, devemos tentar minimizar tanto a função objetivo quanto a sua inviabilidade. Contudo, nem sempre isso será possível, assim é necessário estabelecer uma função de mérito para priorizar o decrescimento de uma das duas funções (objetivo ou inviabilidade). Através de uma combinação convexa entre $f(x)$ e $\varphi(x)$, estabelecemos uma função de mérito $\psi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ dada por

$$\psi(x, \theta) = \theta f(x) + (1 - \theta)\varphi(x), \qquad \theta \in [0, 1]. \tag{8}$$

O critério de aceitação ou rejeição do passo s_c é estipulado através das *reduções reais* (A_{red}) e *prevista* (P_{red}), de maneira que

$$A_{red}^{opt} = f(x) - f(x+s) \quad e \quad A_{red}^{fct} = \varphi(x) - \varphi(x+s) \tag{9}$$

são respectivamente as reduções reais para a função objetivo e para a inviabilidade do problema (2). Seja também

$$P_{red}^{opt} = -\nabla f(x)^T s \quad e \quad P_{red}^{fct} = M(x,0) - M(x,s) = \frac{1}{2}\|C(x)\|_2^2 - \frac{1}{2}\|A(x)s + C(x)\|_2^2 , \tag{10}$$

respectivamente, as reduções previstas para a função objetivo e inviabilidade do subproblema linearizado (6).

Finalmente, por meio dos itens (9) e (10), definimos A_{red} e P_{red} que são, respectivamente, as reduções reais e previstas pela função de mérito de maneira que

$$A_{red} = \theta A_{red}^{opt} + (1-\theta)A_{red}^{fct} \quad e \quad P_{red} = \theta P_{red}^{opt} + (1-\theta)P_{red}^{fct} .$$

O parâmetro θ é recalculado para cada iteração. Para determiná-lo utilizamos, em cada iteração k,

$$\theta_k^{min} = \min\{1, \theta_0, \ldots, \theta_{k-1}\} \quad e \quad \theta_k^{large} = \left[1 + \frac{N}{(k+1)^{1.1}}\right]\theta_k^{min} , \tag{11}$$

sendo $N \geq 0$ um parâmetro que garante que θ tenha um decréscimo não monótono no decorrer das iterações. Esta propriedade será utilizada adiante para garantir a convergência do algoritmo. Definimos também

$$\begin{aligned} \theta_k^{sup} &= \sup\left\{\theta \in [0,1] \,|P_{red} \geq 0.5 P_{red}^{fct}\right\} \\ &= \begin{cases} 0.5\left(\dfrac{P_{red}^{fct}}{P_{red}^{fct} - P_{red}^{opt}}\right), & \text{se } P_{red}^{opt} \leq \dfrac{1}{2}P_{red}^{fct} \\ 1, & \text{caso contrário.} \end{cases} \end{aligned} \tag{12}$$

Por fim, θ_k é determinado através de

$$\theta_k = \min\left\{\theta_k^{sup}, \theta_k^{large}\right\} . \tag{13}$$

Agora será apresentado o Algoritmo PLS, que resolve o problema (2) de maneira aproximada.

Algoritmo PLS

Dado um ponto inicial $x^0 \in X$, um raio de confiança δ_0, tal que $\delta_0 \geq \delta_{\min} \geq 0$, um θ_0 inicial tal que $\theta_0 = \theta_{\max} = 1$, e $k = 0$.

1. Tentar encontrar um s_n que satisfaça

$$A(x^k)s_n = -C(x^k)$$
$$\max\{-0.8\delta_k, x^{\min} - x^k\} \leq s_n \leq \min\{0.8\delta_k, x^{\max} - x^k\}$$

2. Caso não seja possível encontrar s_n

   (a) $d_n \leftarrow -\nabla\varphi(x^k)$

   (b) Determinar $\bar{\alpha}$, solução de

$$\begin{aligned} \min \quad & M(x^k, \alpha d_n) \\ \text{s.a.} \quad & \max\{-0.8\delta_k, x^{\min} - x^k\} \leq \alpha d_n \leq \min\{0.8\delta_k, x^{\max} - x^k\} \;\; e \;\; \alpha \leq 0 \end{aligned}$$

   (c) $s_n^d \leftarrow \bar{\alpha} d_n$

   (d) Determinar s_c tal que $M(x^k, s_c) \leq M(x^k, s_n^d)$.

3. Caso contrário, partindo de s_n, encontrar s_c, solução de

$$\begin{aligned} \min \quad & \nabla f(x^k)^T s \\ \text{s.a.} \quad & A(x^k)s = -C(x^k) \\ & \max\{-\delta_k, x^{\min} - x^k\} \leq \alpha d_n \leq \min\{\delta_k, x^{\max} - x^k\} \end{aligned}$$

4. Determinar $\theta_k \in [0, \theta^{\max}]$

5. Se $A_{red} \geq 0.1P_{red}$, então $x^{k+1} \leftarrow x^k + s_c$

   (a) Se $A_{red} \geq 0.5P_{red}$, então $\delta_{k+1} = \min\{2.5\delta_k, \|x^{\max} - x^{\min}\|_\infty\}$

   (b) Senão $\delta_{k+1} \leftarrow \delta_{\min}$

   (c) $\theta^{\max} \leftarrow 1$

6. Senão $\delta_{k+1} \leftarrow 0.25\|s_c\|_\infty$, $x^{k+1} \leftarrow x^k$, $\theta^{\max} \leftarrow \theta^k$.

O algoritmo está apresentado em 6 etapas. Partindo das condições iniciais previamente estabelecidas, a primeira dessas etapas consiste em procurar por um ponto viável para o problema dentro de uma região de confiança para qual a linearização seja uma boa representação do problema não-linear. Em outras palavras, achar um ponto que satisfaça todas as suas restrições e esteja dentro da região delimitada por um raio de $0.8\delta_k$.

A segunda etapa é empregada caso não seja possível determinar um ponto viável dentro da região de confiança. É definido d_n, como a direção de descida para a função $\varphi(x)$. Assim,

conhecida essa direção o algoritmo minimiza a função de inviabilidade para o problema linearizado sujeita a restrição do conjunto X e de uma região de confiança de $0.8\delta_k$, de maneira que os pontos definidos pela restrição sejam apenas os que se encontram ao longo da direção d_n.

Se através da etapa 1 for possível encontrar um ponto viável o algoritmo vai para a etapa 3, resolvendo o subproblema apresentado em (6).

Na etapa 4 é calculado o parâmetro θ_k através de (11) a (13), de maneira que este seja um valor entre 0 e $\theta^{\max}$.

A quinta etapa se refere aos procedimentos quanto à aceitação do passo s_c encontrado. A condição inicial $A_{red} \geq 0.1P_{red}$ determina se o passo é aceito. Se a desigualdade se verifica, o passo s_c é aceito e o valor de x^k é atualizado para x^{k+1}. Além disso, se A_{red} for suficientemente maior que uma fração de P_{red} (neste caso $0.5P_{red}$), δ_k é aumentado, sendo este definido pelo menor valor entre $2.5\delta_k$ e a norma máxima entre $x^{\max}$ e $x^{\min}$. Caso a diferença entre A_{red} e P_{red} não seja suficientemente grande, δ_k continua o mesmo para a próxima iteração. Finalmente, o parâmetro $\theta^{\max}$ recebe 1 como seu valor.

Caso não seja satisfeito o critério de aceitação do passo s_c, o algoritmo entra na sexta etapa, em que o raio da região de confiança da próxima iteração, δ_{k+1}, é definido por $0.25\,\|s_c\|_\infty$, enquanto que os valores de x^k permanecem os mesmos e o parâmetro $\theta^{\max}$ recebe o valor de θ_k.

3 Estudo de convergência do algoritmo

Apesar de não ser demonstrado neste trabalho, o algoritmo pode encerrar suas iterações de três maneiras diferentes (SENNE, 2009). Partindo disso, será demonstrado que todo ponto de acumulação (ou ponto limite) da sequência de $\{x^k\}$ é φ-estacionário, ou seja, que atende as condições de KKT para o problema (7). Ao longo desta seção são utilizados alguns conceitos e propriedades de Lima (1976).

O Lema a seguir demonstra que, se uma sequência gerada pelo Algoritmo PLS admite um ponto limite não φ-estacionário, A_{red} está afastado de zero.

Lema 1. *Suponha que x^* não seja φ-estacionário e que K_1 seja um conjunto infinito de índices tal que $\lim_{k\in K_1} x^k = x^*$. Então, as componentes do conjunto $\{\delta_k \mid k \in K_1\}$ estão afastadas do zero. Além disso, existe $c_2 \geq 0$ tal que, para $k \in K_1$ suficientemente grande, temos que $A_{red} \geq c_2$.*

Prova. Vide a prova do Lema 3.5 em Gomes e Senne (2011). □

Hipótese H1. A sequência $\left\{x^k\right\}_{k=1}^{\infty}$ gerada pelo Algoritmo PLS é limitada.

Utilizando o Lema 1 e a Hipótese H1, o Teorema a seguir demonstra que todo ponto limite de uma sequência $\{x^k\}_{k=1}^{\infty}$ gerada pelo Algoritmo PLS é φ-estacionário.

Teorema 2. *Seja $\{x^k\}_{k=1}^{\infty}$ a sequência (infinita) gerada pelo algoritmo PLS. Suponha que a Hipótese H1 seja atendida. Então, todo ponto limite de $\{x^k\}_{k=1}^{\infty}$ é φ-estacionário.*

Prova. Suponha que $x^* \in X$ é um ponto de acumulação de $\{x^k\}$ e que não seja φ-estacionário. Seja também $L_k = L(x^k, s)$, $\varphi_k = \varphi(x^k)$, $\psi_k = \psi(x^k, \theta_k)$, para todo $k \in \mathbb{N}$. Assim, para todo $k \in \mathbb{N}$ temos de (8) que

$$\begin{aligned}\psi_{k+1} &= \theta_{k+1}L_{k+1} + (1-\theta_{k+1})\varphi_{k+1} \\ &= \theta_{k+1}L_{k+1} + (1-\theta_{k+1})\varphi_{k+1} - [\theta_k L_{k+1} + (1-\theta_k)\varphi_{k+1}] + [\theta_k L_{k+1} + (1-\theta_k)\varphi_{k+1}] \\ &= (\theta_{k+1}-\theta_k)L_{k+1} + (\theta_k - \theta_{k+1})\varphi_{k+1} + [\theta_k L_{k+1} + (1-\theta_k)\varphi_{k+1}]. \end{aligned} \tag{14}$$

Seja $\beta_k \equiv A_{red}(x^k, s, \theta_k) \geq 0$, ou seja,

$$\beta_k \equiv \theta_k(L_k - L_{k+1}) + (1-\theta_k)(\varphi_k - \varphi_{k+1}).$$

Assim, aplicando a distributiva e subtraindo β_k de (14)

$$\begin{aligned}\psi_{k+1} &= (\theta_k - \theta_{k+1})(\varphi_{k+1} - L_{k+1}) + [\theta_k L_k + (1-\theta_k)\varphi_k] - \beta_k \\ &= (\theta_k - \theta_{k+1})(\varphi_{k+1} - L_{k+1}) + \psi_k - \beta_k. \end{aligned} \tag{15}$$

Pelo Lema 1, $\beta_k \geq c_2 \geq 0$ para um conjunto infinito de índices. Utilizando $k = k+1$ para θ^{min} e θ^{large} definidos em (11), temos

$$\begin{aligned}\theta_{k+1}^{min} = \theta_{k+1}^{large}(1 + \frac{N}{(k+2)^{1.1}})^{-1} \leq \theta_k \\ \implies \theta_{k+1}^{large} \leq \theta_k(1 + \frac{N}{(k+2)^{1.1}}). \end{aligned} \tag{16}$$

Sabemos também que

$$\theta_{k+1} = \min\left\{\theta_{k+1}^{large}, \theta_{k+1}^{sup}\right\}. \tag{17}$$

Por (16) e (17)

$$\theta_{k+1} \leq \theta_k(1 + \frac{N}{(k+2)^{1.1}}) \implies \theta_k + \frac{\theta_k N}{(k+2)^{1.1}} - \theta_{k+1} \geq 0 \tag{18}$$

para todo $k \in \mathbb{N}$. Então, por (15) e (18),

$$\begin{aligned}\psi_{k+1} &= (\theta_k - \theta_{k+1} + \frac{\theta_k N}{(k+2)^{1.1}})(\varphi_{k+1} - L_{k+1}) + \psi_k - \beta_k - \frac{\theta_k N}{(k+2)^{1.1}}(\varphi_{k+1} - L_{k+1}) \\ &\leq (\theta_k - \theta_{k+1} + \frac{\theta_k N}{(k+2)^{1.1}})c + \psi_k - \beta_k + \frac{\theta_k N}{(k+2)^{1.1}}c = (\theta_k - \theta_{k+1})c + \psi_k - \beta_k + 2\frac{\theta_k N}{(k+2)^{1.1}}c. \end{aligned}$$

Escrevendo a inequação acima para $k = 1, 2, \ldots$, e adicionando-as entre si, temos que

$$\sum_{j=0}^{k-1} \psi_{j+1} \leq \sum_{j=0}^{k-1} (\theta_j - \theta_{j+1})c + \sum_{j=0}^{k-1} \psi_j - \sum_{j=0}^{k-1} \beta_j + \sum_{j=0}^{k-1} \frac{\theta_j N}{(j+2)^{1.1}} c$$

$$\implies \psi_k \leq (\theta_0 - \theta_k)c + \psi_0 - \sum_{j=0}^{k-1} \beta_j + \sum_{j=0}^{k-1} \frac{2\theta_j N}{(j+2)^{1.1}} c \leq 2c + \psi_0 - \sum_{j=0}^{k-1} \beta_j + \sum_{j=0}^{k-1} \frac{2cN}{(j+2)^{1.1}} \qquad (19)$$

para todo $k \in \mathbb{N}$. Como $\sum_{j=0}^{\infty} \frac{2cN}{(j+2)^{1.1}}$ é convergente e $\beta_k \geq 0$ para todo k, (19) implica que ψ_k é ilimitada abaixo, contradizendo a Hipótese H1. □

A partir deste momento apresentaremos os resultados que mostram que o algoritmo encontra um ponto estacionário do tipo KKT. Os Lemas de 3 a 5 servirão de base para a demonstração do Lema 6. Finalmente, o Teorema 7 condensará os resultados obtidos pelo Teorema 2 e pelo Lema 6.

Hipótese H2. Seja $\tilde{s}_n$ o vetor obtido no passo 2c do Algoritmo PLS. Neste caso, devemos ter

$$\|\tilde{s}_n\|_2 \leq \mathcal{O}(\|C(x^k)\|_2).$$

Essa hipótese pode ser atendida conforme exemplificado na página 59 de Senne (2009).

Lema 3. *Seja $\{x^k\}_{k=1}^{\infty}$ uma sequência gerada pelo algoritmo PLS. Suponha que a subsequência $\{x^k\}_{k \in K_1}$ convirja para o ponto viável e regular x^*, que não é estacionário para o problema* (2)*. Então, existem $c_1, k_1, \delta' \geq 0$ tais que, para $x \in \{x^k \mid k \in K_1, k \geq k_1\}$, temos*

$$L(x, s_n) - L(x, s_c) \geq c_1 \min\{\delta, \delta'\}.$$

Prova. Vide a prova do Lema 3.8 de Gomes e Senne (2011). □

Lema 4. *Suponha que a Hipótese H2 seja satisfeita e que sejam válidas as hipóteses do Lema 3. Então, existem $\beta, c_2, k_2 \geq 0$ tais que, se $x \in \{x^k \mid k \in K_1, k \geq k_2\}$ e $\|C(x)\|_2 \leq \beta\delta_k$, então*

$$L(x, 0) - L(x, s_c) \geq c_2 \min\{\delta, \delta'\} \quad e \quad \theta^{sup}(x, \delta) = 1,$$

onde θ^{sup} é definido em (12) *e δ' é definido no Lema 3.*

Prova. Vide a prova do Lema 3.9 de Gomes e Senne (2011). □

O Lema a seguir ilustra outra consequência da existência de um ponto viável e regular que não é estacionário para o problema (2). A ideia geral é que se θ_k estivesse afastado de zero, seria possível gerar uma sequência de A_{red} também afastada de zero, contradizendo a Hipótese H1.

Lema 5. *Suponha que as hipóteses dos Lemas 3 e 4, e a Hipótese H1 são antedidas. Então* $\lim_{k\to\infty} \theta_k = 0$.

Prova. Vide a prova do Lema 3.10 de Gomes e Senne (2011). □

O Lema a seguir mostrará que se todos os pontos limites da sequência gerada pelo Algoritmo PLS são viáveis e regulares, então um deles será necessariamente estacionário para o problema 2.

Lema 6. *Seja* $\{x^k\}_{k=1}^{\infty}$ *uma sequência infinita gerada pelo algoritmo PLS. Suponha que todos os pontos limite de* $\{x^k\}_{k=1}^{\infty}$ *sejam factíveis e regulares e que as Hipóteses H1 e H2 sejam válidas. Então, existe um ponto limite da sequência* $\{x^k\}_{k=1}^{\infty}$ *que é estacionário para o problema* (2).

Prova. Vide a prova do Lema 3.13 de Gomes e Senne (2011). □

Teorema 7. *Seja* $\{x^k\}_{k=1}^{\infty}$ *uma sequência infinita gerada pelo Algoritmo PLS. Suponha que as Hipóteses H1 e H2 sejam válidas. Então, todos os pontos de acumulação de* $\{x^k\}_{k=1}^{\infty}$ *são* φ*-estacionários. Além disso, se todos os pontos de acumulação* $\{x^k\}_{k=1}^{\infty}$ *são viáveis e regulares, existe um ponto de acumulação* x^* *que é estacionário para o problema* (2). *Em particular, se todos os pontos* φ*-estacionários são viáveis e regulares, existe uma subsequência de* $\{x^k\}_{k=1}^{\infty}$ *que converge para um ponto estacionário regular de* (2).

Prova. Este teorema é consequência direta do Lema 6 e do Teorema 2. □

4 Um teste numérico

Para exemplificar a aplicação do Algoritmo PLS, buscamos resolver o seguinte problema

$$\begin{aligned} \min \quad & f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\ \text{s.a.} \quad & x_1^2 + x^2 - 1 \\ & -2 \leq x_1 \leq 2 \\ & -3 \leq x_2 \leq 3. \end{aligned}$$

O Algoritmo PLS foi implementado em Matlab e os subproblemas foram resolvidos com o método dos Pontos Interiores. Os parâmetros iniciais foram $x^0 = [1 \quad 1.5]^T$, $\delta_0 = 1$, $\delta_{\min} = 10^{-8}$ e $N = 1$. A Figura 1 apresenta as curvas de nível da função objetivo, a restrição dada pelo círculo de raio 1, o ponto $[0.787 \quad 0.617]^T$ que é a solução aproximada para o problema e o ponto inicial representado pelo asterisco. O eixo horizontal representa a variável x_1 e o vertical x_2. O valor da função para a solução aproximada é $f(0.787, 0.617) = 0.046$. Consideramos como critério de parada que dois pontos aceitos pelo algoritmo possuam uma diferença menor que 10^{-7} e a inviabilidade deve ser menor que 10^{-5}.

As demais figuras mostram a evolução do algoritmo ao longo de suas iterações.

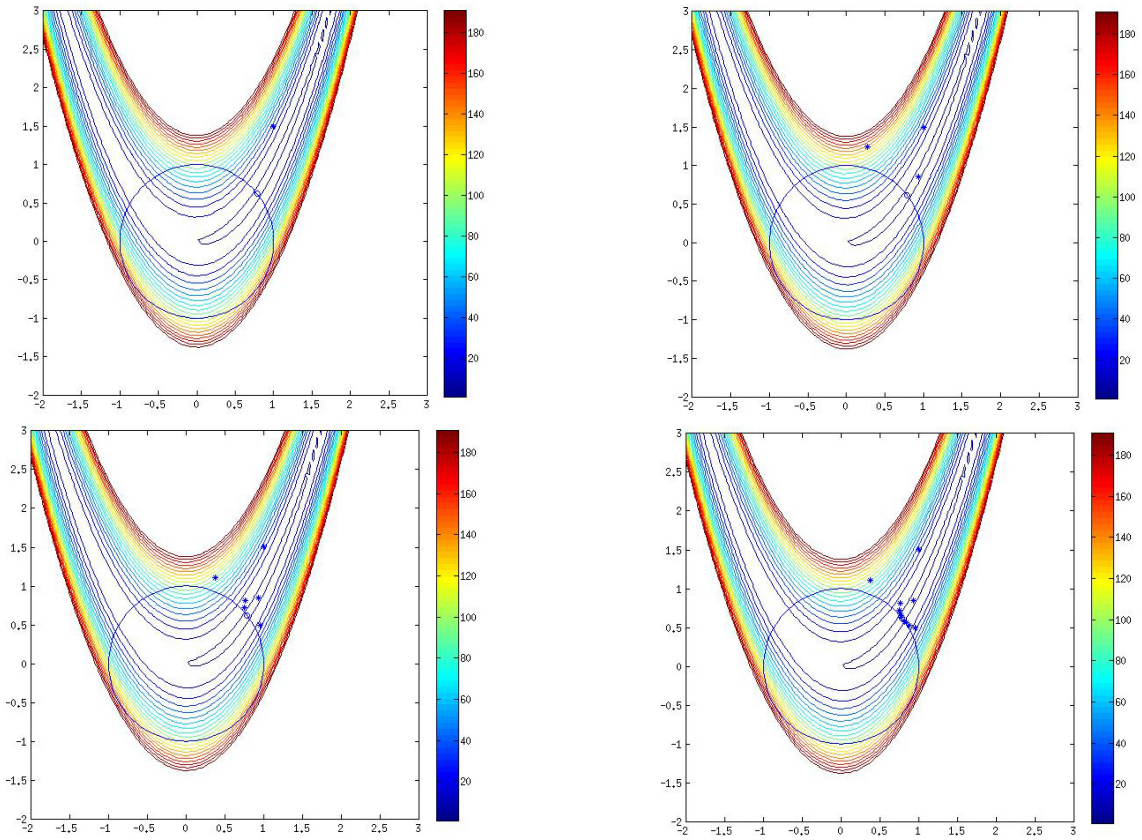


Figura 1: Curvas de nível da função objetivo, restrição e evolução das iterações

Assim, concluímos que com teorias não muito avançadas é possível demonstrar que o algoritmo converge para uma solução, ficando evidenciado com o teste numérico acima. No teste, os critérios de parada são atingidos na iteração $k = 50$, sendo que o algoritmo encontra como solução para o problema o ponto $x^{50} = [0.789 \quad 0.614]^T$, que é uma boa aproximação da solução real, visto que quando aplicado à função temos que $f(0.789, 0.614) = 0.052$.

Referências

GOMES NETO, Francisco de Assis Magalhães; SENNE, Thadeu Alves. An SLP algorithm and its application to topology optimization. CAM, São Carlos, v.30, p.53-89, 2011.

KUHN, H W; TUCKER, A W. Nonlinear Programming. Proceedings Of The Second Berkeley Symposium On Mathematical Statistics And Probability, Berkeley, p.481-492, 1951.

LIMA, Elon Langes. Curso de Análise. Rio de Janeiro, IMPA, vol. 2, 1976.

RIBEIRO, Ademir Alves; KARAS, Elizabeth Wegner. Otimização Contínua – Aspectos Teóricos e Computacionais. São Paulo: Cengage Learning, 2013. 300 p.

SENNE, Thadeu Alves. Otimização Topológica de Mecanismos Flexíveis. 2009. 141 f. Dissertação (Mestrado), Departamento de Matemática Aplicada, Universidade Estadual de Campinas, Campinas, 2009.

Trabalho de reparação e catalogação dos livros antigos existentes no Laboratório de Ensino de Matemática

Brenda Rex Machado[1]
Universidade Estadual do Oeste do Paraná
bre.rex@hotmail.com

Jaqueline do Nascimento
Universidade Estadual do Oeste do Paraná
jaque_nasci@hotmail.com

Edevaldo das Neves Marques
Universidade Estadual do Oeste do Paraná
edevalneves@hotmail.com

Resumo: Esse texto tem por objetivo apresentar alguns dos processos que têm sido realizados com os livros didáticos antigos de Matemática existentes no Laboratório de Ensino de Matemática (LEM), do curso de Matemática do campus de Cascavel. O acervo de livros antigos do LEM é constituído por obras que já existiam nesse local, como por outras recebidas de doações da biblioteca do campus de Cascavel, de professores da Universidade e de colégios da cidade. Estão descritos o processo de higienização dos livros, algumas formas utilizadas para o reparo dos exemplares que apresentam problemas nas capas ou miolo, bem como, o modo como foram classificados, de acordo com os níveis e séries de ensino, com os conteúdos abordados, com as diferentes épocas e denominações do ensino brasileiro, entre outros aspectos. Por fim são apresentados alguns resultados já alcançados depois dos processos de higienização, reparação e catalogação dos livros.

Palavras-chave: História do livro didático; levantamento e classificação de livros; restauração de livros.

1 Introdução

Ao se realizar pesquisas acerca da História da Educação, o livro didático apresenta-se como um objeto de estudo de suma importância, em consequência de fazer parte do universo da cultura escolar, revelando-se aí a importância da sua utilização para a compreensão das práticas escolares ao longo da história da educação.

Todavia, com o passar das décadas os livros didáticos sofrem alterações resultantes de agentes tais como microrganismos, insetos, roedores e até mesmo por meio da poluição. A umidade, a temperatura e a luminosidade inadequadas também causam a sua degeneração. Mas os maiores danos que podem ser ocasionados aos livros e aos documentos são aqueles de acidentes e dos maus tratos que recebem por parte das pessoas que se utilizam deles. Desta forma, os

[1]À Fundação Araucária, pelo apoio financeiro para o desenvolvimento do projeto.

agentes exteriores que por vezes os danificam, podem ser classificados em quatro categorias: os físicos, que englobam a luminosidade, a temperatura e a umidade; os químicos, dentre os quais estão a acidez do papel e a poluição atmosférica; os biológicos, insetos, fungos e roedores; e, ainda os ambientais, como má ventilação, poeira e agentes humanos (MARTINI, 2007).

Sabendo disso e compreendendo a relevância dos livros didáticos para as pesquisas ou investigações no ensino, a partir de arrecadações vindas da biblioteca do campus, de professores, estudantes e escolas da cidade montou-se o acervo e, na sequência, realizou-se a higienização e a catalogação de 100% dos exemplares. Além disso, os exemplares que se encontravam em estado de conservação débil estão passando por um processo de restauração.

2 Os livros didáticos como objeto de pesquisa

Os livros carregam em si muito mais do que apenas o conteúdo didático. Trazem aspectos que nos ajudam a estudar e compreender sobre a história do período de sua publicação, seu autor, editora, além de fazer parte e contribuir para a história do próprio local onde se encontravam. Dessa forma, alinha-se com Carvalho:

> O livro-texto tem história e o papel que desempenha e sua influência estão sempre ligados à sociedade de sua época, talvez até para tentar modificar alguns de seus aspectos, à maneira como essa sociedade, e não somente o autor do livro, vê a ciência, a cultura e o ensino (CARVALHO, 2003, p. 1-2).

Sob mesma óptica temos a interpretação de Oliveira:

> Concebemos, portanto, o livro didático de matemática como uma forma simbólica que exerce grande influência nas salas de aula de matemática, mas, como toda forma simbólica, se abre a uma pluralidade incontrolável de interpretações e possibilidades de usos, conforme nossas próprias concepções de Educação Matemática Escolar, o que nos orienta, por exemplo, na análise dessas obras, a verificar, na medida do possível, alguns desses usos (OLIVEIRA, 2008, p. 60).

Conforme ressaltado por Antonio Vicente Marafioti Garnica em uma entrevista que compõe o trabalho de Hirata (2009), esse olhar para os livros didáticos é um dos elementos que compõe uma forma de se escrever a História da Educação Matemática no Brasil. Assim, não apenas a formação do professor, as políticas educacionais, o comportamento dos alunos, a origem das famílias e a infraestrutura da escola servem como elementos de elaboração dessa história possível.

Para ser possível a realização de estudos sobre esses livros didáticos antigos e de elaborar considerações acerca da História da Educação Matemática brasileira, almeja-se a realização de um levantamento dos livros didáticos antigos de Matemática do Laboratório de Ensino de

Matemática (LEM) e da Biblioteca da Unioeste do campus de Cascavel. Para tanto, na sequência discorre-se sobre a origem dos livros didáticos existentes no acervo e sobre os processos para melhor conservação e preservação que estes necessitam.

3 A constituição do acervo e os processos de higienização e reparos

O LEM possui uma grande quantidade de livros que abrangem várias áreas do conhecimento. Dentre eles, alguns estavam separados com a alcunha de "livros antigos", por serem obras anteriores à década de 1990. Sabendo que na biblioteca do campus de Cascavel também existiam livros antigos, mas desconhecidos, os professores Dulcyene Maria Ribeiro e Jean Sebastian Toillier, resolveram desenvolver um projeto com esses livros didáticos antigos.

Para a ampliação do acervo que existia no LEM, foi elaborada e executada uma proposta de campanha para doação de livros didáticos de Matemática antigos. A campanha se propagou pelo contato pessoal e via internet, por meio de redes sociais e o site da própria universidade. Como resultado, foi recebido um total de 367 livros doados, vindos do Colégio Estadual Wilson Joffre, do Colégio Estadual Padre Pedro Canísio Henz, da Biblioteca Central da Unioeste, campus Cascavel, e de doações de professores do curso de matemática da Unioeste. Além destes exemplares, o acervo contempla os 76 livros antigos pré-existentes no Laboratório de Ensino de Matemática.

Os livros recebidos precisavam passar por um processo de higienização, para isso, fomos instruídos pelo pessoal do departamento de Restauração na biblioteca da Unioeste a manusear os livros corretamente e os materiais necessários para tal.

Foi necessário lixar todos os livros, além de fazer a higiene da capa com sabonete neutro e esponja, colocando capas de plástico embaixo da capa do exemplar que estava sendo higienizado, para que elas evitassem que as folhas do livro molhassem. Em seguida as capas eram secadas com pano. Para finalizar a higiene, algumas folhas do exemplar eram limpas com pano levemente umedecido.

A limpeza dos livros doados foi concluída, bem como, a limpeza dos livros antigos, estes que já pertenciam ao acervo do LEM. Os que se encontravam em estado mais delicado foram separados para reparação e os livros que encontravam mais de um exemplar foram destinados à doação para o uso dos alunos do curso de Matemática.

Após a etapa de higienização, iniciamos o processo de reparação[2] dos livros que apresentavam algum dano, que, possivelmente, foram causados por estocagem ou manuseio inadequado. Além disso, a ação do tempo, também, fragiliza a estrutura do livro e pode levar à necessidade de recuperação.

Para melhor compreensão dos processos de reparação, deve-se antes conhecer a estrutura de um livro, como mostra a figura a seguir:

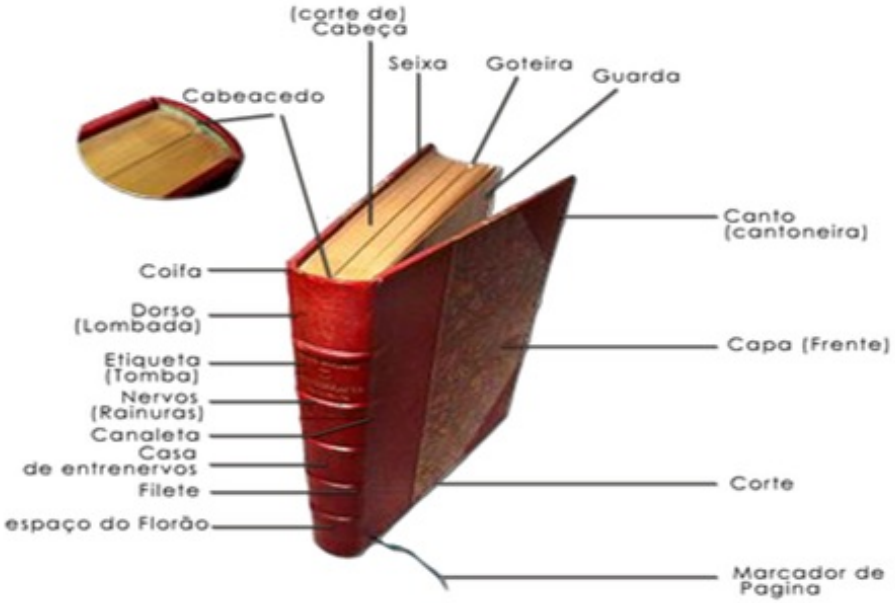


Figura 1: Livro de capa dura[3]

A imagem mostra um livro de capa dura, contudo, os livros com capa flexível (brochura) apresentam estrutura semelhante.

O objetivo de reparar um livro é prolongar a vida útil da obra. Nesse contexto, busca-se conservar as condições de uso do exemplar, assim pode-se dizer que o processo de reparação é um

> conjunto de medidas destinadas à correção de danos causados às obras. Na Biblioteca Nacional, a conservação é entendida como um conjunto de procedimentos que tem por objetivo melhorar o estado físico do suporte, aumentar sua permanência e prolongar-lhe a vida útil, possibilitando, desta forma, o seu acesso por parte das futuras gerações. (ANTUNES, 2010, p. 16)

Ainda sobre o significado de reparar um livro, segundo Milevsky (2001, p. 45) reparo é uma maneira de remediar dano feito a um item, normalmente acrescentando novo material para

[2]Essa etapa do processo foi conduzida por um dos autores desse texto, Edevaldo das Neves Marques, a partir dos seus conhecimentos e experiências anteriores com reparação e restauração de livros.

[3]OLIVEIRA, Gilberto. Encadernação e restauração de livros. 2014. Disponível em: <http://tingcultura.blogspot.com.br/2014_01_24_archive.html> Acessado em: 13 set. 2016.

substituir o material estragado ou deteriorado.

Assim, conclui-se que o processo de reparação visa manter o máximo possível da originalidade do livro, permitindo a substituição de algumas de suas partes, como a lombada e as folhas de guardas, por novos materiais, preservando a obra e possibilitando seu manuseio. Esse processo é recomendado para livros que serão manuseados com alguma frequência.

Já o processo de restauração é muito mais complexo, pois necessita de um especialista em restauração de papel e exige tempo superior ao procedimento anterior e, além disso, é mais custoso. Esse procedimento é recomendado para obras raras com valor histórico insubstituível. As obras que passam por processos de restauração devem ser armazenadas em locais específicos e o manuseio deve ser controlado. Nesse sentido, Antunes (2010, p. 16) afirma que restauração é o conjunto de intervenções, em obras que sofreram danos, para recuperar seu estado original, o máximo possível.

O trabalho envolvendo reparação de livros exige inicialmente a seleção dos exemplares a serem consertados, sendo possível a partir de então organizar as obras de acordo com os processos que serão realizados posteriormente, visto que, há livros apenas com as capas danificadas e outros exemplares com a capa e o corpo com algum dano.

Após a classificação dos livros a serem reparados, por ordem de complexidade do trabalho a ser efetuado, foi iniciado o processo de reparação. Para esse processo foram utilizadas as seguintes ferramentas: estilete, tesoura, bisturi, "peso", placas de madeira, régua, pincéis, agulha, recipientes de vidro e plástico para preparo e conservação da cola, clips de metal, entre outros. Também foram utilizados os materiais a seguir: papel toalha, linha de algodão, cola poliacetato de vinila (PVA), cola carbox metil celulose (CMC), água filtrada, papel cartão, papel percalux, papelão NR28, papel *colorplus* 80 gramas e 120 gramas, papel *colorplus* TX 180 gramas, papel japonês 10 gramas, papel kraft, viés, tecido mourim, entre outros.

No processo de reparação, primeiramente foram desmontados alguns livros, utilizando estilete e bisturi, para separar a capa e contracapa do corpo do livro. Assim foi possível verificar se o lombo/dorso do livro estava em boas condições para ser reutilizado, visto que em alguns livros devido ao estado de conservação não foi possível reutilizar. Contudo as capas e contracapas foram reutilizadas com o intuito de manter o máximo possível a originalidade do livro.

Após a desmontagem de alguns exemplares, foi necessário a preparação da cola para a próxima etapa da reparação. Para isso utilizamos uma mistura de CMC e PVA, sendo 70% e 30% respectivamente, podendo variar a porcentagem devido a especificidade de cada caso.

Na sequência, medimos e recortamos de acordo com as medidas específicas de cada livro,

o papel *color plus*, deixando alguns centímetros de sobra em cada lado do mesmo, estes serviram como novas folhas de guarda e a falsa folha de rosto. Estas foram coladas na frente e no verso do livro e então foi posto um peso em cima deste para a melhor fixação, utilizando papel toalha para absorção da umidade. O peso é feito de *paver*, revestido de papel *kraft* para melhor manuseio do mesmo.

Para reforçar a lombada e as folhas de guarda, utilizou-se papel kraft, o qual foi fixado com cola e, para otimização do trabalho foi destinado novamente um período para a secagem da cola, utilizando novamente o peso e o papel toalha com a mesma finalidade anterior. Depois da secagem, realizamos o corte das sobras que foram deixadas nas folhas de guarda e na falsa folha de rosto.

Posteriormente, utilizamos o papel cartão, este já adequado para servir como suporte da capa e da contracapa do livro. Fixamos neste um novo lombo feito com o papel *color plus* TX 180 gramas, para então colocarmos a capa retirada anteriormente. A capa original antes de ser colada novamente, foi adaptada, isto é, foram feitos reparos necessários, como cortar alguns centímetros desta, para então ocupar o novo espaço reestruturado, sendo fixada com cola. Para conclusão do trabalho, o livro foi posto novamente para secar, com o peso em cima e com o papel toalha para absorção da umidade.

De modo geral, os passos descritos anteriormente são utilizados para a todos os livros que foram, estão sendo ou serão reparados, sendo que durante a separação dos exemplares, a serem consertados, busca-se não desmontar livros com poucos dados. Para essas obras menos danificadas, muitas vezes, basta apenas o procedimento de colagem de algumas de suas partes. Lembrado que um livro só deve passar pelo processo de reparação, ou em caso extremo restauração, se esse for o último caso, ou seja, não tiver outra solução.

4 A catalogação

O trabalho de Hirata (2009) Catalogação de Livros Antigos: Um Exercício em Educação Matemática foi nosso alicerce para adotarmos os princípios de catalogação. Segundo Hirata, o acervo do Grupo de História Oral e Educação Matemática (GHOEM[4]) se baseou no método Classificação Decimal de Dewel (CDD), o qual associa os livros às categorias de acordo com os

[4]O Grupo "História Oral e Educação Matemática" – GHOEM foi criado no ano de 2002. Sua intenção inicial foi reunir pesquisadores em Educação Matemática interessados na possibilidade de usar a História Oral como recurso metodológico. Desde então, essa configuração foi alterada, ampliando-se, de modo a incorporar discussões sobre outros temas e outras abordagens teórico-metodológicas. Pode-se dizer, hoje, que o interesse central do grupo é o estudo da cultura escolar e o papel da Educação Matemática nessa cultura.

conteúdos neles abordados. Mas eles não adotaram esse método em sua forma original, o qual foi adequado de acordo com as necessidades do material que o grupo tinha na época. Logo, partimos do modelo desenvolvido pelo GHOEM e acrescentamos categorias que consideramos pertinentes, por causa de quantidade de livros que tínhamos dessas áreas que não apareciam na classificação do GHOEM. As categorias acrescentadas são: trigonometria, matemática financeira, estatística e a subcategoria "outros" dentro dos "diversos".

Portanto, os livros foram separados por área, sendo elas: 1 – conteúdo matemático (contempla mais de um conteúdo, por exemplo: álgebra e geometria, geometria e aritmética, etc.); 2 – Teoria dos Conjuntos e Lógica; 3 – Álgebra; 4- Aritmética; 5 – Topologia; 6 – Análise; 7 – Geometria; 8 – Probabilidade; 9 – Trigonometria; 10 – Matemática Financeira; 11 – Estatística; 12 – Diversos (que contempla Anais, Atas e Relatórios, Curiosidades, Paradidáticos e de Apoio Pedagógico, História da Matemática, Dicionários, Exemplares "para o ensino", Preparatório e Outros); 13. – Didáticos de Outras Disciplinas e 14. – Literatura de Referência.

Levando em consideração esses dados, foi elaborado um formulário online para possibilitar a catalogação dos livros. Tal formulário especifica o assunto abordado pelo livro (categoria da obra de acordo com a classificação acima), o nível de destinação da obra (que contempla o ginasial; colegial; técnico; normal; Primeiro Grau; Segundo Grau; Ensino Fundamental; Ensino Médio e Ensino Superior); Numeração da obra (de acordo com o método de catalogação adotado); Numeração do autor (de acordo com a tabela P.H.A[5]); Tombo (Sequenciação do acervo); Autor(es) da obra; Título da obra; Edição do exemplar; Local da edição; Editora da obra; Local de impressão; Impressor da obra; Tradutor da obra; Número de páginas; Volume do exemplar; Observações (exemplos: dedicatórias, anotações, local da doação); *Ex-libris*[6]; Coleção a qual a obra pertence; Idioma da obra; Gênero (exemplos: livro do professor, manual); e Preço do exemplar.

5 Algumas considerações até o momento

A catalogação foi concluída e obtivemos 190 exemplares (42,9%) pertencentes à categoria de Conteúdo Matemático, 57 exemplares (12,9%) pertencentes aos Diversos, 53 exemplares (12%) pertencentes à Análise, 39 exemplares (8,8%) pertencentes à Álgebra, 29 exemplares

[5]Tabela PHA (Prado, Heloísa de Almeida) é uma tabela para catalogação de livros em bibliotecas, que permite que sejam consideradas mais de três letras no sobrenome do autor e evita a formação de conjuntos estranhos de letras.

[6]Consideramos como *Ex-Libris* todas as peças encontradas dentro dos livros, como cartas, papeis, fotos, convites, entre outros.

(6,5%) pertencentes à Geometria, 21 exemplares (4,7%) pertencentes à Matemática Financeira, 15 exemplares (3,4%) à Estatística, 11 exemplares (2,5%) pertencentes à Didáticos de Outras Disciplinas, 9 exemplares (2%) pertencentes à Aritmética, 7 exemplares (1,6%) pertencentes à Trigonometria, 6 exemplares (1,4%) pertencentes à Literatura de Referência, 3 exemplares (0,7%) pertencentes à Topologia, 2 exemplares (0,5%) pertencentes à Teoria dos Conjuntos e Lógica e 1 exemplar (0,2%) pertencente à Probabilidade. Então geramos uma etiqueta para cada exemplar, contendo a categoria da obra, número do autor, edição, volume e número do exemplar, tal como é feito para a maioria das bibliotecas.

Para as categorias "Conteúdo Matemático", "Álgebra", "Aritmética" e "Geometria" tivemos uma subcategoria chamada "Nível de destinação da obra", com a intenção de facilitar a organização do acervo. Dessa categoria 5 exemplares (1,9%) são pertencentes ao Ensino Primário, 54 exemplares (20,2%) ao Ensino Secundário, 79 exemplares (29,6%) ao Primeiro Grau, 56 exemplares (21%) ao Segundo Grau, 19 exemplares (7,1%) ao Ensino Fundamental, 5 exemplares (1,9%) ao Ensino Médio e 49 exemplares (18,4%) ao Ensino Superior. Dos exemplares que entraram em ensino secundário obtivemos o "Subnível de destinação da obra" no qual 6 exemplares (11,1%) pertencem ao Secundário (contemplam conteúdos matemáticos tanto para os cursos ginasiais quanto para os cursos colegiais), 33 exemplares (61,1%) ao Ginasial, 12 exemplares (22,2%) ao Colegial, nenhum exemplar (0%) pertence ao Técnico e 3 exemplares (5,6%) pertencem ao Normal.

Durante a campanha de doação de livros foram recebidos 110 livros da Biblioteca Central da Unioeste, campus Cascavel; 88 do Colégio Estadual Padre Pedro Canísio Henz; 135 do Colégio Estadual Wilson Joffre; e 34 livros doados por professores do curso de matemática. 76 livros já existiam coma alcunha de livros antigos no acervo do Laboratório de Ensino de Matemática Prof^a Silvia Gomes Vieira Fabro - Unioeste, Campus Cascavel.

Do total de 443, 72 livros datam sua publicação até o ano de 1970, 226 livros datam sua publicação de 1970 a 1990, 66 livros datam sua publicação a partir do ano de 1990 e 79 livros não apresentam data de publicação da obra. No quadro abaixo, esses dados estão divididos entre os doadores.

Quadro 1: Análise referente ao ano de publicação das obras

Doador	Quantidade de livros que datam até 1970	Quantidade de livros que datam entre 1970 e 1990	Quantidade de livros que datam a partir de 1990	Quantidade de livros que não apresentam data	Quantidade total de livros
Biblioteca Central da Unioeste, *campus* Cascavel	13	85	06	06	110
Colégio Estadual Padre Pedro Canísio Henz	03	21	38	26	88
Colégio Estadual Wilson Joffre	26	67	16	26	135
Laboratório de Ensino de Matemática, Unioeste, *campus* Cascavel	21	40	04	11	76
Professores	09	13	02	10	34

Diante disso, podemos observar que a maior parte dos livros do acervo foi publicada entre os anos de 1970 e 1990.

Incluídos na categoria "Diversos", 4 exemplares (7%) pertencem à sua subcategoria Anais, Atas e Relatórios, 20 exemplares (35,1%) à Curiosidades, Paradidáticos e de Apoio Pedagógico, 2 exemplares (3,5%) à História da Matemática, 2 exemplares (3,5%) à Dicionários, 15 exemplares (26,3%) à Exemplares "para o ensino", 9 exemplares (15,8%) à Preparatórios (Admissão) e 5 exemplares (8,8%) à Outros.

Muitos exemplares não possuíam respostas para todas as perguntas do catálogo. Dos livros que têm data de publicação o mais antigo é *DíeTurnftundeín der Rnabenfchule*, de Dr. Edmund Neuendorff, publicado em 1927 e o mais atual é *Novo Olhar da Matemática*, de Joamir Roberto de Souza, publicado 2013.

6 Considerações finais

Em relação aos procedimentos de reparação dos livros, pode-se dizer que está sendo satisfatório o rendimento do trabalho realizado, visto que, inicialmente foi necessário adquirir os materiais e as ferramentas necessárias, sendo que nem todos os itens utilizados nos procedimentos foram encontrados na região, exigindo sua compra em outras localidades.

Ainda há muitas obras que deverão passar pelo processo de reparação, uma vez que o trabalho realizado até o momento visou apenas os livros de capa flexíveis, portanto, há ainda todos os livros de capa dura e os restantes de capas flexíveis a serem reparados. Contudo, estamos com o estoque de materiais e ferramentas supridos. Portanto, acreditamos que, a partir da prática já realizada, é possível melhorar qualidade e a produtividade do trabalho de reparação de livros.

Dessa forma, esperamos compreender aspectos acerca da História da Educação Matemática e entendemos que os livros didáticos que compõe o nosso acervo podem ser fundamentais para esse processo.

Referências

ANTUNES, Margaret Alves. **Pequenos Reparos em Material Bibliográfico** Notas de biblioteca 2. Secretaria de Estado da Cultura de São Paulo, São Paulo, 2010.

CARVALHO, João Bosco Pitombeira de. Apresentação. In: SCHUBRING, G. **Análise histórica de livros de matemática:** notas de aula. Trad. Maria Laura Magalhães Gomes. Campinas: Autores Associados, 2003.

HIRATA, Vinicius. **Catalogação de livros antigos:** um exercício em educação matemática. 2009. 70 f. Monografia (Iniciação Científica). Universidade Estadual Julio de Mesquita Filho, campus de Bauru. 2009.

MARTINI, Augusto. **Fatores Ambientais e Agentes Causadores de Deteriorização de Documentos**. 2007. Disponível em: <https://asimplicidadedascoisas.wordpress.com/2007/06/14/fatores-ambientais-e-agentes-causadores-de-deteriorizacao-de-documentos/> Acessado em: 03 set. 2016.

MILEVSKY, Robert J. **Manual de pequenos reparos em livros**. 2.ed. Rio de Janeiro: Projeto Conservação Preventiva em Bibliotecas: Arquivo Nacional, 2001. 49p.(Conservação Preventiva em Bibliotecas e Arquivos, 13. Conservação). Disponível em: <http://www.portalan.arquivonacional.gov.br/media/CPBA%2013%20Manual%20Peq%20Reparos%20Livros.pdf> Acesso em: 02 set. 2016.

OLIVEIRA, Fábio Donizeti. **Análise de textos didáticos:** três estudos. Dissertação (Mestrado em Educação Matemática). Instituto de Geociências e Ciências Exatas (IGCE). UNESP, Rio Claro, 2008.

O perfil da violência no Paraná

André Guilherme Unfried[1]
Universidade Estadual do Oeste do Paraná
andre_unfrich@hotmail.com

Rosangela Villwock
Universidade Estadual do Oeste do Paraná
rosangela.villwock@unioeste.br

Resumo: O aumento da violência em todo o mundo está se tornando um problema que acarreta vários danos à sociedade tais como danos materiais, sociais, psicológicos e físicos. Devido à importância do estado do Paraná no cenário nacional, torna-se relevante conhecer o perfil da violência nos municípios do estado do Paraná. Este trabalho tem como objetivo delinear o perfil da violência nos municípios do Estado do Paraná, por meio do Processo de Descoberta de Conhecimento em Bases de Dados. O Agrupamento de Dados foi a tarefa de Mineração de Dados utilizada neste trabalho. Os dados para o desenvolvimento deste trabalho foram coletados no Instituto de Desenvolvimento Econômico e Social – IPARDES, referentes ao ano de 2013. Para aplicação da análise de agrupamentos foi escolhido o método de k-Médias, utilizando-se o o software WEKA. Utilizando o algoritmo k-Médias para a base de dados criada, considerando como critério o SQE, obteve-se três grupos. O primeiro grupo foi caracterizado como o mais violento, o segundo grupo foi caracterizado como o menos violento e o terceiro grupo foi caracterizado por ser o grupo com violência moderada. Os resultados mostram que a maioria dos municípios paranaenses está num grupo com os menores indicadores de violência.

Palavras-chave: Mineração de Dados; Agrupamento de Dados; k-Médias.

1 Introdução

Todos os dias os meios de comunicação divulgam casos de violência remetendo-nos a pensar que esse tipo de atitude seja algo banal, o que não é o caso. O aumento da violência em todo o mundo está se tornando um problema que acarreta vários danos à sociedade tais como danos materiais, sociais, psicológicos e físicos. Devido à importância do estado do Paraná no cenário nacional, torna-se relevante conhecer o perfil da violência nos municípios neste estado.

A evolução dos homicídios no Paraná possui três períodos distintos. No primeiro, entre 1980 e 1992, o índice no estado cresceu menos que a média nacional. Enquanto os homicídios no país aumentaram 63,3% no período, no Paraná a elevação foi de 18,7%. De 1992 a 2000, houve um primeiro "boom"dos assassinatos no estado: as taxas passaram a aumentar acima da média

[1]Os autores agradecem ao CNPq por bolsa de iniciação científica concedida ao primeiro autor e à Fundação Araucária pelo apoio financeiro ao projeto de pesquisa "Identificação do Perfil dos Municípios do Estado do Paraná por meio do Processo de Descoberta de Conhecimento em Bases de Dados".

nacional, impulsionadas principalmente pela região metropolitana de Curitiba (RMC). Neste período, a média de aumento dos casos em todo o país foi de 39,9%; no Paraná, de 44,6% e só na RMC, de 77,7%. No terceiro período, de 2000 até os dias atuais, os índices se acentuaram ainda mais no Paraná, enquanto houve a estagnação das taxas em todo o país. É como se o Paraná tivesse corrido na contramão no que diz respeito à segurança pública. Esse último aumento mais incisivo é puxado tanto pela RMC, quanto pelo interior do estado (Geron, 2011).

O Estado do Paraná fechou o ano de 2014 com uma taxa de homicídios de 22,6 para cada 100 mil habitantes (AGÊNCIA DE NOTÍCIAS DO PARANÁ, 2015), enquanto a taxa de homicídios do Brasil no ano de 2014 foi de 25,8 para cada 100 mil habitantes (STOCHERO, 2015). O índice é utilizado mundialmente para medir eficiência e eficácia nas políticas públicas adotadas pelos governos. Visando uma melhor alocação dos recursos investidos em segurança para precaver o acontecimento de novos crimes, é de suma importância conhecer o perfil de cada município do estado. Para a análise minuciosa dos dados serão utilizadas técnicas e ferramentas computacionais para auxiliar na tomada de decisões. Os resultados que serão obtidos podem ser utilizados como subsídio para elaboração de projetos adaptados aos perfis encontrados, bem como identificação de mudanças necessárias, servindo de instrumento para tomadas de decisão mais acertadas.

Assim, este trabalho tem como objetivo delinear o perfil da violência nos municípios do estado do Paraná, por meio do Processo de Descoberta de Conhecimento em Bases de Dados ou "*Knowledge Discovery in Databases – KDD*".

Dentre as muitas definições existentes para *KDD*, pode-se citar a de Fayyad et al. (1996), que define *KDD* como "o processo, não trivial, de extração de informações implícitas, previamente desconhecidas e potencialmente úteis, a partir dos dados armazenados em um banco de dados". Segundo estes autores, a principal vantagem do processo de descoberta é que não são necessárias hipóteses, sendo que o conhecimento é extraído dos dados sem conhecimento prévio (FAYYAD, 1996).

As tarefas de Mineração de Dados podem ser preditivas ou descritivas. As preditivas usam variáveis para predizer valores futuros ou desconhecidos, enquanto que as descritivas encontram padrões para descrever os dados. As principais tarefas de Mineração de Dados estão relacionadas à Classificação, Associação e Agrupamento de padrões (FAYYAD, 1996). O Agrupamento de Dados foi a tarefa utilizada neste trabalho.

O Agrupamento de dados (*Clustering*) procura grupos de padrões tal que padrões pertencentes a um mesmo grupo são mais parecidos uns com os outros e divergentes a padrões em

outros grupos. A análise de agrupamentos é uma técnica analítica para desenvolver subgrupos significativos de objetos. Seu objetivo é classificar os objetos em um pequeno número de grupos mutuamente excludentes (HAIR JR, 2005).

Os algoritmos de agrupamento podem ser divididos em categorias de diversas formas de acordo com suas características. As duas principais classes de algoritmos de agrupamento são: os métodos hierárquicos e os métodos de particionamento (VILLWOCK, 2009). Os métodos hierárquicos abrangem técnicas que buscam de forma hierárquica os grupos e, por isso, admitem obter vários níveis de agrupamento. Métodos não-hierárquicos ou de particionamento procuram uma partição sem a necessidade de associações hierárquicas. Seleciona-se uma partição dos elementos em k grupos, otimizando algum critério (DINIZ, 2000). O método mais conhecido entre os métodos de particionamento é o das k-médias. Geralmente os k grupos encontrados são de melhor qualidade do que os k grupos gerados pelos métodos hierárquicos (JOHNSON, 2007).

2 Metodologia

Os dados para o desenvolvimento deste trabalho foram coletados no Instituto de Desenvolvimento Econômico e Social – IPARDES (IPARDES). Foram consideradas 15 variáveis para cada um dos 399 municípios paranaenses, referentes ao ano de 2013. São elas: Produto Interno Bruto per Capita (R$ 1,00); Vítimas em Acidentes de Trânsito Mortos no Local (100 mil habitantes); Vítimas em Acidentes de Trânsito com Morte Posterior (100 mil habitantes); Vítimas de Homicídio Doloso (100 mil habitantes); Vítimas de Roubo com Resultado de Morte – Latrocínio (100 mil habitantes); Vítimas de Lesão Corporal com Resultado de Morte (100 mil habitantes); Vítimas de Homicídio Culposo no Trânsito (100 mil habitantes); Taxa de Mortalidade Causas Externas – Acidentes de Trânsito (100 mil habitantes); Taxa de Mortalidade Causas Externas – Agressões (100 mil habitantes); Despesas Municipais por Função – Segurança Pública (R$ 1,00 por mil habitantes); Taxa de Reprovação no Ensino Fundamental (%); Taxa de Reprovação no Ensino Médio (%); Taxa de Abandono no Ensino Fundamental (%) e Taxa de Abandono no Ensino Médio (%).

Para aplicação da análise de agrupamentos foi escolhido o método de k-Médias. O k-Médias possui um parâmetro de entrada k, que corresponde à quantidade de grupos. Inicialmente o método atribui como centroides iniciais, por exemplo, k pontos do conjunto de dados. Em seguida, para cada ponto, calcula-se a distância do ponto para cada um dos centroides. Atribui-se cada ponto ao grupo representado pelo centroide cuja distância for menor. Desta forma, cada ponto do conjunto de dados fica associado a um e apenas um dos k grupos. Após a alocação

inicial, o método segue iterativamente, por meio da atualização dos centroides de cada grupo e da realocação dos pontos aos grupos cujo centroide seja mais próximo. O novo centroide de cada grupo é calculado a cada iteração, sendo o vetor de médias dos pontos alocados ao grupo. O processo iterativo termina quando os centroides dos grupos param de se modificar ou após um número preestabelecido de iterações ter sido realizado (GOLDSCHIMIDT, 2015).

Para aplicação do k-Médias foi utilizado o Weka (disponível em http://www.cs.waikato.ac.nz/ml/weka/). Foram utilizados os parâmetros padrão da ferramenta. A definição do número de grupos foi feita pela avaliação do SQE – Soma do Quadrado do Erro.

3 Resultados e Discussão

Utilizando o algoritmo k-Médias para a base de dados criada, considerando como critério o SQE, obteve-se três grupos. O primeiro grupo é caracterizado como o mais violento dado que apresenta maior número médio de vítimas de acidente de trânsito com mortes no local, de vítimas de homicídio doloso, de vítimas de roubo com resultado de morte (latrocínio), de vítimas de lesão corporal com resultado de morte e maior taxa média de mortalidade em homicídios. Além disso, este grupo tem valores intermediários para número médio de vítimas em acidentes de trânsito com morte posterior, número médio de vítimas de homicídio culposo no trânsito e taxa média de mortalidade em acidentes de trânsito.

Este grupo tem o maior PIB per capita médio, a maior taxa média de reprovação no ensino fundamental e médio e a maior taxa média de abandono no ensino fundamental e médio. Além disso, as despesas municipais com segurança pública, em média, não foram as maiores no estado. Neste grupo há 99 municípios.

Já o segundo grupo é caracterizado como o menos violento dado que apresenta menor número médio de vítimas de acidente de trânsito com mortes no local, de vítimas em acidentes de trânsito com morte posterior, de vítimas de homicídio doloso, de vítimas de lesão corporal com resultado de morte, de vítimas de homicídio culposo no trânsito, de mortalidade em acidentes de trânsito e menor taxa média de mortalidade em homicídios. Além disso, este grupo é o intermediário em número médio de vítimas de roubo com resultado de morte (latrocínio).

Este grupo também tem o menor PIB per capita médio, é o intermediário em taxa média de reprovação no ensino fundamental e médio, em taxa média de abandono no ensino médio e também possui a menor taxa média de abandono no ensino fundamental. Além disso, as despesas municipais para segurança pública, em média, são as menores. Neste grupo há 233

municípios.

O terceiro grupo é caracterizado por ser o grupo com violência moderada dentre os três grupos dado que apresenta valores médios intermediários de vítimas de acidente de trânsito com mortes no local, de vítimas de homicídio doloso, de vítimas de lesão corporal com resultado de morte e de taxa de mortalidade em homicídios. Além disso, possui o maior número médio de vítimas em acidentes de trânsito com morte posterior, de vítimas de homicídio culposo no trânsito e de taxa de mortalidade em acidentes de trânsito. Também é o menor em número médio de vítimas de roubo com resultado de morte (latrocínio).

Este grupo também é o intermediário no PIB per capita médio e em taxa média de abandono no ensino fundamental. Além disso, possui a menor taxa média de reprovação no ensino fundamental e médio e a menor taxa média de abandono no ensino médio. Este grupo tem a maior despesa média municipal em segurança pública. Neste grupo há 67 municípios.

4 Considerações Finais

O objetivo do presente trabalho foi delinear o perfil da violência nos municípios do Estado do Paraná, por meio do Processo de Descoberta de Conhecimento em Bases de Dados.

Os resultados mostram que a maioria dos municípios paranaenses está num grupo com os menores indicadores de violência. Alguns dos municípios deste grupo são: Iguatu, Nova Prata do Iguaçu, Toledo, Marechal Cândido Rondon, Londrina, Ponta Grossa e Pato Branco.

Sugere-se como trabalhos futuros a investigação de outras características dos municípios que fazem parte de um mesmo grupo, como por exemplo, São municípios de fronteira?, Qual o porte dos municípios?, entre outras.

Referências

GERON, V.; ANÍBAL, F. Em 11 anos, taxa de homicídios no Paraná aumenta 86%. 2011. Disponível em: http://www.gazetadopovo.com.br/especiais/paz-tem-voz/em-11-anos-taxa-de-homicidios-no-parana-aumenta-86-aoz7f741e8bfzevqxgngp96q6. Acesso em: 18/01/2016.

AGÊNCIA DE NOTÍCIAS DO PARANÁ. Paraná reduz homicídios e atinge meta estabelecida para o período. 2015. Disponível em: http://www.aen.pr.gov.br/modules/noticias/article.php?storyid=83301. Acesso em: 18/01/2016.

STOCHERO, T. Taxa de homicídios e latrocínios no Brasil em 2014. 2015. Disponível em: http://especiais.g1.globo.com/politica/2015/taxa-de-homicidios-e-latrocinios-no-brasil-em-2014/. Acesso em: 17/02/2016.

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHRUSAMY, R. Advances in knowledge Discovery & Data Mining. California: AAAI/MIT, 1996.

HAIR JR, J.F.; ANDERSON, R.E.; TATHAM, R.L.; BLACK, W.C. Análise Multivariada de Dados. Tradução de: SANTANNA, A. S.; CHAVES NETO. 5 ed. A. Porto Alegre: Bookman, 2005.

VILLWOCK, R. Técnicas de agrupamento e de hierarquização no contexto de KDD – Aplicação a dados temporais de instrumentação geotécnica-estrutural da Usina Hidrelétrica de Itaipu. Tese (Doutorado em Métodos Numéricos em Engenharia). Universidade Federal do Paraná, Curitiba, 2009.

DINIZ, C. A. R.; LOUZADA NETO, F. (2000). Data mining: uma introdução. São Paulo: ABE.

JOHNSON, R.A.; WICHERN, D.W. Applied Multivariate Statistical Analysis. New Jersey: Prentice Hall, 2007.

IPARDES. Instituto Paranaense de Desenvolvimento Econômico e Social. Disponível em: http://www.ipardes.gov.br/. Acesso em: 17/02/2016.

GOLDSCHIMIDT, R; PASSOS, E; BEZERRA, E. Data Mining: Conceitos, técnicas, algoritmos, orientações e aplicações. 2 ed. Rio de Janeiro: Elsevier Editora Ltda, 2015.

Investigação de características relevantes para avaliação do IDHM

Diessica Aline Quinot
Universidade Estadual do Paraná
diessicaquinot@gmail.com

Rosangela Villwock
Universidade Estadual do Paraná
rosangela.villwock@unioeste.br

Resumo: O critério para analisar a qualidade de vida de um determinado local é o Índice de Desenvolvimento Humano (IDH), que consiste na média obtida entre três aspectos: Renda Nacional Bruta (RNB) per capita, grau de escolaridade e saúde. O Índice de Desenvolvimento Humano Municipal (IDHM) também considera essas dimensões, porém alguns indicadores são diferentes. No entanto, podem haver outros fatores relevantes na melhoria do bem-estar econômico e social, que não são levados em conta no cálculo do IDHM. Com base em dados atuais e características que podem proporcionar melhoria da qualidade de vida, buscou-se a identificação de fatores relevantes, que poderiam ser considerados no cálculo do IDHM, bem como identificar áreas de investimentos para aumentar esse índice. Para análise de dados dos 399 municípios do Paraná utilizou-se o processo de descoberta do conhecimento em base de dados (KDD), sendo que a tarefa de mineração de dados foi a classificação. O trabalho foi realizado no software Weka (Waikato Environment for Knowledge Analysis), aplicando-se o algoritmo J48, criando-se árvore de decisão. Os resultados mostraram que fatores como renda média domiciliar (utilizado no cálculo) e proporção da população com dez anos ou mais podem estar relacionados a um bom índice.

Palavras-chave: Mineração de Dados; Classificação; Árvore de decisão.

1 Introdução

Havendo uma busca constante por medidas sociais abrangentes que incluam fatores fundamentais na qualidade de vida, não necessariamente pela dimensão econômica, criou-se no início da década de 1990 o Índice de Desenvolvimento Humano − IDH (Torres, Ferreira e Dini, 2003).

O IDH propõe verificar o grau de desenvolvimento de um país utilizando como indicadores de desempenho: a longevidade (condições de saúde e esperança de vida ao nascer, dentre outros); a Renda Nacional Bruta (RNB) per capita (expressa em poder de paridade de compra (PPC) que exclui diferenças entre a valorização de diferentes moedas dos países); e o grau de escolaridade (taxa de alfabetização de adultos e a expectativa de anos de escolaridade medida pela taxa de matrícula nos níveis de ensino fundamental, médio e superior) (PNUD, 2010).

O IDHM também considera educação, longevidade e renda, porém alguns indicadores são diferentes, como mostra o quadro 1. Embora meçam os mesmos fenômenos, os indicadores levados em conta são mais adequados para avaliar as condições de núcleos sociais menores (PNUD, 2016).

Subíndices	**Indicador**	**IDH**	**IDHM**
Renda	Renda Nacional Bruta (per capita)	x	
	Renda Média Domiciliar (per capita)		x
Longevidade	Esperança de Vida ao Nascer	x	x
Educação	Taxa de Alfabetização	x	x
	Taxa de Matrícula	x	
	Taxa de Frequência à Escola		x

Quadro 1: Indicadores de Desempenho (IDH e IDHM).

Mas serão apenas essas variáveis importantes para se considerar no cálculo desse índice? O IDHM é um bom indicador para avaliar o desenvolvimento de diferentes e diversas comunidades? É claro que o IDH e o IDHM são índices importantes para se identificar algumas situações semelhantes em diferentes lugares ou, por outro lado, opulência e ótimas condições de vida em alguns locais e péssima qualidade de vida e situações humanas inaceitáveis em outros. Porém, para considerar um índice mais abrangente, indicando níveis precisos de desenvolvimento, é necessário, no mínimo, que este seja questionado.

Neste trabalho, teve-se o objetivo de investigar outras características relevantes para melhoria do IDHM. Com análise de dados dos municípios do Paraná, referente ao ano de 2010, no Instituto de Desenvolvimento Econômico e Social − IPARDES. Foram consideradas 22 variáveis para cada município, obtendo-se um grande volume de dados. Para isso foi necessário o uso de tecnologias para o processamento dos dados e extrair informações.

Para atender as necessidades, vem se destacando o KDD - Knowledge Discovery in Databases (Descoberta do Conhecimento em Bases de Dados) e o Data Mining (Mineração de Dados) sendo uma tarefa do KDD (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

Segundo Fayyad et al. (1996) o KDD "é um processo não trivial de descoberta de padrões válidos, novos, úteis e acessíveis. A principal vantagem do processo é que não são necessárias hipóteses, sendo que o conhecimento é extraído dos dados sem conhecimento prévio".

Ainda segundo Fayyad et al. (1996), o KDD refere-se ao processo global de descoberta de conhecimento útil a partir dos dados, e mineração de dados refere-se a uma etapa específica neste

processo. A mineração de dados é a aplicação de algoritmos específicos para extrair padrões a partir de dados. Os passos adicionais no processo KDD, tais como preparação de dados, seleção de dados, limpeza de dados, incorporação de conhecimento prévio adequado, e uma interpretação correta dos resultados na mineração, são essenciais para garantir se o conhecimento derivado a partir dos dados pode ser útil.

Com análise de dados dos municípios do Paraná, tem-se o objetivo de investigar características relevantes para melhoria do IDHM por meio de Árvores e Regras da Classificação.

Na Classificação, cada padrão contém um conjunto de atributos e um dos atributos é denominado classe. Dado um conjunto de registros, seja o registro X_i e seu rótulo categórico Y_i, o objetivo é descobrir uma função que mapeie um conjunto de registros X_i em um conjunto de dados a um único rótulo categórico Y_i, denominada classe (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

Ao remeter uma concepção implícita ao aprendizado supervisionado, considera-se um conjunto de pares ordenados da forma $(x, f(x))$, onde x é a entrada e $f(x)$ a saída de uma função f, desconhecida, aplicada a x. Um dos grupos contém os atributos de predição e o outro o atributo para o qual se deve fazer a predição de um valor. O aprendizado ocorre a partir de vários exemplos de f, obtendo uma função (hipótese) h que se aproxime de f (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

O aprendizado da função h ocorre no conjunto de treinamento, que é representado por registros cujos rótulos sejam conhecidos e devem ser fornecidos. Então, a função h é aplicada ao conjunto de dados com rótulos de classe omitida, sendo denominado conjunto de teste (TAN; STEINBACH; KUMAR, 2009).

Há muitas técnicas para a construção de classificadores, como classificadores baseados em regras, árvore de decisão, classificadores Bayesianos, Redes Neurais, entre outros. Segundo Tan; Steinbach; Kumar (2009) a árvore de decisão é uma técnica simples, porém muito utilizada.

Segundo Rodrigues (2005) uma das principais características da árvore de decisão é sua representação, em que há uma estrutura hierárquica composta de raiz e nós de decisão, em que se tomam decisões acerca da estrutura do nível seguinte. Ao final do processo, cada classe será rotulada de acordo com os nós terminais (nós folha). Trata-se de uma pesquisa que se desenvolve do geral para o particular, em que a cada nó de decisão se particulariza o valor de mais um atributo explicativo.

São exemplos de tarefa de classificação a detecção de mensagens de spam em e-mails (baseada no cabeçalho e conteúdo da mensagem), a categorização de células como malignas

ou benignas (baseada nos resultados de varreduras MRI), dentre outros (TAN; STEINBACH; KUMAR, 2009).

A classificação, por ser uma generalização das informações constantes em um acontecimento, necessita de uma avaliação para expressar sua veracidade. Segundo Tan, Steinbach, Kumar (2009), o modelo gerado pelo algoritmo de aprendizagem deve se adaptar bem aos dados de entrada e prever corretamente os rótulos de classes de registros do conjunto de teste.

Uma medida de desempenho muito utilizada é a acurácia, que mostra os conflitos que existem entre as classes, que ocorre em função de decisões inadequadas do algoritmo de aprendizado. Outra medida de desempenho é a completude, que avalia a capacidade em classificar (apresentar uma resposta) a todos os exemplos do conjunto de dados. Também existe a Matriz de Confusão, uma avaliação detalhada, em que as classificações correspondem às variáveis assumidas como acertos quando $i = j$ e erro se $i \neq j$ (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

2 Metodologia

Inicialmente realizou-se a coleta de dados, referente ao ano de 2010, considerando-se 22 variáveis: Abastecimento de água (unidades atendidas); PIB (per capita); distância da sede municipal à capital (em km); atendimento de esgoto (unidades atendidas); renda média domiciliar per capita; despesas de capitais municipais; população com dez anos ou mais e despesas municipais por função de: Assistência social, previdência social, saúde, trabalho, educação, cultura, direitos da cidadania, urbanismo, habitação, saneamento, agricultura, indústria, comércio e serviços, comunicações, desporto e lazer.

Algumas variáveis foram ajustadas de modo a facilitar sua interpretação bem como permitir a comparação de municípios de diferentes portes: As variáveis "Abastecimento de água", "Atendimento de esgoto" e "População com 10 anos ou mais" foram relativizadas à população; as variáveis relativas à despesa foram relativizadas ao PIB; e o valor 0 foi atribuído às informações não apresentadas.

Foi então construída a matriz D de ordem 399×22, consistindo pelos 399 municípios e por 22 variáveis. Em seguida foi aplicada a normalização por desvio padrão de acordo com a equação:

$$A' = \frac{A - \mu}{\sigma}. \tag{1}$$

Nessa equação, A' é o valor padronizado; A é o valor da entrada; μ é a média do atributo

de entrada; e σ é o desvio padrão do atributo de entrada.

Para atribuir a classe correspondente a cada município, utilizou-se somente o IDHM já estabelecido para cada município. Houve classes divididas por meio da distância do desvio padrão em relação à média ou dividida por intervalos de mesmo tamanho.

Para esse trabalho foi utilizado o algoritmo J48 que é uma implementação do algoritmo C4.5 no Weka e é considerado o mais popular.

O algoritmo C4.5 é um método baseado na indução de árvore de decisão a partir de uma abordagem recursiva de particionamento do conjunto de dados. O conjunto de treinamento é separado em diferentes conjuntos quando é definido algum predicado em relação aos valores de cada atributo preditor. Essa divisão ocorre até que todos ou a maioria dos exemplos de um subconjunto pertençam a uma classe (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

Durante o processo de utilização do algoritmo é interessante conhecer alguns parâmetros que podem ser modificado como, por exemplo, o uso ou não de podas na árvore e o número mínimo de instâncias por folha.

Para explicar esse algoritmo, consideremos um conjunto X cuja a representação é dada por X $(A_1, A_2, A_3, ..., A_n, C)$ em que cada atributo A_i é um atributo previsor e C é o atributo alvo que assume valores do domínio $\{C_1, C_2, ..., C_k\}$ que são as classes definidas (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

No início da classificação, a raiz da árvore contém todo o conjunto de dados com vários exemplos das várias classes. Em seguida, um ponto de separação é escolhido como sendo a melhor condição para separação de classes. Um predicado envolve um dos atributos previsores e induz uma divisão do conjunto de dados em dois ou mais subconjuntos disjuntos, particionado até que cada subconjunto associado a cada nó folha consista predominantemente de registros de uma mesma classe (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

Há duas operações principais durante a construção das árvores: a avaliação para escolher o melhor ponto de separação e a criação das partições usando o melhor ponto de separação. Uma vez determinado o ponto de separação de um nó, as partições podem ser facilmente escolhidas, aplicando o critério de separação identificado (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

Para identificar o melhor ponto de separação para um determinado nó da árvore, considere um conjunto T de registros em que o algoritmo C4.5 calcula $H(T)$, denominada a entropia de T por meio da equação:

$$H(T) = \sum_{j=1}^{k} \frac{freq\, c_j T}{|T|} \cdot log_2 \frac{freq\, c_j T}{|T|} \tag{2}$$

Nesta equação, T representa o conjunto de entrada; $freq\, c_j T$ corresponde à quantidade de registros da classe em T; $|T|$ denota o total de registros do conjunto T; k indica o número de classes distintas que ocorrem em registros de T.

A entropia de um conjunto de dados é um valor entre 0 e 1. Por exemplo, se todos os registros são de uma mesma classe, então a entropia será igual à zero; e se todos os registros são de classes diferentes, a entropia será igual a 1 (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

O próximo passo é determinar $Info_{A_i}(T)$, o valor esperado pela entropia (já que T é dividido pelos valores de A_i). Considerando os valores no domínio A_i (valores de uma variável A_i) que conduz uma partição sobre T, o que resulta em subconjuntos $\{T_i\}$, o cálculo é dado por:

$$Info_{A_i}(T) = \sum_{j=1}^{n} \frac{|T_j|}{|T|} \cdot H(T_j) \tag{3}$$

Nesta equação, T representa o conjunto de registros de entrada; T_j é cada subconjunto da partição induzida por A_i sobre T; $|T_j|$ representa a quantidade de registros em T_j.

Por fim, calcula-se o ganho de informação de um atributo previsor A_i com relação a uma coleção de exemplos T, dado por:

$$GInfo(A_i, T) = H(T) - info_{A_i}(T) \tag{4}$$

O ganho de informação de um atributo A_i em relação a um conjunto T representa a redução da impureza na situação em que os valores de A_i são usados para dividir T (GOLDSCHIMIDT; PASSOS; BEZERRA, 2015).

3 Resultados e discussões

A fim de encontrar regras com uma boa precisão foram testadas diferentes quantidades de classes (4, 5, 6, 8 e 10). As classes foram definidas por meio da distância em relação à média ou por intervalos de mesmo tamanho. Nos diferentes ensaios as variáveis renda média domiciliar per capita e população com 10 anos ou mais foram relevantes. Deve-se considerar aqui que a renda média é utilizada no cálculo do IDHM.

As árvores de decisão analisadas foram as que apresentaram as melhores precisões. As divisões das classes foram feitas de duas formas: na Árvore 1 foram atribuídas quatro classes, sendo que a definição das classes ocorreu a partir da média e do desvio padrão; na Árvore 2, com cinco classes, a definição das mesmas se deu a partir de intervalos de mesmo tamanho.

As classes referentes à Árvores 1 foram as seguintes: Classe 4: IDHM muito acima da média; Classe 3: IDHM intermediário, porém acima da média; Classe 2: IDHM intermediário, porém abaixo da média e Classe 1: IDHM muito abaixo da média. E para a Árvore 2 foram: Classe 5: IDHM muito alto; Classe 4: IDHM alto; Classe 3: IDHM intermediário; Classe 2: IDHM baixo e Classe 1: IDHM muito baixo, sempre relativamente ao Estado do Paraná.

Para a Árvore 1, a precisão da árvore foi de $69,7\%$, sendo que a poda considerou a restrição de 15 objetos para cada folha. Devido à impossibilidade de apresentação da árvore (pela limitação no tamanho da figura) são apresentadas as regras de classificação:

Se a renda média domiciliar é maior então Classe 4, classificando 45 municípios, sendo classificados errados;

Se a renda média domiciliar for a segunda maior então Classe 3, classificando 96 municípios, sendo 17 deles, classificados errados;

Se a renda média domiciliar for a segunda maior e a população for maior, então Classe 3, classificando 80 municípios, sendo 21 deles errados;

Se a renda média domiciliar for a segunda maior, a população for maior e a despesa em saneamento for maior então Classe 3, obtendo 18 municípios, sendo 6 deles, classificados errados;

Se a renda média domiciliar for a terceira maior, a distância entre a sede do município e a capital for maior e o PIB for maior então Classe 3, classificando 21 municípios, sendo 7 deles classificados errados;

Se a renda média domiciliar for a segunda maior, a população for maior e a despesa em saneamento for menor então Classe 2, obtendo 19 municípios, sendo 6 deles, classificados errados;

Se a renda média domiciliar for a terceira maior, a distância entre a sede do município e a capital for menor então Classe 2, classificando 18 municípios, sendo 4 deles classificados errados;

Se a renda média domiciliar for a terceira maior, a distância entre a sede do município e a capital for maior e o PIB for menor então Classe 2, classificando 35 municípios, sendo 6 deles classificados errados;

Se a renda média domiciliar for a quarta maior e a população for maior então Classe 2, classificando 15 municípios, sendo 3 deles classificados errados;

Se a renda média domiciliar for a quarta maior e a população for menor então Classe 1, classificando 29 municípios, sendo 7 deles classificados errados;

Se a renda média domiciliar for a menor então classe 1, com 23 municípios classificados corretamente.

Para a Árvore 2, a precisão da árvore foi de $75,6\%$, sendo que a poda considerou a restrição de 10 objetos para cada folha. Abaixo são apresentadas as regras de classificação:

Se a renda média domiciliar for a maior então Classe 5, classificando 14 municípios, sendo 6 deles classificados errados;

Se a renda média domiciliar for a segunda maior então Classe 4, classificando 53 municípios corretamente;

Se a população com dez anos e mais for menor e a renda média domiciliar for a terceira maior então Classe 4, classificando 19 municípios e destes, 4 deles classificados errados;

Se a população com dez anos e mais for maior a despesa com agricultura for menor então Classe 4, classificando 20 municípios corretamente;

Se a população com dez anos e mais for maior, a despesa com agricultura for maior e a despesa com cultura for menor então Classe 4, obtendo 64 municípios, sendo 12 mal classificados;

Se a renda média domiciliar for a quinta maior, o PIB for maior e a população com dez anos e mais for maior então Classe 4, classificando 20 municípios e destes, 3 classificados errados;

Se a população com dez anos e mais for maior e a despesa com agricultura e cultura for maior então Classe 3, classificando 11 municípios e destes, 4 classificados errados;

Se a população com dez anos e mais for menor e a renda média domiciliar for a quarta maior então Classe 3, classificando 27 municípios, sendo 7 deles classificados errados;

Se a renda média domiciliar for a quinta maior e o PIB for menor então Classe 3, classificando 25 municípios corretamente;

Se a renda média domiciliar for a quinta maior, o PIB for maior e a população com dez anos e mais for menor então Classe 3, classificando 88 municípios, com 16 mal classificados;

Se a renda média domiciliar for a sexta maior e a população com dez anos e mais for maior então Classe 3, classificando 13 municípios, com 2 mal classificados;

Se a renda média domiciliar for a sexta maior e a população com dez anos e mais for menor então Classe 2, classificando 31 municípios, com 8 mal classificados;

Se a renda média domiciliar for a sétima maior então Classe 2, classificando 23 municípios, com 4 mal classificados;

Se a renda média domiciliar for menor e a despesa com agricultura for menor então Classe 1, sendo 4 municípios classificados nessa regra.

Na árvore 2, a Classe 1 não foi identificada pois haviam somente dois objetos atribuídos.

Para simplificar a apresentação dos atributos relevantes para cada classe são apresentados os quadros abaixo.

	RMD - Renda	**População**	**Distância à Capital**	**PIB**	**Saneamento**
Classe 1	Baixa	Baixa	–	–	–
Classe 2	Baixa	Mediana	Baixa	Baixa	Baixa
Classe 3	Mediana	Alta	Alta	Alta	Alta
Classe 4	Alta	–	–	–	–

Quadro 2: Resumo – Árvore 1

Os atributos distância à capital, PIB e saneamento (despesas municipais) são previsores somente das classes 2 e 3, sendo que, da mesma forma, todos se mostraram diretamente relacionados ao IDHM. Cabe observar que a distância à capital, portanto, não foi relevante para garantir melhor IDHM (já que quanto menor a distância, mais baixo em relação à média do estado está o IDHM e vice-versa).

O quadro 02 apresenta um resumo das variáveis relevantes na predição de cada classe para a árvore 2. Os atributos que predizem a classe 1 são a renda e a agricultura (despesas municipais), a classe 2 foi predita somente pela renda e pela população e a classe 5 foi predita somente pela renda.

	RMD - Renda	**População**	**Agricultura**	**PIB**	**Cultura**
Classe 1	Baixa	–	Baixa	–	–
Classe 2	Baixa	Baixa	–	–	–
Classe 3	Mediana	Mediana	Alta	Baixa	Alta
Classe 4	Mediana	Alta	Baixa	Alta	Baixa
Classe 5	Alta	–	–	–	–

Quadro 3: Resumo – Árvore 2

O atributo população é previsor somente das classes 2, 3 e 4, sendo diretamente relacionado ao IDHM. O atributo PIB é previsor somente das classes 2 e 3, sendo que, da mesma forma, se mostrou diretamente relacionados ao IDHM.

O atributo cultura (despesas municipais) é previsor somente das classes 3 e 4, sendo este relacionado, porém inversamente, ao IDHM, ou seja, quanto maior o investimento em cultura, menor o IDHM e vice-versa. Isso pode indicar não retorno imediato do investimento ou mau investimento, entre outros motivos. O atributo agricultura não se revelou relacionado (diretamente ou inversamente) ao IDHM.

Considerações Finais

Considerando que o objetivo era contribuir para a identificação de fatores relevantes, que poderiam ser considerados no cálculo do IDHM ou sugerir investimentos para aumentar o índice, observaram-se como fatores relacionados ao IDHM a renda média domiciliar, o percentual da população com 10 anos ou mais, o PIB per capita, as despesas municipais com saneamento e a distância entre a sede municipal e a capital (esta inversamente).

Fatores como renda média domiciliar e PIB já eram esperados já que estas variáveis estão inseridas no cálculo do IDHM.

Espera-se que os resultados apresentados neste trabalho possam servir como auxilio na melhoria dos índices relativos à qualidade de vida e como replanejamento de investimentos feitos em cada município.

Referências

FAYYAD, U. M.; PIATETSKY-SHAPIRO, G.; SMYTH, P.; UTHRUSAMY, R. **Advances in Knowledge Discovery Data Mining.** California: AAAI/MIT, 1996.

GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. **Data Mining.** Rio de Janeiro: Elselvier, 2015.

IPARDES - Instituto Paranaense de Desenvolvimento Econômico e Social. Disponível em: http://www.ipardes.gov.br/. Acesso em: 24 de março de 2016.

PNUD Brasil - Programa das Nações Unidas para o Desenvolvimento. Disponível em: http://www.pnud.org.br/IDH/DH.aspx. Acesso em: 12 de fevereiro de 2016.

RODRIGUES, M. A. S. **Árvores de Classificação.** Portugal: Universidade dos Açores, 2005.

TAN, P.N.; STEINBACH, M.; KUMAR. V. **Introdução ao DATA MINING Mineração de Dados.** Rio de Janeiro: Ciência Moderna Ltda, 2009.

TORRES, H. da G.; FERREIRA, M. P.; DINI, N. P. Indicadores sociais: por que construir novos indicadores como o IPRS. **São Paulo Perspectiva.** vol.17, n.3-4, pp.80-90, 2003.

Agrupamento de dados a partir de mapas auto-organizáveis: Um estudo sobre a configuração de parâmetros

André Guilherme Unfried[1][2]
UNIOESTE - Universidade Estadual do Oeste do Paraná
andre_unfrich@hotmail.com

Weverton Rodrigo Verica
UNIOESTE - Universidade Estadual do Oeste do Paraná
wevertonverica@hotmail.com

Rosangela Villwock
UNIOESTE - Universidade Estadual do Oeste do Paraná
rosangela.villwock@unioeste.br

Resumo: Teuvo Kohonen desenvolveu em 1982 o método de mapas de característica auto-organizáveis que faz uso de uma estrutura topológica para agrupar as unidades (padrões). "*Self Organizing Map*"(*SOM*; ou "Mapas auto-organizáveis") é baseado no aprendizado competitivo, onde os neurônios de saída da grade competem entre si para serem ativados. O objetivo deste trabalho foi melhorar a eficácia do algoritmo SOM para a tarefa de agrupamento de dados, pela configuração de seus parâmetros. Os parâmetros estudados neste trabalho foram o tamanho da grade, o raio de vizinhança inicial, a taxa de aprendizagem inicial e o número de iterações. As bases de dados utilizadas foram: Pima Indians Diabetes, Iris e Wine. Com a realização deste trabalho foi possível melhorar a eficácia do algoritmo, apresentando propostas de configurações para as bases de dados estudadas.

Palavras-chave: Agrupamento de Dados; Mapas Auto-Organizáveis; Redes Neurais Artificiais.

1 Introdução

Dentre as muitas definições existentes para *KDD*, pode-se citar a de Fayyad *et al.* (1996), que define *KDD* como "o processo, não trivial, de extração de informações implícitas, previamente desconhecidas e potencialmente úteis, a partir dos dados armazenados em um banco de dados". Segundo estes autores, a principal vantagem do processo de descoberta é que não são necessárias hipóteses, sendo que o conhecimento é extraído dos dados sem conhecimento prévio.

O processo *KDD* é um conjunto de atividades contínuas que são compostas, basicamente, por cinco etapas: seleção dos dados, pré-processamento, formatação ou transformação, Mineração de Dados e interpretação e avaliação do conhecimento, que tem como objetivo a

[1]O autor agradece o apoio financeiro oferecido pelo CNPq através da bolsa de inicição científica.
[2]Os autores agradecem à Fundação Araucária pelo apoio financeiro ao projeto de pesquisa "Identificação do Perfil dos Municípios do Estado do Paraná por meio do Processo de Descoberta de Conhecimento em Bases de Dados".

descoberta de padrões válidos, novos, úteis e acessíveis, como ilustrado na figura 1. Cada fase possui uma ligação com as demais, melhorando assim a cada resultado.

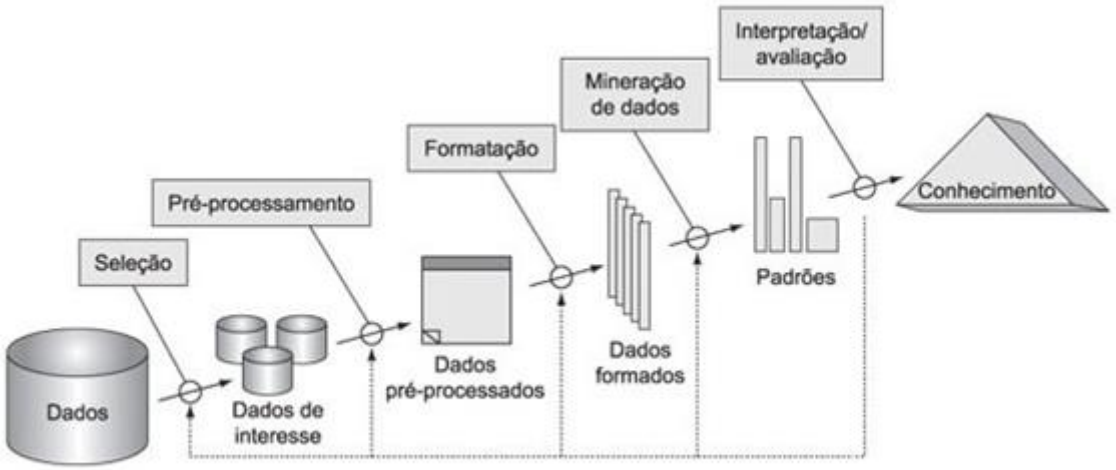


Figura 1: Etapas do processo *KDD*, adaptada de Fayyad *et al.* (1996).

Primeiramente deve-se ter domínio da aplicação e objetivos claros. Na primeira etapa são selecionados e coletados os dados necessários. Na etapa de pré-processamento, esta fase determina a qualidade dos dados, verificam-se os dados faltantes ou inconsistentes. Na etapa de transformação/formatação há uma preparação dos dados visando à aplicação da Mineração de Dados, usando métodos de redução de dimensionalidade dos dados, por exemplo. A etapa de Mineração de Dados é o núcleo do processo, onde são aplicados os algoritmos para extrair padrões dos dados. Nesta etapa deve-se escolher o método a ser utilizado. A etapa de Interpretação dos resultados consiste em validar o conhecimento extraído (FAYYAD *et al.*, 1996).

As tarefas de Mineração de Dados podem ser preditivas ou descritivas. As preditivas usam variáveis para predizer valores futuros ou desconhecidos, enquanto que as descritivas encontram padrões para descrever os dados. As principais tarefas de Mineração de Dados estão relacionadas à Classificação, Associação e Agrupamento de padrões (FAYYAD *et al.*, 1996).

Na Classificação, cada padrão contém um conjunto de atributos e um dos atributos é denominado classe. O objetivo da classificação é encontrar um modelo para predição da classe como função dos outros atributos (TAN; STEINBACH; KUMAR, 2005).

Já na Associação, o objetivo é produzir regras de dependência que irão predizer a ocorrência de um atributo baseado na ocorrência de outros atributos (TAN; STEINBACH; KUMAR, 2005).

O Agrupamento ou Segmentação (*Clustering*) procura grupos de padrões tal que padrões

pertencentes a um mesmo grupo são mais parecidos uns com os outros e divergentes a padrões em outros grupos. Segundo Hair Jr *et al.* (2005), a análise de agrupamentos é uma técnica analítica para desenvolver subgrupos significativos de objetos. Seu objetivo é classificar os objetos em um pequeno número de grupos mutuamente excludentes.

Os algoritmos de agrupamento podem ser divididos em categorias de diversas formas de acordo com suas características. As duas principais classes de algoritmos de agrupamento são: os métodos hierárquicos e os métodos de particionamento (VILLWOCK, 2009).

Os métodos hierárquicos abrangem técnicas que buscam de forma hierárquica os grupos e, por isso, admitem auferir vários níveis de agrupamento (DINIZ; LOUZADA-NETO, 2000).

Métodos não-hierárquicos ou de particionamento procuram uma partição sem a necessidade de associações hierárquicas. Seleciona-se uma partição dos elementos em k grupos, otimizando algum critério (DINIZ; LOUZADA-NETO, 2000).

Outros autores utilizam ainda outras classes os algoritmos de Agrupamento de Dados. Tan, Steinbach e Kumar (2005) classificam os diversos algoritmos de Agrupamento de Dados em Método de Partição, Método Hierárquico, Método Baseado em Densidade e Método Baseado em Malhas. Neste trabalho o algoritmo SOM foi utilizado.

A metodologia de agrupamento de dados a partir do SOM consiste em métodos de visualização e de agrupamentos. Como métodos de visualização podemos citar a Matriz de Densidade. Para agrupamento de dados a partir do SOM uma metodologia utilizada é denominada Agrupamento por Matriz de Densidade (Xu & Li, 2002). Esta metodologia foi utilizada neste trabalho.

O objetivo deste trabalho foi melhorar a eficácia do algoritmo SOM para a tarefa de agrupamento de dados, pela configuração de seus parâmetros.

2 Mapas Auto-Organizáveis - SOM

Segundo Fausett (1994), Teuvo Kohonen, em 1982, desenvolveu o método de mapas de característica auto-organizáveis que faz uso de uma estrutura topológica para agrupar as unidades (padrões). "*Self Organizing Map*" (*SOM*; ou "Mapas auto-organizáveis"), também conhecidos por Redes Neurais de Kohonen, formam uma classe de Redes Neurais Artificiais em que a aprendizagem é não supervisionada.

Segundo Haykin (2001) o principal objetivo do SOM é transformar padrões de entrada de

dimensão arbitrária em um mapa discreto. Os neurônios são colocados nos nós de uma grade, que pode ter qualquer dimensionalidade. Normalmente são utilizadas grades bidimensionais (chamado de 2D-SOM).

Há três passos essenciais envolvidos na formatação do mapa auto-organizável, competição, cooperação e adaptação. Na competição, os neurônios da grade calculam seus respectivos valores de uma função discriminante. Esta função discriminante fornece a base para a competição entre os neurônios. O neurônio particular com o maior valor da função discriminante é declarado vencedor da competição. Na cooperação, o neurônio vencedor determina a localização espacial de uma vizinhança topológica de neurônios excitados, fornecendo assim a base para a cooperação entre os neurônios vizinhos. A adaptação permite que os neurônios excitados aumentem seus valores individuais da função discriminante em relação ao padrão de entrada através de ajustes adequados aplicados a seus pesos sinápticos. A descrição dos processos, baseada em Haykin (2001), é dada a seguir.

2.1 O processo competitivo

Considere um objeto (vetor de entrada) selecionado aleatoriamente do espaço de entrada, representado por:

$$\mathbf{x} = [x_1, x_2, \ldots, x_m]^T \tag{1}$$

onde m é o número de variaveis.

Considere que o vetor peso sináptico do neurônio j seja representado por:

$$\mathbf{w}_j = [w_{j1}, w_{j2}, \ldots, w_{jm}]^T,\ j = 1, 2, 3, \ldots, l \tag{2}$$

onde l é o número total de neurônios da grade.

Para cada um dos n indivíduos, encontra-se o neurônio com o melhor casamento (vencedor) $i(\mathbf{x})$ para determinar a localização onde a vizinhança topológica dos neurônios excitados deve ser centrada, usando o critério da mínima distância euclidiana:

$$i(\mathbf{x}) = arg\ min_j\ ||\mathbf{x} - \mathbf{w}_j||,\ j = 1, 2, 3, \ldots, l \tag{3}$$

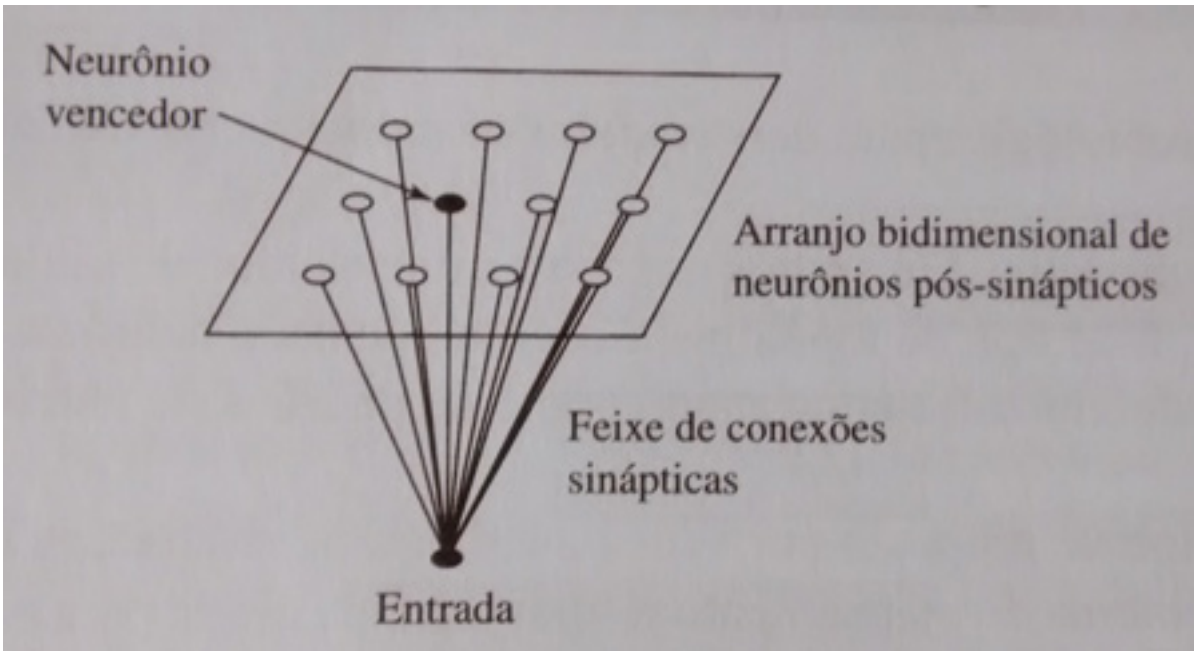


Figura 2: Modelo de Kohonen

2.2 O processo Cooperativo

O neurônio vencedor $i(\mathbf{x})$ localiza o centro de uma vizinhança topológica de neurônios cooperativos. Considere a função de vizinhança topológica variável no tempo $h_{j,i(\mathbf{x})}(n)$.

$$h_{j,i(\mathbf{x})}(n) = exp\left(-\frac{d_{j,i}^2}{2\sigma^2(n)}\right),\ n = 0, 1, 2, \ldots, \tag{4}$$

onde $d_{j,i}$ representa a distância lateral entre o neurônio vencedor i e o neurônio excitado j. O parâmetro σ é a "largura efetiva" da vizinhança topológica, ele mede o grau com o qual neurônios excitados na vizinhança do neurônio vencedor participam do processo de aprendizagem, n representa o tempo discreto.

Para que a cooperação entre neurônios vizinhos se mantenha, é necessário que a vizinhança topológica $h_{j,i}$ seja dependente da distância lateral $d_{j,i}$ entre o neurônio vencedor i e o neurônio excitado j no espaço de saída. No caso de uma grade bidimensional, ela é definida por:

$$d_{j,i}^2 = ||\mathbf{r}_j\ -\ \mathbf{r}_i||^2 \tag{5}$$

onde o vetor discreto $\mathbf{r}_j$ define a posição do neurônio excitado j e $\mathbf{r}_i$ define a posição discreta do neurônio vencedor i, sendo ambos medidos no espaço de saída discreto.

Outra exigência é que o tamanho da vizinhança topológica diminua com o tempo, ou seja, a largura σ da função de vizinhança topológica diminua com o tempo, onde τ_1 é uma

constante de tempo e σ_0 é o raio de vizinhança inicial.

$$\sigma(n) = \sigma_0 \; exp\left(-\frac{n}{\tau_1}\right), \; n = 0, 1, 2, \ldots, \tag{6}$$

Assim, quando o tempo n aumenta, a largura $\sigma(n)$ decresce a uma taxa exponencial e a vizinhança topológica diminui de uma maneira correspondente.

2.3 O processo Adaptativo

Para que o mapa seja auto-organizável, é necessário que o vetor de peso sináptico $\mathbf{w}_j$ do neurônio j da grade se modifique em relação ao vetor de entrada $\mathbf{x}$. Ajustam-se os vetores de peso sináptico de todos os neurônios usando a fórmula de atualização

$$\mathbf{w}_j(n+1) = \mathbf{w}_j(n) + \eta(n) h_{j,i(\mathbf{x})}(n)(\mathbf{x} \; - \; \mathbf{w}_j(n)) \tag{7}$$

onde $\eta(n)$ é o parâmetro da taxa de aprendizagem e $h_{j,i(\mathbf{x})}(n)$ é a função de vizinhança centrada em torno do neurônio vencedor $i(\mathbf{x})$; ambos $\eta(n)$ e $h_{j,i(\mathbf{x})}$ são variados dinamicamente durante a aprendizagem para obter melhores resultados. Essa equação tem o efeito de mover o vetor peso sináptico $\mathbf{w}_i$ do neurônio vencedor i em direção ao vetor de entrada $\mathbf{x}$. Através da apresentação repetida dos dados de treinamento, os vetores de peso sináptico tendem a seguir a distribuição dos vetores de entrada devido à atualização da vizinhança.

O parâmetro taxa de aprendizagem $\eta(n)$ começa em um valor inicial η_0 e então diminui gradualmente com o tempo, n, mas nunca vai a zero.

$$\eta(n) = \eta_0 \; exp\left(-\frac{n}{\tau_2}\right), \; n = 0, 1, 2, \ldots, \tag{8}$$

onde n representa o tempo discreto e τ_2 é uma constante de tempo do algoritmo SOM. Os valores de η_0 e τ_2 utilizados por Haykin (2001) são $\eta_0 = 0,1$ e $\tau_2 = 1000$.

3 Metodologia

Inicialmente foi realizado um estudo sobre o algoritmo e uma pesquisa na literatura por valores utilizados nos parâmetros do algoritmo. Os parâmetros estudados neste trabalho foram o tamanho da grade, o raio de vizinhança inicial, a taxa de aprendizagem inicial e o número de iterações.

Foram escolhidos alguns valores para cada parâmetro e foram feitas composições com os valores definidos. Estas composições foram avaliadas em bases de dados reais e públicas já

rotuladas. As bases de dados utilizadas foram: Pima Indians Diabetes, Iris e Wine, retiradas do repositório da UCI, no endereço: http://archive.ics.uci.edu/ml/datasets.html. O fato das bases de dados já serem rotuladas permitiu avaliar a qualidade do agrupamento obtido, utilizando a porcentagem de acerto como métrica.

Com relação ao parâmetro taxa de aprendizagem inicial, foram avaliados os valores 0.1; 0.5 e 0.8. Para o tamanho da grade foram avaliadas as grades 5x5, 10x10, 15x15 e 20x20. Para os valores do raio de vizinhança inicial foram avaliados os raios 2, 3 e 5. Referente ao número de iterações, foram avaliadas 200, 500 e 1000 iterações.

A ferramenta utilizada para o trabalho foi a YADMT – *Yet Another Data Mining Tool* (Benfatti *et al.*, 2010), em desenvolvimento na Universidade Estadual do Oeste do Paraná desde 2010. Foram efetuadas dez execuções do algoritmo em cada base de dados. No caso de divergências entre os resultados executou-se o cálculo da média das medidas de avaliação (porcentagem de acerto, considerando que cada base de dados já é rotulada).

4 Resultados e Discussão

Para a base de dados Iris, os melhores resultados, tendo como base a porcentagem de acerto, foram obtidos com a utilização dos seguintes parâmetros: aprendizagem inicial: 0.1; grade 10x10; raio inicial: 5; número de iterações: 1000. Com esta configuração, a porcentagem de acerto média foi de 82,9 %. Com a configuração padrão apresentada na ferramenta, a porcentagem de acerto média para esta base de dados foi de 74,5 %. Observa-se um aumento de 8,4% na porcentagem de acerto, melhorando a eficácia do algoritmo.

Para a base de dados Pima, os melhores resultados, tendo como base a porcentagem de acerto, foram obtidos com a utilização dos seguintes parâmetros: aprendizagem inicial: 0.5; grade 5x5; raio inicial: 3; número de iterações: 200. Com esta configuração, a porcentagem de acerto média foi de 71,1 %. Com a configuração padrão apresentada na ferramenta, a porcentagem de acerto média para esta base de dados foi de 65,3 %. Observa-se um aumento de 5,8% na porcentagem de acerto, melhorando a eficácia do algoritmo.

Para a base de dados Wine, os melhores resultados, tendo como base a porcentagem de acerto, foram obtidos com a utilização dos seguintes parâmetros: aprendizagem inicial: 0.8; grade 5x5; raio inicial: 2; número de iterações: 500. Com esta configuração, a porcentagem de acerto média foi de 80,3 %. Com a configuração padrão apresentada na ferramenta, a porcentagem de acerto média para esta base de dados foi de 66,6 %. Observa-se um aumento de 13,7% na

porcentagem de acerto, melhorando a eficácia do algoritmo.

Observando os melhores resultados, a melhor configuração considerando a média das porcentagens de acerto para as três bases de dados, foi obtida com a utilização dos seguintes parâmetros: aprendizagem inicial: 0.1; grade 10x10; raio inicial: 5; número de iterações: 1000.

Para cada parâmetro foi fixado cada valor atribuído a ele e calculada a média das porcentagens de acerto considerando todas as combinações nos parâmetros restantes, isto é, para os demais parâmetros foram considerados todos os valores atribuídos a eles. Assim, a melhor configuração para a base de dados Iris foi: aprendizagem inicial: 0.1; grade 10x10; raio inicial: 2; número de iterações: 500. Com esta configuração, a porcentagem de acerto média foi de 76,5%.

Já a melhor configuração para a base de dados Pima foi: aprendizagem inicial: 0.8; grade 5x5; raio inicial: 2; número de iterações: 200. Com esta configuração, a porcentagem de acerto média foi de 67,1 %. A melhor configuração para a base de dados Wine foi: aprendizagem inicial: 0.5; grade 5x5; raio inicial: 5; número de iterações: 500. Com esta configuração, a porcentagem de acerto média foi de 78,1 %. Observa-se que esta estratégia também melhora a eficácia do algoritmo relativamente à configuração padrão apresentada na ferramenta. Como era de se esperar, o desempenho desta estratégia foi inferior ao da primeira, porém com diminuição máxima de 8% na porcentagem de acerto.

Observando os resultados considerando os parâmetros globalmente, fixado cada valor no parâmetro e combinados os demais, a melhor configuração, considerando os maiores percentuais de acerto para as três bases de dados, foi: aprendizagem inicial: 0.5; grade 5x5; raio inicial: 5; número de iterações: 500. Com esta configuração, a porcentagem de acerto média para as bases de dados analisadas foram de 69,7 % (Iris), 78,1 % (Wine) e 66,7 % (Pima). Observa-se que esta estratégia melhora a eficácia do algoritmo relativamente à configuração padrão apresentada na ferramenta para duas das três bases de dados.

5 Conclusões

Com a realização deste trabalho foi possível melhorar a eficácia do algoritmo, apresentando propostas de configurações para as bases de dados estudadas, bem como uma configuração global. Sugere-se como trabalhos futuros ampliar o estudo utilizando bases de dados diferentes para avaliar a configuração global.

Referências

Benfatti, E. W.; Bonifacio, F. N.; Girardello, A. D.; Boscarioli, C. **Descrição da Arquitetura e Projeto da Ferramenta YADMT - Yet Another Data Mining Tool**. Relatório Técnico nº 01 do Curso de Ciência da Computação, UNIOESTE, Campus de Cascavel, 2010.

Diniz, C. A. R.; Louzada Neto, F. **Data mining: uma introdução**. São Paulo: ABE, 2000.

Fausett, L. **Fundamentals of Neural Networks – Architectures, Algorithms, and Applications**. New Jersey: Prentice Hall, 1994.

Fayyad, U. M.; Piatetsky-Shapiro, G.; Smyth, P.; Uthrusamy, R. (1996). **Advances in knowledge Discovery & Data Mining**. California: AAAI/MIT, 1996.

Hair Jr, J. F.; Anderson, R. E.; Tatham, R. L.; Black, W. C. **Análise Multivariada de Dados**. Tradução de: Santanna, A. S.; Chaves Neto, A. Porto Alegre: Bookman, 2005.

Haykin, S. **Redes neurais: princípios e prática**. Tradução: Engel, P. M. Porto Alegre: Bookman, 2001.

Tan, P. N.; Steinbach, M.; Kumar, V. **Introduction to Data Mining**. Inc. Boston, MA, USA: Addison-Wesley Longman Publishing Co., 2005.

Villwock, R. **Técnicas de agrupamento e de hierarquização no contexto de KDD – Aplicação a dados temporais de instrumentação geotécnica-estrutural da Usina Hidrelétrica de Itaipu**. Tese (Doutorado em Métodos Numéricos em Engenharia). Universidade Federal do Paraná, Curitiba, 2009.

Xu, B.; Li, S. (2002). Automatic Color Identification in Printed Fabric Images by a Fuzzy Neural Network. **AATCC Review**, v. 2, p. 42-45, 2002.

Convergência do método de Região de Confiança com Gradientes Conjugados para otimização irrestrita

Camila Frank Hollmann[1]
Universidade Estadual do Oeste do Paraná
camila_frank@hotmail.com

Paulo Domingos Conejo
Universidade Estadual do Oeste do Paraná
paulo.conejo@unioeste.br

Resumo: Otimização irrestrita consiste em minimizar ou maximizar uma função sem nenhuma restrição sobre as variáveis. Neste trabalho, estamos interessados em minimizar uma função utilizando o método de Região de Confiança com Gradientes Conjugados. Este algoritmo consiste em aproximar a função objetivo por um modelo quadrático e minimizá-lo dentro de uma região chamada região de confiança. Para minimizar o modelo, é utilizada a estratégia de gradientes conjugados. O método de região de confiança irrestrita com a estratégia de gradientes conjugados modificado foi implementada em Scilab e um teste foi realizado.

Palavras-chave: Região de confiança; gradientes conjugados; otimização irrestrita.

1 Introdução

Neste trabalho consideramos o problema de minimização irrestrita

$$\begin{aligned} &\text{minimizar} \quad f(x) \\ &\text{sujeita a} \quad x \in \mathbb{R}^n, \end{aligned} \tag{1}$$

onde $f : \mathbb{R}^n \to \mathbb{R}$ é limitada e duas vezes continuamente diferenciável em $\mathbb{R}^n$.

Usaremos o método de regiões de confiança para resolver o problema (1), definindo a cada iteração o modelo quadrático $m_k(p)$, que é uma aproximação aceitável da função objetivo f na região em torno do ponto corrente x^k, sendo essa chamada de região de confiança. Então minimizamos o modelo sujeito à região de confiança para encontrar o minimizador $\overline{p}$. Se o ponto encontrado fornece uma boa redução na função objetivo dentro da região de confiança, então este ponto é aceito pelo método e a região pode ser aumentada na próxima iteração; por outro lado, se a aproximação é ruim, a região é diminuída e o modelo é minimizado dentro da nova região de confiança.

Cada subproblema consiste em

$$\begin{aligned} &\text{minimizar} \quad m_k(p) = f\left(x^k\right) + \nabla f\left(x^k\right)^T p + \tfrac{1}{2}(p)^T \nabla^2 f\left(x^k\right) p \\ &\text{sujeita a} \quad \|p\| \leq \Delta_k, \end{aligned} \tag{2}$$

[1]Agradeço ao CNPq pela bolsa concedida durante este projeto.

onde Δ_k é o raio da região de confiança, $\nabla f(x^k) \in \mathbb{R}^n$ é o vetor gradiente, $\nabla^2 f(x^k) \in \mathbb{R}^{n\times n}$ é a matriz Hessiana da função objetivo avaliada no ponto x^k e $\|p\|$ é a norma euclidiana no ponto p. A estratégia utilizada para resolver (2) será o Método de Gradientes Conjugados Modificado, que minimiza o modelo em, no máximo, n passos. O minimizador $\overline{p}$ obtido em (2) será o passo para obter o próximo ponto para o problema (1), $x^{k+1} = x^k + \overline{p}$.

2 Método de Região de Confiança

O método de região de confiança utiliza um modelo que aproxima a função objetivo, e então define uma região em torno do ponto corrente na qual é possível confiar no modelo. Com o raio da região definido, calculamos um minimizador aproximado do modelo nesta região. Verificamos, então, se o ponto também fornece um decréscimo na função objetivo. Caso forneça, o ponto é aceito e o processo é repetido, definindo uma região de confiança em torno deste novo ponto corrente. Caso contrário, o raio da região de confiança é reduzido, pois é possível que o modelo não represente adequadamente a função objetivo. Então, procuramos um novo minimizador para o modelo nesta região (RIBEIRO; KARAS, 2013).

Vamos considerar uma função $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^2. Dado um ponto $x^k \in \mathbb{R}^n$, o modelo quadrático de f em torno de x^k é definido por

$$q_k(x) = f\left(x^k\right) + \nabla f\left(x^k\right)^T \left(x - x^k\right) + \frac{1}{2}\left(x - x^k\right)^T B_k \left(x - x^k\right), \tag{3}$$

onde $B_k \in \mathbb{R}^{n\times n}$ pode ser a Hessiana $\nabla^2 f(x^k)$ ou qualquer outra matriz simétrica tal que $\|B_k\| \leq \beta$, para alguma constante $\beta > 0$. A importância de f ser de classe C^2 é para a prova de convergência do método de região de confiança que não apresentaremos neste trabalho e pode ser encontrada em Ribeiro e Karas (2013).

O modelo definido em (3) aproxima a função f em uma vizinhança de x^k. Sendo $\Delta_k > 0$ o raio da região de confiança, a região em que confiamos no modelo é

$$\left\{x \in \mathbb{R}^n \mid \left|x - x^k\right| \leq \Delta_k\right\}.$$

Agora, considere $p = x - x^k$ e $m_k(p) = q_k\left(x^k + p\right)$. Temos, então, o subproblema

$$\begin{aligned} &\text{minimizar} \quad m_k(p) = f\left(x^k\right) + \nabla f\left(x^k\right)^T p + \tfrac{1}{2} p^T B_k p \\ &\text{sujeito a} \quad \|p\| \leq \Delta_k. \end{aligned} \tag{4}$$

No entanto, a solução $\overline{p}$ de (4) deve ser avaliada. O ponto $x^k + \overline{p}$ deve fornecer uma boa redução na função objetivo. Para verificar se houve boa redução, utilizamos a redução real na

função objetivo deniminada *ared* e a redução do modelo, *pred*, por

$$ared = f\left(x^k\right) - f\left(x^k + \overline{p}\right) \text{ e } pred = m_k\left(0\right) - m_k\left(\overline{p}\right). \tag{5}$$

A razão entre a redução na função objetivo e a redução do modelo é dada por

$$\rho_k = \frac{ared}{pred}. \tag{6}$$

Assim, $\overline{p}$ será aceito quando ρ_k for maior que uma constante $\eta \geq 0$, o que significa que a redução real na função objetivo corresponde à pelo menos uma fração da redução no modelo. Neste caso, o processo é repetido com o próximo ponto corrente, $x^{k+1} = x^k + \overline{p}$. Além disso, se $\overline{p}$ está na fronteira da região de confiança e a redução na função objetivo é grande, podemos aumentar o tamanho do raio Δ_k para a próxima iteração. Se $\overline{p}$ está dentro da região, consideramos que o raio atual Δ_k pode ser mantido na próxima iteração. Observe ainda que caso a função objetivo seja quadrática, o modelo que melhor a aproxima é a própria função, então neste caso teremos $\rho_k = 1$ e o raio da região de confiança será aumentado. Caso $\rho_k \leq \eta$, a redução na função objetivo não foi boa o suficiente e $\overline{p}$ é recusado. O processo é repetido resolvendo o subproblema (4) com o raio da região de confiança Δ_k reduzido (RIBEIRO; KARAS, 2013).

Algoritmo do método de Região de Confiança (RC)

Dados: $x^0 \in \mathbb{R}^n$, $\overline{\Delta} > 0$, $\Delta_0 \in \left(0, \overline{\Delta}\right)$, $\eta \in [0, \frac{1}{4})$ e $\epsilon > 0$

$k = 0$

REPITA enquanto $\left\|\nabla f\left(x^k\right)\right\| \neq 0$

Obtenha $\overline{p}$, solução do subproblema (4)

Calcule ρ_k usando (6)

SE $\rho_k > \eta$ então $x^{k+1} = x^k + \overline{p}$ e $\Delta_{k+1} = \Delta_k$

SENÃO $x^{k+1} = x^k$ e $\Delta_{k+1} = \dfrac{\Delta_k}{2}$

$k = k + 1$

Definido o algoritmo do método de região de confiança, nos deparamos com um novo problema: solucionar o subproblema (4) para encontrar $\overline{p}$, ou seja, minimizar o modelo quadrático definido em torno do ponto corrente x^k. Podemos utilizar um algoritmo interno para solucionar o subproblema e o algoritmo anterior como um algoritmo externo, apenas para definir o modelo quadrático, controlar o raio da região de confiança e verificar se a solução obtida pelo algoritmo interno pode ser aceita. Como algoritmo interno, utilizaremos o método de gradientes conjugados, apresentado na seção a seguir (GARDENGHI; SANTOS, 2012).

3 Método de Direções Conjugadas

Nesta seção apresentaremos o método interno que resolve os subproblemas em n passos, com n a dimensão do subproblema. O texto segue o artigo de Gardenghi e Santos (2012).

Definição 1. Seja $A \in \mathbb{R}^{n\times n}$ uma matriz definida positiva. Dizemos que os vetores $d^0, d^1, ..., d^k \in \mathbb{R}^n \backslash \{0\}$ são A-conjugados se $(d^i)^T A d^j = 0$, para todos $i,\ j = 0, 1, ..., k$, com $i \neq j$.

Definição 2. Um subconjunto S de um espaço vetorial V é linearmente independente (L.I.) se, para qualquer subconjunto finito F de S e constantes λ_v tais que $v \in F$, tem-se $\sum_{v\in F} \lambda_v v = 0$ se, e somente se, $\lambda_v = 0\ \forall\ v \in F$.

Lema 3. *Seja $A \in \mathbb{R}^{n\times n}$ uma matriz definida positiva. Um conjunto qualquer de vetores A-conjugados é linearmente independente.*

A prova do Lema 3 pode ser encontrada em Gardenghi e Santos (2012).

Considere agora uma função quadrática $g : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por

$$g(x) = \frac{1}{2} x^T A x + b^T x + c, \tag{7}$$

com $A \in \mathbb{R}^{n\times n}$ definida positiva, $b \in \mathbb{R}^n$ e $c \in \mathbb{R}$. Veremos que o conhecimento de direções conjugadas permite encontrar o minimizador de g.

A função g é convexa e portanto possui apenas um minimizador $\overline{x}$, que é global e satisfaz

$$A\overline{x} + b = \nabla g(\overline{x}) = 0. \tag{8}$$

Dado um conjunto $\{d^0, d^1, ..., d^{n-1}\}$ de direções A-conjugadas, definimos uma sequência tomando $x^0 \in \mathbb{R}^n$ arbitrário e, para $k = 0, 1, ..., n-1$,

$$x^{k+1} = x^k + t_k d^k, \tag{9}$$

onde $t_k = \underset{t\in\mathbb{R}}{\arg\min} \left\{ g\left(x^k + t d^k\right) \right\}$.

Como d^k pode não ser uma direção de descida para a função no ponto x^k, a minimização é calculada sobre toda a reta e não apenas para valores positivos de t. É possível obter uma fórmula explícita para t_k, já que g é uma função quadrática. Seja a função $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}$ dada por $\varphi(t) = f\left(x^k + t d^k\right)$. Assim,

$$\nabla g\left(x^{k+1}\right)^T d^k = \nabla g\left(x^k + t_k d^d\right)^T d^k = \varphi'(t_k) = 0. \tag{10}$$

Agora, de (8), temos

$$\nabla g\left(x^{k+1}\right) = A(x^k + t_k d^k) + b = \nabla g\left(x^k\right) + t_k A d^k. \tag{11}$$

Substituindo (11) em (10), obtemos $\left(\nabla f\left(x^k\right) + t_k A d^k\right)^T d^k = 0$, de onde tiramos a fórmula para t_k

$$t_k = -\frac{\nabla g\left(x^k\right)^T d^k}{\left(d^k\right)^T A d^k}. \tag{12}$$

O teorema a seguir mostra que o algoritmo do método de direções conjugadas em (9) minimiza a quadrática definida em (7) com no máximo n passos.

Teorema 4. *Considere a função quadrática dada por (7) e seu minimizador $\overline{x}$, definido em (8). Dado $x^0 \in \mathbb{R}^n$, a sequência finita definida em (9) cumpre $x^n = \overline{x}$.*

Prova. Pelo Lema 3, o conjunto $\{d^0, d^1, ..., d^{n-1}\}$ é uma base de $\mathbb{R}^n$, pois um conjunto com n vetores linearmente independentes forma uma base de $\mathbb{R}^n$. Logo, existem escalares $a_i \in \mathbb{R}$, $i = 0, 1, ..., n-1$, tais que

$$\overline{x} - x_0 = \sum_{i=0}^{n-1} a_i d^i. \tag{13}$$

Multiplicando a igualdade acima por $\left(d^k\right)^T A$, sendo $k \in \{0, 1, ..., n-1\}$ arbitrário, temos

$$\left(d^k\right)^T A(\overline{x} - x_0) = \sum_{i=0}^{n-1} a_i \left(d^k\right)^T A d^i.$$

Como as direções são A-conjugadas, $\left(d^k\right)^T A d^i = 0$, para $k \neq i$. Portanto,

$$\left(d^k\right)^T A(\overline{x} - x_0) = a_k \left(d^k\right)^T A d^k.$$

Isolando a_k, obtemos

$$a_k = \frac{\left(d^k\right)^T A(\overline{x} - x_0)}{\left(d^k\right)^T A d^k}. \tag{14}$$

No entanto, de (9), temos

$$x^k = x^0 + t_0 d^0 + t_1 d^1 + ... + t_{k-1} d^{k-1}.$$

Multiplicando por $\left(d^k\right)^T A$ novamente e sabendo que as direções são A-conjugadas,

$$\left(d^k\right)^T A x^k = \left(d^k\right)^T A x^0.$$

Substituindo em (14) e usando (8) e a fórmula para t_k em (12),

$$a_k = \frac{(d^k)^T(A\overline{x} - Ax_0)}{\left(d^k\right)^T A d^k} = -\frac{(d^k)^T(b + Ax^k)}{\left(d^k\right)^T A d^k} = -\frac{(d^k)^T \nabla g\left(x^k\right)}{\left(d^k\right)^T A d^k} = t_k. \tag{15}$$

Substituindo (15) em (13), temos

$$\overline{x} - x_0 = \sum_{i=0}^{n-1} t_i d^i = x^n.$$

□

Dado $x^0 \in \mathbb{R}^n$, vamos definir $d^0 = -\nabla g(x^0)$ e, para $k = 0.1, ..., n-2$,

$$d^{k+1} = -\nabla g\left(x^{k+1}\right) + \beta_k d^k, \tag{16}$$

onde x^{k+1} é dado por (9) e β_k é calculado de modo que d^k e d^{k+1} sejam A-conjugadas, logo

$$\left(d^k\right)^T A d^{k+1} = \left(d^k\right)^T A\left(-\nabla g\left(x^{k+1}\right) + \beta_k d^k\right) = 0. \tag{17}$$

Portanto,

$$\beta_k = \frac{\left(d^k\right)^T A \nabla f(x^{k+1})}{\left(d^k\right)^T A d^k}.$$

O algoritmo de Gradientes Conjugados (GC) é apresentado a seguir

Dado $x^0 \in \mathbb{R}^n$, faça $d^k = -\nabla g\left(x^0\right)$

$k = 0$

REPITA enquanto $\nabla g\left(x^k\right) \neq 0$

$$t_k = -\frac{\nabla g\left(x^k\right)^T d^k}{\left(d^k\right)^T A d^k}$$

$$x^{k+1} = x^k + t_k d^k$$

$$\beta_k = \frac{\left(d^k\right)^T A \nabla f(x^{k+1})}{\left(d^k\right)^T A d^k}$$

$$d^{k+1} = -\nabla g(x^{k+1}) + \beta_k d^k$$

$$k = k + 1$$

Caso $\nabla g\left(x^k\right) \neq 0$, o novo ponto pode ser calculado pois teremos $d^k \neq 0$. Utilizando (10) e a definição de d^k, temos

$$\nabla g\left(x^k\right)^T d^k = \nabla g\left(x^k\right)^T \left(-\nabla g\left(x^k\right) + \beta_{k-1} d^{k-1}\right) = -\left\|\nabla g\left(x^k\right)\right\|^2,$$

onde vemos que o Algoritmo de GC gera direções de descida, característica que não é válida para todos os métodos de direções conjugadas. Logo, o algoritmo está bem definido.

Além disso, as direções d^k geradas pelo algoritmo são A-conjugadas (GARDENGHI; SANTOS, 2012).

4 Método de Gradientes Conjugados e a Estratégia da Região de Confiança

O Algoritmo de GC deve passar por algumas modificações para resolver o subproblema (4). Isto porque a estratégia do método de gradientes conjugados minimiza uma função como a definida em (7), na qual a matriz A, que é a Hessiana da função, é uma matriz definida positiva. Porém, o método não soluciona o subproblema (4) quando a Hessiana da matriz quadrática não for definida positiva. Também há outros critérios de parada que devem ser adicionados ao algoritmo para que ele esteja de acordo com o algoritmo externo, que utiliza o método de região de confiança.

Se a norma do passo p for maior que a região de confiança, temos uma solução na borda.

Segundo este critério, se na k-ésima iteração tivermos $\left\|p^{k+1}\right\| \geq \Delta_k$, basta encontrar τ de forma que

$$\left\|p^k + \tau d^k\right\| = \Delta_k,$$

e a solução de (4) será

$$\overline{p} = p^k + \tau d^k.$$

Se a curvatura do modelo quadrático for negativa, temos uma solução na borda.

Sabemos que o algoritmo de gradientes conjugados minimiza o modelo apenas quando a direção de curvatura for positiva. Se a Hessiana $\nabla^2 f\left(x^k\right)$ não for definida positiva, ou seja, $\left(d^k\right)^T \nabla^2 f\left(x^k\right) d^k \leq 0$, o algoritmo irá encontrar uma direção de curvatura negativa ao longo da qual a função decresce. Neste caso, o minimizador dentro da região de confiança estará na borda da região. Portanto, basta calcular τ, tal que

$$\left\|p^k + \tau d^k\right\| = \Delta_k,$$

e a solução de (4) será

$$\overline{p} = p^k + \tau d^k.$$

Se a norma do resíduo r^{k+1} (ver algoritmo a seguir) estiver muito próxima de zero, temos uma solução interna.

O resíduo corresponde ao valor de quão longe estamos de encontrar a solução de um problema e, nesse caso, equivale ao oposto do gradiente da função objetivo. Portanto, para um dado $\epsilon > 0$, se na k-ésima iteração tivemos $\left\|r^{k+1}\right\| \leq \epsilon$, então o valor obtido já está muito próximo do minimizador real da função e a solução de (4) será

$$p^{k+1}.$$

Com estas modificações, podemos definir o algoritmo do método dos gradientes conjugados modificado.

Algoritmo do Método de Gradientes Conjugados Modificado com Estratégia de Região de Confiança (GCMRC)

Consideramos a Hessiana $\nabla^2 f\left(x^k\right) \in \mathbb{R}^{n\times n}$ e o gradiente $\nabla f\left(x^k\right) \in \mathbb{R}^n$ da função objetivo f, avaliados no ponto corrente x^k e uma tolerância $\epsilon > 0$,

Dados: $x^0 \in \mathbb{R}^n$, $\overline{\Delta} > 0$, $\Delta_0 \in \left(0, \overline{\Delta}\right)$, $\eta \in [0, \frac{1}{4})$ e $\epsilon > 0$

$k = 0$

REPITA enquanto $\left\|\nabla f\left(x^k\right)\right\| \geq \epsilon$

Defina $p^0 = 0$, $r = -\nabla f\left(x^k\right)$, $d^0 = r^0$ e $\delta_0 = \left(r^0\right)^T r^0$

para $i = 0, ..., n$ faça

$h_i = \nabla^2 f\left(x^k\right) d^i$

$\lambda_i = \left(d^i\right)^T h_i$

SE $\lambda_i \leq 0$ então

Encontre τ tal que $\left\|p^i + \tau d^i\right\| = \Delta_k$

retorna $\overline{p} = p^i + \tau d^i$

$t_i = \dfrac{\delta_0}{\eta_i}$

$p^{i+1} = p^i + t_i d^i$

SE $\left\|p^{i+1}\right\| > \Delta_k$ então

Encontre τ tal que $\left\|p^i + \tau d^i\right\| = \Delta_k$

retorna $\overline{p} = p^i + \tau d^i$

$r^{i+1} = r^i + t_i h_i$

$\delta_1 = \left(r^{i+1}\right)^T r^{i+1}$

SE $\sqrt{\delta_i} < \epsilon$ então retorna p^{i+1}

$\beta_i = \dfrac{\delta_1}{\delta_0}$

$d^{i+1} = r^{i+1} + \beta_i d^i$

$\delta_0 = \delta_1$

Calcule $ared$ e $pred$ usando (5)

SE $ared > \eta pred$

$x^{k+1} = x^k + d^k$

SE $\rho_k > \dfrac{3}{4}$ e $\left\|d^k\right\| = \Delta_k$ então $\Delta_{k+1} = 2\Delta_k$

SENÃO $\Delta_{k+1} = \Delta_k$

SENÃO

$x^{k+1} = x^k$

$\Delta_{k+1} = \dfrac{\Delta_k}{2}$

$k = k + 1$

5 Implementação do algoritmo

O GCMRC foi implementado em Scilab, para minimizar a função de Beale através do método de região de confiança, resolvendo o subproblema com a estratégia de gradientes conjugados.

Seja a função de Beale $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ dada por $f(x_1, x_2) = T_1^2 + T_2^2 + T_3^2$, sendo $T_1 = 1,5E_0 - x_1S_1$, $T_2 = 2,25E_0 - x_1S_2$, $T_3 = 2,625E_0 - x_1S_3$, $S_1 = (1,0E_0 - x_2)$, $S_2 = \left(1,0E_0 - x_2^2\right)$ e $S_3 = \left(1,0E_0 - x_2^3\right)$. Este exemplo é um clássico na literatura e a solução é $(3; 0,5)$. O gradiente é

$$\nabla f(x_1, x_2) = \begin{pmatrix} 2T_1(-x_1 + x_2) + 2T_2\left(-x_1 + x_2^2\right) + 2T_3\left(-x_1 + x_2^3\right) \\ 2T_1(x_1) + 2T_2(2x_1x_2) + 2T_3\left(3x_1x_2^2\right) \end{pmatrix},$$

e a Hessiana é dada por

$$\nabla^2 f(x_1, x_2) = \begin{pmatrix} \nabla^2 f(x_1, x_2)_{11} & \nabla^2 f(x_1, x_2)_{12} \\ \nabla^2 f(x_1, x_2)_{21} & \nabla^2 f(x_1, x_2)_{22} \end{pmatrix},$$

com

$$\begin{aligned}
\nabla^2 f(x_1, x_2)_{11} &= 2E_0\left((S_1)^2 + (S_2)^2 + (S_3)^2\right) \\
\nabla^2 f(x_1, x_2)_{12} &= 2E_0(T_1 - x_1S_1) + 4E_0x_2(T_2 - x_1(S_2)) + 6E_0(T_3 - x_1S_3)x_2^2 \\
\nabla^2 f(x_1, x_2)_{21} &= 2E_0(T_1 - x_1S_1) + 4E_0x_2(T_2 - x_1(S_2)) + 6E_0(T_3 - x_1S_3)x_2^2 \\
\nabla^2 f(x_1, x_2)_{22} &= 2E_0x_1\left(x_1 + 2E_0T_2 + 4E_0x_1x_2^2 + 6E_0T_3x_2 + 9E_0x_1x_2^4\right).
\end{aligned}$$

Sendo $x^0 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\Delta_0 = 1$, $\eta = 0,2$ e $\epsilon = 10^{-6}$, o algoritmo executa 251 iterações,

encontrando

$$x^{251} = \begin{pmatrix} 2,999967 \\ 0,499992 \end{pmatrix} \text{ e } f\left(x^{251}\right) = 0.$$

Os pontos encontrados a cada iteração estão ilustrados na figura abaixo.

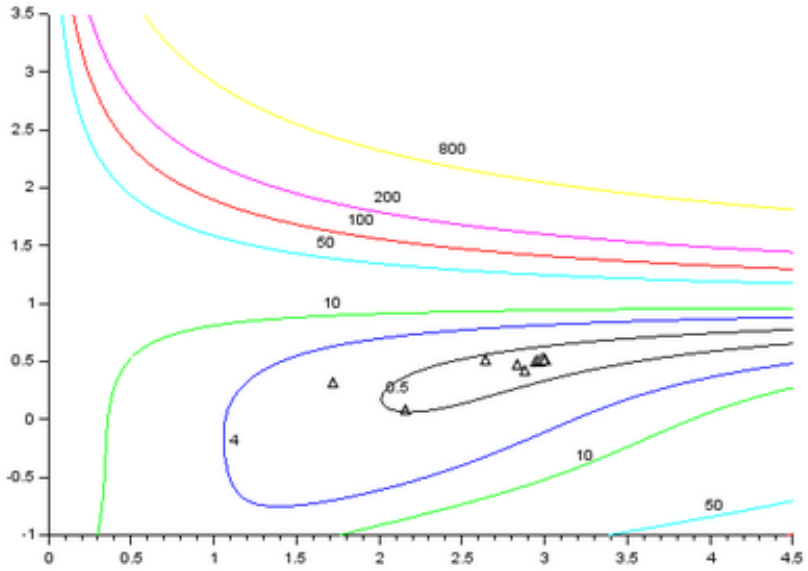


Figura 1: Curvas de nível da função de Beale com os pontos obtidos a cada iteração

Os resultados obtidos foram comparados pelos obtidos por Gardenghi e Santos (2012), que utilizaram o CAS *Maxima* para minimizar a função de Beale. A semelhança entre os resultados obtidos nos faz ver que o Algoritmo GCMRC converge e que sua implementação foi boa, já que o número de iterações realizadas foi bem próximo ao número de iterações necessários no trabalho de Gardenghi e Santos (2012).

Referências

GARDENGHI, John Lenon C.; SANTOS, Sandra Augusta. **Minimização irrestrita usando gradientes conjugados e regiões de confiança.** Campinas, 2012.

RIBEIRO, Ademir Alves; KARAS, Elizabeth Wegner. **Otimização Contínua:** Aspectos Teóricos e Computacionais. São Paulo: Cengage Learning, 2013.

SCILAB. 2015. Disponível em: www.scilab.org. Acesso em: 27 abr. 2015.

A matemática na educação infantil e nos anos iniciais

Fernanda Tomazi
Universidade estadual do oeste do Paraná / campus Cascavel
Fernandatomazi06@gmail.com

Jaqueline Zdebski da Silva Cruz
Universidade estadual do oeste do Paraná / campus Cascavel
jaque_zdebski@hotmail.com

Luiz Augusto Peregrino
Universidade estadual do oeste do Paraná / campus Cascavel
Luiz_Peregrino@hotmail.com

Resumo: Durante a disciplina de Tópico de Ensino de Matemática na Educação Infantil e Anos Iniciais, cursado no ano de 2015 na Unioeste/Cascavel, com o objetivo de conhecer tanto o desenvolvimento da criança, quanto os métodos utilizados para estimulá-la a aprender matemática. O trabalho foi fundamentado sobre a metodologia do construtivismo. Foram realizados estudos e pesquisas nos currículos da educação infantil de Cascavel e da AMOP, feitas discussões sobre leituras de artigos e livros, uma entrevista com professora dos anos iniciais, além de desenvolvimento de atividade na Rede Municipal de Educação. Com os resultados obtidos, foi desenvolvido o artigo que apresenta de forma geral os conceitos que devem ser trabalhados durante os primeiros anos de escolarização da criança e cuidados que o educador deve ter durante o processo de ensino-aprendizado, assim como as especificidades da criança desta faixa etária. Contém também o relato de uma atividade ministrada na rede pública de educação, colocando em prática os conhecimentos obtidos, relacionando a teoria com a prática educativa.

Palavras-chave: Matemática; educação infantil; anos iniciais.

1 Introdução

Todos estão sujeitos a ações educacionais seja, em casa, na rua, na igreja ou nas escolas, de muitos modos utilizamos princípios dela em momentos da nossa vida. Não há uma forma ou um modelo único de educação. Ela existe no imaginário e na ideologia dos grupos sociais e a sua missão é transformar sujeitos capazes de viver e compartilhar o ambiente em que vivem.

Alguns autores defendem que a Matemática é uma ciência que tem a potencialidade de gerar sujeitos com capacidade crítica, e por isso devemos ter certeza de que essa ciência está sendo aprendida com essa finalidade. Assim, surge a necessidade de transmitir esse conhecimento às futuras gerações e a escola é onde o conhecimento construído historicamente pelo homem é formalizado, caracterizando-se como saber científico. Para que esse objetivo de construir o saber científico seja alcançado, a escola sofre mudanças constantes em seus métodos de ensino. Sendo

assim independentemente da idade dos estudantes, os conceitos matemáticos devem ser construídos pelo próprio sujeito e não ensinados, já que quando são ensinados apenas memorizam, sem que atribuam um real significado a eles. Para que o aluno construa o próprio conhecimento, o papel do professor é proporcionar situações que possibilitem essa construção. É essa ideia de "ensinar" que torna a Matemática uma disciplina complicada para a maioria das pessoas, pois o que aprendem são apenas amontoados de fórmulas e procedimentos que são utilizados em situações abstratas e desvinculados da vida real, procedimentos que aparentemente só resolvem problemas dentro da própria Matemática, fazendo com que os estudantes a vejam como uma ciência sem utilidade social e sem envolvimento com as demais áreas do conhecimento. Esses problemas com a matemática podem ser vistos em todos os níveis de escolaridade, mas ficam evidentes a partir dos anos finais do ensino fundamental, quando o aluno tem um professor com formação específica em matemática, e que compreende a Matemática de um ponto de vista diferente dos professores dos anos iniciais.

Porém as lacunas na construção do conhecimento matemático podem começar muito antes das primeiras incursões nos anos iniciais, isso se deve a muitos fatores, mas principalmente a visão social distorcida que se tem da matemática.

2 A importância na Matemática e considerações sobre seu ensino na infância.

A infância é um período de curiosidades, de exploração e de descoberta. A criança que tem possibilidades de contato com brinquedos, jogos de montar, quebra-cabeça, jogo da memória, dentre outros, tem ao brincar, um pensamento em ação, favorecendo o estabelecimento de relações cada vez mais complexas. Como não "sabe" contar, ela precisa, inicialmente, construir noções de "bastante, nada, muito, pouco, igual, mais, menos, maior, menor", entre outros significados que são construídos a partir das comparações que estabelece. Essas comparações também contribuem para a construção do conhecimento lógico-matemático. Assim, quanto mais o educador e o meio oportunizarem ações e recursos que possibilitem maiores serão as oportunidades de desenvolvimento.

A importância da Matemática é reconhecida como essencial ao desenvolvimento do sujeito, desde a mais tenra idade, como é descrito no currículo da educação infantil de Cascavel: "A Matemática, como parte do conjunto de conhecimentos científicos, constitui-se como uma disciplina fundamental, indispensável na construção dos currículos escolares" (CASCAVEL, 2008). E, além disso, os currículos descrevem que a Matemática deve ser mostrada ao aluno como uma

necessidade humana, construída ao longo da história, e ainda deve-se dar a esta ciência um caráter sociocultural, que a aproxime da realidade dos alunos.

Mas durante o processo de ensino da Matemática para crianças, tanto na educação infantil quanto nos anos iniciais, deve-se levar em consideração que elas ainda estão em desenvolvimento e possuem especificidades que devem ser consideradas para que realmente exista essa proximidade da ciência Matemática com a realidade do aluno.

Sendo assim é importante considerar que entre os 2 e 7 anos a criança está desenvolvendo a linguagem e construindo seu raciocínio. Por esse motivo, costumam pedir o nome dos objetos e suas funções, além dos rotineiros "porquês?", típicos da idade. Esse comportamento deve ser incentivado ao máximo, para que a criança tenha um desenvolvimento pleno, com seu raciocínio lógico em construção apresenta ideias equivocadas e constantemente se contradiz. A criança é extremamente criativa e gosta de brincar, em geral são muito ativas e dominam os movimentos do corpo, porém sua concentração nas atividades é de curta duração, além disso, é egocêntrica e por esse motivo pode apresentar dificuldade em interagir com grupos. Também é comum que tenham um ou dois amigos com quem interajam com mais facilidade geralmente do mesmo sexo.

De acordo com Piaget, a criança adquire o conhecimento ao construí-lo a partir de seu interior, em vez de internalizá-lo diretamente de seu meio ambiente, elas podem pensar sobre o conhecimento ensinado por um momento, mas elas não conseguem reter o que e entronado em suas cabeças, a criança em fase de desenvolvimento escolar aprende modificando as velhas ideias, incorporando a uma nova, fazendo assim uma ponte entre o conhecimento passado e presente. Para que o aprendizado realmente aconteça o professor deve utilizar essas características próprias da criança a seu favor e atividades extensas devem ser evitadas. Deixar os alunos exporem suas ideias, principalmente durante atividades lúdicas e incentivá-los a propor soluções para os problemas que surgem ao decorrer das atividades é uma forma de estimular o raciocínio deles.

Além disso, o educador deve ter em mente que a Matemática não se reduz a manipulações numéricas, ela envolve um raciocínio esquematizado, percepção de relações, domínio e compreensão do espaço, entre outras características que segundo Lorenzato (2007) são mais facilmente desenvolvidas, mesmo antes da criança ser capaz de contar objetos, o que em muitas ocasiões, de forma errônea, é compreendido como o primeiro passo para utilizar a Matemática.

O ensino da matemática contempla situações que estimulam o desenvolvimento de sete processos mentais, que são fundamentais para a concepção do sistema de numeração e posteriormente outros conteúdos matemáticos, pois elas facilitam a compreensão dos conceitos básicos e das relações numéricas. Sendo assim é importante que o professor proponha situações diversas

nas quais os alunos possam desenvolver esse tipo de raciocínio. Segundo Lorenzato (2007) esses processos são:

A correspondência é a capacidade de relacionar, por exemplo, para cada aluno uma mochila, para cada criança um, nenhum ou muitos irmãos, posteriormente que para cada quantidade existe um símbolo.

A comparação permite estabelecer relações como maior, menor, igual, diferente, semelhante.

A classificação consiste em organizar objetos através de características comuns ou diferenças dos mesmos.

Na sequenciação estabelece-se uma ordem onde um objeto sucede o outro, sem considerar características próprias do objeto. É como ordenar os alunos em ordem de chegada, nesse ato não se considera, por exemplo, a altura dos mesmos, mas sim um fator externo as características dos alunos.

Já na seriação a disposição dos elementos é preestabelecida, e em geral estão relacionadas com as diferenças entre eles.

A inclusão é a ideia que um objeto ou conjunto está contido em outro maior que o primeiro.

A conservação é a concepção de que a disposição espacial do objeto não altera características como peso, tamanho e quantidade.

Além de indicados por vários autores como a melhor forma de iniciar o estudo da Matemática, esses conteúdos também são indicados pelos currículos como conteúdos a serem trabalhados desde o maternal. Porém esses conteúdos não devem ser trabalhados dissociados de contextos, da mesma forma que a Matemática não precisa ser trabalhada como uma disciplina isolada, pois nessa fase as disciplinas são integradas.

3 Entrevista com professora de anos iniciais

Durante a disciplina de Tópico de Ensino de Matemática na Educação Infantil e Anos Iniciais, houve um momento em que se realizou entrevista com uma professora da educação infantil. O objetivo era analisar e investigar se os objetivos do currículo infantil de Cascavel estavam sendo desenvolvidos e alcançados, além de conhecer características da prática educativa dos anos iniciais. Nessa entrevista, percebeu-se como a interdisciplinaridade é presente nessa

prática, pois a professora apresentou extrema dificuldade em nos falar apenas sobre os aspectos relacionados à forma como trabalha os conteúdos de Matemática que estão previstos no currículo. Temos o melhor exemplo disso considerando o trabalho que ela descreveu para o qual utiliza uma história infantil para trabalhar português, simultaneamente conta e ordena os personagens, para trabalhar Matemática e utiliza-se de outros aspectos para trabalhar as demais disciplinas.

Porém ao decorrer da entrevista foi possível constatar que a professora compreende a Matemática como basicamente aplicação de algoritmos, ou como visto no exemplo citado a realização de contagem. Em momento algum falou sobre construção do conceito de número, mas sim sobre contagem de objetos, realização de cálculos. Quando questionada sobre como as crianças reagiam a problemas sem palavras que pudessem ser relacionadas diretamente com algoritmos, tais como: ganhar, perder, pagar, entre outras, ela falou que essas palavras devem obrigatoriamente aparecer nos problemas.

Sobre as dificuldades no ensino-aprendizagem de Matemática não apresentou um conteúdo que fosse difícil de ser compreendido pelos alunos. Segundo a professora a maior dificuldade dos alunos é na disciplina de português.

Porém se pensarmos na forma como a professora compreende a matemática os alunos dificilmente terão dificuldades, pois eles apenas deverão repetir procedimentos apresentados pela professora, raramente tendo que refletir sobre eles. Sendo assim realmente não há dificuldades, pois, os alunos simplesmente decoram os conteúdos. O que a afasta do ensino pensado por Fiorentini (1995, p. 31):

> A Matemática, sob uma visão histórico-crítica, não pode ser concebida como um saber pronto e acabado, mas, ao contrário, como um saber vivo, dinâmico e que, historicamente, vem sendo construído, atendendo a estímulos externos (necessidades sociais) e internos (necessidades teóricas de ampliação dos conceitos).

A proposta da professora se afasta dessa ideia, pois, em geral, esses problemas por ela citados não são interessantes aos alunos, já que narram situações que estão distantes da realidade dos mesmos, sendo assim não tornam visível à utilização da Matemática no cotidiano, dando aos alunos a falsa impressão de que não precisam dela. Mesmo que nessa fase essa ideia não seja tão clara para os alunos, com o passar do tempo à distância entre a matemática aprendida na escola e a utilizada por ele durante situações cotidianas fica cada vez mais evidente.

4 Relato de aplicação em sala de aula

A aula foi ministrada na rede pública de ensino do município de Cascavel, na sala de pré-escola, com dezessete alunos de aproximadamente cinco anos. O objetivo da aula aplicada é ampliar os conceitos de comparação e correspondência, a partir da história dos números. O conteúdo foi trabalhado a partir do cordel: "Contando a história dos números" de Ana Raquel Campos. Segundo Smole, Rocha, Cândido e Stancanelli (2007, p.2):

> Integrar literatura nas aulas de matemática representa uma substancial mudança no ensino tradicional da matemática, pois em atividades deste tipo, os alunos não aprendem primeiro a matemática para depois aplicar na história, mas exploram a matemática e a história ao mesmo tempo.

Antes mesmo de qualquer atividade feita com os alunos nota-se a curiosidade nos novos objetos e professores na sala, essa curiosidade deve ser usada como ferramenta de interiorização do objetivo da atividade, desde que trabalhada de forma correta.

A aula iniciou apresentando os ministrantes e o cenário utilizado para a contação da história, os cubos do material dourado, representando as pedras, e o saquinho.

Figura 1: Cenário

Os alunos sentaram-se em círculo no chão, durante leitura as ovelhas eram movimentas para fora e para do curral e as pedrinhas colocadas e tiradas do saquinho, como narra a história. E as após o término da leitura questionou-se sobre o que significava se sobrava uma pedra depois que as ovelhas entrassem no curral e o que o pastor deveria fazer, eles não responderam. Então

foi feita a simulação. Primeiramente tirando as ovelhas do curral e para cada uma colocando uma pedra no saquinho, depois se colocava cada ovelha no curral tirando uma "pedra" do saquinho para cada uma, e no processo escondeu-se uma ovelha, portanto sobrou uma "pedra", então os alunos responderam que uma ovelha estava faltando e o pastor tinha que procurá-la.

Em seguida foi proposto que cada aluno colorisse e montasse uma ovelha, para ser usada na atividade seguinte, onde novamente os alunos sentaram em círculo, então um aluno foi escolhido como pastor e para cada ovelha que os colegas colocassem no centro do círculo ele deveria colocar uma "pedra" no saquinho, logo após os alunos retiram, cada um sua ovelha, e juntamente retiraram uma "pedra" do saquinho, ao fim sobraram duas "pedras", cujas ovelhas correspondentes foram retiradas sem que os alunos notassem, novamente pediu-se aos alunos o que deveriam fazer, sugeriram que deviam procurá-las e assim o fizeram.

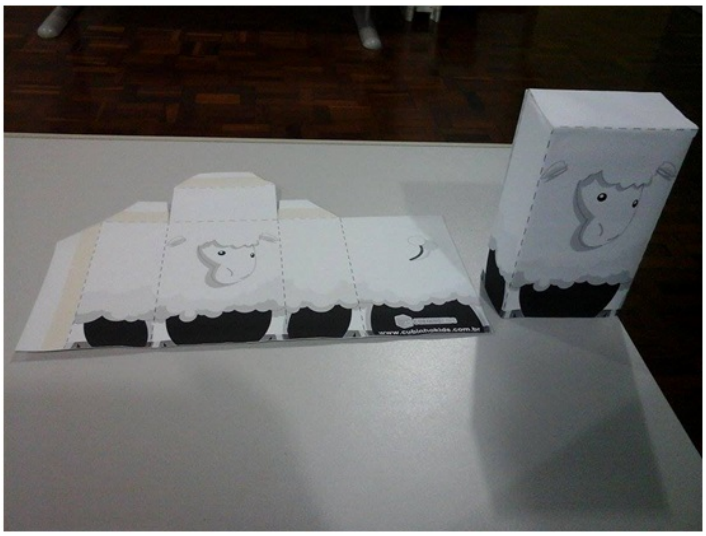


Figura 2: Modelo de ovelha montada pelos alunos

O tempo não foi suficiente para que se realizasse a outra atividade planejada, que envolveria o mesmo conceito de correspondência, juntamente com a ideia de mudança de base. Mas as atividades realizadas envolveram os alunos, e aparentemente eles compreenderam os conceitos envolvidos, sabendo aplicá-los nas situações que surgiram.

Nesse caso mesmo não envolvendo situações que os alunos vivem, foi criada uma atmosfera que os submeteu a uma situação que no momento foi interessante para eles, pois por um tempo fizeram parte da história, como descrito por Smole, Rocha, Cândido e Stancanelli (2007, p. 3):

> Sendo assim, através da conexão entre literatura e matemática, o professor pode criar situações na sala de aula que encorajem os alunos a compreenderem e se familiarizarem mais com a linguagem matemática, estabelecendo ligações cognitivas entre a linguagem materna, conceitos da vida real e a linguagem matemática formal, dando oportunidades para eles escreverem e falarem sobre o vocabulário matemático, além de desenvolverem habilidades de formulação e resolução de problemas enquanto desenvolvem noções e conceitos matemáticos.

5 Conclusão

A criança é dotada de capacidades desconhecidas, que podem levar a um futuro de sucesso e felicidade. Caso o objetivo seja mesmo uma escolarização adequada, devemos primeiramente desenvolver essas capacidades para que o desenvolvimento da criança não deva ser entregue ao acaso ou a métodos de ensino que não atinjam os objetivos propostos.

Devemos saber que não somos nós professores os construtores do conhecimento da criança e sim os colaboradores desse processo. Os currículos apresentam-se como um material base para conhecer os conteúdos e suas peculiaridades, técnicos como o professor Sergio Lorenzatto, contribuem na formação de professores, porém isso não é suficiente. É dever do professor conhecer processos de esses materiais e com eles produzir momentos didáticos. Sabemos que cada escola tem sua própria realidade, o que faz com que os conhecimentos necessários para os alunos de cada uma difiram. Não há como garantir o sucesso de uma aula planejada. Para isso um apoio ao professor bem como a necessidade de uma formação continuada pode contribuir para o sucesso da aprendizagem. Trabalhar com a educação infantil, fase da aprendizagem é importante, requer que o professor deva estar em dia com o desenvolvimento das propostas de ensino e atualizado profissionalmente.

Referências

CASCAVEL, AMOP. Associação dos Municípios do Oeste do Paraná. Departamento de Educação. Currículo Básico para a Escola Pública Municipal: Educação Infantil e Ensino Fundamental. Cascavel: ASSOESTE, 2010.

CASCAVEL (PR). Secretaria Municipal de Educação. Currículo para a rede pública municipal de ensino de Cascavel. Cascavel, PR: Ed. Progressiva, 2008.

LORENZATO, Sergio. **Coletânea formação de professores:** Educação Infantil e percepção matemática . 6ª ed. São paulo: Autores associados , 2007. 202 p.

ROSSINI, Maria Augusta Sanches. **Aprender:** tem que ser gostoso.... 2ª ed. Pretópolis: Vozes, 2003. 226 p.

Conteúdos de Matemática e o "desinteresse" dos alunos: reflexões sobre a nossa experiência

Camila Rodrigues dos Santos
Universidade Estadual do Oeste do Paraná, Cascavel, PR, Brasil
rds.camila@hotmail.com

Eliandra de Oliveira
Universidade Estadual do Oeste do Paraná, Cascavel, PR, Brasil
eliandra.oliveiraaa@gmail.com

Dulcyene Maria Ribeiro
Universidade Estadual do Oeste do Paraná, Cascavel, PR, Brasil
dulcyene.ribeiro@unioeste.br

Resumo: O texto "Conteúdos de Matemática e o "desinteresse" dos alunos: reflexões sobre a nossa experiência" apresenta ideias a serem discutidas com relação ao ensino da Matemática e o efeito de nossas práticas. Discutimos alguns fatores que influenciam na prática pedagógica e no ensino de determinados conteúdos considerados com pouca aplicação no cotidiano dos nossos alunos. A substituição do método tradicional – no qual o professor é o agente transmissor e o aluno o passivo – por práticas mais criativas e eficientes em termos da relação entre ensino e aprendizagem não é fácil. Então nunca devemos utilizar aulas tradicionais? Será que os alunos não aprendem nada desse modo? Todos os conteúdos podem ser abordados por uma metodologia diferenciada? Será que o desinteresse dos alunos é culpa do professor? Essas e mais perguntas nos deixaram preocupadas com relação ao que é realmente ensinar e se existe uma forma "correta" disso ser feito. Por isso, esse texto é composto mais de perguntas do que de respostas. São as nossas reflexões iniciais sobre a ação docente, sobre os papéis desempenhados pelos professores, em especial, o de matemática.

Palavras-chave: Educação Matemática; desinteresse dos alunos; prática escolar.

1 Introdução

Durante o estágio realizado nesse ano na disciplina de Metodologia e Prática de Ensino - Estágio Supervisionado II, ministramos aulas na turma do 2^o ano "A" durante 20 horas-aulas. As aulas aconteceram no período da manhã, na Escola Estadual Jardim de Santa Felicidade. Durante essa experiência tivemos várias oportunidades de refletirmos.

Estávamos diante de um conteúdo que é visto como tendo pouca aplicação prática no cotidiano dos alunos, determinantes. Com isso não tivemos oportunidade de ministrar aulas mais dinâmicas, envolvendo jogos, atividades práticas e nem um material manipulável para que esse estudo fosse de uma melhor forma concretizado.

Num primeiro momento, pensamos em trocar de turma. Estagiar, por exemplo, em uma turma de primeiro ano do Ensino Médio, seria mais fácil, pois de modo geral, em boa parte do ano, nessa série se é trabalhado com o conteúdo de funções, que em nossa opinião é bem mais fácil de adequar às diferentes metodologias, como resolução de problemas, o uso de mídias e teconologias, etc. Mas, quando se define uma turma para realizar o estágio, vários fatores são levados em consideração, como o horário das aulas, que precisa ser adequado para o orientador e para os estagiários, as conversas e os acordos estabelecidos com os professores e coordenadores da escola, etc. Isso quer dizer que não é o conteúdo o fator determinante.

Além disso, mesmo que durante o ano letivo vários conteúdos sejam ensinados, é estabelecido a priori, pelos professores da escola, uma ordem para eles, o que não nos permitiu trabalhar com sistemas sineares, antes de determinantes, nem escolher trabalhar com Análise Combinatória ou Trigonometria, por exemplo, conteúdos que a nosso ver favoreceriam a utilização de atividades práticas e a utilização de várias formas de encaminhamentos metodológicos.

Então durante o período do estágio começamos a sentir que nossas aulas não estavam sendo atraentes e não sentíamos que os alunos estavam realmente interessados nas aulas. Logo começamos a conversar com a nossa orientadora para analisar o que estávamos fazendo de errado e começamos a ver que as coisas não têm uma única resposta. Foi então que começamos uma reflexão a respeito do desinteresse dos alunos com a disciplina de Matemática e os métodos utilizados por nós professores de matemática e licenciados, em conteúdos que consideramos com pouca aplicação no dia a dia da maioria dos alunos da Educação Básica.

Por isso, esse texto é composto mais de perguntas do que de respostas. São as nossas reflexões iniciais sobre a ação docente, sobre os papeis desempenhados pelos professores, em especial, o de matemática, para as quais, em sua maioria, ainda não temos respostas.

2 Aulas tradicionais e o ensino de determinantes

Nas atividades do estágio aconteceram aulas expositivas e dialogadas, nas quais buscamos criar uma situação para expor o conceito pretendido utilizando exemplos ou definições. Além disso, os alunos resolviam exercícios, em sala, contando com a nossa ajuda, em explicações diretamente a eles nas carteiras e também no quadro, quando as explicações eram a todos eles. Esse tipo de aula (tradicional) pôde ser evidenciado durante a maioria das aulas, quando ensinamos determinantes e sistemas lineares.

Segundo Machado (2013), aula tradicional é uma aula em que o professor passa o tempo

todo - ou quase todo - expondo oralmente a matéria, cabendo ao aluno um papel passivo na relação entre ensino e aprendizagem. Sabemos que nem sempre esse método funciona, afinal o professor passa a maior parte do tempo falando e o aluno, passivamente, escutando ou fingindo escutar, o que não é nada satisfatório.

Mas nunca devemos utilizar aulas tradicionais? Será que os alunos não aprendem nada desse modo? E os conteúdos, todos podem ser abordados por meio de uma metodologia diferenciada? Essas e mais perguntas começaram a nos deixar preocupadas com relação ao que é realmente ensinar e se existe uma forma "correta" disso ser feito.

Determinantes é um conteúdo matemático diretamente relacionado ao estudo de matrizes. Por isso, é difícil encontrar alguma referência a eles sem que junto apareçam reflexões sobre o ensino de matrizes, como nas Diretrizes Curriculares Paranaenses, documento que orienta o que deve ser ensinado de Matemática no Estado do Paraná. Determinantes e matrizes, são possíveis de serem ensinados no Ensino Médio de modo que tenha sentido para os alunos?

> De fato, as matrizes são importantes não apenas como forma de organizar dados numéricos (tabelas), mas também pelas diversas propriedades das operações que definem a álgebra matricial. Todavia, na forma tradicional de ensinar matrizes e determinantes no ensino médio, como tabelas de números e um número associado a uma matriz, respectivamente, as definições das operações matriciais e determinantes, assim como as suas propriedades, tornam-se bastante artificiais. (JAHN, 2013, p. 4).

A autora sugere que o uso da Geometria pode ajudar a tornar estes conceitos mais naturais, como também desenvolver a intuição sobre o assunto. Justifica que os aspectos geométricos das matrizes e determinantes estão presentes em aplicações tecnológicas tais como em Computação Gráfica, por exemplo. Mas essas aplicações práticas estão distantes da realidade dos alunos do Ensino Médio, o que de certa forma, continua a não lhes fazer qualquer sentido.

Nas Diretrizes Curriculares da Educação Básica da Secretaria do Estado da Educação do Paraná (DCE), matrizes e determinantes são citados no conteúdo estruturante Números e Álgebra, que segundo as diretrizes, deve ser aprofundado no Ensino Médio:

> [...] de modo a ampliar o conhecimento e domínio deste conteúdo para que o aluno: conceitue e interprete matrizes e suas operações, conheça e domine o conceito e as soluções de problemas que se realizam por meio de determinantes. (PARANÁ, 2008, p.52).

No texto dos Parâmetros Curriculares Nacionais do Ensino Médio (PCN 1997), não há uma menção direta do conteúdo das matrizes e determinantes. Seria a intenção de eliminar esse conteúdo do currículo?

Há quem defenda que o aluno se sente motivado a aprender quando ele percebe que o conteúdo a ser ensinado é necessário para a resolução de uma situação problema. Há autores

que defendem que a situação problema precisa ser algo da realidade do aluno. Outros dirão que esses problemas podem ser da própria Matemática. Nesse caso, uma das aplicações dos determinantes, seria em possibilitar o cálculo da área de uma região triangular. Mas isso seria uma motivação para os alunos?

3 O desinteresse e algumas estratégias para o professor

Um fator observado por nós estagiárias foi o fato dos alunos não demonstrarem muito interesse com o estudo da disciplina de Matemática. Num primeiro momento atribuímos isso somente às nossas aulas, mas pudemos verificar que não eram só com elas. As suas atitudes já aconteciam quando ainda não éramos as responsáveis pelas aulas. A maioria dos alunos não se importavam com as tarefas e faltavam bastante às aulas, inclusive em dias de avaliações. No dia da avaliação que realizamos, dos 34 alunos da turma, apenas 14 compareceram à escola e realizaram a avaliação.

Os PCNs (BRASIL, 1997) enfatizam que a Matemática costuma provocar duas sensações contraditórias, tanto por parte de quem ensina como por parte de quem aprende: de um lado, a comprovação de que se trata de uma área de conhecimento importante; de outro, a insatisfação diante dos resultados negativos obtidos com muita frequência em relação à sua aprendizagem.

De acordo com o censo comum, sempre escutamos que uma das grandes dificuldades na educação do nosso país é o desinteresse por parte de muitos alunos, por qualquer atividade escolar. Alunos que frequentam às aulas por obrigação, sem, contudo, se preocupar com as atividades básicas, como tarefas e avaliações.

Portanto, perante este problema, o que restaria ao professor, nos casos em que o aluno apresenta-se desinteressado quanto ao saber? Ao refletirmos nessa pergunta começamos a analisar algumas coisas que ocorrem durantes às aulas de Matemática e recordamos de algumas falas dos alunos. Na sequência, estão algumas delas: “A Matemática é importante para o futuro, para conseguir um bom emprego”; “A Matemática ajuda a raciocinar”; “Nós precisamos da matemática no nosso dia-a-dia”; “Tudo é Matemática”; “Eu gosto de Matemática porque eu entendo”; “As aulas de Matemática são chatas”; “Aulas de Matemática são cansativas”; “Quando eu gosto do professor eu gosto da matéria”; “Não sei pra quê tantas aulas de Matemática na semana”. Por essas frases, é possível observarmos diversos pontos de vista com relação à Matemática que é ensinada na escola.

Pezzini e Szymanski (s/d) citando Kupfer (1995, p. 79), afirmam que “o processo de

aprendizagem depende da razão que motiva a busca de conhecimento". Para as autoras:

> Os alunos precisam ser provocados, para que sintam a necessidade de aprender, e não os professores "despejarem" sobre suas cabeças noções que, aparentemente, não lhes dizem respeito. A forma de apresentar o conteúdo, portanto, pode agir em sentido contrário, provocando a falta de desejo de aprender que seria, para os alunos, o distanciamento que se coloca entre o conteúdo e a realidade de suas vidas. Quando o aluno não percebe de que modo o conhecimento poderá ajudá-lo, como desejará algo que lhe parece inútil? (PEZZINI; SZYMANSKI, s/d., p.2).

A Matemática é, de certo modo, rigorosa em suas aplicações e demonstrações, ela precisa chegar bem próxima da exatidão para dar confiabilidade ao fenômeno estudado e isso com certeza dificulta a ideia de provocar a necessidade de aprendê-la para a maioria dos alunos. Talvez por ser tão rígida provoca certo medo aos alunos que a acham difícil criando assim uma relação ríspida, às vezes, até traumática que pode resultar em dificuldade, falta de interesse ou a rejeição. Isso influencia o processo de ensinar.

> Sem dúvida, ensinar é algo muito difícil e trabalhoso. E mais difícil se torna quando as condições atrapalham. Mas é preciso que [...] o exercício de ensinar permaneça vinculado ao intento de promover as condições necessárias para, transcendendo o instruir e o adestrar, auxiliar o encontro da inteligência do educando com a vida, o encontro de sua sensibilidade com a pluralidade rica do viver. (MORAIS, 1986, p. 6).

Entendemos que é preciso adotar métodos que nos auxiliem a mudar o ponto de vista dos alunos com relação à Matemática. Alguns deles estão relacionados à postura e atitudes do professor e não diretamente a alguma metodologia específica.

Sabemos que os alunos gostam de professores que explicam bem a matéria, que os tratam bem e com respeito. Alguns alunos gostam do atendimento individual (na carteira) dado por alguns professores. Com esse atendimento individual o aluno cria a coragem necessária para fazer perguntas e tirar dúvidas, que ele não faria em público, por medo de perguntar (PEZZINI; SZYMANSKI). Isso deve ser aproveitado pelos professores.

Segundo Freire e Faundez, (1985, p. 46) a pergunta é o início da aprendizagem. "[...] o que o professor deveria ensinar [...] seria, antes de tudo, ensinar a perguntar. Porque o início do conhecimento, repito, é perguntar. E somente a partir de perguntas é que se deve sair em busca de respostas." Sendo assim, é necessário que se incentive o aluno a fazer perguntas e, deste modo, podemos tentar sanar algumas dúvidas e preencher algumas lacunas que acabam ficando no processo de ensino e aprendizagem.

De todo modo, precisamos ver os nossos erros e saber que boa parte do desinteresse dos alunos pode ser reflexo de aulas desconexas com a realidade, de conteúdos sem ligações práticas ao dia a dia, que muitas vezes proporcionamos aos alunos.

Lógico que mudar isso não é uma tarefa fácil e que isso nem sempre é possível atender (como ocorrido durante nossa regência). Porém oportunidades de aplicações matemáticas podem surgir naturalmente de observações de situações e ambientes que envolvem o cotidiano de todas as pessoas. Nós como professores e licenciados atentos devemos reunir condições básicas de estabelecer relações e desenvolver aplicações, afinal é nossa função apresentar na escola uma Matemática como uma ferramenta útil e por mais que a tarefa seja ardua não podemos desisitir.

Talvez num primeiro momento, essas aplicações pareçam difíceis de serem percebidas, mas a prática continuada levará ao nosso aperfeiçoamento e das atividades propostas. Acreditamos que isso leve à diminuição do desinteresse, uma vez que estabelecida as relações currículo/realidade, logo os alunos tendem a compreender a importância da Matemática aceitando-a mais facilmente, mesmo que continuem não "gostando" da disciplina.

4 Mais algumas reflexões

Consideramos que os alunos percebem que, nem sempre, seus mestres agem como esperado e que isso interfere em seu aprendizado. Muitas vezes não demonstramos aos nossos alunos o interesse em ensinar, o interesse de estar ali junto com eles aprendendo coisas novas e crescendo.

Com isso aprendemos que o desinteresse dos alunos nem sempre é "culpa" do conteúdo a ser ensinado ou mesmo da própria Matemática, sabemos que nem todos os conteúdos nos permite criar aulas diferenciadas (como determinantes), porém devemos fazer o possível para que os alunos estejam em contato com a matemática/realidade.

Num primeiro momento achávamos que o desinteresse dos alunos era devido ao conteúdo, mas pudemos notar que haviam outros fatores envolvidos nessa problemática e que isso ocorre em outras disciplinas também. Nossas aulas não tiveram muito sucesso -digamos não atraiu os nossos alunos - não somente por conta da metodologia utilizada ou mesmo do conteúdo a ser ensinado. Contudo em alguns casos esse desinteresse pode ser o resultado de anos em contato com o ensino da matemática como algo sem significado e sem aplicação.

Percebemos que ser educador não é somente expor conhecimentos, não é apenas passar informações, ensinar tabuadas, expor definições, apresentar o mundo sem significados.

Segundo Sanches ser educador é ter o sabor cuidadoso do afeto, é construir um ambiente de discussões, de curiosidade e reflexão sobre o mundo, é abrir espaço para o criar, fazer e refazer, é aprender a pensar, para aprender a ser o aprendiz/educador que ainda não somos.

Sanches (2005) observa:

> Para desvelar e recuperar a escola é preciso [...] buscar respostas nas teorias, nas pesquisas, nas experiências bem sucedidas, [...] para ajudar na construção de caminhos para a superação, transformação e construção da qualidade educacional. (SANCHES, 2005, p.81).

Aprendemos que devemos estar em constante reflexão acerca de nossa prática pedagógica e assim melhorar a cada dia e estar em constante processo de desconstrução e construção da prática educativa, das teorias e saberes que fundamentam o trabalho educativo.

5 Agradecimentos

Gostaríamos de dirigir os nossos sinceros agradecimentos à orientação da professora Dulcyene Maria Ribeiro que nos incentivou durante a elaboração do presente relato. Obrigada por insistir no nosso crescimento, foi um privilégio sermos suas orientandas. Prestamos também o nosso agradecimento à professora Arleni Elise Sella Langer, por se mostrar disponível em nos aconselhar e nos ajudar a ter forças diante das dificuldades, não apenas, nesta fase do estágio, mas também durante a Licenciatura. Obrigada por todos os conselhos. E ainda agradecemos ao corpo docente e não docente da Escola Estadual Jardim de Santa Felicidade.

Referências

BRASIL. Secretaria de Educação. **Parâmetros Curriculares Nacionais para o Ensino Médio**. Brasília: MEC/SEF, 1997. Disponível em: <http://portal.mec.gov.br/seb/arquivos/pdf/ciencian.pdf>. Acesso em: 01 ago. 2016.

FREIRE, Paulo; FAUNDEZ, Antonio. **Por uma pedagogia da pergunta**. Rio de Janeiro: Paz e Terra, 1985.

JAHN, Marta Lena. A geometria de matrizes e determinantes. Dissertação (Mestrado Profissional em Matemática). Setor de Ciências Exatas e Naturais. UEPG. Ponta Grossa, 2013.

MACHADO, Luiz Alberto. **Coragem: da aula tradicional à aula criativa**. 2013. Disponível em: <http://www.fbcriativo.org.br/pt/site/publicacoes/artigos/criatividade/> Acesso em: 01 ago. 2016.

MORAIS, Regis de. **O que é Ensinar?** São Paulo: EPU, 1986.

PARANÁ. Secretaria de Estado da Educação. **Diretrizes Curriculares da Educação Básica - Matemática**. Curitiba: SEED/DEB. 2008.

PEZZINI, Clenilda Cazarin; SZYMANSKI. Maria Lidia Sica. **Falta de desejo de aprender:** causas e consequências. s/d. Disponível em: <http://www.diaadiaeducacao.pr.gov.br/portals/pde/arquivos/853-2.pdf> Acesso em: 01 ago 2016.

SANCHES, Cláudio Castro. **Descontruir construindo um caminho para uma nova escola:** recuperação da escola - pensar o pensado. Petrópolis: Vozes, 2005.

A Transformada de Laplace e as equações de Bessel e Legendre

Rodrigo Luiz Langaro[1]
Universidade Estadual do Oeste do Paraná - UNIOESTE
rodrigollangaro@hotmail.com

Sandro Marcos Guzzo
Universidade Estadual do Oeste do Paraná - UNIOESTE
smguzzo@gmail.com

Resumo: Neste trabalho realizamos estudos acerca de equações diferenciais, especialmente sobre o uso da Transformada de Laplace na resolução de equações diferenciais com coeficientes não constantes. Direcionamos nosso foco para duas equações específicas, a equação de Bessel, em que obtivemos sucesso a partir da transformada e apresentemos a solução, e a equação de Legendre, em que houve êxito a partir da técnica desenvolvida por Pierre Simon Laplace.

Palavras-chave: Transformada de Laplace; Transformada de Laplace Inversa; Equação de Bessel.

1 A Transformada de Laplace

A Transformada de Laplace, ferramenta muito útil na resolução de equações diferenciais, é o foco principal de nosso trabalho. Desenvolvida pelo matemático francês Pierre Simon Laplace há aproximadamente 200 anos, esse método batizado com o sobrenome de Pierre nada mais é do que uma transformada integral com algumas propriedades interessantíssimas que auxiliam na resolução de equações diferenciais.

Definição 1. Seja uma função f definida para $t \geq 0$. A integral

$$\mathcal{L}\{f(t)\}(s) = F(s) = \int_0^\infty e^{-st} f(t)dt$$

será chamada de Trasnformada de Laplace de f, desde que a integral convirja.

Uma de suas principais propriedades é a linearidade pois

$$\int_0^\infty e^{-st}[\alpha f(t) + \beta g(t)]dt = \alpha \int_0^\infty e^{-st} f(t)dt + \beta \int_0^\infty e^{-st} g(t)dt,$$

segue que

$$\mathcal{L}\{\alpha f(t) + \beta g(t)\} = \alpha \mathcal{L}\{f(t)\} + \beta \mathcal{L}\{g(t)\}$$

Dessa forma podemos determinar a função $f(t)$ tal que $\mathcal{L}\{f(t)\} = F(s)$ através da Transformada de Laplace Inversa, outra transformação linear útil na resolução de equações diferenciais.

[1]Rodrigo Luiz Langaro foi bolsista de Iniciação Científica da Fundação Araucária.

Definição 2. Se $F(s)$ representa a Transformada de Laplace de uma função $f(t)$, isto é, $\mathcal{L}\{f(t)\} = F(s)$, dizemos então que $f(t)$ é a Transformada de Laplace Inversa de $F(s)$ e escrevemos $f(t) = \mathcal{L}^{-1}\{F(s)\}(t)$.

Tanto para a Transformada de Laplace quanto para a Transformada Inversa de Laplace apresentamos uma tabela de Transformadas e Transformadas Inversas das funções básicas, que segue abaixo.

Tabela de Transformadas

$$(a)\mathcal{L}\{1\} = \frac{1}{s}$$
$$(b)\mathcal{L}\{t^n\} = \frac{n!}{s^{n+1}}, n = 1, 2, 3, \cdots$$
$$(c)\mathcal{L}\{e^{at}\} = \frac{1}{s-a}$$
$$(d)\mathcal{L}\{\operatorname{sen} kt\} = \frac{k}{s^2+k^2}$$
$$(e)\mathcal{L}\{\cos kt\} = \frac{s}{s^2+k^2}$$
$$(f)\mathcal{L}\{\operatorname{senh} kt\} = \frac{k}{s^2-k^2}$$
$$(g)\mathcal{L}\{\cosh kt\} = \frac{s}{s^2-k^2}$$

Tabela de Transformadas Inversas

$$(a)1 = \mathcal{L}^{-1}\left\{\frac{1}{s}\right\}$$
$$(b)t^n = \mathcal{L}^{-1}\left\{\frac{n!}{s^{n+1}}\right\}$$
$$(c)e^{at} = \mathcal{L}^{-1}\left\{\frac{1}{s-a}\right\}$$
$$(d)\operatorname{sen} kt = \mathcal{L}^{-1}\left\{\frac{k}{s^2+k^2}\right\}$$
$$(e)\cos kt = \mathcal{L}^{-1}\left\{\frac{s}{s^2+k^2}\right\}$$
$$(f)\operatorname{senh} kt = \mathcal{L}^{-1}\left\{\frac{k}{s^2-k^2}\right\}$$
$$(g)\cosh kt = \mathcal{L}^{-1}\left\{\frac{s}{s^2-k^2}\right\}$$

Nosso objetivo é usar a Transformada de Laplace na resolução de equações diferenciais. Para isso, apresentamos dois resultados que podem ser obtidos através de uma construção recursiva. O primeiro trata-se de uma potência n de t, em que a Trasnformada de Laplace de um

produto de uma função t pode ser obtida diferenciando-se a Transformada de Laplace de $f(t)$. Se $F(s) = \mathcal{L}\{f(t)\}$, então

$$\begin{aligned}\frac{d}{ds}F(s) &= \frac{d}{ds}\int_0^\infty e^{-st}f(t)dt \\ &= \int_0^\infty \frac{\partial}{\partial s}[e^{-st}f(t)]dt \\ &= -\int_0^\infty e^{-st}tf(t)dt = -\mathcal{L}\{tf(t)\},\end{aligned}$$

isto é

$$\mathcal{L}\{tf(t)\} = -\frac{d}{ds}\mathcal{L}\{f(t)\}.$$

Em decorrência disso, temos

$$\mathcal{L}\{t^2f(t)\} = \mathcal{L}\{t.tf(t)\} = -\frac{d}{ds}\mathcal{L}\{tf(t)\} = -\frac{d}{ds}\left(-\frac{d}{ds}\mathcal{L}\{f(t)\}\right) = \frac{d^2}{ds^2}\mathcal{L}\{f(t)\}.$$

Dessa forma, podemos sugerir:

Teorema 3. *Se $F(s) = \mathcal{L}\{f(t)\}$ e $n = 1, 2, 3, \cdots$, então*

$$\mathcal{L}\{t^nf(t)\} = (-1)^n\frac{d^n}{ds^n}F(s).$$

O segundo resultado é de suma importância para nossos estudos, pois envolve a transformada da derivada de ordem n, algo natural para quem busca resolver equações diferenciais utilizando a Transformada de Laplace. Começando pela derivada de ordem 1 e integrando por partes, temos

$$\begin{aligned}\mathcal{L}\{f'(t)\} &= \int_0^\infty e^{-st}f'(t)dt \\ &= e^{-st}f(t)|_0^\infty + s\int_0^\infty e^{-st}f(t)dt \\ &= -f(0) + s\mathcal{L}\{f(t)\} = sF(s) - f(0).\end{aligned}$$

Levando em consideração que $e^{-st}f(t) \to 0$ quando $t \to \infty$ (desde que $s > 0$) e com a ajuda do que acabamos de concluir podemos encontrar uma expressão para Transformada de uma derivada segunda, sendo

$$\begin{aligned}\mathcal{L}\{f''(t)\} &= \int_0^\infty e^{-st}f''(t)dt \\ &= e^{-st}f'(t)|_0^\infty + s\int_0^\infty e^{-st}f'(t)dt \\ &= s\mathcal{L}\{f'(t)\} - f'(0) \\ &= s[sF(s) - f(0)] - f'(0),\end{aligned}$$

ou seja

$$\mathcal{L}\{f''(t)\} = s^2F(s) - sf(0) - f'(0).$$

De forma análogam podemos mostrar que

$$\mathcal{L}\{f'''(t)\} = s^3F(s) - s^2f(0) - sf'(0) - f''(0),$$

o que nos leva a sugerir que

Teorema 4. *Se $f, f', \cdots, f^{(n-1)}$ forem contínuas em $[0,\infty)$ e de ordem exponencial, e se $f^{(n)}(t)$ for contínua por partes em $[0,\infty)$, então*

$$\mathcal{L}\{f^{(n)}(t)\} = s^nF(s) - s^{n-1}f(0) - s^{n-2}f'(0) - \cdots - f^{(n-1)}(0),$$

onde $F(s) = \mathcal{L}\{f(t)\}(s)$.

A definição de função de ordem exponencial pode ser encontrada em Zill (2011), bem como mais resultados acerca do tema em estudo.

Com essas ferramentas em mãos escolhemos duas equações específicas para aplicarmos a Transformada de Laplace em sua resolução. Essas equações são a equação de Bessel e a equação de Legendre, sendo que a segunda não é satisfeita utilizando tal método.

2 A equação de Bessel

A equação de Bessel é do tipo

$$ty'' + y' + ty = 0.$$

Aplicando a Transformada de Laplace e isolando $Y(s) = \mathcal{L}\{y(t)\}(s)$, temos que

$$\begin{gathered}
\mathcal{L}\{ty''\} + \mathcal{L}\{y'\} + \mathcal{L}\{ty\} = 0 \\
-\frac{d}{ds}\mathcal{L}\{y''\} + sY(s) - y(0) - \frac{d}{ds}\mathcal{L}\{y\} = 0 \\
-\frac{d}{ds}\mathcal{L}\{y''\} + sY(s) - y(0) - Y'(s) = 0 \\
-\frac{d}{ds}(s^2Y(s) - sy(0) - y'(0)) + sY(s) - y(0) - Y'(s) = 0 \\
-2sY(s) - s^2Y'(s) + sY(s) - Y'(s) = 0 \\
-(1+s^2)Y'(s) - sY(s) = 0 \\
(1+s^2)Y'(s) + sY(s) = 0
\end{gathered}$$

$$(1+s^2)^{\frac{1}{2}}Y'(s)+\frac{1}{2}\frac{2s}{(1+s^2)^{\frac{1}{2}}}Y(s)=0$$
$$((1+s^2)^{\frac{1}{2}}Y(s))'=0$$
$$(1+s^2)^{\frac{1}{2}}Y(s)=C$$
$$Y(s)=C(1+s^2)^{-\frac{1}{2}}.$$

Podemos reescrever $Y(s)$ como sendo $Y(s)=Cs^{-1}(1+s^{-2})^{-\frac{1}{2}}$. Vamos expandir o termo $(1+s^{-2})^{-\frac{1}{2}}$ em uma série binomial válida para $s>1$. Sendo a série binomial um somatório representado pela expressão

$$(1+x)^k=\sum_{n=0}^{\infty}\binom{k}{n}x^n,$$

onde

$$x=s^{-2},\qquad k=-\frac{1}{2}\quad\text{e}\quad\binom{k}{n}=\frac{k(k-1)(k-2)\cdots(k-n+1)}{n!}.$$

Então,

$$(1+s^{-2})^{-\frac{1}{2}}=\sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}(s^{-2})^n,$$

sendo

$$\begin{aligned}\binom{-\frac{1}{2}}{n}&=\frac{(-\frac{1}{2})(-\frac{3}{2})\cdots(\frac{1-2n}{2})}{n!}\\&=\frac{(-1)^n}{n!}\frac{(1\cdot3\cdot5\cdot\cdots\cdot(2n-1))}{2^n}\\&=\frac{(-1)^n}{n!}\frac{1\cdot2\cdot3\cdot\cdots\cdot(2n-1)\cdot(2n)}{2^n(2\cdot4\cdot6\cdot\cdots\cdot2n)}\\&=\frac{(-1)^n}{n!}\frac{(2n)!}{2^n2^n(1\cdot2\cdot3\cdot\cdots\cdot n)}\\&=\frac{(-1)^n}{n!}\frac{(2n)!}{2^{2n}n!}.\end{aligned}$$

Substituindo isto, temos

$$(1+s^{-2})^{-\frac{1}{2}}=\sum_{n=0}^{\infty}\binom{-\frac{1}{2}}{n}(s^{-2})^n=\sum_{n=0}^{\infty}\frac{(-1)^n}{n!}\frac{(2n)!}{2^{2n}n!}(s^{-2})^n,$$

e desta forma,

$$\begin{aligned}Y(s)&=Cs^{-1}\sum_{n=0}^{\infty}(s^{-2n})\frac{(-1)^n(2n)!}{(n!)^2(2^{2n})}\\&=C\sum_{n=0}^{\infty}\frac{(s^{-2n})(s^{-1})(-1)^n(2n)!}{(n!)^2(2^{2n})}\\&=C\sum_{n=0}^{\infty}\frac{(s^{-(2n+1)})(-1)^n(2n)!}{(n!)^2(2^{2n})}.\end{aligned}$$

Aplicando a Transformada Inversa temos que

$$(\mathcal{L}^{-1}\{Y(s)\})(t) = \mathcal{L}^{-1}\left\{C\sum_{n=0}^{\infty}\frac{(s^{-(2n+1)})(-1)^n(2n)!}{(n!)^2(2^{2n})}\right\}.$$

Podemos comutar a Transformada Inversa com o somatório (infinito), obtendo

$$\begin{aligned}(\mathcal{L}^{-1}\{Y(s)\})(t) &= C\sum_{n=0}^{\infty}\mathcal{L}^{-1}\left\{\frac{(s^{-(2n+1)})(-1)^n(2n)!}{(n!)^2(2^{2n})}\right\}\\ &= C\sum_{n=0}^{\infty}\frac{(-1)^n}{(n!)^2(2^{2n})}\mathcal{L}^{-1}\left\{\frac{(2n)!}{s^{2n+1}}\right\}\end{aligned}$$

Olhando para a tabela de Transformadas Inversas, temos que

$$\mathcal{L}^{-1}\frac{(2n)!}{s^{2n+1}} = t^{2n},$$

e portanto,

$$\begin{aligned}y(t) &= C\sum_{n=0}^{\infty}\frac{(-1)^n}{(n!)^2(2^{2n})}\mathcal{L}^{-1}\left\{\frac{(2n)!}{s^{2n+1}}\right\}\\ &= C\sum_{n=0}^{\infty}\frac{(-1)^n}{(n!)^2(2^{2n})}t^{2n} = C\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n}}{(n!)^2(2^{2n})}.\end{aligned}$$

3 A Equação de Legendre

A equação de Legendre de ordem α, com $\alpha > -1$ é uma equação linear de segunda ordem definida por

$$(1-t^2)y'' - 2ty' + \alpha(\alpha+1)y = 0,$$

pode ser reescrita da seguinte forma,

$$y'' - t^2y'' - 2ty' + \alpha^2y + \alpha y = 0,$$

com as condições iniciais, $y(0) = 0$ e $y'(0) = 1$.

Aplicando a Transformada de Laplace, temos que

$$\mathcal{L}\{y''\} - \mathcal{L}\{t^2y''\} - \mathcal{L}\{2ty'\} + \mathcal{L}\{\alpha^2y\} + \mathcal{L}\{\alpha y\} = 0,$$

donde

$$\begin{aligned}[s^2Y(s) - sy(0) - y'(0)] - \left[\frac{d^2}{ds^2}(s^2Y(s) - sy(0) - y'(0))\right]\\ - 2\left[\frac{d}{ds}(sY(s) - y(0))\right] + \alpha^2Y(s) + \alpha Y(s) = 0.\end{aligned}$$

Substituindo nas condições iniciais, temos

$$s^2Y(s) - 1 - \frac{d^2}{ds^2}(s^2Y(s) - 1) - 2\frac{d}{ds}(sY(s)) + \alpha^2 Y(s) + \alpha Y(s) = 0.$$

Resolvendo $\frac{d^2}{ds^2}(s^2Y(s) - 1)$ separadamente, obtemos

$$\begin{aligned}\frac{d^2}{ds^2}(s^2Y(s) - 1) &= \frac{d}{ds}\left(\frac{d}{ds}s^2Y(s)\right) \\ &= \frac{d}{ds}(2sY(s) + s^2Y'(s)) \\ &= 2Y(s) + 2sY'(s) + 2sY'(s) + s^2Y''(s).\end{aligned}$$

Resolvendo a derivada da segunda, encontramos novamente o termo $s^2Y''(s)$. Se olharmos para a equação de Legendre, veremos que saímos de uma expressão desse tipo, na variável t, e quando aplicamos a Transformada de Laplace voltamos a uma expressão assim. Logo, podemos verificar que a Transformada de Laplace nem sempre é útil para se resolver equações diferenciais com coeficientes variáveis, a menos que todos os coeficientes sejam, no máximo, funções lineares da variável independente.

Referências

Zill, Dennis. **Equações diferenciais com aplicações em modelagem**. Tradução da 9^a edição norte-americana. 2 ed. São Paulo: Cengage Learning, 2011.

Índice de autores