

Anais da

XXXVI Semana
Acadêmica da
Matemática
2022

Anais da
**XXXVI Semana Acadêmica
da Matemática**

21/11/2022 a 25/11/2022



Realização
Curso de
Matemática



Apoio
CCET
Centro de Ciências
Exatas e Tecnológicas

Comissão organizadora da XXXVI Semana Acadêmica da Matemática:

Jean Sebastian Toillier (Coordenador)
Cássia Delizandra Bonazza
Clezio Aparecido Braga
Francieli Cristina Agostinnetto Antunes
Maíri Poeta Cortina Castilho da Silva
Ricardo Mondini Ferrazza
Paulo Domingos Conejo

Comitê científico:

Amarildo de Vicente	Luciana Pagliosa Carvalho Guedes
Andre Vicente	Pamela Gonçalves
Andrea Büttner Ciani	Paulo Domingos Conejo
Arlení Elise Sella Langer	Pedro Pablo Durand Lazo
Carla Melli Tambarussi	Plínio Lucas Dias Andrade
Clezio Aparecido Braga	Raquel Lehrer
Driely Santos Nobre	Rogério Luis Rizzi
Fabiana Magda Garcia Papani	Rosângela Villwock
Felipe Leandro da Silva Costa	Sandro Marcos Guzzo
Flávio Roberto Dias Silva	Simone Aparecida Miloca
Francieli Cristina Agostinnetto Antunes	Tiago Emanuel Klüber
Jean Sebastian Toillier	
Jesus Marcos Camargo	

Arte da Capa:

Clezio Aparecido Braga

Diagramação:

Clezio Aparecido Braga

Semana Acadêmica de Matemática (36.: 2022 Cascavel-PR)
Anais da XXXVI Semana Acadêmica de Matemática/ Jean
Sebastian Toillier (Coordenador). ---Cascavel (PR):UNIOESTE,
2022.
110 p.

ISSN 2526-0804
Vários autores
Versão on-line
21/11/2022 a 25/11/2022

1. Matemática- estudo e ensino. 2. Pesquisa científica. 3. Ensino superior-Evento. I. Toillier, Jean Sebastian. II. Título.

Apresentação

A Semana Acadêmica da Matemática está na sua XXXVI 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 XXXVI 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, 11 trabalhos foram inscritos e aceitos para apresentação oral e publicação nos anais do evento. São em geral trabalhos resultados das pesquisas de Iniciação Científica e de Monografia desenvolvidos por alunos do curso de Matemática. Registramos também 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 os conhecimentos adquiridos pelos alunos e professores nos seus respectivos projetos. O registro destes trabalhos 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.

Índice de Trabalhos

O uso da estatística como base norteadora na avaliação da aprendizagem ..	4
Condições de otimalidade para problemas de otimização	16
Convergência do método de região de confiança para otimização irrestrita	28
Tópicos de Álgebra Linear e o Teorema de Convergência da Rede Neural Perceptron	40
O uso da Unity para criação de jogos didáticos	51
O uso da metodologia Resolução de problemas no estágio supervisionado .	60
Aspectos de um curso de formação continuada para professores da rede muni- cipal de ensino no município de Cafelândia: um enfoque na resolução de problemas como metodologia	69
Implementação do algoritmo de Dijkstra com apoio do software Geogebra	76
Aproximações de irracionais por meio de racionais	84
Um estudo sobre os fatores que promovem a evasão e a permanência dos discentes no curso de licenciatura em matemática da Unioeste – campus Cascavel no período de 2016 - 2022	90
PNLD: Dados históricos, seu funcionamento e algumas considerações rele- vantes acerca da aquisição	97

O uso da estatística como base norteadora na avaliação da aprendizagem.

Thays Perin¹

Universidade Estadual do Oeste do Paraná - UNIOESTE
thaaysperin@gmail.com

Dr. Alessandra dos Santos²

Universidade Estadual do Oeste do Paraná - UNIOESTE
Alessandra.santos23@unioeste.br

Resumo: Atualmente, uma das políticas mais comentadas no país refere-se à correção do fluxo escolar, ou seja, da diminuição das evasões, desistências e repetências. Assim, esse artigo analisa os dados escolares com uso de ferramentas estatísticas a fim de fornecer conhecimentos de tais técnicas para que os profissionais da área possam melhor inferir e detectar problemas e soluções no desenvolvimento do seu trabalho. Essas avaliações podem ter diversas finalidades, mas, nesse trabalho, nos limitamos a exemplificar duas situações. A primeira refere-se a avaliar se os alunos absorveram um conteúdo específico. Para tal, comparou-se suas notas em dois períodos, sendo antes e após a aplicação do conteúdo. Neste caso é como se o professor aplicasse a avaliação diagnóstica e após, tentasse sanar certas dificuldades apresentadas na primeira avaliação. O segundo exemplo compara duas turmas com igual desenvolvimento de conhecimento e sob condições parecidas entre os alunos, porém assume-se duas diferentes metodologias de ensino empregadas pelo mesmo professor. Assim, comparou-se as notas médias das turmas para avaliar se tais procedimentos realmente se diferiam. Ambas as situações apresentaram dados fictícios, sendo esses simulados e criados para servir como base de exemplo. Para realizar as análises, utilizou-se medidas de estatística descritiva e teste T de *Student*.

Palavras-chave: Estatística básica; Avaliação da aprendizagem; Teste T de *Student*; *Software R*.

1 Introdução

O presente artigo procura utilizar de estudos estatísticos como artefato de auxílio aos professores na autoavaliação dos modelos de ensino aplicados em sala, e com isso, pode-se identificar possíveis desníveis e impasses quanto a aprendizagem. De acordo com Luckesi (1996, p.165), o processo avaliativo deve ir além de impor notas, esse significa “identificar

impasses e buscar soluções”. Com o intuito de demonstrar e exemplificar o uso da estatística como base norteadora na avaliação do ensino/aprendizagem, foi realizado um estudo que avalia amostras de notas escolares de duas turmas fictícias (A e B). Para que este estudo seja possível, torna-se necessário identificar e trabalhar os conceitos e características relacionadas a aprendizagem.

Para tal, o presente artigo foi organizado em cinco partes. Na primeira foi apresentado o embasamento teórico, no qual fez-se uma abordagem sobre a aprendizagem, procurando evidenciar a necessidade de utilizar das tecnologias como aliadas na avaliação do ensino. O cenário analisado, os dados e análises estatísticas são apresentados na metodologia. A terceira parte foi dedicada para a apresentação dos resultados obtidos, no qual se pretende responder aos objetivos deste artigo. A quarta parte destina-se à conclusão e a última apresenta a programação utilizada para a análise dos dados por meio do software R.

2 Fundamentação Teórica

2.1 Aprendizagem

A aprendizagem é construída através de mudanças comportamentais que decorrem de experiências vividas. Ela ocorre desde o nascimento até o final da vida. A teoria de Piaget (1990) defende essa ideia, afirmando que, ao ser inserido no meio (que ocorre ao nascer), a criança é exposta a novas ideologias e costumes, que contribuem e influenciam na sua mudança comportamental, ou seja, a aprendizagem ocorre da interação do sujeito com o meio no qual foi inserido. Um exemplo é a linguagem verbal, visto que a criança tende a compreender e a reproduzir o idioma do meio que está inserido, ou seja, os pais e familiares atuam como mediadores.

Na escola, os professores são os mediadores do conhecimento, sendo a ligação entre o estudante e o conhecimento, de modo que o faça pensar de forma crítica, não apenas recebendo as informações e aceitando-as sem questioná-las.

De acordo com Bulgræn (2010), o docente possui

(...) a responsabilidade de agir como sujeito em meio ao mundo e de ensinar para seus educandos o conhecimento acumulado historicamente, dando-lhes a oportunidade de também atuarem como protagonistas na sociedade. (B ULGRAEN, 2010, pg. 31)

Seguindo o raciocínio de Bulgræn (2010), é indubitável que os educadores possuem a capacidade de transformar uma sociedade alienada em uma sociedade crítica e pensante. Sendo assim, acredita-se que um dos papéis do educador é adequar a sua prática pedagógica e avaliativa ao contexto atual da sociedade (BULGRAEN, 2010).

O conhecimento, principalmente o matemático, não deve estar vinculado somente a verdades absolutas e imutáveis, na qual o professor é o centro que impõe o conteúdo; e sim

da interação em sala voltada para o objetivo de ensinar o aluno e compreender as melhores formas de ensinar o conteúdo a ser trabalhado (BARALDI, 1999). Por este motivo é necessário mediar o conteúdo visando diversos aspectos, particularidades e culturas individuais dos alunos.

É de suma importância que o conhecimento que o aluno traz consigo seja levado em consideração no processo ensino-aprendizagem. Para Lorenzato (2010, p.24), “deve-se apoiar no saber vivenciado, pois é adaptando os novos conhecimentos aos já adquiridos que o aluno aprende”. Desta forma, torna-se necessário que o professor conheça estratégias para explorar os conhecimentos prévios adquiridos pelo aluno ao longo do tempo.

A interação aluno-professor também é de extrema importância para que ocorra a aprendizagem. Segundo Berbel (2011, p. 29), a empatia do aluno com o professor “[...] facilita a identificação pessoal com aquilo que ele apresenta em sala de aula, possibilitando a valorização das atividades e conteúdos propostos e a internalização das exigências ou demandas externas”. Para Freire (1986), é o diálogo entre professor e aluno que torna possível construir o saber científico, e assim, oportuniza o aluno a construir o seu próprio saber.

É necessário destacar necessidade de dispor ao educando um ambiente acolhedor e agradável, tornando este, um lugar propício para que ele possa desenvolver o melhor de si. Quando a avaliação dos conhecimentos adquiridos pelo aluno, deve-se procurar diversificar os métodos avaliativos e buscando identificar o desenvolvimento das habilidades e competências apresentadas.

2.2 Avaliação da aprendizagem

Avaliação da aprendizagem é o instrumento que avalia a evolução dos alunos no decorrer do processo ensino-aprendizagem. Quando se refere a esse assunto, evidencia-se que o processo não se refere apenas a aplicação de testes avaliativos, mas sim ao acompanhamento do aluno nas diferentes fases da aprendizagem.

De acordo com o dicionário Michaelis (2021), a avaliação é definida como “a apreciação, cômputo ou estimativa da qualidade de algo ou da competência de alguém”, ou seja, é o ato de avaliar se algo ou alguém é capaz de exercer o que lhe foi deferido.

Quando vista de forma correta, a avaliação pode impactar positivamente no trabalho do professor, sendo utilizada como diagnóstico ao docente sobre o desenvolvimento da aprendizagem. O que corrobora com Luckesi (1996, p. 43) ao dizer: “para não ser autoritária e conservadora, a avaliação terá de ser diagnóstica, ou seja, deverá ser o instrumento dialético do avanço, terá de ser o instrumento da identificação de novos rumos”. Logo, é através da avaliação que se pode evidenciar possíveis problemas na aprendizagem, e assim, sendo deferido ao professor, o papel de buscar alternativas para saná-los.

2.3 A utilização da estatística na avaliação da aprendizagem

É inevitável avaliar o desenvolvimento de uma turma com relação ao método de ensino e o processo ensino-aprendizado em si, não é uma tarefa fácil. Em alguns momentos, as reflexões sobre os resultados são filosóficas e norteiam resultados empíricos na comparação de turmas e alunos. Nesta perspectiva, acredita-se que a utilização de técnicas estatísticas inferenciais, como base norteadora, pode ser uma alternativa que contribui positivamente na avaliação sobre o desenvolvimento dos estudantes.

Técnicas de estatística descritiva e inferencial podem auxiliar o professor na avaliação do seu trabalho em uma ou mais turmas. O gráfico boxplot, por exemplo, é uma ferramenta gráfica utilizada para sintetizar e analisar dados quantitativos. O principal objetivo ao se trabalhar com este recurso é a possibilidade de “descrever didaticamente a estrutura, interpretação, construção, modificações e aplicações deste recurso visual” (VALLADARES *et al.* 2017, p. 1). O gráfico boxplot “é capaz de fornecer diversas medidas estatísticas como mínimo, primeiro quartil, segundo quartil (mediana), terceiro quartil, o máximo e os possíveis outliers (valores discrepantes)” (VIZONÁ; MIGUEL, 2021, s/p). Já o histograma é uma espécie de gráfico de barras que mostra a distribuição de uma amostra ou população, os quais são organizados e agrupados em intervalos ou em valores absolutos. Com o histograma é possível resumir grandes bases de dados de forma simplificada e de fácil visualização e ainda, comparar resultados (SIQUEIRA, 2021). No histograma é possível obter os dados de centralidade, amplitude e simetria.

Como técnicas inferenciais, o teste T de *Student* é um teste de hipótese. Atualmente muito utilizado na estatística, pois sua função é comparar médias. É possível comparar uma média amostral com valor hipotético ou ainda duas médias, sendo entre dois grupos distintos ou um mesmo grupo em dois momentos (esse caso é chamado dado pareado).

3 Metodologia

Para o desenvolvimento desse trabalho criamos dois cenários com objetivos diferentes, descritos nas subseções 3.1 e 3.2. Para cada situação, os dados foram simulados e analisados estatisticamente. Inicialmente, utilizou-se técnicas de análises descritivas tais como gráficos (boxplot e histograma), média aritmética, nota máxima e mínima, medidas separatrizes e desvio padrão a fim de avaliar quantitativamente os resultados. Depois, utilizou-se análises inferenciais de comparação de médias tais como: testes T-Student para amostras pareadas e independentes, suposto que os dados seguem distribuição normal. Essa condição será avaliada por meio do teste de Shapiro-Wilk com hipótese de nulidade de que os dados seguem distribuição normal. Para o estudo apresentado em 3.2, a situação de homogeneidade de variâncias será avaliada por meio do teste F, com hipótese de nulidade que as variâncias são homogêneas (iguais).

Para as inferências (testes de hipótese) será assumido o nível de significância de 5%; assim, se o nível descritivo (p-valor) do teste aplicado for menor que 0,05, então há evidências estatísticas para se rejeitar a hipótese de nulidade do mesmo.

Todas as estatísticas desenvolvidas nesse material foram realizadas no software estatístico R (R CORE TEAM, 2021) por meio da interface gráfica do software RStudio. A programação encontra-se em anexo.

3.1 Primeiro estudo

Inicialmente foi simulado a situação de que o professor deseja analisar se sua metodologia de ensino está surtindo efeito significativo nos alunos. Para isso, é como se duas “avaliações iguais” fossem desenvolvidas em momentos diferentes com alunos de uma turma fictícia denominada “A”. No primeiro momento (t_0), é como se o professor estivesse aplicando uma prova diagnóstica antes de iniciar qualquer aula sobre o conteúdo e, no segundo (t_n), a avaliação fosse reaplicada após o conteúdo ser explorado em aula. Na estatística, chama-se pareada as informações coletadas, visto que o mesmo indivíduo participa da coleta nos momentos t_0 e t_n .

Na simulação assumimos hipoteticamente uma turma com 35 alunos, sendo que as notas de cada aluno provinham de uma distribuição Normal Multivariada com médias para os instantes t_0 e t_n iguais a 50 e 80, respectivamente. Já as variâncias, t_0 e t_n , foram 30 e 25, respectivamente, com correlação de 0,5 entre os momentos; caso alguma nota fosse simulada com resultado menor que zero ou maior que 100, o resultado era truncado no valor mínimo ou máximo do intervalo, respectivamente 0 ou 100.

Para considerar que o método de ensino usado está sendo eficiente, é necessário que a nova média da turma seja estatisticamente superior a anterior. Assim, caso a condição de normalidade (Teste Shapiro-Wilk) seja satisfeita, o teste T de *Student* pareado é empregado. Nessa situação, a hipótese de nulidade é a de que as médias nos momentos t_0 e t_n são iguais; assumindo como hipótese alternativa que, após o processo de ensino do conteúdo, a turma apresente uma média superior a encontrada anteriormente (Tabela 1). Caso a hipótese de nulidade seja rejeitada, isso indicará que a professora conseguiu influenciar positivamente no aprendizado dos alunos; caso contrário, a diferença não é significativa o que indica que o desempenho não foi satisfatório. Este é, portanto, um teste denominado unilateral.

Tabela 1: Hipótese da primeira análise

Hipótese	Descrição da hipótese
Nulidade	$H_0: \mu_0 = \mu_n \iff \mu_0 - \mu_n = 0$
Alternati-	$H_a: \mu_0 < \mu_n \iff \mu_n - \mu_0 > 0$

va

Fonte: acervo da autora.

3.2 Segundo Estudo

No segundo exemplo, será analisado se duas turmas fictícias, A e B, absorveram com a mesma intensidade o conteúdo, ou seja, será analisado se, em média, as notas das turmas A e B são semelhantes após a aplicação do conteúdo (assumindo que as avaliações aplicadas fossem idênticas). Para tal, simulamos as notas das duas turmas de modo independente, supondo que ambas continham valor de notas seguindo a distribuição Normal e restrição de valores entre 0 a 100. Para a turma A, assumiu-se 35 alunos, com nota média igual a 60 e o desvio padrão de 25. Já a turma B continha 37 alunos, média igual a 70 e desvio padrão igual a 10.

Neste caso, as hipóteses a serem avaliadas é se há ou não diferença entre as médias dos valores observados na turma A (μ_A) e B (μ_B), assumindo como hipótese de nulidade a situação de que não há diferença entre as médias das turmas versus a hipótese alternativa de que as médias são diferentes (Tabela 2). Assim, para tal comparação, foi utilizado o teste bilateral T de *Student* para amostras independentes (dado a condição de normalidade dos dados e homogeneidade de variâncias). Se a hipótese de nulidade for rejeitada, tem-se que as notas médias das turmas diferem; caso contrário a diferença não é significativa o que indica que o desempenho dos alunos foi semelhante. Vale salientar que para essa análise estamos supondo que a professora é a mesma e sua metodologia e avaliação também são iguais em ambas as turmas.

Tabela 2: Hipótese da segunda análise

Hipótese	Descrição da hipótese
Nulidade	$H_0: \mu_A = \mu_B \iff \mu_A - \mu_B = 0$
Alternativa	$H_a: \mu_A \neq \mu_B \iff \mu_A - \mu_B \neq 0$

Fonte: acervo da autora.

4 Resultados obtidos

Inicialmente foram analisados os resultados obtidos ao realizar os estudos estatísticos na turma A antes e depois da aplicação do conteúdo (primeiro estudo).

Com o estudo descritivo, observou-se que a turma A, antes da aplicação do conteúdo possuía 38,9 e 61,6 como menores e maiores notas respectivamente, sua nota mediana era 49,4, sua média era 49,99, com desvio padrão de 5,10. Após a aplicação do conteúdo, a turma obteve 72,4 e 86,9 como menores e maiores notas respectivamente, a nota mediana foi 80,7, sua média passou a ser 80,68, com desvio padrão de 3,34. Como podemos corroborar na Figura 1, os resultados da avaliação após a apresentação do conteúdo (momento t_n) foram

notas superiores e com menor variabilidade do que no primeiro momento (t_0). Além disso, pode-se observar que no primeiro momento, alguns alunos da turma apresentaram notas discrepantes quando comparado com as notas centrais. Esses resultados, também denominados de outlier, são apresentados na Figura 1 como bolinhas.

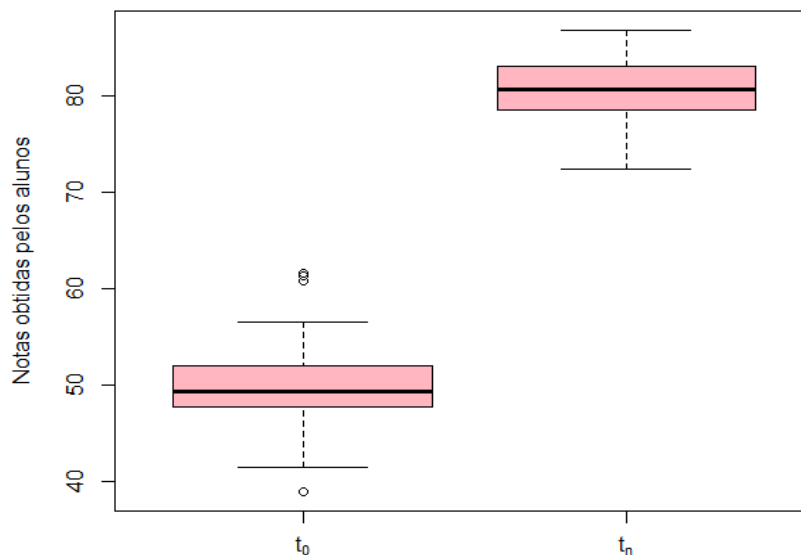


Figura 1: Gráfico boxplot com dados da turma A no instante t_0 e t_n respectivamente

Fonte: Acervo dos autores.

Após a análise descritiva foram realizadas as análises de normalidade. Os dados apresentaram distribuição normal ($p > 0,05$) Na Figura 2 observa-se o histograma dos dados.

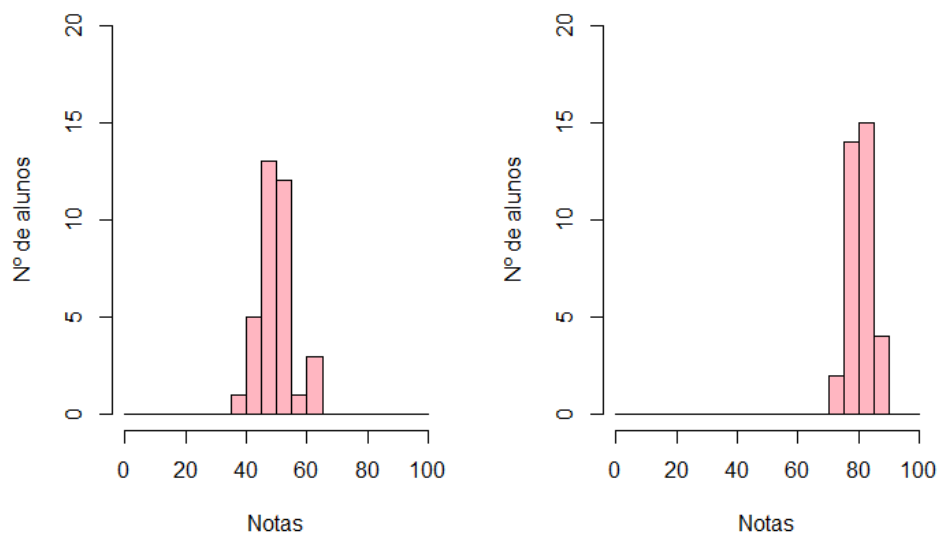


Figura 2: Histograma das notas da turma A no instante t_0 (à direita) e t_n (à esquerda).

Fonte: Acervo dos autores.

Ao realizar o teste T-Student pareado encontrou-se que o valor $P(T_{\text{calc}} < -32,97) < 0,001$. Com isso, baseado nessa amostra e ao nível de significância de 5%, o professor pode assumir que há evidências estatísticas para rejeitar H_0 . Portanto, é possível concluir que a média das notas da turma A após a aplicação do conteúdo é estatisticamente superior à média da inicial. Então, o professor pode avaliar que seu trabalho foi significativo, pois ele conseguiu influenciar positivamente no aprendizado dos alunos.

Com o segundo estudo de simulação, observou-se que a turma A, obteve 10,8 e 100,0 como menores e maiores notas, respectivamente; sua nota mediana foi 63,5, sua média 59,77 e o desvio padrão 21,75. A turma B obteve 54,1 e 87,5 como menores e maiores notas, respectivamente; a nota mediana foi 69,2, a média 69,45 e o seu desvio padrão foi 8,40. Estes dados podem ser facilmente observados no gráfico boxplot (Figura 3). Observe que existem notas discrepantes (muito baixas quando comparadas com os valores centrais) e a variabilidade da turma A é superior a turma B sendo que ambas apresentam uma leve assimetria com relação ao valor da mediana.

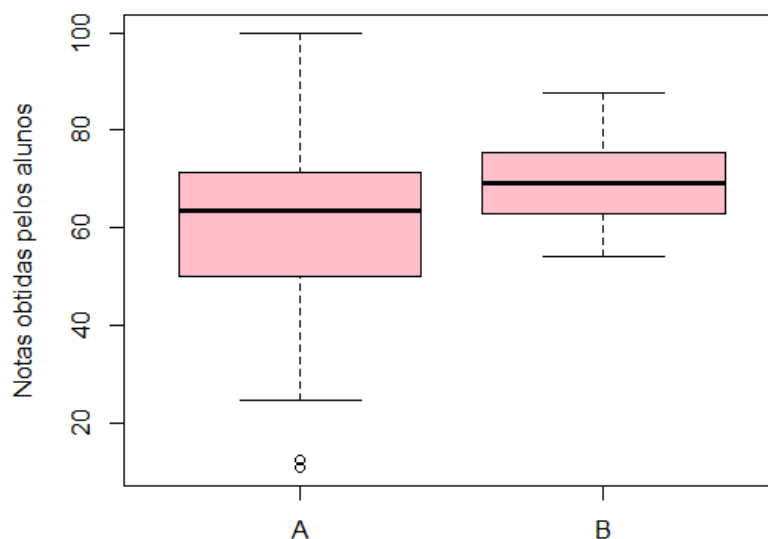


Figura 3: Gráfico boxplot com dados das turmas A e B respectivamente.

Fonte: Acervo dos autores.

Com o teste de normalidade e homogeneidade verificamos que os dados seguem uma distribuição Normal ($p_A=0,38$ e $p_B=0,62$) e as variâncias são heterogêneas ($p<0,01$). Na Figura 4, observa-se a distribuição dos dados de acordo com a turma.

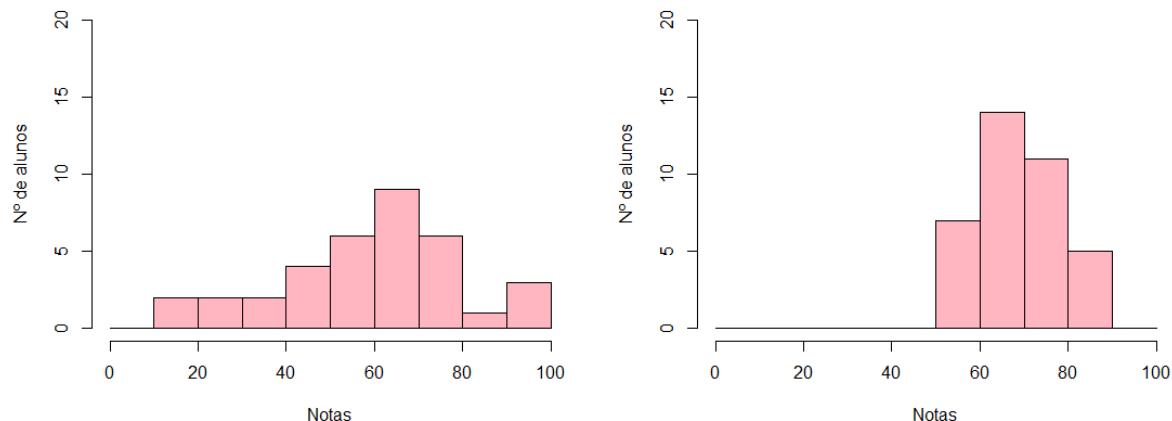


Figura 4: Histograma das notas da turma A e B respectivamente.

Fonte: Acervo dos autores.

Comparando a média das turmas, pelo teste T-Student com variâncias desiguais, encontrou-se que o valor $P(T_{\text{calc}} > 2,46 \text{ ou } T_{\text{calc}} < -2,46) = 0,02$. Com isso, dada essa amostra e o nível de significância de 5%, há evidências para se rejeitar H_0 . Portanto, é possível concluir que a média das notas das turmas A e B são estatisticamente diferentes. Pelo boxplot pode entender que alguns alunos da turma A precisam de um reforço maior no conteúdo do que a turma B.

Conclusões

As análises feitas servem para evidenciar possíveis dificuldades e evolução na aprendizagem dos estudantes. De modo geral, pelo gráfico boxplot, os professores podem identificar se alguns resultados são específicos de alguns alunos ou se o comportamento central é semelhante, detectando resultados considerados discrepantes ou não. Além disso, o comportamento geral pode ser observado juntamente com as medidas descritivas e o gráfico do histograma. Assim, é possível verificar a necessidade de um trabalho planejado de acompanhamento e de atividades específicas para os alunos e turmas, com o objetivo de fortalecer os conhecimentos básicos.

No primeiro estudo, apresentado nesse trabalho, concluiu-se que na turma fictícia A os resultados foram positivos com a aplicação do conteúdo, pois a evolução dos alunos ficou estatisticamente significativa, ou seja, a metodologia aplicada foi satisfatória.

No segundo estudo, ao comparar as notas de duas turmas fictícias A e B, foi possível concluir que ambas as turmas não absorveram, em média, a mesma quantidade do conteúdo aplicado. Assim, o professor deve continuar avaliando essas metodologias e entender melhor

porque a turma B apresentou a média com aproximadamente 10 pontos acima do resultado da turma A.

De acordo com o supracitado, conclui-se que atrelar a estatística básica e inferencial a avaliação da aprendizagem é um artefato que pode auxiliar os professores. Assim, todo o corpo docente, pode deixar de fazer afirmações empíricas e apresentar afirmações científicas. Portanto, realizar análises periódicas pautadas em métricas estatísticas podem contribuir para um melhor desenvolvimento do trabalho do professor e conseqüentemente uma aprendizagem significativa; sanando possíveis dificuldades com precedência.

Referências

- AVALIAÇÃO. Dicionário Michaelis. **Editora Melhoramentos**: 2021. Disponível em: <https://michaelis.uol.com.br/busca?id=3QoY>. Acesso em 19 set. 2021
- BARALDI, Ivete Maria. Refletindo sobre as concepções matemáticas e suas implicações para o ensino diante do ponto de vista dos alunos. **Mimesis**, Bauru, v. 20, n. 1, p. 07- 18, 1999.
- BERBEL, N. As metodologias ativas e a promoção da autonomia dos estudantes. **Semina: Ciências Sociais e Humanas**, Londrina, v. 32, n. 1, p. 25-40, jan./jun. 2011.
- BULGRAEN, V. C. O papel do professor e a sua mediação nos processos de elaboração do conhecimento. **Revista Conteúdo**, Capivari, v. 1, n. 4, p. 30-38, ago/dez. 2010.
- FREIRE, P. **Pedagogia da autonomia**: saberes necessários à prática educativa. São Paulo: Paz e Terra, 1986. Ano de digitação: 2002.
- LORENZATO, S. **Para aprender matemática**. 3. ed. rev. Campinas: Autores Associados, 2010.
- LUCKESI, C. C.. Avaliação da aprendizagem escolar. 3 ed. São Paulo: Cortez, 1996.
- PIAGET, J. **Seis estudos de Psicologia**. 24. ed. Rio de Janeiro: Forense Universitária Ltda, 1990.
- VALLADARES NETO, J; SANTOS, C. B; TORRES, E. M; ESTRELA, C. Boxplot: um recurso gráfico para a análise e interpretação de dados quantitativos. **Revista Odontológica do Brasil Central**, v. 26, n. 76, 2017.
- R CORE TEAM. **R**: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2021.
- SIQUEIRA, D. **Histograma**: o que é, exemplos, gráficos e tipos. [S.I], Alura, 2021. Disponível em: https://www.alura.com.br/artigos/o-que-e-um-histograma?gclid=CjwKCAjwKmaBhBMEiwAy INuwNymp-iwxM_ahab9zBLEZpbN2f2LIGmEtsjSIO8NJZA5GpufkvAewxoCDBEQAvD_BwE. Acesso em: 15 out. 2022

VIZONÁ, A; MIGUEL, L. **Como interpretar um gráfico Boxplot**. São Paulo, ICMC Júnior, 2021. Disponível em: https://icmcjunior.com.br/como-interpretar-um-grafico-boxplot/?gclid=CjwKCAjwKmaBhBMEiwAyINuwObk3pGaa33R4P2WbZA2yBe3nIagug6PwSk2oKbxUALz34pKpiUc5RoCizcQAvD_Bw. Acesso em 15 out. 2022.

Apêndice: Programação utilizada no Software R (Script)

```
## Situação 1 --- Teste Antes e Depois
summary(notas_antes) #medidas resumo: mínimo, quartis, mediana,
média, máximo
summary(notas_depois)
sd(notas_antes) ## desvio padrão de A
sd(notas_depois) ## desvio padrão de D
boxplot(cbind(notas_antes, notas_depois),ylab="Notas obtidas pelos
      alunos", names=c(expression("t"[0]), expression("t"[n])),
      col=" light pink")      # gráfico
par(mfrow=c(1,2))
hist(notas_antes, xlim = c(0,100), breaks = seq(0,100,5), main="",
      ylim = c(0,20), freq = T, ylab = "Nº de alunos", xlab= "Notas",
      col = "light pink")
hist(notas_depois, xlim = c(0,100), breaks = seq(0,100,5), main="",
      ylim = c(0,20), freq = T, ylab = "Nº de alunos", xlab= "Notas",
      col = "light pink")
diferença <- notas_antes - notas_depois
shapiro.test(diferença) # teste de normalidade
t.test(notas_antes, notas_depois, var.equal = T,paired=T,
      alternative = "less",conf.level=0.95) # teste T

## Situação 2 - Teste comparando turma A e B
summary(A)
summary(B)
sd(A)
sd(B)
boxplot(cbind(A,B),ylab="Notas obtidas pelos alunos",
      names=c(expression("A"),expression("B")), col="light pink")
hist(A, xlim = c(0,100), breaks = seq(0,100,10), main="", ylim =
      c(0,20),freq = T, ylab = "Nº de alunos", xlab= "Notas", col =
      "light pink")
hist(B, xlim = c(0,100), breaks = seq(0,100,10), main="", ylim =
```

```
      c(0,20),freq = T, ylab ="Nº de alunos", xlab= "Notas", col =  
      "light pink")  
var.test(A,B) # teste de homogeneidade de variâncias  
shapiro.test(A) # teste de normalidade  
shapiro.test(B) # teste de normalidade  
  
t.test(A,B, var.equal = F,paired=F, conf.level=0.95) # teste T
```

Condições de otimalidade para problemas de otimização

Andre Luiz Zanin da Cruz
Universidade Estadual do Oeste do Paraná
andrezanincruz@hotmail.com

Paulo D. Conejo
Universidade Estadual do Oeste do Paraná
paulo.conejo@unioeste.br

Resumo: Na maioria dos problemas de otimização é muito difícil encontrar pontos minimizadores (maximizadores) globais. Para contornar essa dificuldade, são buscados pontos que satisfazem determinadas propriedades que também seriam verificadas pelos extremos globais, essas propriedades são chamadas de condições de otimalidade. Este trabalho é dedicado ao estudo das condições de otimalidade na otimização irrestrita e com restrição de igualdade pelo método dos multiplicadores de Lagrange com uma restrição. Além disso, aproveitaremos para apresentar as condições de Karush-Kuhn-Tucker à luz da Condição de Qualificação de Independência Linear da otimização restrita com restrição de desigualdade que muito se relaciona com as condições de Lagrange, e revisaremos alguns conceitos e resultados básicos da Álgebra Linear e da otimização.

Palavras-chave: Otimização; Condições de otimalidade; Multiplicadores de Lagrange.

1 Introdução

Em um problema de otimização, buscamos a princípio encontrar um ponto que minimize ou maximize globalmente a função objetivo f . O conjunto Ω formado pelos pontos a serem aplicados na função é chamado de conjunto viável, podendo este estar sobre determinadas restrições em seu interior. Usualmente, um problema de otimização é escrito na forma

$$\begin{aligned} &\text{minimizar} && f(x) \\ &\text{sujeito a} && x \in \Omega, \end{aligned} \tag{1}$$

sendo o conjunto viável definido neste caso por $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0 \text{ e } g(x) \leq 0\}$, com $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ sendo funções diferenciáveis.

Se o conjunto $\Omega = \mathbb{R}^n$, dizemos que estamos trabalhando com a otimização irrestrita, isto é, se trata de um problema irrestrito. Quando Ω é formado por pontos que satisfazem apenas igualdades, por exemplo, $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0 \text{ e } g(x) = 0\}$ significa que estamos trabalhando com a otimização restrita com restrições de igualdades. No problema 1, estamos definindo um problema da otimização restrita com restrição de desigualdade.

Quando minimizamos uma função, nem sempre é uma tarefa fácil encontrar minimizadores, visto que por conta da profundidade do conjunto viável, podem existir pontos que apresentam valores ainda menores. Segundo Haeser (2009), para contornar essa dificuldade, buscamos por pontos que satisfaçam determinadas propriedades que também seriam satisfeitas

pelo minimizador global. Essas propriedades são chamadas de condições de otimalidade e esses pontos de pontos estacionários.

Na otimização irrestrita as condições de otimalidade necessárias e suficientes caracterizam o minimizador local como solução para este tipo de problema, e elas podem ser verificadas apenas utilizando as derivadas da função objetivo (RIBEIRO & KARAS, 2013). Já na otimização restrita, Joseph-Louis Lagrange (1736–1813) apresenta uma condição de otimalidade necessária para determinar pontos estacionários de f restrita aos pontos de $h(x) = 0$, denominada condição de Lagrange ou também de Teorema dos multiplicadores de Lagrange (MARCHAND, 2016; LUENBERGER, 2008).

Esse método mostra que se x^0 é minimizador do problema de otimização restrita com restrições de igualdades, sobre algumas hipóteses, vai existir vetores $\lambda \in \mathbb{R}^m$ e $\mu \in \mathbb{R}^p$ tais que podemos escrever o vetor $\nabla f(x^0)$ como uma combinação linear dos vetores gradientes dessas restrições, $\nabla h(x^0)$ e $\nabla g(x^0)$, isto é,

$$\nabla f(x^0) + \sum_{i=1}^m \lambda_i \nabla h_i(x^0) + \sum_{j=1}^p \mu_j \nabla g_j(x^0) = 0.$$

Neste trabalho, estamos interessados em desenvolver um estudo sobre as condições de otimalidade da otimização irrestrita e da otimização restrita com restrição de igualdade pelo método dos multiplicadores de Lagrange que é o mais difundido na literatura.

Aproveitaremos para também apresentar a condição de otimalidade de Karush-Kuhn-Tucker (KKT) sobre a Condição de Qualificação de Independência Linear, considerado por muitos o teorema central da otimização e que muito se assemelha a condição de Lagrange. Além disso, vamos discutir alguns resultados e conceitos necessários da Álgebra Linear e da Otimização para todo esse estudo.

Por questão notacional, durante o texto vamos nos referir aos pontos do \mathbb{R}^n com o índice acima, por exemplo, x^0, x^3, x^5 , entre outros. E para as n coordenadas desses pontos vamos usar o índice abaixo, por exemplo, $x^0 = (x_1^0, x_2^0, \dots, x_n^0)$.

2 Conceitos iniciais

2.1 Resultados da Álgebra Linear

Nesta seção nosso objetivo é apresentar alguns resultados da Álgebra Linear que serão utilizados nas seções futuras. Ao estudarmos as condições de otimalidade da otimização irrestrita, é importante sabermos sobre as matrizes simétricas, autovetores e matrizes positivas definidas, já para a condição de Lagrange é fundamental conhecer o conceito de combinação linear e vetores linearmente independentes. Todos esses resultados estão presentes na Álgebra Linear e podem ser encontrados em Anton & Rorres (2012), Ribeiro & Karas (2013) e Steinbruch (1987).

Inicialmente, vamos aceitar que já conhecemos o que é um espaço e subespaço vetorial no \mathbb{R}^n , mais os conceitos de subconjuntos fechados e limitados.

Definição 1 (Combinação linear). Dizemos que um vetor $u \in \mathbb{R}^n$ é uma combinação linear dos

vetores v^1, v^2, \dots, v^k em \mathbb{R}^n se u puder ser expresso na forma

$$u = \alpha^1 v^1 + \alpha^2 v^2 + \dots + \alpha^k v^k$$

em que $\alpha^1, \alpha^2, \dots, \alpha^k$ são escalares. Esses escalares são denominados coeficientes da combinação linear. No caso em que $k = 1$, a equação anterior se torna $u = \alpha^1 v^1$, de que combinação linear de um vetor é simplesmente um múltiplo escalar desse vetor.

Definição 2. Seja $V \subset \mathbb{R}^n$ um espaço vetorial e $B = \{v^1, v^2, \dots, v^k\} \subset V$. Dizemos que B é Linearmente Independente (LI) se, e somente se, nenhum vetor de B é combinação linear dos demais. Caso um vetor de B seja combinação linear dos demais, B é dito ser Linearmente Dependente (LD).

Teorema 3. *Seja V um espaço vetorial e $B = \{v^1, v^2, \dots, v^k\} \subset V$. Então B é LI se, e somente se, $\alpha^1 v^1 + \alpha^2 v^2 + \dots + \alpha^k v^k = 0$ tem apenas a solução trivial $\alpha^i = 0$ para todo $i \in \{1, 2, \dots, k\}$.*

Prova. Em (STEINBRUCH, 1987). □

Definição 4 (Autovalor e Autovetor). Seja A uma matriz real quadrada de ordem $n \times n$, então um vetor não nulo $x \in \mathbb{R}^n$ é denominado autovetor de A , se Ax for múltiplo escalar de x , isto é, existe ρ tal que $Ax = \rho x$. O escalar ρ é chamado de autovalor de A e dizemos que x é um autovetor associado a ρ .

Da equação $Ax = \rho x$, conseguimos reescrevê-la como $Ax - \rho Ix = 0$ ou $(\rho I - A)x = 0$, sendo I a matriz identidade de ordem n , e assim obtemos o seguinte teorema.

Teorema 5. *Se $A \in \mathbb{R}^{n \times n}$, então ρ é um autovalor de A se, e somente se, ρ satisfaz a equação*

$$\det(\rho I - A) = 0.$$

Prova. Em (ANTON & RORRES, 2012). □

A última equação é denominada equação característica de A . Se o determinante $\det(\rho I - A)$ é expandido, resulta no polinômio $P(\rho)$ de grau n que denominamos de polinômio característico.

Definição 6 (Matriz positiva definida). Seja A uma matriz simétrica e $Q(x) = x^T Ax$ sua forma quadrática associada. Dizemos que a matriz A é

I) Positiva definida se $Q(x) = x^T Ax > 0$, para todo $x \in \mathbb{R}^n - \{0\}$.

II) Semidefinida positiva se $Q(x) = x^T Ax \geq 0$, para todo $x \in \mathbb{R}^n - \{0\}$.

III) Negativa definida se $Q(x) = x^T Ax < 0$ para todo $x \in \mathbb{R}^n - \{0\}$.

IV) Semidefinida negativa se $Q(x) = x^T Ax \leq 0$, para todo $x \in \mathbb{R}^n - \{0\}$.

V) A é indefinida se existem vetores não nulos $x, z \in \mathbb{R}^n$ tal que $Q(x) = x^T Ax > 0$ e $Q(z) = z^T Az < 0$.

O teorema logo abaixo é utilizado para indicar quando uma matriz é positiva definida, negativa definida ou indefinida. Caso algum dos autovalores seja igual a zero, teremos o caso de semidefinida positiva ou negativa.

Teorema 7. *Seja A uma matriz simétrica. Valem as afirmações:*

- I) $x^T Ax$ é positiva definida se, e somente se, todos os autovalores de A são positivos.
- II) $x^T Ax$ é negativa definida se, e somente se, todos os autovalores de A são negativos.
- III) $x^T Ax$ é indefinida se, e somente se, A tem autovetores positivos e negativos.

Prova. Em (ANTON & RORRES, 2012). □

Quando A é simétrica, podemos provar que seus autovalores são reais. Como consequência podemos demonstrar o teorema dos extremos condicionados. Esse teorema é utilizado no estudo da otimização usando as formas quadráticas e na demonstração do Teorema 19.

Teorema 8 (Extremos condicionados). *Seja A uma matriz simétrica $n \times n$ cujos autovalores em ordem crescente de tamanho são $\rho_1 \leq \rho_2 \leq \dots \leq \rho_n$. Então*

I) *A forma quadrática $x^T Ax$ atinge um valor máximo e um valor mínimo no conjunto de vetores tais que $\|x\| = 1$.*

II) *O valor máximo atingido na parte I ocorre em um autovetor unitário associado ao autovalor ρ_n .*

III) *O valor mínimo atingido na parte I ocorre em um autovetor unitário associado ao autovalor ρ_1 .*

Prova. Em (ANTON & RORRES, 2012). □

2.2 Resultados da Otimização

Nesta seção nosso objetivo é introduzir alguns dos principais resultados da otimização, que serão necessários para os estudos das condições de otimalidade. Essas informações podem ser encontradas mais detalhadamente em Friedlander (1994), Haeser (2009), Lima (2004), Ribeiro & Karas (2013) e William (2013).

Como já comentado na introdução, as condições de otimalidade são propriedades satisfeitas pelos minimizadores globais do problema. Como na maioria dos problemas, encontrar esses pontos é usualmente difícil, procuramos por pontos que não são esses minimizadores globais, mas que satisfazem também essas propriedades. Os pontos que satisfazem essas propriedades são denominados de pontos estacionários (ou pontos críticos).

Definição 9 (Condição de Otimalidade). Uma propriedade P sobre os pontos de \mathbb{R}^n é dita uma condição de otimalidade se

$$x \text{ minimizador global} \Rightarrow x \text{ satisfaz } P.$$

Uma forma de classificar as condições de otimalidade é pela sua força. De acordo com Haeser (2009), as condições de otimalidade fortes são aquelas que são satisfeitas por uma pequena

quantidade de pontos, mas que não são minimizadores globais. Já as condições fracas possuem uma grande quantidade de pontos que as satisfazem e que não são minimizadores globais. As condições de otimalidade mais fortes são aquelas que caracterizam x^0 como minimizador global.

Mas como já comentado, encontrar ou verificar que x^0 é esse extremo global é difícil, por isso estamos interessados em propriedades mais fracas. Um exemplo de condição mais fraca que essa anterior é a de ser minimizador local.

Definição 10. Considere uma função $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x^0 \in \Omega$. Dizemos que x^0 é um minimizador local de f em Ω quando existe $\delta > 0$, tal que $f(x^0) \leq f(x)$, para todo $x \in B(x^0, \delta) \cap \Omega$. Caso $f(x^0) \leq f(x)$, para todo $x \in \Omega$, x^0 é dito minimizador global de f em Ω .

Por mais que encontrar minimizadores locais seja mais fácil, sua verificação na maioria das vezes é também complicada. Na otimização restrita, a verificação acaba sendo mais simples por dispormos de conceitos e resultados que garantem a existência destes pontos, como é o caso da compactidade do conjunto com o teorema de Weierstrass.

Definição 11. Um conjunto $\Omega \subset \mathbb{R}^n$ é compacto quando Ω é fechado e limitado.

O teorema de Weierstrass pode garantir a existência de extremos globais de f caso essa seja contínua e Ω compacto. De maneira análoga, o teorema a seguir pode ser demonstrado para o caso de maximizador global, mas aqui focaremos apenas para o valor de mínimo global.

Teorema 12 (Teorema de Weierstrass). *Sejam $f : \mathbb{R}^n \rightarrow \mathbb{R}$ contínua e $\Omega \subset \mathbb{R}^n$ compacto e não vazio. Então existe minimizador global de f em Ω .*

Prova. Em (RIBEIRO & KARAS, 2013). □

Com o resultado seguinte podemos compreender a funcionalidade do teorema dos multiplicadores de Lagrange quando Ω é uma curva ou superfície, isso se deve ao fato do vetor gradiente da equação que forma a curva ou a superfície ser um vetor normal em qualquer um de seus pontos.

Definição 13. Um vetor que é ortogonal a todo vetor que seja tangente a alguma curva C da superfície S no ponto p é chamado de vetor normal a S no ponto p .

Teorema 14. *Se a equação de uma superfície S for dada por $q(x) = 0$, com todas as derivadas parciais contínuas e nem todas nulas no ponto $p \in S \subset \mathbb{R}^n$, então o vetor normal a S no ponto p será $\nabla q(p)$.*

A seguir apresentamos um dos resultados mais necessários para a demonstração do teorema dos multiplicadores de Lagrange. Esse teorema apresenta as condições locais para que de uma relação do tipo $H(x, y) = 0$ consigamos definir $x = \gamma(y)$. Em outras palavras, considerando uma curva ou superfície, caso um ponto da mesma satisfaça a hipótese, será garantida na vizinhança daquele ponto uma função $\gamma(y)$ de classe C^1 . Mais detalhes em (LIMA, 2004).

Teorema 15 (Teorema da Função Implícita). *Seja $H : \mathbb{R}^k \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ uma função de classe C^1 . Suponha que $H(x^0, y^0) = 0$ e*

$$\left| \frac{\partial H(x^0, y^0)}{\partial y} \right| \neq 0.$$

Então, existe um aberto $W \subset \mathbb{R}^k$ e $\gamma : W \rightarrow \mathbb{R}^m$ função de classe C^1 tais que

$$I) x^0 \in W \text{ e } \gamma(x^0) = y^0$$

$$II) H(x, \gamma(x)) = 0, \forall x \in W.$$

Prova. Em (ELON, 2004). □

3 Otimização Irrestrita e suas Condições de Otimalidade

Vamos agora estudar as condições necessárias de primeira e segunda ordem da otimização irrestrita. Essas condições apresentam o que é necessário para que um ponto x^0 seja um minimizador local de f . Informações mais detalhadas podem ser encontradas em Flor (2016), Friedlander (1994) e Ribeiro & Karas (2013).

Teorema 16 (Condição necessária de primeira ordem). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável no ponto $x^0 \in \mathbb{R}^n$. Se x^0 é um minimizador local de f , então $\nabla f(x^0) = 0$, ou seja, $\frac{\partial f(x^0)}{\partial x_i} = 0$ para todo $i \in \{1, 2, \dots, n\}$.*

Prova. Em (RIBEIRO & KARAS, 2013). □

Definição 17 (Ponto estacionário). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função continuamente diferenciável, em $x^0 \in \mathbb{R}^n$ a condição de $\nabla f(x^0) = 0$, então x^0 é dito ser um ponto crítico (ou estacionário) de f .*

De acordo com Flor (2016), uma condição de otimalidade não é utilizada apenas para encontrar os pontos candidatos a minimizadores locais, mas sim para verificar um algoritmo numérico iterativo, estabelecendo seu critério de parada, o tipo de convergência e a sua robustez.

Importante comentar que a recíproca desse teorema não é válida, ou seja, nem todo ponto estacionário é um minimizador local. Considere o exemplo abaixo verificando a recíproca falsa.

Exemplo 1. Considere a função $f(x) = x^5$. Fazendo $f'(x) = 5x^4$ e igualando a zero vamos ter $f'(x) = 0 \Rightarrow 5x^4 = 0 \Rightarrow x^0 = 0$, no entanto esta função não possui mínimo e nem máximo local em $x = 0$.

A próxima condição de otimalidade é um pouco mais forte que a anterior, trazendo agora uma nova característica para o minimizador local x^0 que é ter a matriz Hessiana aplicada nesse ponto $\nabla^2 f(x^0)$ sendo semidefinida positiva, isto é, todos os autovalores dessa matriz devem ser maiores ou iguais a zero.

Teorema 18 (Condição necessária de segunda ordem). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes diferenciável no ponto $x^0 \in \mathbb{R}^n$. Se x^0 é um minimizador local de f , então $\nabla f(x^0) = 0$ e a matriz Hessiana de f no ponto x^0 é semidefinida positiva, isto é,*

$$d^T \nabla^2 f(x^0) d \geq 0,$$

para todo vetor direcional $d \in \mathbb{R}^n - \{0\}$.

Prova. Em (RIBEIRO & KARAS, 2013). □

Assim como a condição necessária de primeira ordem, sua recíproca também não é verdadeira, ou seja, mesmo um ponto x^0 satisfazendo essas condições, não garante que ele será um minimizador local. Tome novamente como exemplo a função $f(x) = x^5$.

Exemplo 2. De $f(x) = x^5$, sabemos que o ponto estacionário é $x^0 = 0$ por $f'(x) = 0$. A matriz Hessiana neste caso será uma matriz unitária, formada apenas por $\nabla^2 f(x) = f''(x) = 20x^3$. Quando aplicada em x^0 temos $f''(0) = 0$. O único autovalor possível seria o zero, satisfazendo o Teorema 18, mas esta função não possui máximo/mínimo local.

Com as condições necessárias de primeira e segunda ordem podemos obter pontos que são candidatos a minimizadores e maximizadores locais, porém nem todos eles serão esses extremos locais. Agora veremos a condição suficiente de segunda ordem que garante que o ponto encontrado é um extremo local. No entanto, nem todos os pontos de extremos locais da função objetivo serão identificados por esse teorema.

Teorema 19 (Condição suficiente de segunda ordem). *Seja $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função duas vezes diferenciável no ponto $x^0 \in \mathbb{R}^n$. Se x^0 satisfaz a condição de ser um ponto crítico e de ter a matriz Hessiana positiva definida, ou seja, $\nabla f(x^0) = 0$ e $d^T \nabla^2 f(x^0) d > 0$ com $d \in \mathbb{R}^n - \{0\}$, então x^0 é um minimizador local estrito da função.*

Prova. Em (RIBEIRO & KARAS, 2013). □

Exemplo 3. Considere $f(x) = x^6$, temos que $x^0 = 0$, a matriz Hessiana aplicada nesse ponto é $\nabla^2 f(0) = 0$ e o único autovalor possível é $\rho = 0$, então o Teorema 19 não é satisfeito, mas $x^0 = 0$ é ponto de mínimo local.

3.1 Otimização restrita e o método dos multiplicadores de Lagrange

As principais referências neste tópico foram Evandro (2014), Friedlander (1994), Marchand (2016), Oliveira (2012), Santos (2015) e William (2013).

O teorema dos multiplicadores de Lagrange é o método mais difundido na otimização restrita com restrição de igualdade, sendo desenvolvido por Joseph-Louis Lagrange (1736 – 1813) um importante matemático e físico francês de origem italiana do século XVIII. Em sua principal obra *Mécanique Analytique* desenvolveu os estudos sobre a mecânica, também contribuiu com soluções numéricas e algébricas de equações na teoria dos números, dentre outras realizações nas áreas da matemática, física e astronomia (SANTOS, 2015).

Com o método dos multiplicadores de Lagrange podemos modificar um problema de otimização restrita de n variáveis e m restrições de igualdade, em um problema da otimização irrestrita com $n + m$ variáveis. Essa transformação de problema restrito para irrestrito é feita com a adição dos multiplicadores de Lagrange (MARCHAND, 2016). Importante comentar que não existe um método geral para resolver todo os tipos de problemas da otimização restrita com restrição de igualdade. Podemos por exemplo também resolver a equação de restrição para uma das variáveis em termo das outras, mas esse caminho é muito complicado, por isso a preferência pelo método.

Durante a demonstração abaixo, por questão de organização textual, foram usadas três representações distintas para derivadas parciais, sendo elas $D_x f$, f_x e $\frac{\partial f(x)}{\partial x}$.

Teorema 20 (Multiplicadores de Lagrange - Uma restrição). *Sejam $f, g : D \subset \mathbb{R}^n \rightarrow \mathbb{R}$ funções de classe C^1 . Suponha $n > 1$. Se x^0 é um ponto de extremo local de f sujeita a restrição $g(x) = 0$ e vale a condição de regularidade $g_{x_r}(x^0) \neq 0$ para algum $r \in \{1, 2, 3, \dots, n\}$, então há uma constante λ de tal forma que*

$$f_{x_i}(x^0) - \lambda g_{x_i}(x^0) = 0, \quad \text{com } 1 \leq i \leq n. \quad (2)$$

Assim, x^0 é um ponto crítico de $f - \lambda g$.

Prova. Por conveniência notacional, tomemos $r = 1$ a derivada parcial diferente de zero no ponto x^0 , e vamos denotar $u = (x_2, x_3, \dots, x_n)$ e $u^0 = (x_2^0, x_3^0, \dots, x_n^0)$. Como $g_{x_1}(x^0) \neq 0$, o Teorema da Função Implícita afirma que existe uma única função com derivada contínua $\gamma = \gamma(u)$ definida numa vizinhança $W \subset \mathbb{R}^{n-1}$ de u^0 , de tal forma que $(\gamma(u), u) \in D$ para todo $u \in W$, com $\gamma(u^0) = x_1^0$, e

$$g(\gamma(u), u) = 0, \quad u \in W.$$

Agora, vamos definir

$$\lambda = \frac{f_{x_1}(x^0)}{g_{x_1}(x^0)}, \quad (3)$$

isso é permitido porque $g_{x_1}(x^0) \neq 0$. Isso implica em (2) ser válida para $i = 1$.

Agora se $2 \leq i \leq n$, derivando a equação $g(\gamma(u), u) = 0$ pela regra da cadeia vamos obter

$$\frac{\partial g(\gamma(u), u)}{\partial x_i} + \frac{\partial g(\gamma(u), u)}{\partial x_1} \frac{\partial \gamma(u)}{\partial x_i} = 0. \quad (4)$$

E também,

$$D_x f(\gamma(u), u) = \frac{\partial f(\gamma(u), u)}{\partial x_i} + \frac{\partial f(\gamma(u), u)}{\partial x_1} \frac{\partial \gamma(u)}{\partial x_i}. \quad (5)$$

Tomando $(\gamma(u^0), u^0) = x^0$, em (4) vamos ter

$$\frac{\partial g(x^0)}{\partial x_i} + \frac{\partial g(x^0)}{\partial x_1} \frac{\partial \gamma(u^0)}{\partial x_i} = 0. \quad (6)$$

Se x^0 for um ponto de extremo local de f sujeito a restrição $g(x) = 0$, então u^0 é um ponto de extremo local irrestrito de $f(\gamma(u), u)$. Portanto, de (5) temos que

$$\frac{\partial f(x^0)}{\partial x_i} + \frac{\partial f(x^0)}{\partial x_1} \frac{\partial \gamma(u^0)}{\partial x_i} = 0. \quad (7)$$

Das equações (6) e (7), vamos ter a seguinte matriz com seu determinante igual a zero

$$\begin{vmatrix} \frac{\partial f(x^0)}{\partial x_i} & \frac{\partial f(x^0)}{\partial x_1} \\ \frac{\partial g(x^0)}{\partial x_i} & \frac{\partial g(x^0)}{\partial x_1} \end{vmatrix} = 0.$$

Como o determinante de uma matriz A é igual ao determinante de sua transposta A^T , temos o sistema

$$\begin{pmatrix} \frac{\partial f(x^0)}{\partial x_i} & \frac{\partial g(x^0)}{\partial x_i} \\ \frac{\partial f(x^0)}{\partial x_1} & \frac{\partial g(x^0)}{\partial x_1} \end{pmatrix} \begin{pmatrix} v \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Como por hipótese $g_{x_1}(x^0) \neq 0$, u deve ser diferente de zero em uma solução não trivial. Assim, assumindo $v = 1$, obtemos

$$\begin{pmatrix} \frac{\partial f(x^0)}{\partial x_i} & \frac{\partial g(x^0)}{\partial x_i} \\ \frac{\partial f(x^0)}{\partial x_1} & \frac{\partial g(x^0)}{\partial x_1} \end{pmatrix} \begin{pmatrix} 1 \\ w \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (8)$$

Em particular, $\frac{\partial f(x^0)}{\partial x_1} + w \frac{\partial g(x^0)}{\partial x_1} = 0$, então $-w = \frac{f_{x_1}(x^0)}{g_{x_1}(x^0)}$, que de (3) implica em $-w = \lambda$.

Assim, $\begin{pmatrix} \frac{\partial f(x^0)}{\partial x_i} & \frac{\partial g(x^0)}{\partial x_i} \\ \frac{\partial f(x^0)}{\partial x_1} & \frac{\partial g(x^0)}{\partial x_1} \end{pmatrix} \begin{pmatrix} 1 \\ -\lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$. Calculando a multiplicação entre as duas matrizes, chegamos em $f_{x_i}(x^0) - \lambda g_{x_i}(x^0) = 0$, com $1 \leq i \leq n$. □

Suponhamos que $n = 2$, podemos definir a função lagrangeana como $l : \mathbb{R}^2 \times \mathbb{R} \rightarrow \mathbb{R}$ dada por $l(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda \cdot g(x_1, x_2)$. Essa é a função irrestrita $f - \lambda g$, com x^0 sendo o ponto estacionário candidato a extremo local de f sujeita aos pontos de $C = \{x \in \mathbb{R}^n \mid g(x) = 0\}$, satisfazendo as condições $\nabla f(x) = \lambda \cdot \nabla g(x)$ e $g(x) = 0$. Com o Teorema 12, se C for compacto então será verificada a existência de extremos globais para f .

Na figura abaixo a curva $g(x_1, x_2) = 0$ é a nossa restrição de igualdade e em algum momento ocorre o primeiro contato com as curvas de nível da função f . Esse ponto x^0 é o mínimo local da f restrita aos pontos de $g(x_1, x_2) = 0$. Por valer a condição de regularidade, podemos utilizar o Teorema 15 que garante a existência de uma função de derivada contínua $\gamma(u)$ sobre $g = 0$. Como $\nabla g(x_1^0, x_2^0)$ é ortogonal ao vetor tangente de $g = 0$ no ponto x^0 , vamos ter que $\nabla g(x_1^0, x_2^0) \perp \gamma'(u^0)$. Como u^0 é um extremo local irrestrito de $f(\gamma(u), u)$, isso implica em $\nabla f(x^0) \cdot \gamma'(u^0) = 0$, ou seja, $\nabla f(x^0) \perp \gamma'(u^0)$. Portanto, os vetores $\nabla f(x^0)$ e $\nabla g(x^0)$ são paralelos.

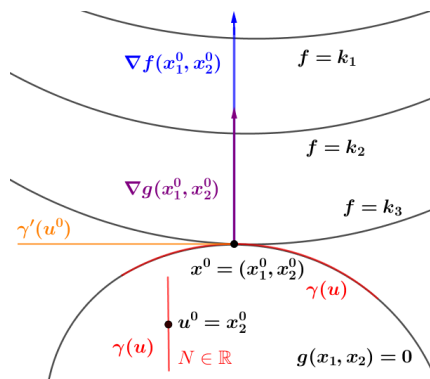


Figura 1: Visualização do Teorema 20 no \mathbb{R}^2
 Fonte: Autor (2022)

4 Otimização restrita e as condições de KKT

Nesta seção, vamos apresentar brevemente as condições de otimalidade de Karush-Kuhn-Tucker (KKT) na otimização com restrição de desigualdade. Os resultados e conceitos dessa seção foram retirados de Luenberger (2008), Silva (2020) e Silva (2022).

O Teorema de KKT é uma modificação do Teorema 20 para a resolução de problemas da otimização com restrições de desigualdade. Foi desenvolvido primeiro por William Karush em 1939, na sua dissertação de mestrado. No entanto, seu reconhecimento pelo teorema só veio na década de 70, depois que o mesmo teorema foi apresentado por Harold W. Kuhn e Albert W. Tucker (HAESER, 2015).

Esse teorema mistura condições de otimalidade da otimização irrestrita, quando o ponto de mínimo local x^0 está no interior do conjunto Ω onde a restrição é dita inativa, e a condição de Lagrange quando o ponto está na restrição ativa. No entanto, as condições de KKT sozinhas não são consideradas condições de otimalidade, é preciso de um hipótese em seu enunciado que permita sua demonstração, chamamos essa hipótese de condição de qualificação.

Existem diversas condições de qualificação, mas a mais famosa é a Condição de Qualificação de Independência Linear. Um ponto satisfaz essa condição de qualificação se os vetores gradientes das restrições de igualdade são linearmente independentes nesse ponto. E essa é também a condição de regularidade do Teorema 20.

Definição 21. Uma restrição de desigualdade $g(x) \leq 0$ com $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ é dita ativa se $g_i(x) = 0$ e inativa se $g_i(x) < 0$, para todo $i \in \{1, 2, \dots, p\}$.

Definição 22. (Condição de qualificação de Independência Linear) Dizemos que a condição de qualificação de independência linear (LICQ) é satisfeita em x^0 quando o conjunto formado pelos vetores gradientes das restrições de igualdade e das restrições de desigualdade ativas é linearmente independente, isto é, $\{\nabla h_i(x^0), i = 1, 2, \dots, m\} \cup \{\nabla g_k(x^0), k \in J(x^0)\}$ é LI, sendo $J(x^0) = \{j \mid g_j(x^0) = 0\}$ o conjunto das restrições ativas de g em x^0 .

Teorema 23 (Condições de Karush-Kuhn-Tucker via LICQ). *Sejam $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$ e $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ funções de classe C^1 , com $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0 \text{ e } g(x) \leq 0\}$. Tomando x^0*

como ponto de mínimo local para o problema

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & h(x) = 0 \\ & g(x) \leq 0 \end{array}$$

e supondo que x^0 é um ponto regular para as restrições. Então, existem os vetores $\lambda \in \mathbb{R}^m$ e $\mu \in \mathbb{R}^p$ com $\mu \geq 0$ tais que

$$\nabla f(x^0) + \lambda^T \nabla h(x^0) + \mu^T \nabla g(x^0) = 0 \quad (9)$$

$$\mu^T g(x^0) = 0. \quad (10)$$

Conclusões

As condições de otimalidade são fundamentais no estudo de algoritmos numéricos iterativos, contribuindo com a verificação aproximada de pontos para os critérios de parada, no estudo da convergência e da robustez. Na otimização irrestrita, as condições necessárias da otimização irrestrita podem ser usadas para encontrar e verificar pontos candidatos a minimizadores locais, no entanto, nem todos eles serão extremos locais. Já a condição suficiente de segunda ordem caso satisfeita, garante que o ponto é um minimizador local, porém nem todo os extremos locais da função satisfazem esse teorema.

O teorema dos multiplicadores de Lagrange permite determinar pontos candidatos a minimizadores locais da função f restrita aos pontos de $g = 0$ através da função lagrangeana. Com a adição dos multiplicadores de Lagrange, podemos modificar o problema com restrição de igualdade para um problema da otimização irrestrita com essa nova função. É semelhante as condições de Lagrange, temos as condições de KKT que busca unir as condições de otimalidade da otimização irrestrita com as da otimização com restrição de igualdade pelo método dos multiplicadores. Porém, KKT necessita de condições de qualificação para ser uma condição de otimalidade, mas assim como as condições de lagrange, podemos usar como hipótese a regularidade do ponto.

Essas condições são importantes no estudo da otimização, principalmente KKT que é considerado o teorema central dessa área científica. Portanto, é fundamental estudar e compreender suas características para a aplicação em algoritmos numéricos.

Referências

ANTON, H.; RORRES, C. **Álgebra linear com aplicações**. 10. ed. Porto Alegre: Bookman, 2012.

FRIEDLANDER, A. **Elementos da Programação não Linear**. Rio de Janeiro: Editora Unicamp, 1994.

FLOR, J. A. R. **Tópicos em Condições de Otimalidade para Otimização não linear**. 2016. Tese (Doutorado em Matemática Aplicada) - Instituto de Matemática e Estatística, Universidade de São Paulo - USP, São Paulo, 2016.

HAESER, G. **Condições de otimalidade de primeira e segunda ordem em otimização não linear**. Instituto de Matemática e Estatística, Universidade de São Paulo - USP. 2015. Disponível em: <https://www.ime.usp.br/~ghaeser/tese.pdf>. Acesso em: 23 set. 2022.

HAESER, G. **Condições sequenciais de otimalidade**. 2009. Tese (Doutorado em Matemática Aplicada) - Instituto de Matemática Estatística e Computação Científica, Universidade Estadual de Campinas - UNICAMP, Campinas, 2009.

LIMA, E. L. **Análise Real**. Rio de Janeiro: IMPA, 2004. v. 2.

LUENBERGER, D. G. **Introduction to Linear and Nonlinear Programming**. 3. ed. Cambridge: Addison-Wesley, 2008.

MARCHAND, L. **Multiplicadores de Lagrange: uma aplicação em problemas de otimização global restrita**. 2016. Trabalho de Conclusão de Curso (Bacharelado em Matemática Aplicada) - Instituto de Matemática, Estatística e Física, Universidade Federal do Rio Grande - FURG, Rio Grande, 2016.

RIBEIRO, A. A.; KARAS, E. W. **Otimização Contínua Aspectos Teóricos e Computacionais**. São Paulo: Cengage Learning, 2013.

SANTOS, J. Grandes Matemáticos: Joseph-Louis Lagrange. **Matemática é Fácil**. 2015. Disponível: <https://www.matematicaefacil.com.br/2018/06/grandes-matematicos-joseph-louis-lagrange.html>. Acesso em: 10 out. 2022.

SILVA, E. J. **Caracterizações integrais para otimalidade global casos mono e multiobjetivo**. 2020. Dissertação (Mestrado em Matemática Aplicada) - Setor de Ciências Exatas, Universidade Federal do Paraná - UFPR, Curitiba, 2020.

SILVA, V. C. **Condições sequenciais de otimalidade e condições de qualificação em otimização não-linear**. 2022. Dissertação (Mestrado em Matemática Aplicada) - Instituto de Biociências, Letras e Ciências Exatas, Universidade Estadual Paulista - UNESP, São José do Rio Preto, 2022.

STEINBRUCH, A. **Álgebra Linear**. São Paulo: McGraw-Hill, 1987.

WILLIAM, F. T. **The Method of Lagrange Multipliers**. São Antônio: Person Education, 2013. Disponível em: http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_LAGRANGE_METHOD.PDF. Acesso em: 13 mar. 2022.

Convergência do método de região de confiança para otimização irrestrita

Alexsandro A. A. de Freitas
Universidade Estadual do Oeste do Paraná
alexsandro.freitas@unioeste.br

Paulo D. Conejo
Universidade Estadual do Oeste do Paraná
paulo.conejo@unioeste.br

Resumo: Neste trabalho apresentamos a demonstração da convergência de um algoritmo simplificado da classe região de confiança irrestrito.

Palavras-chave: Otimização; região de confiança.

1 Introdução

A Otimização é a área da Matemática que trata de problemas cujo interesse consiste em encontrar pontos de máximo ou de mínimo de funções. Um problema de otimização consiste em minimizar ou maximizar uma determinada função, f , denominada função objetivo, sujeita a determinadas condições, Ω , denominadas restrições como em RIBEIRO e KARAS (2013). Um problema de otimização,

$$\begin{array}{ll} \text{minimizar} & f(x) \\ \text{sujeito a} & x \in \Omega, \end{array} \quad (1)$$

em que o denominado conjunto viável $\Omega = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$, com $h : \mathbb{R}^n \rightarrow \mathbb{R}^p$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^q$ e $f : \mathbb{R}^n \rightarrow \mathbb{R}$ funções diferenciáveis. Quando conjunto Ω é o espaço n -dimensional, o problema de otimização é dito irrestrito.

Existe uma classe de algoritmos denominados região de confiança e que resolvem (1). Neste trabalho, apresentamos a convergência de um algoritmo simplificado de região de confiança para problemas de otimização irrestrita.

2 Conceitos Preliminares

Nesta seção inicial definimos alguns elementos indispensáveis para o estudo da convergência do algoritmo de região de confiança a ser apresentado neste texto. Consideramos alguns conceitos relacionados à Álgebra Linear e Topologia como conhecidos pelo leitor. Ao longo do texto, os vetores tomados em \mathbb{R}^n foram sempre considerados vetores colunas. Mais detalhes sobre tais conceitos podem ser encontrados em RIBEIRO e KARAS (2013).

Definição 1. Minimizador Global

Considere uma função $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x^* \in \Omega$. Dizemos que x^* é um minimizador local de f em Ω quando existe $\delta > 0$, tal que $f(x^*) \leq f(x)$, para todo $x \in B(x^*; \delta) \cap \Omega$. Caso $f(x^*) \leq f(x)$, para todo $x \in \Omega$, x^* é dito minimizador global de f em Ω .

Uma sequência de vetores em \mathbb{R}^n é uma função $f : \mathbb{N} \rightarrow \mathbb{R}^n$. Usaremos a notação (f^k) para sequência. Um ponto do $x^* \in \mathbb{R}^n$ é dito estacionário para o problema (1) quando $\nabla f(x^*) = 0$. É possível provar que um ponto de mínimo global é estacionário, logo, quando falamos em convergência de um algoritmo, estamos nos referindo à convergência de uma sequência formada pelos gradientes da função objetivo do problema (1).

Quando restringimos o domínio de uma sequência a um subconjunto infinito dos naturais, temos o que denominamos subsequência. Os pontos limites de uma subsequência são denominados pontos de acumulação da sequência. O maior deles é denominado limite superior da sequência e denotado por \limsup , e o menor, denominado limite inferior e denotado por \liminf . Estudamos resultados de convergência que envolvem subsequências, e neste caso, dizemos que a convergência é fraca.

Uma sequência (a^k) em \mathbb{R} é dita decrescente quando $a^k \geq a^{k+1}$ para todo $k \in \mathbb{N}$. Análogo para crescente. Nestes dois casos, dizemos simplesmente que (a^k) é monótona. Uma sequência é limitada quando todos seus termos são limitados. Por Weiertrass (HANDLER, p. 4, 2018), sabemos que toda sequência monótona limitada é convergente.

Definição 2. Função contínua e continuidade uniforme

Dizemos que uma função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é contínua num ponto a de seu domínio se, e somente se, para qualquer $\epsilon > 0$ escolhido, existir um $\delta > 0$, tal que $\forall x \in X$ tal que $\|x - a\| < \delta$, implicar em $|f(x) - f(a)| < \epsilon$. É notável que para f ser contínua em um ponto, este ponto deve pertecer ao domínio da f .

Enquanto isso, uma função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uniformemente contínua se, e somente se, $\forall \epsilon > 0, \exists \delta > 0$ tal que, $\forall x, y \in X$ tal que $\|x - y\| < \delta$, implique em $|f(x) - f(y)| < \epsilon$.

Veja que a diferença entre continuidade e continuidade uniforme é sutil. Na continuidade, δ depende de ϵ e do ponto escolhido, ou seja, a continuidade é garantida localmente. Já na continuidade uniforme δ depende apenas de ϵ , o que estabelece uma continuidade global.

Definição 3. Função Lipschitz

Uma função $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é dita Lipschitz contínua se existir uma constante $c > 0$ tal que

$$|f(x) - f(y)| \leq c\|x - y\|, \forall x, y \in X.$$

Se $f : X \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função Lipschitz contínua então f é contínua.

Além disso, uma função f é dita localmente Lipschitz contínua se para cada ponto x do domínio existe uma vizinhança $V(x)$ tal que a restrição de f a $V(x)$ é Lipschitz contínua.

Taylor de primeira e segunda ordem

Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função com derivadas parciais de primeira ordem contínuas e $b \in \mathbb{R}^n$. Então podemos escrever

$$f(x) = f(b) + \nabla f(b)^T(x - b) + r(x), \quad (2)$$

com

$$\lim_{x \rightarrow b} \frac{r(x)}{\|x - b\|} = 0.$$

Aqui o operador ∇ indica gradiente de uma função e T indica a transposta de uma matriz.

O polinômio $p_1(x) = f(b) + \nabla f(b)^T(x - b)$ é chamado polinômio de Taylor de ordem 1 da função f . Dentre os polinômios de grau menor ou igual a 1, ele é o que melhor aproxima f . É também o único que satisfaz $p(b) = f(b)$ e $p'(b) = f'(b)$.

O limite nulo em (2) significa que para x próximo de b , o resto $r(x)$ é muito pequeno e vai para zero mais rápido que $\|x - b\|$.

Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função com derivadas parciais de primeira e segunda ordens contínuas e $b \in \mathbb{R}^n$. Então

$$f(x) = f(b) + \nabla f(b)^T(x - b) + \frac{1}{2}(x - b)^T \nabla^2 f(b)(x - b) + r(x), \quad (3)$$

com

$$\lim_{x \rightarrow b} \frac{r(x)}{\|x - b\|^2} = 0.$$

Aqui $\nabla^2 f(b)$ é a matriz Hessiana da f em b .

Teorema 4. Desigualdade de Cauchy-Schwarz

Sejam u, v em \mathbb{R}^n , então

$$|u^T v| \leq \|u\| \|v\|$$

e a igualdade é válida se, e somente se, os elementos u e v são Linearmente Dependentes (LD).

Demonstração da igualdade e desigualdade. Em (BONELLI, 2017, p. 50).

Teorema 5. Teorema do valor médio

Considere $f : \mathbb{R}^n \rightarrow \mathbb{R}$ uma função de classe C^1 e $b \in \mathbb{R}^n$. Então existe $t \in (0, 1)$ tal que

$$f(b + d) = f(b) + \nabla f(b + td)^T d.$$

Demonstração. Em (BIANCHINI & SANTOS, 2018, p. 231).

3 Método de Região de Confiança

O método de região de confiança gera uma sequência (x^k) que, sob certas hipóteses, converge para a solução do problema irrestrito (1). O método utiliza um modelo quadrático que

aproxima a função objetivo, e então define uma região em torno do ponto na qual é possível confiar no modelo.

Vamos considerar uma função $f : \mathbb{R}^n \rightarrow \mathbb{R}$ 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

$$q_k(x) = f(x^k) + \nabla f(x^k)^T (x - x^k) + \frac{1}{2} (x - x^k)^T B_k (x - x^k), \quad (4)$$

onde $B_k \in \mathbb{R}^{n \times n}$ pode ser a hessiana $\nabla^2 f(x^k)$ (3) ou qualquer outra matriz simétrica que satisfaça $\|B_k\| \leq \beta$, para alguma constante $\beta > 0$, independente de $k \in \mathbb{N}$.

Este modelo aproxima suficientemente bem a função f numa vizinhança de x^k , que escrevemos como

$$\{x \in \mathbb{R}^n \mid \|x - x^k\| \leq \Delta_k\},$$

em que confiamos o modelo.

Simplificando a notação, consideramos

$$d = x - x^k \text{ e } m_k(d) = q_k(x^k + d).$$

Iniciamos a aplicação do método resolvendo (possivelmente de maneira aproximada) o problema

$$\begin{aligned} &\text{minimizar} && m_k(d) = f(x^k) + \nabla f(x^k)^T d + \frac{1}{2} (d)^T B_k d \\ &\text{sujeito a} && \|d\| \leq \Delta_k. \end{aligned} \quad (5)$$

Assim obtemos um passo d^k . Em seguida, analisamos este passo. Esperamos que o ponto $x^k + d^k$ provoque uma redução na função objetivo que seja no mínimo, uma fração da redução do modelo. Para formalizar, podemos definir a redução real na função objetivo e a redução predita pelo modelo como

$$ared = f(x^k) - f(x^k + d) \text{ e } pred = m_k(0) - m_k(d).$$

A razão entre a redução na função objetivo e a redução do modelo é dada por

$$\rho_k = \frac{ared}{pred}. \quad (6)$$

Ao resolver o problema (5), o passo d^k obtido é aceito, ou seja, é bom quando a razão ρ_k for maior que uma constante $\eta > 0$ dada. Neste caso, aceitamos o passo, definimos $x^{k+1} = x^k + d^k$ e repetimos o processo, mantendo o raio da região de confiança. Caso isto não ocorra, ou seja, o passo d^k não é aceito, reduzimos o raio Δ_k pela metade e resolvemos o subproblema novamente, agora com o novo raio.

Apresentamos agora o algoritmo de região de confiança, com simplificações em como são aceitos os passos, e atualizados os raios Δ_k das regiões.

Algoritmo - Região de Confiança

Dados: $x^0 \in \mathbb{R}^n$, $\bar{\Delta} > 0$, $\Delta_0 \in (0, \bar{\Delta})$, $\epsilon > 0$ e $k = 0$

REPITA enquanto $\|\nabla f(x^k)\| \geq \epsilon$

Obtenha d^k , solução do subproblema (5)

Calcule ρ_k usando (6)

SE $\rho_k > \frac{1}{4}$

$$x^{k+1} = x^k + d^k$$

$$\Delta_{k+1} = \Delta_k$$

SENÃO

$$\Delta_{k+1} = \frac{\Delta_k}{2}.$$

Veamos como funciona o método de região de confiança geometricamente:

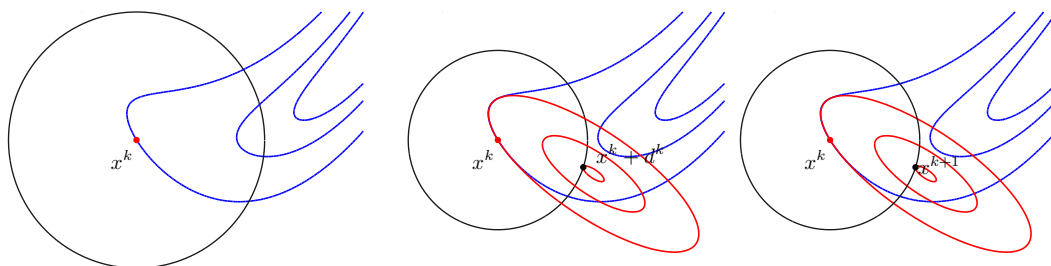


Figura 2: Quando o passo de região de confiança é aceito

Note que na imagem da esquerda na Figura 1, apenas definimos o raio Δ_k e temos o ponto x^k . Na imagem do centro, encontramos o ponto $x^k + d^k$ ao darmos o primeiro passo, veja que este ponto atinge a fronteira da região de confiança e provoca decréscimo na função objetivo, como previsto no algoritmo, chamamos este ponto de x^{k+1} , como exposto na imagem da direita da Figura 1.

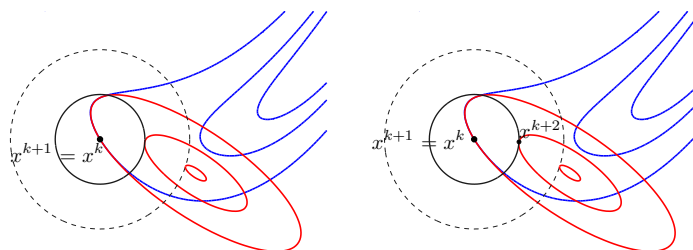


Figura 3: Quando o passo de região de confiança falha

Veja que na imagem da esquerda da Figura 2, o raio escolhido à priori foi grande demais (raio tracejado) e não houve sucesso no passo, logo, o raio foi diminuído pela metade para nova tentativa de passo.

Já na imagem da direita da Figura 2, com o novo raio, que tem metade do tamanho do anterior, o passo teve sucesso e encontramos x^{k+2} . Para prosseguir com os próximos passos, bastaria que repetíssemos esse processo.

4 O passo de Cauchy

O passo de Cauchy é capaz de provocar decréscimo suficiente na função objetivo para termos convergência do algoritmo apresentado. Note que precisamos fazer isso pois nem sempre conseguimos resolver o problema (5) de forma exata, então obtemos uma solução aproximada que seja suficiente para garantir a convergência global do algoritmo.

Denotaremos $g_k = \nabla f(x^k)$. Definimos o passo de Cauchy como o minimizador de m_k ao longo da direção oposta ao gradiente, sujeito a região de confiança, ou seja,

$$d_c^k = -t_k g_k, \quad (7)$$

onde $t_k > 0$ é solução do problema

$$\begin{aligned} \text{minimizar} \quad & m_k(-t g_k) = f(x^k) - t \|g_k\|^2 + \frac{1}{2} t^2 g_k^T B_k g_k \\ \text{sujeito a} \quad & \|t g_k\| \leq \Delta_k. \end{aligned} \quad (8)$$

Iremos supor que a solução aproximada do problema (5) seja ao menos tão boa quanto a solução de Cauchy.

No próximo lema estimamos a redução do modelo no passo de Cauchy.

Lema 6. *Temos que o passo de Cauchy que definimos em (5) satisfaz*

$$m_k(0) - m_k(d_c^k) \geq \frac{1}{2} \|g_k\| \min\{\Delta_k, \frac{\|g_k\|}{\|B_k\|}\}.$$

Demonstração. Inicialmente, obteremos t_k , que é solução de (8), ou seja, o minimizador da função quadrática

$$\xi(t) = f(x^k) - t \|g_k\|^2 + \frac{1}{2} t^2 g_k^T B_k g_k$$

no intervalo $0 \leq t \leq \frac{\Delta_k}{\|g_k\|}$. Para isto consideremos dois casos: $g_k^T B_k g_k > 0$ e $g_k^T B_k g_k \leq 0$.

(i) Se $g_k^T B_k g_k > 0$, então é possível provar que a função ξ é convexa e tem minimizador irrestrito

$$t^* = \frac{\|g_k\|^2}{g_k^T B_k g_k}. \quad (9)$$

Podem ocorrer dois subcasos. O primeiro deles é quando $t^* \leq \frac{\Delta_k}{\|g_k\|}$. Neste caso temos $t_k = t^*$, logo

$$m_k(0) - m_k(d_c^k) = f(x^k) - [f(x^k) + g_k \left(\frac{-\|g_k\|^2}{g_k^T B_k g_k} g_k \right) + \frac{1}{2} \left(\frac{-\|g_k\|^2}{g_k^T B_k g_k} g_k \right) B_k d_c^k] = \frac{1}{2} \frac{\|g_k\|^4}{g_k^T B_k g_k}.$$

Utilizando o Teorema 4, obtemos

$$m_k(0) - m_k(d_c^k) \geq \frac{1}{2} \frac{\|g_k\|^2}{B_k}. \quad (10)$$

Já no segundo subcaso temos $t^* > \frac{\Delta_k}{\|g_k\|}$, o que implica que o minimizador de ξ está na fronteira. Sendo assim, de (9), obtemos

$$t_k = \frac{\Delta_k}{\|g_k\|} < \frac{\|g_k\|^2}{g_k^T B_k g_k},$$

que implica em

$$t_k^2 g_k^T B_k g_k < t_k \|g_k\|^2 = \|g_k\| \Delta_k.$$

Portanto,

$$m_k(d_c^k) < f(x^k) - \|g_k\| \Delta_k + \frac{1}{2} \|g_k\| \Delta_k = f(x^k) - \frac{1}{2} \|g_k\| \Delta_k.$$

Disto, segue que

$$m_k(0) - m_k(d_c^k) > \frac{1}{2} \|g_k\| \Delta_k. \quad (11)$$

(ii) Agora temos o caso em que $g_k^T B_k g_k \leq 0$, que fornece

$$m_k(d_c^k) = f(x^k) - t_k \|g_k\|^2 + \frac{1}{2} t_k^2 g_k^T B_k g_k \leq f(x^k) - t_k \|g_k\|^2. \quad (12)$$

Aqui, a função ξ é decrescente para $t \geq 0$ e assim o ponto de Cauchy também está na fronteira da região de confiança, ou seja, $t_k = \frac{\Delta_k}{\|g_k\|}$. Assim, por (12),

$$m_k(0) - m_k(d_c^k) \geq \|g_k\| \Delta_k \geq \frac{1}{2} \|g_k\| \Delta_k. \quad (13)$$

De (10), (11) e (13) temos que

$$m_k(0) - m_k(d_c^k) \geq \frac{1}{2} \|g_k\| \min \left\{ \Delta_k, \frac{\|g_k\|}{\|B_k\|} \right\}, \quad (14)$$

que demonstra o lema.

5 Convergência

Agora, vamos estabelecer a convergência do método de região de confiança, as demonstrações dos Lemas e Teoremas apresentados nesta seção foram retiradas de (RIBEIRO e KARAS, 2010) e (CORRADI, 1997), mas com as devidas simplificações que sugerimos no algoritmo, em relação a um algoritmo usual de Região de Confiança. Iniciamos supondo que o algoritmo região de confiança gera uma seqüência infinita (x^k) em \mathbb{R}^n e que são satisfeitas as hipóteses listadas abaixo.

H1. A função objetivo f é de classe C^1 , com ∇f Lipschitz.

H2. A solução aproximada d^k de (5) satisfaz

$$pred = m_k(0) - m_k(d^k) \geq c_1 \|\nabla f(x^k)\| \min \left\{ \Delta_k, \frac{\|\nabla f(x^k)\|}{\|B_k\|} \right\},$$

onde $c_1 > 0$ é uma constante.

H3. A função f é limitada inferiormente no conjunto

$$N = \{x \in \mathbb{R}^n | f(x) \leq f(x^0)\}.$$

H4. As matrizes são uniformemente limitadas, ou seja, existe uma constante $\beta > 0$ tal que $\|B_k\| \leq \beta$ para todo $k \in \mathbb{N}$.

Dentre estas, as hipóteses H1, H3 e H4 são comuns em análise de convergência. Perante o Lema 6, a hipótese H2 significa obter um passo cuja redução no modelo seja proporcional a redução obtida pelo passo de Cauchy.

Assim, o primeiro resultado nos dará uma estimativa da razão ρ_k , que definimos em (6).

Lema 7. *Se satisfeitas as hipóteses H1-H5, então existe uma constante $c > 0$ tal que*

$$|\rho_k - 1| \leq \frac{c\Delta_k^2}{\|\nabla f(x^k)\| \min\{\Delta_k, \frac{\|\nabla f(x^k)\|}{\beta}\}}.$$

Demonstração. Pelo Teorema 5, temos

$$f(x^k + d^k) = f(x^k) + \nabla f(x^k + \theta_k d^k)^T d^k,$$

com $\theta_k \in (0, 1)$. Logo,

$$ared - pred = \frac{1}{2}(d^k)^T B_k d^k - (\nabla f(x^k + \theta_k d^k) - \nabla f(x^k))^T d^k.$$

Sabendo que ∇f é Lipschitz e utilizando o Teorema 4, concluímos que existe $c_0 > 0$ tal que

$$|ared - pred| \leq c_0 \Delta_k^2.$$

Assim,

$$|\rho_k - 1| = \left| \frac{ared - pred}{pred} \right| \leq \frac{c_0 \Delta_k^2}{c_1 \|\nabla f(x^k)\| \min\{\Delta_k, \frac{\|\nabla f(x^k)\|}{\beta}\}},$$

que prova o lema para $c = \frac{c_0}{c_1}$.

A seguir, apresentamos um Teorema que nos permite concluir que se uma sequência (x^k) for limitada, segue que ela possui um ponto de acumulação estacionário.

Teorema 8. *Novamente supomos que as hipóteses H1-H4 são satisfeitas, então*

$$\liminf_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0.$$

Demonstração. Supondo por absurdo que isto seja falso, temos que existe $\epsilon > 0$ tal que $\|\nabla f(x^k)\| \geq \epsilon$, para todo $k \in \mathbb{N}$. Considere $\Delta^* = \min\{\frac{\epsilon}{\beta}, \frac{\epsilon}{2c}\}$, onde β e c são as constantes de (12). Logo, se $\Delta_k \leq \Delta^*$, então

$$\Delta_k \leq \frac{\epsilon}{\beta} \leq \frac{\nabla f(x^k)}{\beta} \text{ e } \Delta_k \leq \frac{\epsilon}{2c}.$$

Portanto, pelo Lema 7, temos que

$$|\rho_k - 1| \leq \frac{c\Delta_k}{\epsilon} \leq \frac{1}{2}.$$

Assim, $\rho_k \geq \frac{1}{2} > \frac{1}{4}$ e pelo algoritmo apresentado temos $\Delta_{k+1} \geq \Delta_k$. Logo, sabemos que o raio é reduzido somente se $\Delta_k > \Delta^*$, caso em que $\Delta_{k+1} = \frac{\Delta_k}{2} > \frac{\Delta^*}{2}$. Assim, podemos concluir que

$$\Delta_k \geq \min\{\Delta_0, \frac{\Delta^*}{2}\}, \quad (15)$$

para todo $k \in \mathbb{N}$. Agora, considere o conjunto

$$\lambda = \{k \in \mathbb{N} | \rho_k \geq \frac{1}{4}\}.$$

Dado $k \in \lambda$, pelo mecanismo do algoritmo e pela hipótese H2 temos

$$f(x^k) - f(x^{k+1}) = f(x^k) - f(x^k + d^k) \geq \frac{1}{4}(m_k(0) - m_k(d^k)) \geq \frac{1}{4}c_1\epsilon \min\{\Delta_k, \frac{\epsilon}{\beta}\}.$$

Em vista de (15), temos que existe uma constante $\delta^* > 0$ tal que

$$f(x^k) - f(x^{k+1}) \geq \delta^*, \quad (16)$$

para todo $k \in \lambda$. Por outro lado, a sequência $(f(x^k))$ é não crescente e, por H3, limitada inferiormente, de onde segue que $f(x^k) - f(x^{k+1}) \rightarrow 0$. Portanto, de (16), concluímos que o conjunto λ é finito. Logo, $\rho_k < \frac{1}{4}$, para todo $k \in \mathbb{N}$ suficientemente grande e então Δ_k será reduzido à metade em cada iteração. Isto implica em $\Delta_k \rightarrow 0$, o que contradiz (15). Sendo assim, a afirmação do Teorema é verdadeira.

Agora, podemos provar a convergência global do método de região de confiança.

Teorema 9. *De novo supondo que são satisfeitas as hipóteses H1-H4. Então*

$$\nabla f(x^k) \rightarrow 0.$$

Demonstração. Supondo por absurdo que para algum $\epsilon > 0$ o conjunto

$$\lambda = \{k \in \mathbb{N} | \|\nabla f(x^k)\| \geq \epsilon\}$$

seja infinito. Dado $k \in \lambda$, considere o primeiro índice $\iota_k > k$ tal que $\|\nabla f(x^{\iota_k})\| \leq \frac{\epsilon}{2}$. A existência de ι_k é assegurada pelo Teorema 8. Como ∇f é Lipschitz, temos

$$\frac{\epsilon}{2} \leq \|\nabla f(x^k) - \nabla f(x^{\iota_k})\| \leq L\|x^k - x^{\iota_k}\|, \quad (17)$$

para alguma constante $L > 0$. Portanto,

$$\frac{\epsilon}{2L} \leq \|x^k - x^{\iota_k}\| \leq \sum_{j \in S_k} \|x^j - x^{j+1}\| \leq \sum_{j \in S_k} \Delta_j, \quad (18)$$

onde $S_k = \{j \in \mathbb{N} | k \leq j < \iota_k \text{ e } x^{j+1}x^j\}$. Pelo algoritmo, pela hipótese H2 e pela definição de ι_k , temos

$$f(x^k) - f(x^{\iota_k}) = \sum_{j \in S_k} (f(x^j) - f(x^{j+1})) \geq \sum_{j \in S_k} \eta(m_j(x^j) - m_j(x^{j+1})) \geq \sum_{j \in S_k} \eta c_1 \epsilon \min \left\{ \Delta_j, \frac{\epsilon}{\beta} \right\}.$$

Agora, definindo $\delta^* = \min \left\{ \frac{\eta c_1 \epsilon^2}{2L}, \frac{\eta c_1 \epsilon^2}{\beta} \right\}$ e usando (18), teremos que

$$f(x^k) - f(x^{\iota_k}) \geq \delta^*, \quad (19)$$

para todo $k \in \lambda$. Em contrapartida, a sequência $(f(x^k))$ é não crescente e, por H4, limitada inferiormente, de onde temos que $f(x^k) - f(x^{k+1}) \rightarrow 0$, o que contradiz (19). Sendo assim, o Teorema está provado.

Uma consequência imediata do Teorema 9 é que todo ponto de acumulação de uma sequência gerada pelo algoritmo é estacionário. Veja que, se $x^k \rightarrow \bar{x}$, então a continuidade de ∇f garante que $\nabla f(x^k) \rightarrow \nabla f(\bar{x})$ no universo dos naturais. Por outro lado, pelo 9, temos que $\nabla f(x^k) \rightarrow 0$. Logo, $\nabla f(\bar{x}) = 0$.

Podemos rerepresentar o Lema 7 e o Teorema 8, agora com uma hipótese mais fraca que H1. Faremos isso, no Teorema 11, substituindo a condição de Lipschitz de ∇f pela continuidade uniforme.

Lema 10. *Supondo que f é de classe C^1 e que são satisfeitas as hipóteses H2-H4. Então*

$$|\rho_k - 1| \leq \frac{\Delta_k \left(\frac{\beta}{2} \Delta_k + \limsup_{t \in [0,1]} \{ \|\nabla f(x^k + td^k) - \nabla f(x^k)\| \} \right)}{c_1 \|\nabla f(x^k)\| \min \left\{ \Delta_k, \frac{\|\nabla f(x^k)\|}{\beta} \right\}},$$

onde c_1 e β são as constantes das hipóteses H2 e H4, respectivamente.

Demonstração. Pelo Teorema do valor médio, temos que

$$f(x^k + d^k) = f(x^k) + \nabla f(x^k + \theta_k d^k)^T d^k,$$

com $\theta_k \in (0, 1)$. Logo,

$$ared - pred = \frac{1}{2} (d^k)^T B_k d^k - (\nabla f(x^k + \theta_k d^k) - \nabla f(x^k))^T d^k.$$

Pela desigualdade de Cauchy-Schwarz e pela hipótese H4, temos que

$$|ared - pred| \leq \Delta_k \left(\frac{\beta}{2} \Delta_k + \limsup_{t \in [0,1]} \{ \|\nabla f(x^k + td^k) - \nabla f(x^k)\| \} \right).$$

Note que $|\rho_k - 1| = \left| \frac{ared - pred}{pred} \right|$ e usando H2, provamos o Lema.

Teorema 11. *Supondo f de classe C^1 , com ∇f uniformemente contínua e satisfeitas as hipóteses H2-H4. Então*

$$\liminf_{k \rightarrow \infty} \|\nabla f(x^k)\| = 0.$$

Demonstração. Suponha por absurdo que isto seja falso. Então existe $\epsilon > 0$ tal que $\|\nabla f(x^k)\| \geq \epsilon$, para todo $k \in \mathbb{N}$. Pela continuidade uniforme de ∇f , existe $\delta > 0$ tal que se $\|d^k\| \leq \delta$, então

$$\limsup_{t \in [0,1]} \{\|\nabla f(x^k + td^k) - \nabla f(x^k)\|\} \leq \frac{c_1 \epsilon}{4}. \quad (20)$$

Considere $\Delta^* = \min\{\frac{\epsilon}{\beta}, \delta, \frac{c_1 \epsilon}{2\beta}\}$, onde c_1 e β são as constantes das hipóteses H2 e H4, respectivamente. Se $\Delta_k \leq \Delta^*$, então segue que

$$\Delta_k \leq \frac{\epsilon}{\beta} \leq \frac{\nabla f(x^k)}{\beta}, \Delta_k \leq \delta \text{ e } \frac{\beta \Delta_k}{2c_1 \epsilon} \leq \frac{1}{4}. \quad (21)$$

Portanto, pelo Lema 10, por (20) e (21),

$$|\rho_k - 1| \leq \frac{1}{c_1 \epsilon} \left(\frac{\beta}{2} \Delta_k + \frac{c_1 \epsilon}{4} \right) = \frac{\beta \Delta_k}{2c_1 \epsilon} + \frac{1}{4} \leq \frac{1}{2}.$$

Assim, temos que $\rho_k \geq \frac{1}{2} > \frac{1}{4}$ e pelo algoritmo $\Delta_{k+1} \geq \Delta_k$. Isto quer dizer que o raio é reduzido somente se $\Delta_k > \Delta^*$, caso onde $\Delta_{k+1} = \frac{\Delta_k}{2} > \frac{\Delta^*}{2}$. Disso, concluímos que

$$\Delta_k \geq \min \left\{ \Delta_0, \frac{\Delta^*}{2} \right\}, \quad (22)$$

para todo $k \in \mathbb{N}$. Consideremos agora o conjunto

$$\lambda = \left\{ k \in \mathbb{N} \mid \rho_k \geq \frac{1}{4} \right\}.$$

Dado $k \in \lambda$, pelo algoritmo e pela hipótese H2 temos

$$f(x^k) - f(x^{k+1}) = f(x^k + d^k) \geq \frac{1}{4} (m_k(0) - m_k(d^k)) \geq \frac{1}{4} c_1 \epsilon \min \left\{ \Delta_k, \frac{\epsilon}{\beta} \right\}.$$

Por (22), temos que existe uma constante $\delta^* > 0$ tal que

$$f(x^k) - f(x^{k+1}) \geq \delta^*, \quad (23)$$

para todo $k \in \lambda$. Por outro lado, a sequência $(f(x^k))$ é não crescente e, por H3, limitada inferiormente, de onde temos que $f(x^k) - f(x^{k+1}) \rightarrow 0$. Portanto, concluímos de (23) que o conjunto λ é finito. Assim, $\rho_k < \frac{1}{4}$, para todo $k \in \mathbb{N}$ suficientemente grande e então Δ_k será reduzido à metade em cada iteração. Isto implica em $\Delta_k \rightarrow 0$, o que contradiz (22), provando a afirmação do Teorema.

6 Conclusão

Neste trabalho, estudamos a convergência de um algoritmo simplificado de região de confiança que converge fraca e fortemente (Teoremas 8 e 9) para um ponto estacionário do problema de otimização irrestrita. Ainda, como apresentado no Teorema 11, quando enfraquecemos as hipóteses considerando apenas continuidade uniforme na função objetivo, podemos estabelecer uma convergência fraca a pontos estacionários.

Referências

- BIANCHINI, Waldecir; SANTOS, Angela Rocha dos. **Aprendendo cálculo de uma variável**. Rio de Janeiro: Ufrj, 2018. 404 p. Disponível em: <http://www.im.ufrj.br/waldecir/calculo1/aprendendocalculo/calculo1.pdf>. Acesso em: 15 out. 2022.
- BONELLI, Rebeca Cristina. **Desigualdades Matemáticas e Aplicações**. 2017. 144 f. Dissertação (Mestrado) - Curso de Matemática, Instituto de Geociências e Ciências Exatas, Universidade Estadual Paulista Júlio de Mesquita Filho, Rio Claro, 2017. Cap. 3. Disponível em: https://repositorio.unesp.br/bitstream/handle/11449/151180/bonelli_rebeca_me_rcla.pdf. Acesso: 15 out. 2022.
- CORRADI, G. (1997) A trust region algorithm for unconstrained optimization. **Int. J. Comput. Math.** 65, 109-119.
- RIBEIRO, A. A.; KARAS, E. W. (2013). **Otimização Contínua Aspectos Teóricos e Computacionais**. São Paulo: Cengage Learning.

Tópicos de Álgebra Linear e o Teorema de Convergência da Rede Neural Perceptron

David Hapner Barzotto¹
Universidade Estadual do Oeste do Paraná
dhbarzotto@hotmail.com

Simone Aparecida Miloca
Universidade Estadual do Oeste do Paraná
simone.miloca@unioeste.br

Resumo: A Álgebra Linear tem uma aplicabilidade em um vasto campo das ciências puras e aplicadas, como matemática, física, computação e as engenharias. Este trabalho tem o intuito de mostrar resultados algébricos utilizados na demonstração do teorema de convergência da primeira grande Rede Neural Artificial (RNA), o Perceptron. Mostraremos que o Perceptron sempre irá encontrar um hiperplano separador para a classificação binária de dados.

Palavras-chave: Álgebra Linear; Perceptron; Redes Neurais Artificiais (RNAs).

1 Introdução

As Redes Neurais Artificiais (RNAs) fazem parte do campo de estudo da Inteligência Artificial (IA), porém, também estão presentes em outras áreas do conhecimento tais como computação evolucionária, metaheurística e machine learning, sendo capazes de resolver os mais diversos problemas, como classificação, regressão e clusterização. Possuem inspiração nas Redes Neurais Biológicas (RNBs) e seus principais filamentos, como o dendrito, o corpo celular e o axônio. A Figura 4 exemplifica um modelo biológico e um artificial com neurônios e conexões direcionadas entre eles. Devido esta motivação biológica, os elementos de processamento de uma rede neural são denominados neurônios ou nós, sendo uma unidade de processamento de informações fundamental para o funcionamento da rede. A partir disso, pode-se definir três conceitos básicos que caracterizam uma rede neural artificial:

- Um padrão da rede chamado de pesos sinápticos, também conhecidos por conexões sinápticas que, unidos com uma operação binária, combinam os dados de entrada pertencentes a um vetor \mathbf{x} com um vetor dos pesos sinápticos \mathbf{w} montando um hiperplano separador das classes dos dados na forma de $\mathbf{w}^T \mathbf{x} = 0$;
- Uma regra de agregação de dados que correlaciona as entradas dadas aos neurônios, ponderados com suas respectivas conexões sinápticas;
- Uma função chamada de ativação com o intuito de modelar o cálculo dos dados podendo gerar funções não lineares ou condicionar a saída do neurônio em um intervalo determinado.

¹Agradecemos a Fundação Araucária pela bolsa no projeto PIBIC.

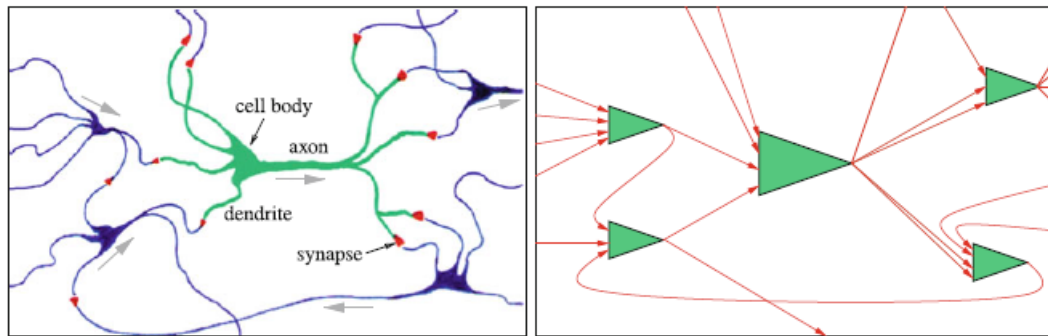


Figura 4: Modelo biológico e modelo artificial de uma rede neural.
Fonte: Extraído de Ertel (2017).

A Figura 5 exemplifica o modelo de um neurônio. Tem-se X_1, X_2, \dots, X_m sinais de entrada conectados ao neurônio k que são multiplicados pelos pesos sinápticos, $W_{k1}, W_{k2}, \dots, W_{km}$, incluindo uma variável polarização (bias) b_k que é adicionada ao somatório da função de ativação φ , com o intuito de aumentar o grau de liberdade desta função e, conseqüentemente, a capacidade de aproximação da rede.

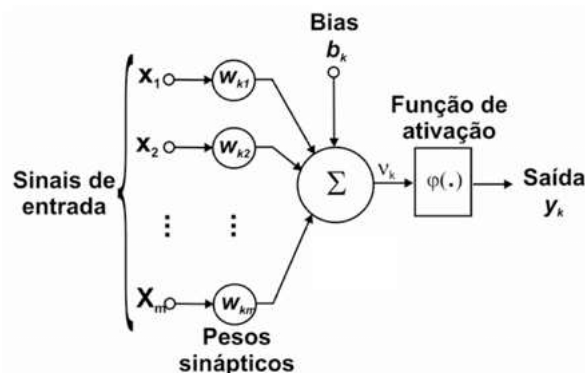


Figura 5: Modelo de um Neurônio Artificial.
Fonte: Extraído de Haykin (2001).

Matematicamente, um neurônio k é escrito como

$$y_k = \varphi\left(\sum_{j=1}^m W_{kj}X_j + b_k\right)$$

onde, $\varphi(\cdot)$ é a função de ativação e y_k são os sinais de saída do neurônio.

Segundo Haykin (2001), o Perceptron é a rede neural mais simples. Foi construído com o foco de solucionar classificações binárias e é uma rede considerada linear, de uma camada apenas, ou seja, supondo que se tenha um conjunto de entradas e se conheça suas respectivas saídas, o Perceptron deve aprender os padrões e classificar corretamente novas entradas. Este trabalho tem por objetivo apresentar o teorema de convergência do Perceptron com a adaptação baseada em Haykin (2001), escrevendo sobre os principais conceitos de Álgebra Linear necessários para a prova do teorema.

2 Tópicos de Álgebra Linear

Esta seção tem por objetivo escrever conceitos e resultados de Álgebra Linear utilizados na parte que envolve o Teorema de Convergência do Perceptron.

O texto que segue refere-se a espaços vetoriais reais. Temos interesse particular no espaço vetorial euclidiano, que é um espaço vetorial real, de dimensão finita, em que está definido um produto interno. As principais referências foram Poole (2003) e Steinbruch e Winterle (1995).

2.1 Dependência Linear em Espaços Vetoriais

Definição 1. Sejam $\{v_1, v_2, v_3, \dots, v_n\}$ um conjunto de vetores de um espaço vetorial V e os escalares $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$, temos que:

$$v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \dots + \alpha_n v_n$$

é dito uma combinação linear dos vetores $v_1, v_2, v_3, \dots, v_n$.

Exemplo 1. Sejam $\mathbf{v}_1 = (1, 5)$ e $\mathbf{v}_2 = (4, -1)$. Então $\mathbf{v} = (14, 7)$ é uma combinação linear de \mathbf{v}_1 e \mathbf{v}_2 pois $\mathbf{v} = (14, 7) = 2\mathbf{v}_1 + 3\mathbf{v}_2 = 2(1, 5) + 3(4, -1)$.

A definição a seguir apresenta o conceito de dependência e independência linear entre vetores.

Definição 2. Seja V um espaço vetorial e $S = \{v_1, v_2, \dots, v_n\} \subset V$. S é dito linearmente dependente (LD) se existirem escalares $\alpha_1, \alpha_2, \dots, \alpha_n$, com pelo menos um $\alpha_i \neq 0$ tais que

$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0.$$

S é dito linearmente independente (LI) se não for linearmente dependente.

2.2 Produto Interno e Norma

Definição 3. Seja V um espaço vetorial real. Um produto interno é uma função

$$\begin{aligned} \langle \cdot, \cdot \rangle : V \times V &\longrightarrow \mathbb{R} \\ (u, v) &\longmapsto \langle u, v \rangle \end{aligned}$$

que satisfaz as propriedades, para todos u, v, w em V e $\alpha \in \mathbb{R}$

- i) $\langle u, u \rangle \geq 0$, e $\langle u, u \rangle = 0 \Leftrightarrow u = 0$
- ii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$
- iii) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$
- iv) $\langle u, v \rangle = \langle v, u \rangle$

O Exemplo 2 é um produto interno, também chamado de produto escalar usual de vetores em \mathbb{R}^n .

Exemplo 2. Sejam $\mathbf{u} = (x_1, x_2, \dots, x_n)$ e $\mathbf{v} = (y_1, y_2, \dots, y_n)$. Tem-se

$$\langle \mathbf{u}, \mathbf{v} \rangle = (x_1 y_1 + x_2 y_2 + \dots + x_n y_n) = \mathbf{u} \mathbf{v}^T.$$

Em um espaço vetorial V com produto interno, podemos definir norma de um vetor $v \in V$ que provém deste produto interno como

$$\|v\| = \sqrt{\langle v, v \rangle}. \quad (1)$$

Exemplo 3. Em \mathbb{R}^2 , seja $\mathbf{v} = (x_1, x_2)$. A norma é dada por,

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{x_1^2 + x_2^2}.$$

Assim, temos que dado um vetor $\mathbf{v} = (x_1, x_2, x_3, \dots, x_n)$ em \mathbb{R}^n , a norma euclidiana é definida por

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{x_1^2 + x_2^2 + x_3^2 + \dots + x_n^2}.$$

2.3 Desigualdade de Cauchy-Schwarz

Teorema 4. *Seja V um espaço vetorial com produto interno, então:*

$$|\langle u, v \rangle| \leq \|u\| \|v\|, \forall u, v \in V.$$

Prova. Os vetores u e v serão LD ou LI. Faremos a prova considerando cada uma destas situações.

Caso 1. Sejam u e v linearmente dependentes, logo temos $\alpha \in \mathbb{R}$ tal que $u = \alpha v$. Então,

$$|\langle u, v \rangle| = |\langle \alpha v, v \rangle| = |\alpha| \langle v, v \rangle = |\alpha| (\sqrt{\langle v, v \rangle})^2 = |\alpha| \sqrt{\langle v, v \rangle} \sqrt{\langle v, v \rangle}$$

usando a propriedade (ii) descrita na Definição 3 e a igualdade (1), temos,

$$|\alpha| \sqrt{\langle v, v \rangle} \sqrt{\langle v, v \rangle} = \sqrt{\alpha^2 \langle v, v \rangle} \sqrt{\langle v, v \rangle} = \sqrt{\langle \alpha v, \alpha v \rangle} \sqrt{\langle v, v \rangle} = \sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle} = \|u\| \|v\|.$$

Segue que vale a igualdade,

$$|\langle u, v \rangle| = \|u\| \|v\|.$$

Caso 2. Sejam u e v linearmente independentes, logo, existe $\alpha \in \mathbb{R}$ em que $u + \alpha v \neq \mathbf{0}$, então $\langle u + \alpha v, u + \alpha v \rangle > 0$, com isso,

$$\langle u + \alpha v, u + \alpha v \rangle = \langle u, u \rangle + \langle u, \alpha v \rangle + \langle \alpha u, v \rangle + \langle \alpha v, \alpha v \rangle > 0.$$

Pela Definição 3 teremos

$$\langle u, u \rangle + \alpha \langle u, v \rangle + \alpha \langle u, v \rangle + \alpha^2 \langle v, v \rangle = \langle u, u \rangle + 2\alpha \langle u, v \rangle + \alpha^2 \langle v, v \rangle > 0.$$

A inequação de grau 2 está em função de α e não apresenta raízes reais, apenas complexas, o seu discriminante Δ deve ser menor que 0, ou seja,

$$\Delta < 0 \Rightarrow (2\langle u, v \rangle)^2 - 4(\langle u, u \rangle)(\langle v, v \rangle) < 0 \Rightarrow 4\langle u, v \rangle^2 < 4(\langle u, u \rangle)(\langle v, v \rangle) \Rightarrow \langle u, v \rangle^2 < \langle u, u \rangle \langle v, v \rangle$$

podemos extrair a raiz quadrada da inequação,

$$\sqrt{\langle u, v \rangle^2} < \sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle} \Rightarrow \langle u, v \rangle < \sqrt{\langle u, u \rangle} \sqrt{\langle v, v \rangle}.$$

Por (1) temos,

$$|\langle u, v \rangle| < \|u\| \|v\|.$$

Pelo Caso 1 e Caso 2 temos a validade da desigualdade de Cauchy-Schwarz,

$$|\langle u, v \rangle| \leq \|u\| \|v\|.$$

□

2.4 Hiperplanos

Definição 5. Um hiperplano H do \mathbb{R}^n é o conjunto

$$H = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid a_1x_1 + a_2x_2 + \dots + a_nx_n = b ; a_1, \dots, a_n \in \mathbb{R}\}.$$

As Figuras 6 e 7 exemplificam hiperplanos em \mathbb{R}^2 (retas).

Um hiperplano divide \mathbb{R}^n em dois semiespaços H_1 e H_2 .

$$H_1 = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid a_1x_1 + a_2x_2 + \dots + a_nx_n \geq b ; a_1, \dots, a_n \in \mathbb{R}\}$$

$$H_2 = \{(x_1, x_2, \dots, x_n) \in \mathbb{R}^n \mid a_1x_1 + a_2x_2 + \dots + a_nx_n \leq b ; a_1, \dots, a_n \in \mathbb{R}\}.$$

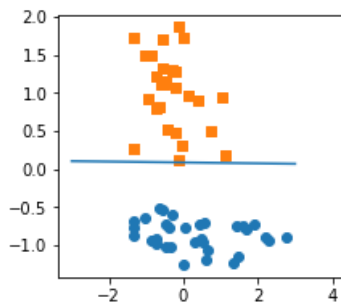


Figura 6: Exemplo de Hiperplano.

Fonte: Produzido pelo autor.

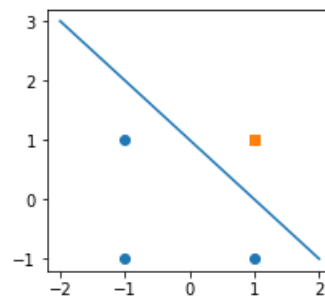


Figura 7: Hiperplano para a função AND booleana.

Fonte: Produzido pelo autor.

Um hiperplano pode ser escrito na notação de produto escalar, como

$$\mathbf{a}^T \mathbf{x} = 0$$

onde $\mathbf{a} = (b, a_1, \dots, a_m)^T$ e $\mathbf{x} = (1, x_1, \dots, x_m)^T$.

A próxima definição traz o conceito de conjuntos linearmente separáveis.

Definição 6. Dois conjuntos $C_1 \subset \mathbb{R}^{m+1}$ e $C_2 \subset \mathbb{R}^{m+1}$ são linearmente separáveis se existir $\mathbf{a} = (b, a_1, \dots, a_m)^T$, onde b, a_1, \dots, a_m são números reais, tais que

$$\mathbf{a}^T \mathbf{x} > 0, \forall \mathbf{x} = [1, x_1, \dots, x_m]^T \in C_1$$

e

$$\mathbf{a}^T \mathbf{x} \leq 0, \forall \mathbf{x} = [1, x_1, \dots, x_m]^T \in C_2$$

em que b é o coeficiente linear do hiperplano, que tem o nome de bias quando se trata de hiperplanos de RNAs.

3 Rede Neural Perceptron

Formalmente descrito por Fausett (1994), Haykin (2001) e Kovács (2006), o Perceptron, criado e idealizado por Frank Rosenblatt, segue os parâmetros da álgebra booleana, ou seja, AND e OR, logo um senso binário $\{0,+1\}$ ou bipolar $\{-1,+1\}$. No caso do Perceptron, seu algoritmo iterativo de ajuste de pesos é calculado como mostra o Algoritmo 1. A ideia principal do Perceptron é classificar duas classes linearmente separáveis, C_1 e C_2 . Conseqüentemente, tal classificação só é possível pois a rede gera um hiperplano separador em m -dimensões descrito como $\sum_{i=0}^m w_i(n)x_i(n) + b(n) = 0$ ou na forma de produto interno $\mathbf{w}^T(n)\mathbf{x}(n) = 0$, em que n é o número de iterações feitas. O funcionamento se dá da seguinte forma:

[1] Caso o n -ésimo termo de \mathbf{x} seja classificado corretamente na n -ésima iteração, então o vetor \mathbf{w} não é corrigido.

$$\mathbf{w}(n+1) = \mathbf{w}(n), \text{ se } \mathbf{w}^T \mathbf{x} > 0. \text{ Logo } \mathbf{x}(n) \in C_1 \quad (2)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n), \text{ se } \mathbf{w}^T \mathbf{x} \leq 0. \text{ Logo } \mathbf{x}(n) \in C_2. \quad (3)$$

[2] Se o passo 1 for falso, faz-se as seguintes atualizações.

$$\mathbf{w}(n+1) = \mathbf{w}(n) - \eta(n)\mathbf{x}(n), \text{ se } \mathbf{w}^T \mathbf{x} > 0. \text{ Logo } \mathbf{x}(n) \in C_2 \quad (4)$$

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \eta(n)\mathbf{x}(n), \text{ se } \mathbf{w}^T \mathbf{x} \leq 0. \text{ Logo } \mathbf{x}(n) \in C_1 \quad (5)$$

onde η é um parâmetro denominado taxa de aprendizado e geralmente é um valor entre zero e um.

3.1 Exemplo de Funcionamento do Perceptron

Exemplo 4. Utilize a rede neural Perceptron para classificar a função lógica ?AND?, cujas entradas e saídas são mostradas na Tabela 1. Considere a taxa de aprendizado η como sendo fixo e igual a 1, o valor de entrada para o bias também 1, com peso $b(n) = 0$.

Primeira Iteração

1. Primeiro padrão E1 é apresentado à rede.

$$y(1) = b + \sum_{i=1}^2 w_i(1)x_i(1) = 0 + 0 \times 0 + 0 \times 0 = 0, \text{ assim, } d(1) = y(1) = 0.$$

Como $d(1) = y(1)$, os pesos não se atualizam, logo,

$$w_1(2) = w_1(1) = 0, w_2(2) = w_2(1) = 0, b(2) = b(1) = 0.$$

2. Segundo padrão E2 é apresentado à rede.

$$y(2) = b(2) + \sum_{i=1}^2 w_i(2)x_i(2) = 0 + 0 \times 0 + 0 \times 1 = 0. \text{ Assim, } d(2) = y(2) = 0.$$

Como $d(2) = y(2)$, os pesos não se atualizam, assim,

$$w_1(3) = w_1(2) = 0, w_2(3) = w_2(2) = 0, b(3) = b(2) = 0.$$

Entrada: Inicializar o vetor de pesos \mathbf{w} e o bias b ;
 Inicie a taxa de aprendizado η ($0 < \eta \leq 1$);
while Critério de Parada não satisfeito **do**
 for Para cada par de dados de treinamento (x_i, d_i) **do**
 Execute ;
 $y = b + \sum_{i=1}^n w_i x_i$;
 if $y = d$ **then**
 ocorre [1] ;
 $w_i(\text{novo}) = w_i(\text{atual})$;
 $b(\text{novo}) = b(\text{atual})$;
 else
 ocorre [2] ;
 Atualizar os pesos e a tendência ;
 $w_i(\text{novo}) = w_i(\text{atual}) + \eta x_i$;
 $b(\text{novo}) = b(\text{atual}) + \eta$;
 end
 end
 Teste a condição de parada.
end

Algoritmo 1: Algoritmo *Perceptron*

Tabela 1: Tabela para a função AND

	Entrada		Saída
	$x_1(n)$	$x_2(n)$	d_i
E1	0	0	0
E2	0	1	0
E3	1	0	0
E4	1	1	1

Fonte: Produzido pelo autor.

3. Terceiro padrão E3 é apresentado à rede.

$$y(3) = b(3) + \sum_{i=1}^2 w_i(3)x_i(3) = 0 + 0 \times 1 + 0 \times 0 = 0, \text{ assim, } d(3) = y(3) = 0.$$

Como $d(3) = y(3)$, os pesos não se atualizam, logo,

$$w_1(4) = w_1(3) = 0, \quad w_2(4) = w_2(3) = 0, \quad b(4) = b(3) = 0.$$

4. Quarto padrão E4 é apresentado à rede.

$$y(4) = b(4) + \sum_{i=1}^2 w_i(4)x_i(4) = 0 + 0 \times 1 + 0 \times 1 = 0, \text{ logo, } d(4) \neq y(4).$$

Como $d(4) \neq y(4)$, os pesos são atualizados, logo temos,

$$w_1(5) = w_1(4) + \eta(4)x_1(4) = 0 + 1 \times 1 = 1$$

$$w_2(5) = w_2(4) + \eta(4)x_2(4) = 0 + 1 \times 1 = 1$$

$$b(5) = b(4) + \eta(4) \times 1 = 0 + 1 \times 1 = 1.$$

A Figura 8 mostra os resultados da primeira iteração. Após 10 épocas (iterações) tem-se o hiperplano separador que é mostrado na Figura 9, cuja equação final fica $2x_1 + x_2 - 2 = 0$.

$d(n)$	$x_1(n)$	$x_2(n)$	$w_1(n)$	$w_2(n)$	$bias$
0	0	0	0	0	0
0	0	1	0	0	0
0	1	0	0	0	0
1	1	1	0	0	0
-	-	-	1	1	1

Figura 8: 1ª Iteração do Perceptron.
 Fonte: Produzido pelo autor.

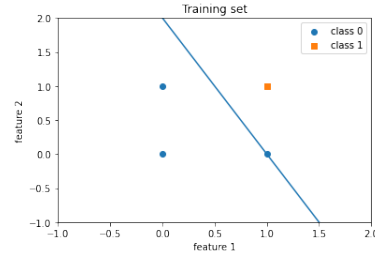


Figura 9: Hiperplano Perceptron - função AND.
 Fonte: Produzido pelo autor.

4 Teorema de Convergência do Perceptron

Teorema 7. *Sejam \mathbf{x} um vetor de entrada, \mathbf{w} o vetor peso, 1 o valor de entrada para o bias e n o passo por iteração. Tome as classes C_1 e C_2 linearmente separáveis por um hiperplano $\mathbf{w}^T \mathbf{x} = 0$. Então o algoritmo Perceptron converge para um n_{max} de iterações para qualquer inicialização do vetor de pesos \mathbf{w} , sem qualquer critério de parada preestabelecido, com isso gerando um hiperplano separador através de um n_{max} :*

$$n_{max} = \frac{\beta \|\mathbf{w}_0\|^2}{\alpha^2}.$$

Nossa notação considera

$$\mathbf{x}(n) = [1, x_1(n), x_2(n), x_3(n), \dots, x_m(n)]^T \quad \text{e} \quad \mathbf{w}(n) = [b(n), w_1(n), w_2(n), w_3(n), \dots, w_m(n)]^T$$

onde n é a n -ésima iteração e a saída da combinação linear é dada por

$$y(n) = \sum_{i=0}^m w_i(n) x_i(n) = \mathbf{w}^T(n) \mathbf{x}(n).$$

Definimos $w_0(n) = b(n)$. Para n fixo, $\mathbf{w}^T \mathbf{x} = 0$ descreve um hiperplano separador de duas classes C_1 e C_2 . Como os dados de entrada das duas classes C_1 e C_2 são linearmente separáveis, então existe um subconjunto de vetores de treinamento $\mathbf{x}_1(n) = X_1 \in C_1$ e outro subconjunto $\mathbf{x}_2(n) = X_2 \in C_2$ sendo $X_1 \cup X_2 = X$, o conjunto X de treinamento completo. Dados tais vetores de entrada, os pesos serão ajustados de modo que as classes sejam linearmente separáveis,

$$\mathbf{w}^T \mathbf{x} > 0, \quad \forall x \in C_1 \quad \text{e} \quad \mathbf{w}^T \mathbf{x} \leq 0, \quad \forall x \in C_2.$$

Prova. Seja $\eta(n) = \eta > 0$, em que η é uma constante independente do número de iterações. Tomemos $\eta = 1$, para termos a regra de adaptação com incremento fixo para o Perceptron. Considere a condição inicial $\mathbf{w}(0) = 0$ e suponha que $\mathbf{w}^T(n) \mathbf{x}(n) < 0$ e $\mathbf{x}(n) \in X_1$. Ou seja, considere por hipótese que o Perceptron classifica incorretamente os vetores de entrada \mathbf{x} . Como a parte 1 do algoritmo de aprendizado é falsa, vamos para a segunda parte do Algoritmo de aprendizado, que é a atualização dos pesos,

$$\mathbf{w}(n+1) = \mathbf{w}(n) + \mathbf{x}(n), \quad \forall \mathbf{x}(n) \in C_1.$$

Dado a condição de início $\mathbf{w}(0) = 0$, temos que

$$\begin{aligned}\mathbf{w}(1) &= \mathbf{w}(0) = 0 \\ \mathbf{w}(2) &= \mathbf{w}(1) + \mathbf{x}(1) \\ \mathbf{w}(3) &= \mathbf{w}(2) + \mathbf{x}(2) = \mathbf{w}(1) + \mathbf{x}(1) + \mathbf{x}(2) = 0 + \mathbf{x}(1) + \mathbf{x}(2)\end{aligned}$$

assim, iterativamente temos,

$$\mathbf{w}(n+1) = \mathbf{x}(1) + \mathbf{x}(2) + \mathbf{x}(3) + \cdots + \mathbf{x}(n) = \sum_{i=1}^n \mathbf{x}(i). \quad (6)$$

Como assumimos que C_1 e C_2 são linearmente separáveis, então existe uma solução \mathbf{w}_0 de $\mathbf{w}^T \mathbf{x} \geq 0$ para $\mathbf{x}(n) \in X_1$. Assim, podemos definir um número positivo α , tal que

$$\alpha = \min_{\mathbf{x}(n) \in X_1} \mathbf{w}_0^T \mathbf{x}(n). \quad (7)$$

Multiplicando ambos os lados da equação 6 por \mathbf{w}_0^T , temos,

$$\mathbf{w}_0^T \mathbf{w}(n+1) = \mathbf{w}_0^T \sum_{i=1}^n \mathbf{x}(i). \quad (8)$$

Pela equação 7 e a equação 8, tem-se a desigualdade

$$\mathbf{w}_0^T \mathbf{w}(n+1) \geq n\alpha. \quad (9)$$

Agora, utilizando a desigualdade de Cauchy-Schwarz (Teorema 4), para os vetores \mathbf{w}_0 e $\mathbf{w}(n+1)$, teremos a desigualdade,

$$\|\mathbf{w}_0\| \|\mathbf{w}(n+1)\| \geq \mathbf{w}_0^T \mathbf{w}(n+1)$$

e elevando ao quadrado ambos os lados, encontramos,

$$\|\mathbf{w}_0\|^2 \|\mathbf{w}(n+1)\|^2 \geq [\mathbf{w}_0^T \mathbf{w}(n+1)]^2. \quad (10)$$

Como vale a desigualdade 9, notando que ambos os membros são não negativos, teremos,

$$\|\mathbf{w}_0\|^2 \|\mathbf{w}(n+1)\|^2 \geq [\mathbf{w}_0^T \mathbf{w}(n+1)]^2 \geq n^2 \alpha^2.$$

Assim encontramos

$$\begin{aligned}\|\mathbf{w}_0\|^2 \|\mathbf{w}(n+1)\|^2 &\geq n^2 \alpha^2 \\ \|\mathbf{w}(n+1)\|^2 &\geq \frac{n^2 \alpha^2}{\|\mathbf{w}_0\|^2}.\end{aligned} \quad (11)$$

A inequação 11 nos diz que toda parte acima da curva da função quadrática é uma possível solução mínima, como exemplifica a parte (a) da Figura 10.

Por outro lado, a partir da equação 4 temos que,

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mathbf{x}(k), \text{ para } k = 1, \dots, n, \mathbf{x}(k) \in X_1. \quad (12)$$

Calculando a norma euclidiana temos

$$\|\mathbf{w}(k+1)\| = \|\mathbf{w}(k) + \mathbf{x}(k)\|. \quad (13)$$

Elevando ao quadrado ambos os lados da equação 13 temos que,

$$\|\mathbf{w}(k+1)\|^2 = \|\mathbf{w}(k) + \mathbf{x}(k)\|^2 = \|\mathbf{w}(k)\|^2 + 2\mathbf{w}^T(k)\mathbf{x}(k) + \|\mathbf{x}(k)\|^2.$$

Da suposição que o Perceptron classifica erroneamente o vetor de entrada $\mathbf{x}(k) \in X_1$, temos $\mathbf{w}^T(k)\mathbf{x}(k) < 0$, logo, o fator $2\mathbf{w}^T(k)\mathbf{x}(k)$ é negativo, assim obtemos,

$$\|\mathbf{w}(k+1)\|^2 \leq \|\mathbf{w}(k)\|^2 + \|\mathbf{x}(k)\|^2, \quad k = 1, 2, 3, \dots, n$$

ou de forma equivalente,

$$\|\mathbf{w}(k+1)\|^2 - \|\mathbf{w}(k)\|^2 \leq \|\mathbf{x}(k)\|^2, \quad k = 1, 2, 3, \dots, n.$$

Como k é definido a partir de n e da hipótese $\mathbf{w}(0) = \mathbf{0}$, temos que existe $\beta > 0$ tal que

$$\|\mathbf{w}(n+1)\|^2 \leq \sum_{k=1}^n \|\mathbf{x}(k)\|^2 \leq n\beta$$

logo,

$$\|\mathbf{w}(n+1)\|^2 \leq n\beta. \tag{14}$$

A equação 14 nos mostra que a solução da curva linear é delimitada pela parte inferior da função, como é possível ver na parte (b) da Figura 10. Assim, define-se que β é um valor positivo dado pela equação 15

$$\beta = \max_{\mathbf{x}(k) \in X_1} \|\mathbf{x}(k)\|^2. \tag{15}$$

Contudo, a inequação 14 é contrária a 11 para valores suficientemente grandes de n , como exemplifica a Figura 10. A inequação 14 diz que os pesos tendem a crescer infinitamente, o que de fato é impossível. A partir disso, inferimos que n não pode ser maior que um número máximo. Assim, ambas as inequações são satisfeitas com o sinal de igualdade, como ilustrado na parte (c) da Figura 10, em que o eixo das ordenadas é definido por $\|\mathbf{w}(n)\|^2$ e o eixo das abscissas é definido pelo número de iterações n da rede. Portanto,

$$\frac{n_{max}^2 \alpha^2}{\|\mathbf{w}_0\|^2} = n_{max} \beta \Leftrightarrow n_{max} = \frac{\beta \|\mathbf{w}_0\|^2}{\alpha^2}. \tag{16}$$

A prova foi feita para para $\eta(n) = 1$ pois facilita os cálculos, já que seu valor não é importante desde que seja positivo, e, caso $\eta \neq 1$, apenas escala os vetores padrões sem afetar seu hiperplano separador. Existindo um vetor solução \mathbf{w}_0 , a atualização dos pesos deve chegar a um valor máximo de iterações n_{max} para gerar um hiperplano para as classes C_1 e C_2 . Provamos para $\eta = 1$. Como a inequação 11 cresce quadraticamente em função de n e a inequação 14 cresce linearmente em função de n , sempre existirá um n para qualquer α ou β pertencente aos reais, em que a sua intersecção satisfaça o critério de parada. \square

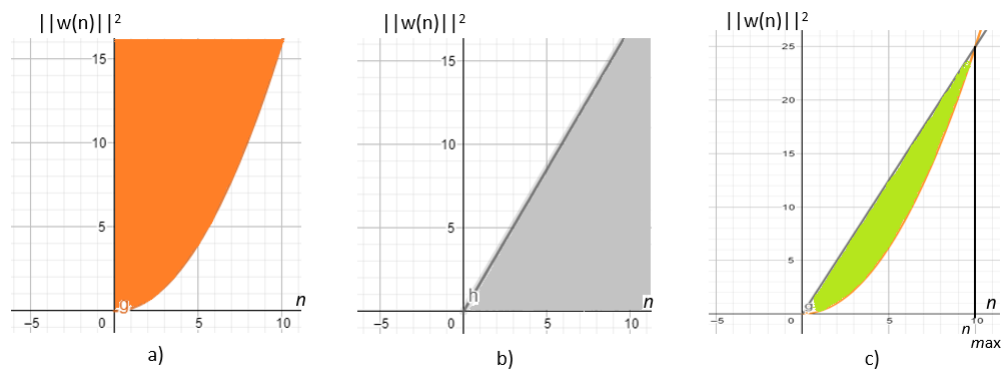


Figura 10: **a** solução da equação 11; **b** solução da equação 14; **c** solução da equação 16.

Fonte: Produzido pelo autor.

Conclusão

Neste trabalho mostramos o teorema de convergência para o Perceptron. Percebemos a utilização de tópicos estudados na Álgebra Linear como a desigualdade de Cauchy-Schwarz, que ao ser definida e usada na prova, limita o número de iterações a um valor mínimo α e a norma euclidiana que limita a um valor máximo β .

Referências

- ERTEL, Wolfgang. **Introduction to Artificial Intelligence**. 2. ed. Londres: Springer, 2017. 356 p.
- FAUSETT, Laurene. **Fundamentals of Neural Networks: architectures, algorithms and applications**. Nova Iorque: Prentice-Hall, 1994. 461 p.
- HAYKIN, Simon. **Redes Neurais: princípios e prática**. 2. ed. São Paulo: Bookman, 2001. 900 p.
- KOVÁCS, Zsolt L.. **Redes Neurais Artificiais: fundamentos e aplicações**. 4. ed. São Paulo: Livraria da Física, 2006. 171 p.
- POOLE, David. **Álgebra Linear**. São Paulo: Cengage Ctp, 2003. 718 p.
- STEINBRUCH, Alfredo; WINTERLE, Paulo. **Álgebra Linear**. São Paulo: Pearson Universidades, 1995. 600 p.

O uso da *Unity* para criação de jogos didáticos

Thais de Souza¹

Universidade Estadual do Oeste do Paraná
thaissouza38@hotmail.com

Flavio Roberto Dias Silva²

Universidade Estadual do Oeste do Paraná
frdsilva@yahoo.com.br

Resumo: Este trabalho tem como objetivo fazer uma breve reflexão sobre o uso e criação de jogos como ferramenta didática, assim realizando uma breve introdução ao motor de criação de jogos *Unity*, introduzindo sua história e apresentando alguma de suas funcionalidades e como esse programa pode ser usado para criação de jogos com fim educativo.

Palavras-chave: matemática, *Unity*, didática, jogos didáticos, atividades práticas, motores de jogos.

1 Introdução

A tecnologia vem ganhando mais espaço a cada dia na vida das pessoas, muitas vezes nos apresentando ferramentas de pesquisa, redes sociais e programas com propósitos de facilitar a vida dos usuários, porém, além de programas úteis, também são apresentados diariamente programas com intuito de entretenimento, como aplicativos de vídeo e jogos de diversos temas e conteúdo.

Como expresso por Camila de Oliveira Amélio (2018) no artigo “A Indústria e o Mercado de Jogos Digitais” o mercado de jogos é um dos que mais cresce no mundo, milhares de jovens estão envolvidos nesse meio, jogando diariamente jogos como *Minecraft*, *Fortnite* e *League of Legends*. O interesse por esse tipo de lazer chega a ser tamanho ao ponto de muitos terem vontade de criar seus próprios jogos, seja apenas por uma brincadeira pessoal e entre amigos, ou até mesmo por incentivo das escolas, onde os alunos são introduzidos ao básico de programação.

Muitas vezes para o desenvolvimento lógico dos alunos quanto para um primeiro contato com esse “mundo” da programação, além do uso de jogos como atividades didáticas

¹ Acadêmica do curso de licenciatura em matemática; voluntária no PIC.

Professor do curso de Licenciatura em Matemática da Unioeste, campus de Cascavel.

que são usadas como reforço ou introdução de algum para ensinar algum determinado conteúdo. Mas por que usar programação e jogos nas escolas? Como observado por Ilka Erondina de Moura:

Na escola os alunos podem ter a oportunidade de utilizar os recursos tecnológicos que auxiliam na sua aprendizagem, interagindo e utilizando diferentes formas de aprender e brincar, de acordo com Silva (2007) as crianças gostam de brincar e nada melhor do que aprender brincando. Quando os alunos brincam jogando é possível que o professor acompanhe a sua evolução e verifique se os mesmos estão aprendendo, de acordo com as propostas curriculares, planejar quais atividades podem desenvolver e propiciar para que o educando avance e aprenda mais (MOURA, ILKA ERONDINA DE,2018).

Com esses fatores que tornam os jogos e sua criação atrativos e que podemos utilizá-los em sala de aula como ferramenta didática, por que não criar nossos próprios jogos? Ou por que não utilizarmos programas de criação de jogos como ferramenta investigativa? A criação de jogos outra hora já foi complexa e bem demorada, porém, hoje há programas que facilitam no desenvolvimento, esses são chamados de Motores de Jogos.

2 A game engine Unity

Motores de Jogos, do inglês *Game Engines*, são programas de computador e/ou o conjunto de bibliotecas, criados para simplificar esse processo de desenvolvimento de jogos. Existem diversos motores de jogos, desde mais conhecidos e habitualmente usados, tanto quanto programas menos conhecidos. Dos mais conhecidos podemos citar: *Game Maker*, *Unity*, *RPG Maker*, *Unreal etc.* Sendo destes a *Unity* um dos motores de jogos mais conhecidos no mercado de, sendo utilizada na criação de jogos de sucesso, e sendo uma das mais poderosas, dando liberdade quase ilimitada para que o desenvolvedor que fizer preferência de uso.

A *Unity* é um motor de jogos americano criado no ano de 2005 que a princípio tinha como foco principal a criação de jogos em 3D para computadores. Com o passar dos anos o programa recebeu atualizações para possibilitar a criação de jogos em 2D e compatibilidade para aparelhos celulares.

Alguns programas criados para o aprendizado de programação, e motores de jogos mais simples possuem ferramentas de programação em blocos. Esse tipo de programa abre porta para o aprendizado lógico de programação, introduz alguma linguagem mesmo que simples. Como estamos apresentando uma ferramenta mais avançada de criação, também estamos nos referindo ao uso de Ambientes de Desenvolvimento Integrado (IDE do inglês *Integrad Development Envrioment*), uma ferramenta para editar códigos e executar *scripts*.

Na maioria dos motores de jogos necessita-se o uso de alguma IDE e na *Unity* não é diferente. Podemos utilizar a que já acompanha o programa, ou vincular alguma outra de

nossa preferência. Como IDE utilizam de linguagem de programação, temos que saber qual é a mais compatível com o sistema que estamos utilizando, para o programa que estamos apresentando, a linguagem mais comumente usada é C# (*Csharp*) que é orientada uma linguagem a objetos. Sabendo como utilizar a IDE dentro do programa, podemos realizar a introdução da interface do programa, essa que é bem ampla e possui várias ferramentas. Faremos apenas uma breve introdução aos itens e conceitos mais utilizados para a base de criação de jogos analisando a documentação disponível o site do próprio programa.

Um dos itens básicos para o início do uso do programa é a criação de objetos em 3D. Podemos facilmente adicionar esferas, cubos, cilindros etc. Todos esses elementos podem ser customizados, mudando seu tamanho, cor, material e funcionalidade (Figura 1). Como são objetos em três dimensões podendo realizar um estudo analítico sobre eles. Claro, que se for utilizado uma ferramenta de modelagem, podemos também criar diferentes tipos e formatos de objetos para adicionar ao programa.

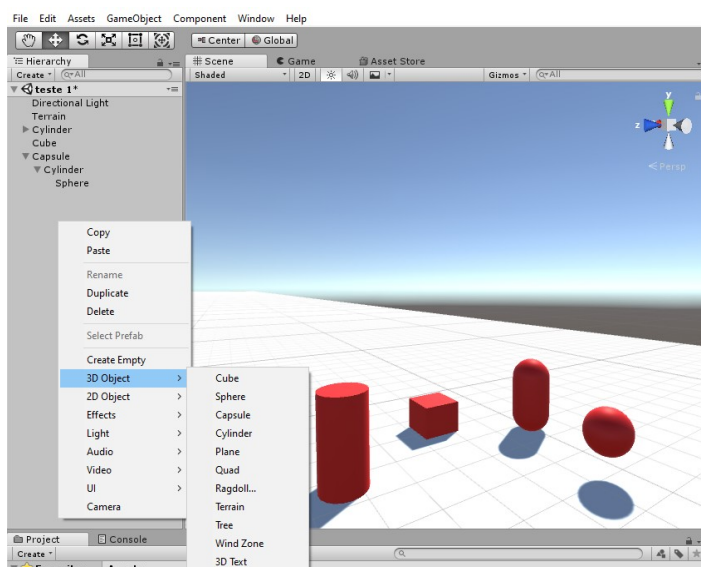


Figura 1: Criação de objetos 3D

Fonte: Acervo da autora

Com a criação de objetos, somos apresentados a outra ferramenta bem interessante presente no programa, a física. A física dos objetos já é pré-programada, para que possa facilitar o desenvolvimento do jogo. Com ela conseguimos colocar solos que possuem gravidade, adicionar rigidez aos objetos no cenário para que assim interajam facilmente com a física presente no programa, sem precisar de muita programação por parte do desenvolvedor. Essa ferramenta é muito utilizada na parte da programação de movimentação de objetos. Apesar de sua utilização ser simples ainda precisamos usar conceitos da física para que ela funcione corretamente, calculando a força que será exercida no objeto, a quantidade de quadros que ele vai se mover por segundos e a sua velocidade por segundo. Além da gravidade, temos propriedades que validam quando um objeto está encostando em outro,

usando conceitos básicos de programação de verdadeiro ou falso, caso verdadeiro uma ação acontece, caso falso, outra ação acontece. Um exemplo de ferramenta que possui essa propriedade são os *Raycast*, que basicamente, são raios ou linhas invisíveis que indicam uma direção e distância determinada, dizendo se o objeto está ou não, interagindo com outro objeto (Figura 2).

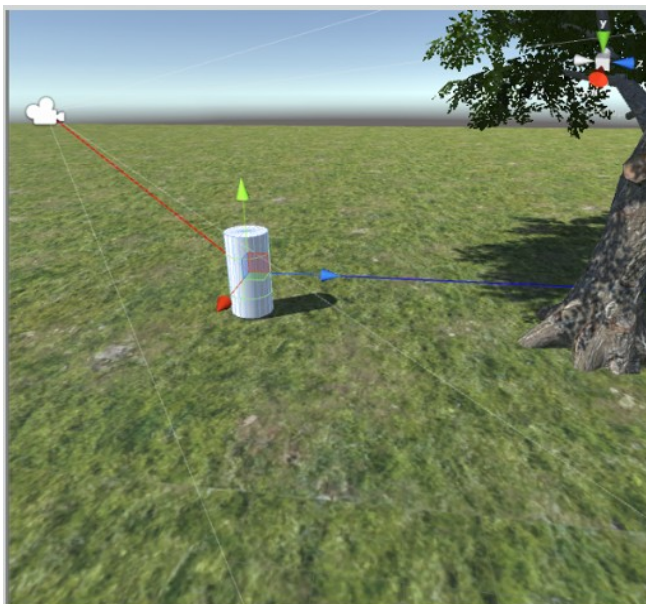


Figura 2: *Raycast* em uso

Fonte: Desenvolvimento de Jogos Eletrônicos com Unity 3D Como mencionado anteriormente, podemos utilizar essa ferramenta para interagir com outros objetos, podendo destruí-los, movimentá-los ou realizar qualquer outra ação que seja do interesse a ser apresentado no jogo. Por mais que muitas dessas funcionalidades estejam pré-programadas. Essas ferramentas precisam que suas funções e realização de interação sejam programadas utilizando a IDE escolhida, podendo-se fazer de uma maneira mais simples, ou mais profissional. O *Raycast*, a física e a inclusão de objetos são as ferramentas mais importantes quando estamos falando de criação em 3D.

A interface também nos possibilita a utilização de ferramentas estéticas, como podemos observar nas Figura 1 e 2, temos um sistema pré-programado de luzes e sombra, com função de dia e noite, que dão profundidade ao ambiente, podemos criar efeitos de tiro, brilho, corrida, salto, adicionar sons e criar cenários utilizando a parte de modelagem presente dentro do programa, além de muitas outras ferramentas pertinentes ao desenvolvimento de jogos. Todas essas funções que aqui estão sendo comentadas para o uso do programa na versão 3D, possuem funções análogas para o 2D.

A *Unity* a princípio era uma ferramenta de criação de jogos exclusivamente em 3D, porém, no ano de 2017 ela recebeu sua primeira atualização para criação de jogos 2D. Com ela pode-se utilizar grades para criar cenários com uso de imagens, inserir personagens e

objetos em 2D análogo ao que já acontecia em suas versões anteriores. Porém, como estamos falando de objetos em duas dimensões, algumas funcionalidades tiveram que ser implementadas para que fosse possível a sua realização. Foram então adicionadas ferramentas de grade, edição e corte de imagem e ferramentas de animação em quadros (Figura 3).

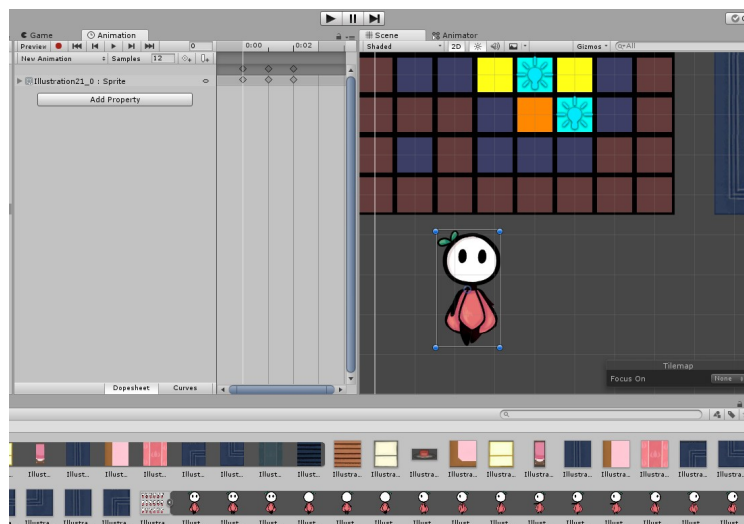


Figura 3: Animação na *Unity*

Fonte: Acervo da autora

Algumas das ferramentas presentes no 3D precisam ser adaptadas para o uso no 2D, como ferramentas de sombras ou de efeitos, mas, todas podem ser facilmente substituídas ou programadas para o uso.

O programa contém também, formas para que o desenvolver possa organizar seu projeto. A *Unity* possibilita a criação de arquivos e pastas dentro do programa que serão vinculados com a máquina em que o programa está sendo utilizado. Com elas, podemos salvar o cenário que está sendo criado, objetos juntamente com todas as suas especificações como: cor, tamanho e código. Com esses arquivos salvos é possível criar pastas com os itens pré-programados para que não seja necessário programar um novo. Esse tipo de pasta é chamado de *prefab*. Essa ferramenta é utilizada em jogos que possuem o mesmo inimigo múltiplas vezes por cenário, em jogos de quebra cabeça em que há itens com as mesmas funcionalidades ou em elementos estéticos, que podem manter o código, porém, com parte visual modificada.

Há muitas outras funções e especificações do programa que não serão citadas aqui. A partir dessa pequena análise do programa podemos fazer sua aplicação como ferramenta didática.

3 Uso da *Unity* como ferramenta didática

Ao utilizar a *Unity* podemos realizar investigações sobre as formas geométricas e uso da física, porém, sua funcionalidade principal é para criação de jogos, desde os mais básicos aos mais avançados. Na iniciação científica desenvolvida durante graduação, foi dado o problema de fazer estudos sobre o programa *Unity*, além de fazer o desenvolvimento de algum jogo com âmbito matemático. Com ajuda do professor orientador desenvolvemos a ideia de criar um jogo baseado no clássico de computador e celular Boxxle, um jogo de 1990, onde o objetivo principal é arrastar todas as caixas para cima de todos os botões presentes no mapa para que se possa avançar para a próxima fase do jogo (Figura 4). Cada fase possui uma quantidade mínima de movimentos que pode ser realizada, assim o jogador deve observar cada mapa, observando qual caminho fará com que realize a menor quantidade de movimentos. Para saber a quantidade de movimentos realizados, há um contador no canto superior direito da tela. Se, ao mover uma caixa para um local em qual não é possível realizar mais nenhum movimento, travando-a, o nível precisa ser reiniciado. Esse jogo precisa de muita observação; em cada fase podemos estimar a quantidade máxima e mínima de movimentos realizados, ou utilizar a lógica para realizar os movimentos necessários.



Figura 4: Fase do jogo Boxxle

Fonte: <https://www.retrogames.cc/gameboy-games/boxxle-usa-europe-rev-a.html>

É possível realizar uma análise de quantos movimentos mínimos são necessários para passar de fase, porém, mesmo não realizando a quantidade mínima de movimentos, ainda é possível terminar o nível sem muitos problemas.

No texto “Desenvolvimento de Softwares Educacionais Matemáticos para o Público Infantil”, da autora Ilka Erondina de Moura, são apresentados alguns jogos com âmbito didático, na área da matemática desenvolvidos na *Unity*, esses jogos são: *Medieval Game* e *Corrida matemática*.

O *Medieval Game* se trata de um jogo de adição matemática destinado ao público infantil. Esse jogo tem o intuito de estimular aos alunos aprender as adições de forma divertida e estimular o cálculo mental. Durante o jogo são realizadas dez jogadas, ou seja, dez operações serão geradas. Os alunos têm um tempo determinado para respondê-las; caso não a respondam a tempo, o Bárbaro chega no baú do tesouro, o jogador perde uma vida e, não marca ponto (Figura 5). O jogador possui três vidas, quando a partida inicia; o Bárbaro sai da torre e anda em direção ao baú, uma operação é gerada. Se o jogador responder a tempo e, corretamente, ele marca ponto e não perde vida, caso erre a resposta, ele perde uma vida e, não marca ponto.



Figura 5: Tela principal do jogo *Medieval Game*

Fonte: MOURA, Ilka Erondina de. (2018, p.35)

As operações são geradas de forma randômica, assim o aluno vai ter uma adição distinta a cada rodada. A fim de felicitá-lo e agilizar a resposta, os alunos podem clicar nos botões que estão em ordem crescente, no centro inferior da tela.

O Jogo Corrida Matemática utiliza as duas formas mais comuns de operações adição e subtração que são geradas randomicamente. Nesse jogo será apresentada uma resposta para a operação e, o jogador deverá classificá-la em verdadeira ou falsa. O jogador deverá realizar a operação de adição ou subtração e verificar se a resposta exibida está correta (Figura 6). Caso esteja correta deverá pressionar o botão verdadeiro, se estiver errada, o botão que deve ser clicado é o falso; tudo isso deve ser feito em um limite de tempo. Se o jogador pressionar a resposta errada ele pode perder uma das três vidas que possui e, não marcará pontos. Esse processo se segue até o final da pista de corrida



Figura 6: Painel exibido com a operação no jogo Corrida Matemática

Fonte: MOURA, Ilka Erondina de. (2018, p.38)

Os jogos acima citados, são todos criados no motor de jogos *Unity*, além destes há muitos outros jogos criados com intuito educativo e de desenvolvimento lógico, sendo de diversas áreas de ensino.

Conclusões

Em sala de aula sempre utilizamos jogos de maneira física para o desenvolvimento cognitivo, lógico ou como maneira de aplicação de algum conteúdo. Com o passar dos anos a atenção dos jovens para a tecnologia vem tomando força, podemos utilizar essa atenção para a aplicação de jogos digitais. Utilizar jogos já conhecidos como referência para estudo, ou ainda, para maneira mais efetiva, criar e incentivar a criação de jogos por parte dos alunos, tendo em vista que esse tipo de atividade envolve ativamente o uso da matemática, seja para calcular a área de alguma figura, o volume de algum sólido, a utilização de matemática básica, geometria analítica, e o uso de conceitos básicos de física, como velocidade, força e atrito.

Referências

AMÉLIO, Camila de Oliveira. **A indústria e o mercado de jogos digitais: Evolução, características e desafios.** 20. 10p – Universidade Federal de Minas Gerais, Belo Horizonte, 2018. Disponível em:
<http://www.sbgames.org/sbgames2018/files/papers/IndustriaFull/188510.pdf>. Acesso em: 30 out. 2022

MOURA, Ilka Erondina de. **Desenvolvimento de *softwares* educacionais matemáticos para o público infantil.** 2018. 48 f. TCC (Graduação) - Curso de Tecnologia em Análise e

DESENVOLVIMENTO de Sistemas, Universidade Tecnológica Federal do Paraná - UTFPR, Medianeira, 2018. Disponível em:

<http://repositorio.utfpr.edu.br/jspui/bitstream/1/13365/1/softwareeducacionaismatematicosinfantil.pdf>. Acesso em: 30 out. 2022.

UNITY USER MANUAL 2021.3 (LTS). **Unity Documentation**, 2022. Disponível em:

<https://docs.unity3d.com/Manual/UnityManual.html>. Acesso em: 31 out. 2022.

O uso da metodologia resolução de problemas no estágio supervisionado

Milena Cristina Heydt

Universidade estadual do Oeste do Paraná - Unioeste
milenaheydt@hotmail.com

Eduardo Ghisi

Universidade estadual do Oeste do Paraná - Unioeste
eduardoghisi66@outlook.com

Fabiana Fatima Delabona

Universidade estadual do Oeste do Paraná - Unioeste
fabi.delabona@hotmail.com

Pamela Gonçalves

Universidade estadual do Oeste do Paraná - Unioeste
pamela.goncalves@unioeste.br

Resumo: A Resolução de Problemas é uma das metodologias usadas no processo de ensino e de aprendizagem de matemática, pois proporciona a participação do aluno na construção do conhecimento matemático, possibilita o desenvolvimento do raciocínio lógico, permite o estudante usar sua criatividade, elaborar estratégias, e, ao encontrar a solução do problema, desfrutar a satisfação da descoberta. Com isso, escolhemos esta metodologia para nossas aulas de regência, parte prática da disciplina de Metodologia e Prática de Ensino de Matemática: Estágio Supervisionado II do curso de Licenciatura em Matemática da Universidade Estadual do Oeste do Paraná (Unioeste), *campus* de Cascavel. Realizamos a regência em duas turmas de 1º anos do Ensino Médio em uma escola pública estadual do município de Cascavel. Utilizamos os problemas como ponto de partida para introduzir o produto de matrizes. Nosso objetivo era que os alunos chegassem intuitivamente no processo de multiplicação de matrizes. Ao fim da atividade concluímos que o objetivo foi alcançado, os discentes acabaram por aplicar a ideia de multiplicação de matrizes sem antes estudar a teoria matemática.

Palavras-chave: Resolução de Problemas; Multiplicação de Matrizes; Estágio Supervisionado.

1 Introdução

É comum em nosso cotidiano nos depararmos com problemas de diferentes tipos a serem resolvidos. Com a matemática não é diferente, segundo Onuchic (1999), problemas

matemáticos têm ocupado lugar central no currículo escolar desde a antiguidade. Ainda de acordo com a autora, são encontrados registros de problemas matemáticos na história egípcia, chinesa e grega, e ainda, são encontrados em livros-texto de matemática dos séculos XIX e XX.

A importância dos problemas matemáticos no ensino pode ser observada nos documentos educacionais brasileiros que, dentre suas orientações, sugerem trabalhar com problemas matemáticos em sala de aula. Neste contexto, a Base Nacional Comum Curricular (BRASIL, 2018), destaca que no Ensino Médio os alunos devem desenvolver habilidades que servirão para resolverem problemas ao longo de suas vidas. Podemos pensar nisso na formação do aluno como cidadão, que no futuro, ao sair da escola, vai se deparar com situações que precisarão encontrar uma solução por conta própria.

Os Parâmetros Curriculares Nacionais do Ensino Médio - PCNEM (2000), consideram necessária a compreensão de conceitos e procedimentos matemáticos tanto para o cidadão ter suas conclusões e fazer argumentações, quanto para agir como consumidor prudente ou tomar decisões em suas atividades em sociedade. Para isso, neste mesmo documento encontramos alguns objetivos que o ensino da matemática possui no Ensino Médio. Dentre eles, dois objetivos envolvem resolver problemas, estes são: “[...] desenvolver as capacidades de raciocínio e resolução de problemas, de comunicação, bem como o espírito crítico e criativo; utilizar com confiança os procedimentos de resolução de problemas para desenvolver a compreensão dos conceitos matemáticos” (BRASIL, 2000, p. 42).

Diante disto, escolhemos a Resolução de Problemas como metodologia de ensino em nossas aulas de regência pois compreendemos que esta metodologia proporciona a participação do aluno na construção do conhecimento, desenvolve o raciocínio lógico, permite o aluno usar sua criatividade, elaborar estratégias, e, ao encontrar a solução do problema, desfrutar a satisfação da descoberta (POLYA, 1997).

2 Resolução de Problemas

Trabalhar com a Resolução de Problemas não pode ser confundido com aplicação e resolução de listas de exercícios. Além do mais, devemos observar que exercício e problema são coisas diferentes, “um problema se diferencia de um exercício na medida em que, neste último caso, dispomos e utilizamos mecanismos que nos levam, de forma imediata, à solução” (ECHEVERRÍA; POZO, 1998, p.16). O que podemos entender é que em um exercício o procedimento a ser adotado é percebido de imediato, já em um problema isso não ocorre dessa forma. As Orientações Complementares aos Parâmetros Curriculares Nacionais PCN+ (2002), também diferencia a metodologia da simples aplicação de exercícios. Este documento diz que,

A resolução de problemas é peça central para o ensino de matemática, pois o pensar e o fazer se mobilizam e se desenvolvem quando o indivíduo está engajado ativamente no enfrentamento de desafios. Esta competência não se

desenvolve quando propomos apenas exercícios de aplicação de conceitos e técnicas matemáticas, pois, neste caso, o que está em ação é uma simples transposição analógica: o aluno busca na memória um exercício semelhante e desenvolve passos análogos aos daquela situação, o que não garante que seja capaz de usar seus conhecimentos em situações diferentes ou mais complexas (BRASIL, 2002, p.112).

Ainda neste sentido, “resolver um problema é encontrar um caminho onde nenhum outro é conhecido de antemão, encontrar um caminho a partir de uma dificuldade, encontrar um caminho que contorne um obstáculo, para alcançar um fim desejado” (POLYA, 1997, p.1-2). Polya (1997) também se preocupa com a forma que um professor trabalha com a Resolução de Problemas. Além dos quatro passos que descreve para resolver um problema, Polya faz algumas sugestões

[...] primeiro, ele deveria estabelecer a classe certa de problemas para os alunos: não muito difíceis, nem muito fáceis demais, naturais e interessantes, que desafiem sua curiosidade, adequados a seu conhecimento. Ele deveria também se permitir algum tempo para apresentar os problemas apropriadamente, de modo que apareça sob ângulo correto. Depois, o professor deveria ajudar seus alunos convenientemente. Não muito pouco, senão não há progresso. Não demais, senão o aluno não terá o que fazer. Não ostensivamente, senão os alunos adquirem aversão ao problema, em cuja solução o professor ficou com maior parte. Entretanto, se o professor auxilia seus alunos apenas o suficiente e discretamente, deixando-lhes alguma independência ou pelo menos alguma ilusão de independência, eles podem se inflamar e desfrutar a satisfação da descoberta (POLYA, 1997, p.3).

Além de trabalharmos com a Resolução de Problemas, utilizamos os problemas como meio para introduzir o conhecimento matemático. De acordo com Schroeder e Lester (1989), há três maneiras diferentes de abordar a Resolução de Problemas: ensinar *sobre* a Resolução de Problemas, ensinar *para* a Resolução de Problemas e ensinar *através* da Resolução de Problemas.

Segundo Schroeder e Lester (1989), ao ensinar sobre a Resolução de Problemas, os alunos são ensinados sobre as fases de resolver um problema, isto é, o professor que tem essa abordagem se baseia no modelo de resolver problemas de Polya (1995), o qual descreve quatro passos para resolver um problema. O primeiro é fazer a compreensão do problema, ou seja, interpretá-lo, entender claramente o que se deseja solucionar. O segundo passo é o estabelecimento de um plano, isto é, elaborar uma estratégia, encontrar alguma relação entre os dados dos problemas e a incógnita, pensar em esquemas e cálculos. O terceiro passo é executar o plano estabelecido, por ele em prática, realizar os cálculos ou desenhos que cheguem em uma conclusão para o problema. Por fim, o quarto passo é fazer o retrospecto da solução, isto é, retomar o problema e analisar o processo de resolução, neste momento podem ser encontrados equívocos na execução do plano, ou até mesmo perceber que o plano estabelecido não faz sentido para aquela situação.

Ainda de acordo com Schroeder e Lester (1989), ao ensinar para a Resolução de Problemas o foco do ensino está na aplicação do conhecimento matemático, ou seja, o professor se concentra em como a matemática ensinada pode ser utilizada para resolver problemas. Nessa abordagem, o propósito principal de aprender matemática é ser capaz de usá-la.

Ao se ensinar através da Resolução de Problemas, para Schroeder e Lester (1989), o problema não é considerado apenas como um propósito para se aprender matemática, é valorizado como um meio de se fazer isso. O ensino do conteúdo matemático começa com um problema e as técnicas matemáticas são desenvolvidas como resposta para ele. Acreditamos que está seja a melhor maneira de abordar a metodologia, pois dessa forma o aluno participa da construção do conhecimento.

3 A utilização da Resolução de Problemas em nossas aulas de regência

Um exemplo da utilização da metodologia em nossas aulas foi para introduzir o produto de matrizes. Para cada turma selecionamos um problema referente ao conteúdo e deixamos que resolvessem individualmente sem antes terem sido apresentados à multiplicação de matrizes.

O problema “Bombons a granel” escolhido para a turma do 1° A, foi adaptado de um vídeo elaborado por Barichello, Camarneiro e Alves (2012) disponível em: <https://m3.ime.unicamp.br/recursos/1055>.

Problema “Bombons a granel”: Dona Ioná vende caixas de bombons. Os tipos de bombons são: ao leite, meio amargo e diet. Com esses três tipos ela monta quatro caixas diferentes. A matriz A abaixo possui para cada tipo de bombom uma coluna e para cada caixa uma linha.

	Ao leite	Meio amargo	Diet	
Caixa 1	20	15	5] = A
Caixa 2	15	20	5	
Caixa 3	5	5	30	
Caixa 4	24	24	2	

Figura 1: Caixas e tipos de bombons.

Fonte: Próprios autores.

Cada tipo de bombom tem um preço e quantidade de calorias diferentes. A matriz B a seguir possui uma coluna para o preço de cada bombom e uma coluna para a quantidade de calorias e uma linha para cada tipo de bombom.

$$\begin{array}{l}
 \text{Preço} \quad \text{Kcal} \\
 \text{Ao leite} \\
 \text{Meio amargo} \\
 \text{Diet}
 \end{array}
 \begin{bmatrix}
 1,00 & 85 \\
 1,50 & 80 \\
 2,00 & 70
 \end{bmatrix} = B$$

Figura 2: Preço e calorias de cada bombom
 Fonte: Próprios autores.

Qual o preço e quantidade de calorias em cada caixa?

Para o 1ºB, o problema escolhido foi retirado de um material teórico escrito por Benevides ([2017]) intitulado “Operações com matrizes” publicado no Portal da OBMEP. O material com o problema está disponível em: https://cdnportaldaoobmep.impa.br/portaldaoobmep/uploads/material_teorico/dixrghlgk5c0s.pdf

Problema “Confeitaria de bolos”: Uma confeitaria produz três tipos de bolos vendidos em duas lojas. A matriz A abaixo possui uma linha para cada loja, uma coluna para cada tipo de bolo e indica quantos bolos de cada tipo foram vendidos por cada loja em uma dada semana.

$$\begin{array}{l}
 \text{Bolo 1} \quad \text{Bolo 2} \quad \text{Bolo 3} \\
 \text{Loja A} \\
 \text{Loja B}
 \end{array}
 \begin{bmatrix}
 5 & 4 & 3 \\
 3 & 2 & 4
 \end{bmatrix} = A$$

Figura 3: Lojas e tipos de bolos.
 Fonte: Próprios autores.

Temos também uma matriz B , onde cada linha corresponde a um dos bolos, cada coluna corresponde a um ingrediente do bolo e as entradas indicam as quantidades de cada ingrediente necessário para fabricar cada bolo.

$$\begin{array}{l}
 \text{Farinha} \quad \text{Açúcar} \quad \text{Leite} \quad \text{Manteiga} \quad \text{Ovos} \\
 (g) \quad (g) \quad (ml) \quad (g) \\
 \text{Bolo 1} \\
 \text{Bolo 2} \\
 \text{Bolo 3}
 \end{array}
 \begin{bmatrix}
 500 & 200 & 500 & 150 & 4 \\
 400 & 100 & 300 & 250 & 5 \\
 300 & 150 & 600 & 0 & 6
 \end{bmatrix} = B$$

Figura 4: Quantidade de ingrediente para cada bolo.

Fonte: Próprios autores.

Qual a quantidade de ingredientes que cada loja precisou para produzir os bolos desta semana?

Aplicamos os problemas no dia 12 de maio, com duração de 2 horas/aula em cada turma. Inicialmente fizemos a leitura do enunciado para auxiliar os alunos na interpretação da situação, para identificarem os dados e a pergunta do problema. Após isso, deixamos que os alunos resolvessem sozinhos para que elaborassem suas próprias estratégias de resolução e auxiliávamos quando necessário.

Durante nossos auxílios e análise da resolução dos problemas, percebemos que os alunos tinham diferentes modos de pensar. No 1ºA, a maioria dos alunos multiplicou as quantidades de cada tipo de bombom (ao leite, meio amargo e diet) em cada caixa (caixa 1, caixa 2, caixa 3 e caixa 4) pelo preço e pela quantidade de calorias, o que podemos ver na figura a seguir (Figura 5), que mostra a resolução de um dos alunos da turma do 1º A.

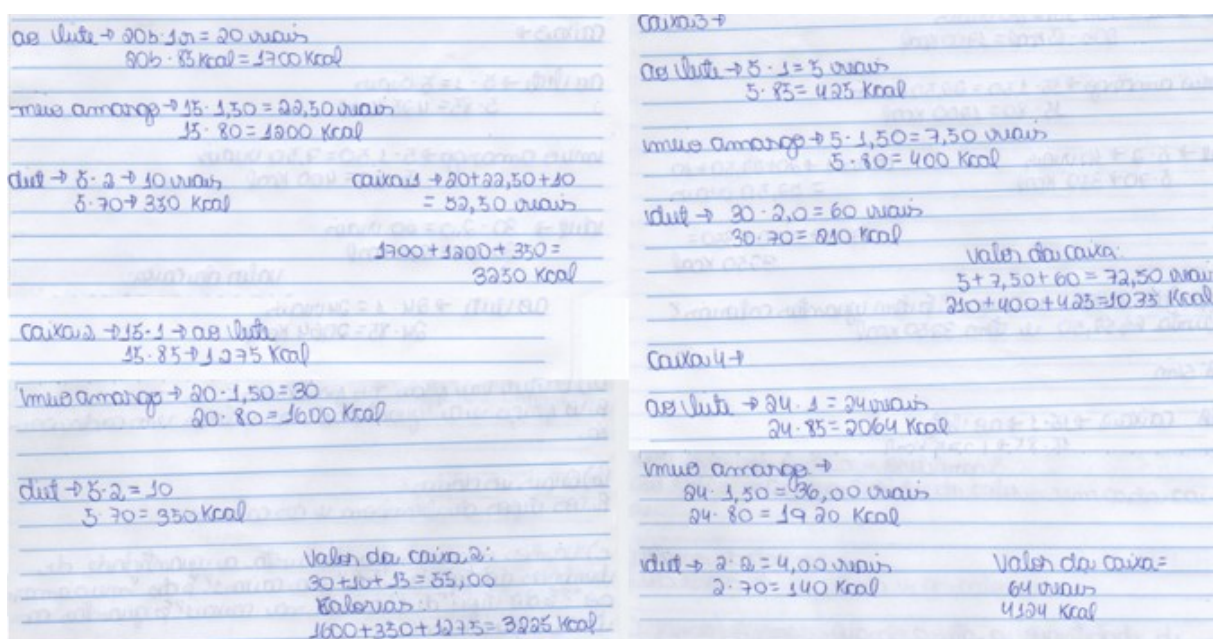


Figura 5: Resolução do problema "Bombons a granel" pelo aluno 1do 1ºA.

Fonte: Próprios autores.

Na turma do 1ºB, a maioria dos alunos resolveu o problema multiplicando as quantidades de cada ingrediente (farinha, açúcar, leite, manteiga e ovos) necessário para a produção de cada tipo de bolo (Bolo 1, Bolo 2 e Bolo 3) pela quantidade de cada tipo de bolo vendido em cada loja (Loja A e Loja B). O que pode ser visto na seguinte figura (Figura 6).

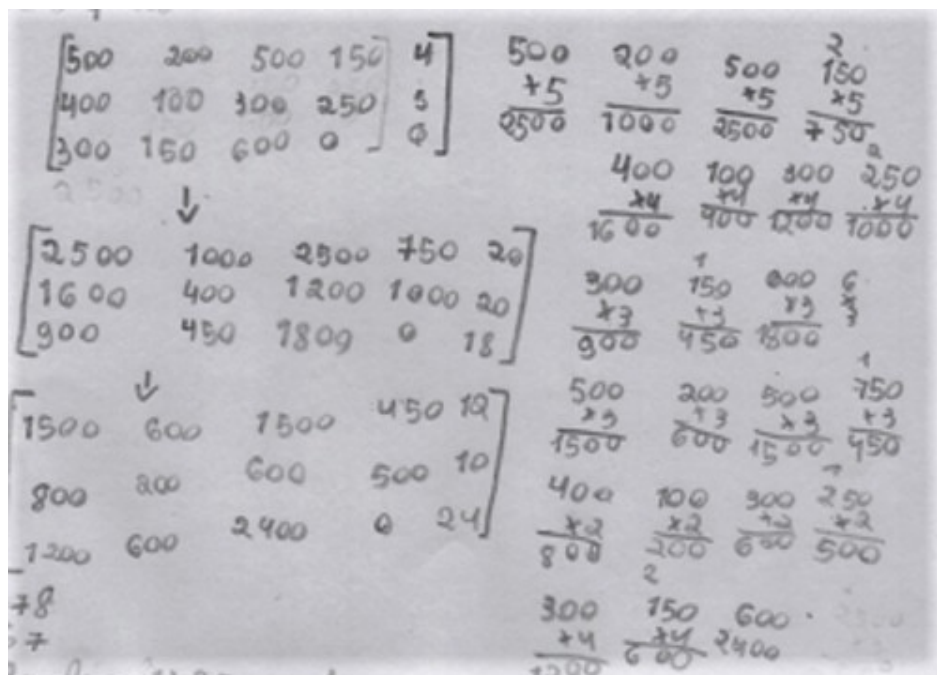


Figura 6: Resolução do problema "Confeitaria de bolo" pelo aluno 2 do 1ºB.

Fonte: Próprios autores.

As formas de registros das respostas também ocorreram de formas distintas em ambas as turmas. Alguns alunos apenas deixaram as contas com os valores encontrados, o que vemos na resolução do aluno 1 (Figura 5), outros construíram tabelas (Figura 7) e até mesmo matrizes (Figura 8) e (Figura 9), outros representaram seus resultados em esquemas (Figura 10).

	Preço	Kalorias
C1	52,50	2250
C2	62,00	3225
C3	72,50	2925
C4	84,00	4100

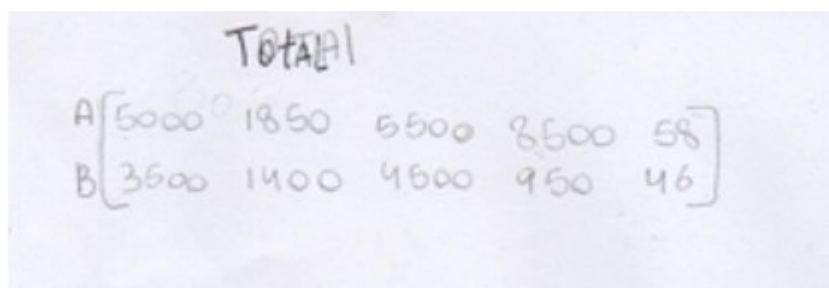
Figura 7: Registro dos resultados encontrados pelo aluno 3 do 1ºA em forma de tabela.

Fonte: Próprios autores.

	Preço	KCAL
	52.50	2250
	45.00	3225
	72.50	3725
	64.00	4100

Figura 8: Registro dos resultados encontrados pelo aluno 4 do 1ºA em forma de matriz.

Fonte: Próprios autores.

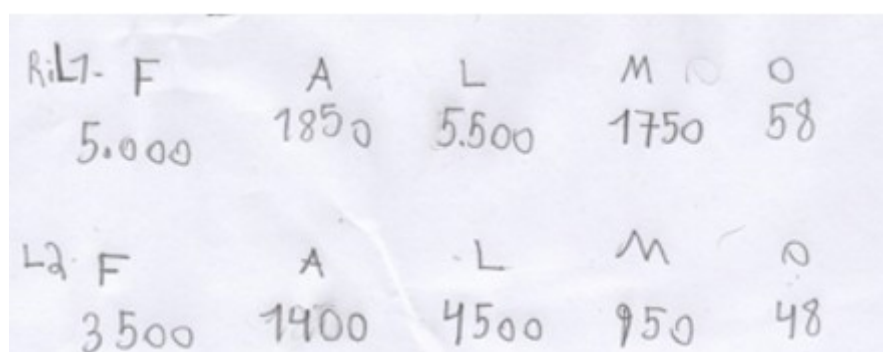


TOTAL

$$A \begin{bmatrix} 5000 & 1850 & 5500 & 8500 & 58 \end{bmatrix}$$
$$B \begin{bmatrix} 3500 & 1400 & 4500 & 950 & 46 \end{bmatrix}$$

Figura 9: Registro dos resultados encontrados pelo aluno 5 do 1ºB em forma de matriz.

Fonte: Próprios autores.



R1	F	A	L	M	O
	5.000	1850	5.500	1750	58
L2	F	A	L	M	O
	3500	1400	4500	950	48

Figura 10: Registro dos resultados encontrados pelo aluno 6 do 1ºB em um esquema.

Fonte: Próprios autores.

Notamos que nem todos os resultados encontrados pelos alunos estão corretos, porém, consideramos que, ao trabalharmos com a metodologia, o mais importante é que os alunos participem da construção do conhecimento matemático, e, observamos isso com as resoluções dos alunos. Mesmo com alguns equívocos em seus cálculos, intuitivamente utilizaram o produto de matrizes para resolver os problemas.

Conclusões

É importante destacar que os alunos estavam acostumados com a metodologia tradicional de ensino, isto é, primeiro a exposição do conceito, depois exemplos resolvidos pelo professor e então listas de exercícios nos quais precisavam apenas seguir os procedimentos já conhecidos.

Com a Resolução de Problemas, os alunos se depararam com algo diferente do que estavam acostumados em sala de aula. Não sabiam de imediato qual conceito aplicar para resolver o problema e tiveram que usar seus conhecimentos para elaborar uma estratégia e encontrar uma solução. Com isso, buscavam constantemente a confirmação dos professores para cada passo da resolução que faziam. Foi necessário tomar cuidado para que nossos auxílios nos momentos de dúvidas dos alunos não interferissem seus raciocínios.

Outro fato a ser destacado é a dificuldade com operações matemáticas. Os alunos tinham uma certa dependência da calculadora para qualquer conta que realizavam, até mesmos cálculos que poderiam resolver mentalmente. Devido a isso, ao fazer as contas sem o uso da calculadora, percebemos diversos erros nos registros das respostas dos problemas.

Com a análise dos problemas propostos, mesmo com alguns obstáculos, podemos considerar que, de acordo com nosso objetivo os alunos utilizaram intuitivamente o processo de multiplicação de matrizes. A metodologia exigiu dos alunos além de somente a aplicação de conceitos, proporcionou autonomia e possibilitou a participação destes alunos na construção do conhecimento matemático.

Referências

- BENEVIDES, F. S. **Operações com Matrizes**. [2017]. Disponível em: https://cdnportaldabmep.impa.br/portaldabmep/uploads/material_teorico/dixrghlgk5c0s.pdf. Acesso em: 19 abr. 2022.
- BARICHELLO, L; CAMARNEIRO, F; ALVES, R. S. de O. **Bombons a granel**. 2012. Disponível em: <https://m3.ime.unicamp.br/recursos/1055>. Acesso em: 19 abr. 2022.
- BRASIL. **Parâmetros Curriculares Nacionais para o Ensino Médio**. Brasília: Ministério da Educação, 2000.
- BRASIL. **Base Nacional Comum Curricular**. Brasília: Ministério da Educação, 2018.
- BRASIL. **Orientações Educacionais Complementares aos Parâmetros Curriculares Nacionais**. Brasília: Ministério da Educação, 2002.
- ECHEVERRÍA, M. del P. P.; POZO, J. I. Aprender a resolver problemas e resolver problemas para aprender. In: POZO, J. I. (org.). **A solução de problemas**. Porto Alegre: Artes Médicas, 1998. p. 13-43.
- ONUCHIC, L. De La R. Ensino-aprendizagem de matemática através da resolução de problemas. In: BICUDO, M. A. V. (Org) **Pesquisa em educação matemática: Concepções e perspectivas**. São Paulo: Editora UNESP, 1999. p. 199-218.
- POLYA, G. **A arte de resolver problemas: um novo aspecto do método matemático**. Rio de Janeiro: Editora Interciência Ltda, 1995.
- POLYA, G. Sobre a resolução de problemas da matemática na High School. In: KRULIK, S; REYS, R. E. **A resolução de problemas na matemática escolar**. São Paulo: Atual, 1997. p. 13.
- SCHROEDER, T. L.; LESTER, F. K., Jr. **Developing Understanding in Mathematics via Problem Solving**. In: TRAFTON, P. R.; SHULTE, A. P. (Eds.). *New directions for elementary school mathematics*. Reston: NCTM, 1989.p.31-42.

Aspectos de um curso de formação continuada para professores da rede municipal de ensino no município de Cafelândia: um enfoque na resolução de problemas como metodologia

Jheniffer Rafaelly Vieira da Silva
Universidade Estadual do Oeste do Paraná
jheniffer.silva2@unioeste.br

Resumo: O presente texto trata de um projeto de pesquisa e com o intuito de apresentar discussões que serão desenvolvidas acerca da problemática: como a resolução de problemas é entendida e utilizada a partir de um curso de formação continuada para os professores alfabetizadores da rede municipal de ensino de Cafelândia – PR. Com o objetivo de estudar as potencialidades da formação continuada ministrada aos professores da rede municipal de ensino, desenvolveremos uma pesquisa de abordagem qualitativa, identificada como uma pesquisa experimental, utilizando questionários como ferramenta para a obtenção de dados. Como base teórica apresentaremos assuntos pertinentes ao ensino de matemática nos anos iniciais, bem como a resolução de problemas como metodologia de ensino e aprendizagem. Ainda buscaremos apresentar uma breve abordagem no que diz respeito a formação continuada de professores.

Palavras-chave: Ensino fundamental I. Formação de professores. Resolução de problemas.

1 Introdução

O presente texto aborda um projeto de pesquisa, desenvolvido na disciplina de Monografia, do curso de Licenciatura em Matemática, da Universidade Estadual do Oeste do Paraná (Unioeste), *campus* Cascavel, sob orientação do professor Jean Sebastian Toillier. Assim, irei apresentar os elementos que me motivam para a realização do projeto, meu problema de pesquisa e os objetivos, uma breve fundamentação teórica e os aspectos metodológicos que cercam o projeto em execução.

O interesse para a realização deste trabalho surgiu a partir de um curso de formação continuada ofertado pela Secretaria Municipal de Educação (SEMED) do município de Cafe-

lândia - Paraná (PR), em que atuo há dois anos como professora dos anos iniciais na etapa de alfabetização. A temática do curso centrava-se na apresentação do Material Complementar de Apoio Didático *Educa Juntos*¹, de Matemática, elaborado por Ana Ruth Starepravo. Ele contém quatro cadernos de atividades elaborados para aulas de matemática na etapa da alfabetização, que abrange os três primeiros anos do ensino fundamental I. A proposta do material consiste em aulas voltadas para a Resolução de Problemas e na utilização de jogos durante aulas de matemática.

Como acadêmica do curso de Licenciatura em Matemática, tinha um conhecimento prévio referente à metodologia de ensino Resolução de Problemas. Então, durante a leitura do material disponibilizado para a formação continuada que seria realizada no mês de agosto, foi-me possível encontrar pontos mencionados pela autora Ana Ruth Starepravo semelhantes à metodologia antes estudada. Desse modo, havia uma expectativa em relação à formação, pois, mesmo com o conhecimento teórico sobre Resolução de Problemas, a prática em sala de aula na etapa de alfabetização, ainda era uma grande barreira.

A formação de professores foi realizada em Cascavel (PR) no dia 24 de agosto de 2022 e, tinha como objetivo apresentar o material acima citado a dois professores de cada município pertencente ao Núcleo Regional de Educação (NRE) de Cascavel. Posteriormente, os participantes deveriam repassar aos demais professores de seus respectivos municípios. Fui escolhida juntamente com outra professora da rede como representante do município de Cafelândia (PR). Durante as oito horas de realização do curso, notava-se que para a maioria ali presente a metodologia de ensino abordada em todo material era desconhecida.

A novidade para muitos despertava curiosidade e dúvidas em relação a essa nova maneira de ensinar e aprender matemática. Comentários como: “dou aula há mais de 25 anos, e nunca tinha visto uma criança de 1º ano resolver problemas desta maneira”; “nunca pensei em ensinar matemática assim” surgiram no decorrer da formação. O novo encantou a todos os presentes e despertou um grande entusiasmo para repensar e mudar algumas práticas no ensino de matemática desde a alfabetização.

Partindo desse cenário, constrói-se a problemática de pesquisa: como a resolução de problemas é entendida e utilizada a partir de um curso de formação continuada para os professores alfabetizadores da rede municipal de ensino de Cafelândia (PR)? Com o objetivo de estudar as potencialidades dessa formação continuada para os professores que deverá ser realizada em fevereiro de 2023, será desenvolvida uma pesquisa de abordagem qualitativa, utilizando questionários para a obtenção de dados que, posteriormente, serão analisados com o intuito de nos ajudar a refletir sobre nosso problema de pesquisa.

¹ Trata-se de um material complementar disponibilizado pela Secretaria da Educação e do Esporte do Paraná (SEED/PR) para os três primeiros anos do ensino fundamental. Sua organização foi estruturada para servir como apoio no trabalho com as defasagens de aprendizagem, sobretudo daquelas oriundas do período em que as crianças ficaram fora do ambiente físico escolar, em decorrência da Pandemia.

2 Objetivo de pesquisa

- Estudar as potencialidades de um curso de formação continuada para professores do ensino fundamental I em que a resolução de problemas será a metodologia de ensino e aprendizagem adotada em aulas de matemática.

2.1 Objetivos específicos

- Identificar a partir de questionários como professores do ensino fundamental I utilizam a resolução de problemas nas aulas de matemática;
- Investigar conhecimentos adquiridos durante a formação continuada, por meio de um questionário;
- Compreender modos de abordar a resolução de problemas como metodologia de ensino de matemática no ensino fundamental I.

3 Referencial teórico

No decorrer da pesquisa, abordaremos assuntos relacionados ao ensino de matemática nos anos iniciais, com enfoque na etapa de alfabetização, compreendida nos três primeiros anos do ensino fundamental I. Estudaremos a resolução de problemas como uma metodologia de ensino e aprendizagem para aulas de matemáticas, bem como assuntos pertinentes à formação continuada de professores.

3.1 Matemática nos anos iniciais.

Estudar sobre o ensino e aprendizagem de matemática nos anos iniciais do ensino fundamental I não é uma tarefa simples, pois envolve discussões que se iniciam com a formação dos professores polivalentes, que para Nacarato, Mengali e Passos (2021) são os profissionais que atuam na educação infantil e anos iniciais do ensino fundamental, sendo os mesmos responsáveis por lecionar todas as disciplinas que compõem o currículo. No entanto, para tal, esses professores recebem uma formação inicial generalizada, podendo ser ela de nível superior – como em cursos de Pedagogia – ou em nível médio – como nos cursos de Formação de Docentes, o antigo Magistério.

Com o objetivo de entender o processo de formação inicial, Nacarato, Mengali e Passos (2021) desenvolvem um estudo abrangendo a trajetória escolar dessas professoras no que diz respeito às reformas curriculares. Podemos observar com esse estudo se essas professoras, enquanto alunas, receberam aulas de matemáticas centradas no detalhamento de conteúdos e uso de algoritmos desprovidos de qualquer significado. E que, posteriormente, sua formação como docente deixou lacunas sobre o ato de ensinar matemática diferente da maneira com que aprendeu, ou seja, surge uma barreira para romper com metodologias de ensino que se mostram ineficazes na aprendizagem dos alunos.

É importante ressaltar que é justamente nas séries de atuação das professoras polivalentes que corre a alfabetização matemática, termo este compreendido “como uma habilidade que transcende à leitura e escrita, carregando um valor crítico” (SKOVSMOSE 2001, p. 66 *apud* NACARATO; MENGALI; PASSOS, 2021, p. 30). Alfabetizar matematicamente é “reconhecer a natureza crítica” (NACARATO; MENGALI; PASSOS 2021, p. 30) da matemática e promover um ensino repleto de possibilidades e significados.

No entanto, para que uma alfabetização matemática crítica realmente seja alcançada é preciso que mudanças ocorram, que velhos costumes sejam superados, que aulas de matemática possam ir além de repetições de exercícios carentes de significados reais para os alunos. É entender que a “aprendizagem matemática não ocorre por repetições e mecanizações, mas que se trata de uma prática social que requer envolvimento do aluno em atividades significativas” (NACARATO; MENGALI; PASSOS, 2021, p. 31).

3.2 Resolução de problemas

Em busca de uma metodologia que promova um ensino com produção de significados, Nacarato, Mengali e Passos (2021) alegam que a resolução de problemas aparece como uma metodologia que potencializa o ensino e aprendizagem de matemática. Em vista disso, podemos destacar três diferentes maneiras de utilizar a resolução de problemas em sala de aula, sendo elas “(1) o ensino sobre Resolução de Problemas, (2) o ensino para resolução de problemas e (3) o ensino através da resolução de problemas” (SCHROEDER; LESTER, 1989 *apud* ONUCHIC; ALLEVATO, 2014, p. 37). A seguir abordaremos características de cada uma dessas três concepções segundo as autoras Onuchic e Allevato (2014).

O ensino sobre a resolução de problemas pode ser entendido como uma maneira de orientar os alunos no que se refere a resolver problemas, mostrando regras e procedimentos independentes dos conteúdos abordados. Na segunda concepção, o ensino para a resolução de problemas configura uma atividade na qual os alunos só podem resolver problemas após uma habilidade treinada ou de algum algoritmo previamente estudado, ou seja, “a matemática é ensinada separada de suas aplicações” (ONUCHIC; ALLEVATO, 2014, p.38). Por fim, a terceira concepção que se refere ao ensino através da resolução de problemas, na qual “a Mate-

mática e a resolução de problemas são consideradas simultaneamente e são construídas mútua e continuamente” (p.38).

Pensando em um ensino de matemática que parte do princípio de que “a aprendizagem se realiza pela construção dos conceitos pelo próprio aluno” (ONUCHIC; ALLEVATO, 2014, p.40), a resolução de problemas como ponto de partida no ensino, colocaria o aluno a confrontar suas concepções e construir novos conceitos. Desta forma, a última concepção abordada, o ensino através da resolução de problemas, é a que será considerada ao pensarmos no ensino de matemática.

3.3 Formação continuada

Processos de formações continuadas caracterizam a carreira docente, pois fazem parte do desenvolvimento profissional dos professores. Dessa maneira Imbernón (2010) diz que:

A formação continuada dos professores, mais do que atualizá-los, deve ser capaz de criar espaços de formação, de pesquisa, de inovação, de imaginação, etc., e os formadores de professores devem saber criar tais espaços para passarem do ensinar ao aprender (IMBERNÓN, 2010, p.11).

Entendemos que formações continuadas devem atuar como potencializadoras, promovendo espaços nos quais os professores tenham a oportunidade de refletir suas práticas com a finalidade de questioná-las e, por fim, reinventá-las. Para tal, se faz necessário que os docentes participem ativamente dos estudos, da elaboração de planejamentos, das ações a serem desenvolvidas, garantindo assim que sua realidade escolar esteja presente, o que possibilita uma representatividade capaz de edificar os conhecimentos abordados nos cursos de formações continuadas.

4 Metodologia

Durante o desenvolvimento desta pesquisa planejamos realizar um curso de formação continuada com professores alfabetizadores que atuam nas séries iniciais do ensino fundamental I no município de Cafelândia – PR. Desse modo optamos pela realização de uma pesquisa de abordagem qualitativa, identificada como pesquisa exploratória, de acordo com Selltiz (1967). Também, Gil (2002, p. 41) explica que “essas pesquisas envolvem: (a) levantamento bibliográfico; (b) entrevistas com pessoas que tiveram experiências práticas com o problema pesquisado; e (c) análise de exemplos que estimulem a compreensão”.

No decorrer do mês de novembro de 2022, os materiais para o desenvolvimento do curso de formação continuada deverão ser elaborados em parceria com duas coordenadoras da SEMED de Cafelândia – PR. Estimamos que 45 professores da rede participem da formação.

Abaixo uma representação do cronograma para realização do curso com carga horária de oito horas, que está sujeito a mudanças em alguns pontos:

Quadro 1 Cronograma da formação continuada.

Município	Data e horário agendados para a formação	Local onde será realizada a formação	Previsão do número de participantes
Cafelândia - PR	15/02/2023 – 07:30 às 17:30	Escola Municipal André Luiz da Silva Pres-tes - EF	Prof. ^a 1 ° ano. 15 participantes.
Cafelândia – PR	16/02/2023 – 07:30 às 17:30	Escola Municipal André Luiz da Silva Pres-tes – EF	Prof. ^a 2 ° ano. 15 participantes.
Cafelândia - PR	17/02/2023 – 07:30 às 17:30	Escola Municipal André Luiz da Silva Pres-tes - EF	Prof. ^a 3 ° ano. 15 participantes.

Fonte: elaborado pela autora.

Planejamos subdividir a carga horária de oito horas de modo que às quatro horas iniciais atuem como uma troca de experiências, na qual, os participantes serão convidados a partilhar suas práticas em sala no que se refere a resolução de problemas em aulas de matemática. Assim, pretendemos realizar a coleta de dados através de um questionário entregue aos professores, com o objetivo de identificar como é entendida e utilizada a resolução de problemas em aulas de matemática. A elaboração do questionário para obtenção de dados buscará abordar questões mais abertas, que segundo Gil (2008, p. 122) caracterizam um questionário, no qual os respondentes tenham a liberdade de oferecer suas próprias respostas.

As demais quatro horas serão organizadas em formato de oficina assim como é proposto no Material Complementar de Apoio Didático *Educa Juntos*, de Matemática elaborado por Ana Ruth Starepravo. Buscaremos trazer elementos e práticas que podem ser utilizadas em aulas de matemática dentro da proposta de ensino através da resolução de problemas. Nesse momento, é importante que as práticas antes compartilhadas pelos professores sejam retomadas, pois para Nacarato, Mengali e Passos (2021, p.34) “As práticas que forem questionadas, refletidas e investigadas poderão contribuir para as mudanças de crenças e saberes”. Ao final pretendemos realizar uma nova coleta de dados, desta vez com o objetivo de identificar os resultados alcançados com a formação.

A pesquisa se desenvolverá durante o ano letivo de 2022, com início no ano civil de 2022 e finalização no civil de 2023. Levantamentos teóricos buscarão aprofundar as ideias apresentadas nesse projeto, com o objetivo de adquirir base teórica para a análise dos dados coletados. Os resultados serão apresentados na monografia que deverá ser defendida até o final do mês de junho de 2023.

Referências

- GIL, A. C. **Como elaborar projetos de pesquisa**. 4ed.. São Paulo: Atlas, 2002.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6 ed.. São Paulo: Atlas, 2008.
- IMBERNÓN, F. **Formação continuada de professores**. Porto Alegre: Artmed, 2010.
- NACARATO, A. M; MENGALI, B. L. D; PASSOS, C. L. B. **A matemática nos anos iniciais do ensino fundamental**: Tecendo fios do ensinar e do aprender. 3ed.. Belo Horizonte: Autêntica, 2021.
- ONUCHIC, L. R; ALLEVATO, N. S. G. Ensino – Aprendizagem – Avaliação de Matemática: por que Através da Resolução de Problemas? In: ONUCHIC, L. R; ALLEVATO, N. S. G. NOUGUTI, F. C. H; JUSTULIN. A. M. **Resolução de problemas**: teoria e prática. Jundiá: Paco editorial, 2014. p. 35-52.
- SELLTIZ, C. **Métodos de pesquisa nas relações sociais**. São Paulo: Herder, 1967.
- STAREPRAVO, A. R. **Material Complementar de Apoio Didático Educa Juntos de Matemática**: Caderno de orientações gerais. – Curitiba: SEED-PR, 2022. Disponível em: https://www.educacao.pr.gov.br/educa_juntos_matematica.

Implementação do algoritmo de Dijkstra com apoio do software Geogebra.

Amarildo de Vicente
Unioeste - Cascavel
amarildo.vicente@gmail.com

Resumo: Em teoria dos grafos um dos recursos de grande importância é a determinação do menor caminho entre dois pontos distintos. A fim de obter tal caminho há vários algoritmos e entre estes se encontra o algoritmo de Dijkstra. O propósito deste trabalho é apresentar a implementação deste algoritmo com apoios do software Geogebra, a fim de possibilitar que, além do custo do percurso, se tenha também uma visão geométrica do caminho a ser percorrido.

Palavras-chave: Grafos; Caminho mínimo; Roteamento.

1 Introdução

A Teoria dos grafos tem uma variedade muito grande de aplicações em diversas áreas, como química, física, matemática, engenharia, computação, etc. Um exemplo bastante comum para ilustrar seu uso é o de rotas feitas por veículos de sistemas de coleta de resíduos, em que estes devem passar por todos os locais evitando, se possível, passar duas vezes no mesmo endereço, a fim de reduzir o tempo de trabalho e custos de manutenção. Essa ferramenta também se faz presente na indústria, onde a fabricação de um dado produto deve passar por vários processos uma única vez seguindo uma rota determinada. Outra situação onde se emprega essa teoria é na localização de centros de atendimento destinados à sociedade, como a de um hospital, que em cidades projetadas possuem a menor distância entre pontos específicos, de escolas, asilos e fábricas, etc. O uso de grafos se faz presente também na educação básica. Cáceres (2021) faz uma abordagem sobre o uso de grafos como recurso para o ensino de matrizes no ensino médio.

A determinação da menor rota entre dois pontos distintos de um grafo é um dos recursos de grande utilidade e um dos principais algoritmos para este fim é o algoritmo de Dijkstra. Neste trabalho será apresentada uma forma de implementar tal algoritmo como o uso do software Geogebra. Na sequência serão apresentados alguns dos conceitos sobre grafos, necessários para o desenvolvimento do trabalho.

2 Matriz de custos

Sob um aspecto computacional a representação de um grafo pode ser feita por meio de matrizes. Uma das representações matriciais muito útil é a matriz de custos. Primeiramente é necessário compreender o significado de custo neste contexto. Se o propósito é avaliar a distância de um vértice v_i a um vértice v_j então o custo é a distância entre eles. Por vezes o interesse não é a distância, mas o tempo de viagem. Neste caso o custo é o tempo gasto no percurso.

Há ainda outros fatores que podem representar custos, ou pesos, como limite de capacidade de escoamento de uma via, riscos de acidentes, etc.

Definição 1. Considere-se um grafo $G = (V, A)$, em que V representa um conjunto de vértices e A representa um conjunto de arestas. Seja w_{ij} o custo da aresta (v_i, v_j) , caso tal aresta exista. Define-se a matriz de custos C por

$$c_{ij} = \begin{cases} w_{ij} & \text{se } (v_i, v_j) \in A \\ \infty & \text{caso contrário} \end{cases}$$

Nesta definição pressupõe-se que o grafo não possui laços, de forma que sendo $i = j$ então $w_{ij} = 0$. Pode parecer estranho o símbolo ∞ na definição anterior, mas ele representa apenas um número positivo grande, que serve como estratégia para o funcionamento do algoritmo de Dijkstra, como será visto posteriormente.

Exemplo 1

Na Figura 11 está a representação de um grafo contendo cinco vértices.

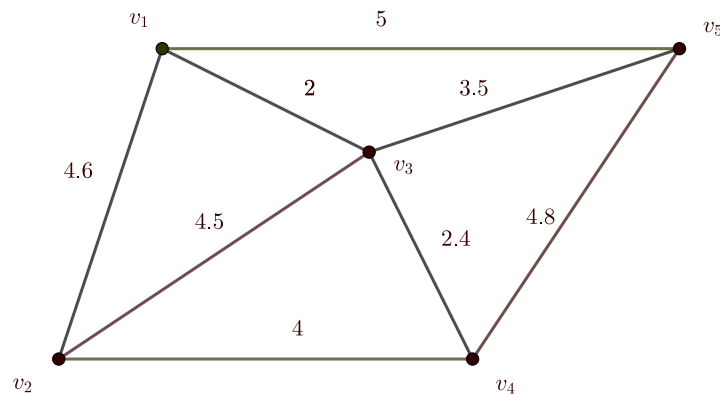


Figura 11: Grafo valorado

De acordo com a Definição 1 sua matriz de custos é

$$C = \begin{bmatrix} 0 & 4.6 & 2 & \infty & 5 \\ 4.6 & 0 & 4.5 & 4 & \infty \\ 2 & 4.5 & 0 & 2.4 & 3.5 \\ \infty & 4 & 2.4 & 0 & 4.8 \\ 5 & \infty & 3.5 & 4.8 & 0 \end{bmatrix}$$

Conforme pode ser observado, o símbolo ∞ representa o custo de arestas não existentes, como (v_1, v_4) .

Em muitas situações há a necessidade de orientar uma ou mais arestas de um grafo. Havendo tal orientação o grafo é chamado de grafo orientado, ou dígrafo. Geometricamente esta orientação pode ser representada por uma seta. Na Figura 12 tem-se um grafo orientado. Neste caso o custo para ir de um vértice v_i a um vértice v_j pode ser diferente do custo para ir de v_j a v_i . Desta forma a matriz de custos de um dígrafo em geral não é simétrica, diferentemente do que ocorre no Exemplo 1 em que o grafo não é orientado.

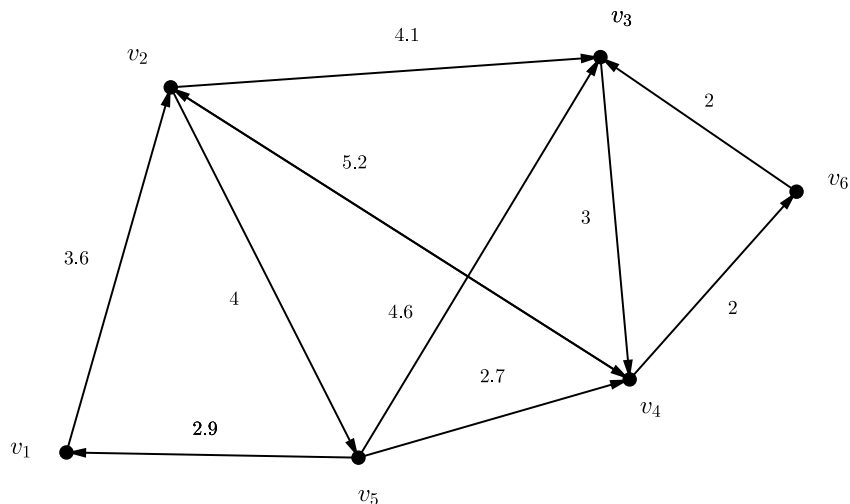


Figura 12: Grafo Direcionado

No grafo da Figura 1 não é difícil perceber que, sendo escolhido o vértice v_1 como origem e v_4 como destino, então o caminho de menor custo consiste na rota v_1, v_3, v_4 , com custo 4.4. Mas em geral o processo não é tão simples. O algoritmo de Dijkstra é um dos procedimentos que podem ser empregados a fim de obter o caminho de menor custo de um vértice inicial v_0 escolhido a qualquer outro vértice v_j do grafo.

3 Algoritmo de Dijkstra

A versão do algoritmo a seguir é devida a Rabuske(1992).

1) Seja $G(V, A)$ um grafo direcionado e v_0 um vértice inicial escolhido como ponto de partida. Seja L uma função que associe cada aresta (v_i, v_j) de G a um número real não negativo c_{ij} dada por

$$L(v_i, v_j) = \begin{cases} \infty & \text{se não existir a aresta } (v_i, v_j) \\ 0 & \text{se } v_i = v_j \\ c_{ij} & \text{se } v_i \neq v_j \text{ e existe a aresta } (v_i, v_j) \end{cases}$$

2) Constrói-se um conjunto S que contém os vértices v_i cuja distância de v_0 a cada v_j seja mínima, desde que esses valores sejam conhecidos.

3) A cada passo deve-se adicionar ao conjunto S um vértice pertencente a $V - S$ tal que o comprimento do caminho v_0 a v_j seja menor que o correspondente de qualquer outro vértice $V - S$. Desta forma, pode-se garantir que o caminho mínimo de v_0 a qualquer vértice v em S contém somente vértices pertencentes a S .

Forma resumida do algoritmo de Dijkstra conforme Rabuske(1992).

Início

$S = \{v_0\};$

$D[v_0] = 0;$

Para cada vértice $v \in V - \{v_0\}$ faça $D[v] = L(v_0, v);$

Enquanto $S \neq \{ \};$

Escolha o vértice $w \in V - S$ tal que $D[w]$ seja mínima;

Coloque w em S , isto é, faça $S = S \cup \{w\};$

Para cada $v \in V - S$ faça

$D[v] = \min(D[v], D[w] + L(w, v));$

Fim

4 Implementação do algoritmo de Dijkstra no Geogebra

O software Geogebra dispõe de comandos que podem ser executados por meio da barra de entrada. Por exemplo, para criar um ponto $A = (a, b)$ pode-se usar a ferramenta “Ponto” disponível no menu ou pode-se digitar $A = (a, b)$ diretamente na barra de entrada. Mas há ferramentas bem mais sofisticadas para criar pequenos programas, como *Se*, *Elemento*, *Sequência*, entre outros.

Exemplo 1

a) Seja a matriz

$$M = \begin{bmatrix} 0 & 6 & 2 \\ 4 & 0 & 1 \\ 2 & 6 & 0 \end{bmatrix}$$

O comando $a = \text{Elemento}(M, 2, 1)$ indica que a irá receber o valor $M(2, 1)$, isto é, $a = 4$;

b) A instrução $S = \text{Sequência}(0, 12, 2)$ atribui a S a lista de números inteiros pares de 0 a 12;

c) Sejam os pontos $A = (1, 1)$ e $B = (4, 2)$. A expressão $\text{Se}(M(1, 1) > 0, \text{Vetor}(A, B))$ ordena ao Geogebra que crie o vetor AB , se $M(2, 1) > 0$.

O primeiro passo para obter o menor caminho de um vértice origem de um grafo a todos os outros vértices é criar o grafo. Para isso é necessário produzir pontos que irão representar os vértices. Depois disso é necessário produzir a matriz de custos, que além de permitir o cálculo do custo total de uma trajetória escolhida também permite visualizar em que vértice uma aresta tem origem e em qual vértice ela incide. O grafo da Figura 12 será usado como forma de validar o procedimento apresentado.

A fim de produzir tal grafo é necessário em primeiro lugar criar os vértices que vão de v_1 a v_6 e uma lista Lv contendo tais vértices. Para isto basta digitar $Lv = \{v_1, v_2, v_3, v_4, v_5, v_6\}$. Também é necessário uma variável Nv para guardar o número de vértices. Isso pode ser feito

digitando-se $Nv = 6$ na barra de comandos ou usando a instrução $Nv = Comprimento(Lv)$. Em seguida é preciso criar a matriz de custos MC , que no caso é dada por

$$MC = \begin{bmatrix} 0 & 3.6 & \infty & \infty & \infty & \infty \\ \infty & 0 & 4.1 & 5.2 & 4 & \infty \\ \infty & \infty & 0 & 3 & 0 & \infty \\ \infty & 5.2 & \infty & 0 & \infty & 2 \\ 2.9 & \infty & 4.6 & 2.7 & 0 & \infty \\ \infty & \infty & 2 & \infty & \infty & 0 \end{bmatrix}$$

Para efeito de programação o símbolo ∞ deve ser trocado por um número positivo P “grande”. No presente caso será utilizado $P = 50$, inserido via barra de comandos.

Para possibilitar a visualização do grafo usa-se a instrução a seguir, inserida via barra de comandos.

*Arcos = Sequência(Sequência(Se(Elemento($M(i, j) < M$,
Vetor(Elemento(Lv, i), Elemento(Lv, j)), ($0, 0$)), $j, 1, n$), $i, 1, n$)*

Não havendo aresta de um vértice v_i para v_j este comando produzirá um vetor nulo $(0,0)$, o qual não terá utilidade para efeitos de cálculos. Deixando-se o comando sem este vetor o Geogebra colocará no seu lugar um ponto de interrogação, mas isso não acarreta erro no processamento. Ao final desta etapa ter-se-á a lista de vetores e o grafo desenhado.

Na hora de executar o programa é preciso especificar os vértices de origem e de destino. Para tornar o processo dinâmico faz-se necessário criar duas variáveis, que serão chamadas de Origem e Destino. No exemplo em questão será estabelecido

$$Origem = 1 \text{ e } Destino = 6,$$

correspondendo aos vértices v_1 e v_6 , respectivamente. Estas atribuições devem ser inseridas na barra de comandos.

Neste ponto é o momento de produzir a matriz de distâncias mínimas prevista no algoritmo de Dijkstra. Os scripts a seguir foram baseados no trabalho de Falcón, Moreno e Ríos (2015). Eles devem ser inseridos em “Propriedades” de qualquer um dos objetos já produzidos, na aba “JavaScript Global”.

```
function distancia(){
  var i,j, d, m, v, a, a0;
  var s=0; n = ggbApplet.getValue("n"); var D = new Array();
  var NM = new Array(); var P = new Array();
  for(i =0; i < n*n; i++){
    D[i] = 100;}
  for(a0 =1; a0 < n+1; a0++){
    for(i =0;i<n;i++){
      NM[i]=0;
      P[i]=-1;
      D[n*n+(a0-1)*n+i]=-1;}
    a=a0;
    D[(a0-1)*n+a-1]=0;
    NM[a-1]=1;
```

```

while(a!=0){
    for(i =0;i<n;i++){
        for(j =0;j<n;j++){
            m=ggbApplet.getListValue("MC",i,j);
            if (NM[i]==0 & m!=0 & D[(a0-1)*n+i]>D[(a0-1)*n+a-1]+m){
                D[(a0-1)*n+i]=D[(a0-1)*n+a-1]+m;
                P[i]=a;
                D[n*n+(a0-1)*n+i]=a;}}
    d=100;
    for(i=0;i<n;i++){
        if (NM[i]==0){
            if (s==0){s=i; d=D[(a0-1)*n+i];}
            else{
                if (D[(a0-1)*n+i]<d){d=D[(a0-1)*n+i];s=i;}}}
    if (d==100){v=0;}
    else{v=s;}
    if (v==0){a=0;}
    else{NM[v]=1;a=v+1;}}
return(D);}
    
```

A execução deste código produz uma matriz linha D que não ficará visível na tela. A fim de visualizar seus elementos pode-se criar uma lista d nas propriedades dos objetos *Origem* e *Destino*. Na aba “Ao Atualizar” de ambos os objetos é preciso digitar

$$ggbApplet.evalCommand("d = \{ + distancia() + \}");$$

Além disso, na parte inferior da janela deve-se selecionar a opção *JavaScript*.

A matriz de distâncias mínimas produzidas pelo programa anterior pode ser visualizada inserindo-se a instrução que segue na barra de entradas.

$$MD=Sequência[Sequência[Se[Elemento[d,(i-1)n+j]==100, \infty, \\ Elemento[d, (i-1)n+j]], j, 1, n], i, 1, n]$$

Neste ponto é necessário criar uma lista auxiliar L com n vértices, inicialmente formada por zeros. Esta lista é atualizada de modo que, ao final do processo, ela contenha os vértices que determinam o menor caminho do vértice *Origem* a qualquer outro vértice. Este procedimento é feito com a instrução que segue.

$$L =Sequência[Se[Elemento[d, n^2 + (Origem - 1)n + i] =? -1, 0, \\ Elemento[d, n^2 + (Origem - 1)n + i]], i, 1, n].$$

O próximo programa, que também deve ser inserido em “Propriedades” na aba “JavaScript Global”, tem por finalidade obter o menor caminho entre os vértices informados.

```

function caminho(){
    var i,j,m;
    var ip = ggbApplet.getValue("Origem");
    var fp = ggbApplet.getValue("Destino");
    var n = ggbApplet.getValue("n");
    var A = new Array();
    
```

```

var C = new Array();
for(i =0;i<n;i++){
    m=ggbApplet.getListValue("L",i+1);A[i]=m;}
    C[0]=A[fp-1]; j=0;
for(i =0;i<n-1;i++)
if (C[i]==0){C[i+1]=0;}
else{
if (C[i]!=ip){
C[i+1]=A[C[i]-1];}
else{
C[i+1]=0;}} return(C);}
    
```

O caminho gerado por esta função também deve ser informado nas propriedades dos objetos *Origem* e *Destino*, na aba “Ao Atualizar”, como segue.

```

ggbApplet.evalCommand("caminho = {" + caminho() + "}");
    
```

A visualização gráfica do menor caminho é produzida com a instrução que segue, inserida na barra de entradas.

```

CamMin = União(Sequência(Se(Elemento(caminho, i)*Elemento(caminho, i + 1) > 0,
Vetor(Elemento(V, Elemento(caminho, i + 1)), Elemento(V, Elemento(caminho, i))), i, 1, n),
Se(Elemento(caminho, 1) > 0, Vetor(Elemento(V, Elemento(caminho, 1)), Elemento(V, Destino)))).
    
```

A Figura 13 mostra o menor caminho quando o vértice origem é v_1 e o destino é v_6 , em relação ao grafo da Figura 12.

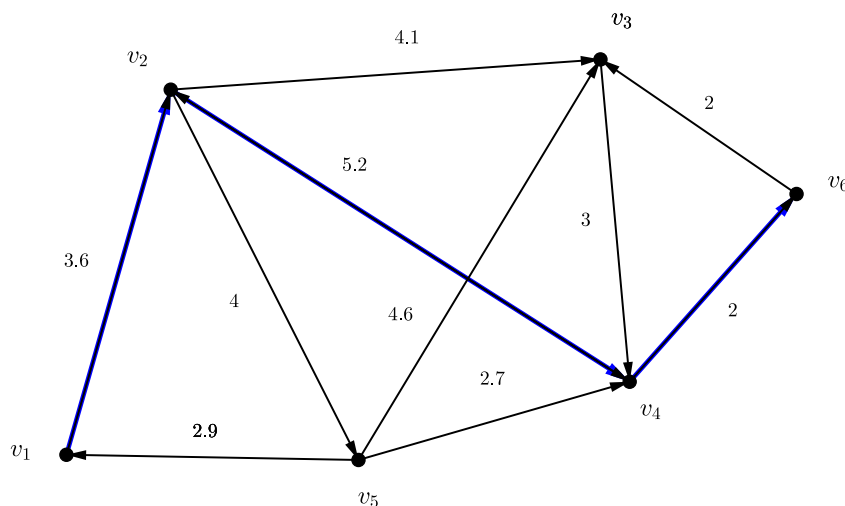


Figura 13: Menor caminho entre os vértices v_1 a v_6

A lista *CamMin* de caminhos obtida foi $CamMin = \{0, 1, 2, 2, 2, 4\}$. Os números que aparecem nesta lista indicam que, partir do vértice que ele representa deve-se ir ao vértice representado pela sua posição na lista. Neste exemplo o vértice 6 (sexto elemento na lista) veio

do vértice 4 (número da posição 6). O vértice 4 por sua vez (quarta posição na lista) veio do vértice 2 (número da quarta posição). Por sua vez o vértice 2 (segundo elemento da lista) veio do vértice 1 (número da segunda posição). O caminho é, portanto, v_1, v_2, v_4, v_6 . A mesma ideia pode ser usada para qualquer outro vértice, desde que o vértice seja considerado como origem.

Considerações finais

Embora a o trabalho desenvolvido tenha apenas um caráter ilustrativo, ele dá uma boa ideia da versatilidade do software Geogebra para tarefas mais complexas. Na ilustração foi apresentado apenas um pequeno problema, mas o processo pode ser facilmente adaptado para um problema de porte maior. A dificuldade em relação ao algoritmo é a produção da matriz de custos, que é um trabalho desgastante, já que a matriz tem que ser produzida por meio de digitação.

Referências

CÁCERES, Francisco Rafael. O uso de grafos como motivação para o estudo de matrizes. Orientador: Amarildo de Vicente. Dissertação (Mestrado) - Unioeste, Cascavel, 2021.

FALCON, Raúl M.; MORENO, Angela; RÍOS, Ricardo **Designing Evacuation Routes with GeoGebra**. Universidade de Sevilla, 2015. Disponível em: <https://idus.us.es/bitstream/handle/11441/69095/Designing%20evacuation%20routes%20with%20GeoGebra.pdf?sequence=1&isAllowed=y>. Acesso em: 12/nov/2021.

RABUSKE, Márcia Aguiar. **Introdução à Teoria dos Grafos**. 1a. ed. Florianópolis: Editora da UFSC, 1992.

Aproximações de irracionais por meio de racionais

Álex Henrique Correa Dias
Universidade Estadual do Oeste do Paraná
alex.dias1@unioeste.br

Prof Dr Rian Lopes de Lima
Universidade Estadual do Oeste do Paraná
rian.lima@unioeste.br

Resumo: Os números irracionais chamam atenção de estudantes e pesquisadores em Matemática há mais de dois mil anos, uma vez que os mesmos se apresentam através de tão somente a negação do que é ser um número racional. Os números irracionais não apresentam padrão na parte decimal e não possuem uma forma simples de representação. Nesse sentido, este trabalho apresenta alguns resultados que fornecem aproximações boas de um número irracional por meio de números racionais, podendo ser tão preciso quanto se desejar.

Palavras-chave: Números Irracionais; Aproximação; Números Racionais.

1 Introdução

Os números irracionais são objetos de estudo para muitos matemáticos justamente por apresentarem características diferenciadas em relação ao conjunto dos racionais, começando pela definição que reside numa negação, provar que algo não se encaixa em determinado formato pode ser trabalhoso e pouco intuitivo. Dentre os muitos números irracionais temos o π e o número de Euler e , cujas provas da irracionalidade requerem conceitos além da matemática básica.

Os números irracionais possuem a característica de não terem uma representação racional $\frac{a}{b}$ com a e b inteiros e $b \neq 0$. Por exemplo, a demonstração de que $\sqrt{2}$ é irracional é um dos primeiros desafios que encontramos num curso de Matemática, e tal prova não costuma ser assimilada com rapidez pelos discentes.

Já que um irracional não pode ser escrito na forma $\frac{a}{b}$ com a e b inteiros e $b \neq 0$, cabe o questionamento se podemos, ao menos, obter uma aproximação racional de forma que o erro pode ser controlado de maneira satisfatória. A resposta a essa pergunta é sim, podemos encontrar delimitações para a diferença de um irracional e um racional de maneira que o erro fique limitado por um racional.

Tal trabalho complementa o estudo apresentado em disciplinas como Análise Real, onde se discute a densidade dos irracionais em \mathbb{R} , porém sem discutir caminhos específicos de como fixar um irracional entre dois racionais, e mais além, escolhendo um erro conveniente.

2 Desenvolvimento

Uma primeira forma de se compreender a aproximação de irracionais por racionais é realizar essa aproximação por números inteiros, como apresentado a seguir. Se for arredondado

um número real, substituindo-o pelo inteiro mais próximo, o erro cometido será, no máximo, igual a $\frac{1}{2}$. Por exemplo, 5,5 por 5, ou 17,8 por 18, o erro será, em cada caso, menor ou igual a $\frac{1}{2}$.

Teorema 1. *Para qualquer número irracional α , existe um único inteiro m tal que*

$$-\frac{1}{2} < \alpha - m < \frac{1}{2}.$$

Prova. Considere α um número irracional. No intervalo $(\alpha - \frac{1}{2}, \alpha + \frac{1}{2})$ deve existir um inteiro m , uma vez que o comprimento do intervalo é 1 e tanto $\alpha - \frac{1}{2}$ quanto $\alpha + \frac{1}{2}$ não são inteiros, ou seja,

$$\alpha - \frac{1}{2} < m < \alpha + \frac{1}{2}.$$

Subtraindo α de ambos os lados da desigualdade, tem-se:

$$-\frac{1}{2} < m - \alpha < \frac{1}{2}.$$

Multiplicando por (-1) segue que:

$$-\frac{1}{2} < \alpha - m < \frac{1}{2}.$$

Que em valor absoluto fica

$$|\alpha - m| < \frac{1}{2}.$$

□

Por exemplo, consideremos $\alpha = \sqrt{2}$. Veja que

$$\begin{aligned} 1 < 2 = \frac{8}{4} < \frac{9}{4} &\Rightarrow 1 < \sqrt{2} < \frac{3}{2} \\ \Rightarrow 0 = 1 - 1 < \sqrt{2} - 1 < \frac{3}{2} - 1 = \frac{1}{2} \\ \Rightarrow -\frac{1}{2} < \sqrt{2} - 1 < \frac{1}{2} \\ \Rightarrow |\sqrt{2} - 1| < \frac{1}{2}. \end{aligned}$$

O próximo teorema fornece uma aproximação com erro controlável:

Teorema 2. *Sejam α um número irracional qualquer e n um número inteiro positivo qualquer. Então existe um número racional $\frac{m}{n}$, tal que*

$$-\frac{1}{2n} < \alpha - \frac{m}{n} < \frac{1}{2n}.$$

Prova. Seja n um número natural. Como $n\alpha$ que é um número irracional, pelo Teorema 1, existe um inteiro m tal que

$$-\frac{1}{2} < n\alpha - m < \frac{1}{2}.$$

Dividindo a desigualdade por n , temos que

$$-\frac{1}{2n} < \alpha - \frac{m}{n} < \frac{1}{2n}.$$

O que fornece um erro controlável por $\epsilon = \frac{1}{2n}$ dependendo apenas do n escolhido:

$$\left| \lambda - \frac{m}{n} \right| < \frac{1}{2n}.$$

□

Um exemplo usando o Teorema 2 é apresentado a seguir:

Determinar uma aproximação racional de $\alpha = \sqrt{3}$ tomando $n = 28$.

Veja que

$$n\alpha = 28\sqrt{3} = \sqrt{2352}.$$

Temos,

$$\begin{aligned} 2304 &< 2352 < 2401 \\ \Rightarrow \sqrt{2304} &< \sqrt{2352} < \sqrt{2401} \\ \Rightarrow 48 &< 28\sqrt{3} < 49. \end{aligned}$$

Suponha que $28\sqrt{3} - 48 > \frac{1}{2}$, daí

$$\begin{aligned} 53\sqrt{3} - 96 &> 1 \\ \Rightarrow 56\sqrt{3} &> 97 \\ \Rightarrow 9408 &> 9409. \end{aligned}$$

O que é um absurdo. Dessa forma, segue que:

$$\begin{aligned} -\frac{1}{2} &< 28\sqrt{3} - 48 < \frac{1}{2} \\ \Rightarrow -\frac{1}{56} &< \sqrt{3} - \frac{48}{28} < \frac{1}{56} \\ \Rightarrow \left| \sqrt{3} - \frac{12}{7} \right| &< \frac{1}{56}. \end{aligned}$$

Para obter melhores aproximações basta tomar um valor de n maior, obtendo assim um erro menor.

O próximo teorema estabelece uma aproximação um pouco melhor:

Teorema 3. *Seja λ um número irracional e k um inteiro positivo, existe um número racional $\frac{m}{n}$, com $n < k$, tal que*

$$-\frac{1}{kn} < \lambda - \frac{m}{n} < \frac{1}{kn}.$$

Prova. Vamos considerar os números irracionais $\lambda, 2\lambda, 3\lambda, \dots, k\lambda$, onde $\alpha_t = [t\lambda]$ (parte inteira) e $\beta_t = t\lambda - \alpha_t$ (parte decimal):

$$\begin{aligned}\lambda &= \alpha_1 + \beta_1 \\ 2\lambda &= \alpha_2 + \beta_2 \\ &\vdots \\ k\lambda &= \alpha_k + \beta_k.\end{aligned}$$

Vamos subdividir o intervalo $[0, 1]$ em k partes iguais: $I_1, I_2, I_3, \dots, I_k$, cada uma com comprimento $\frac{1}{k}$. Com isso, o intervalo I_1 conterá os números entre 0 e $\frac{1}{k}$, I_2 , os números entre $\frac{1}{k}$ e $\frac{2}{k}$, etc.

Vamos analisar dois casos:

Caso 1: *O intervalo I_1 contém um ou mais β 's.* Dessa forma, existe β_n no intervalo I_1 .

O número β_n é igual a $n\lambda - \alpha_n$ e, com isso, temos

$$0 < n\lambda - \alpha_n < \frac{1}{k},$$

porque I_1 é o intervalo de 0 a $\frac{1}{k}$. Daí, segue que

$$-\frac{1}{k} < n\lambda - \alpha_n < \frac{1}{k}$$

e, dividindo por n , tem-se

$$-\frac{1}{kn} < \lambda - \frac{\alpha_n}{n} < \frac{1}{nk},$$

que em valor absoluto fica

$$\left| \lambda - \frac{\alpha_n}{n} \right| < \frac{1}{nk},$$

onde $m = \alpha_n$.

Caso 2: *O intervalo I_1 não contém nenhum dos β 's.* Nesse caso, os β 's estão nos demais intervalos I_2, \dots, I_k . Portanto, terá que existir, pelo menos, um intervalo contendo dois ou mais β 's.

Suponha que $\beta_j > \beta_i$ estejam no mesmo intervalo. Como o comprimento desse intervalo é $\frac{1}{k}$, devemos ter

$$-\frac{1}{k} < 0 < \beta_j - \beta_i < \frac{1}{k}.$$

Mas, $\beta_j = j\lambda - \alpha_j$ e $\beta_i = i\lambda - \alpha_i$, de forma que

$$-\frac{1}{k} < (j\lambda - \alpha_j) - (i\lambda - \alpha_i) < \frac{1}{k},$$

ou ainda,

$$-\frac{1}{k} < (j - i)\lambda - (\alpha_j - \alpha_i) < \frac{1}{k}.$$

Fazendo $j - i = n$ e $\alpha_j - \alpha_i = m$, escrevemos

$$-\frac{1}{k} < n\lambda - m < \frac{1}{k},$$

que dividido por n fica

$$-\frac{1}{nk} < \lambda - \frac{m}{n} < \frac{1}{nk}.$$

Por fim, temos um erro controlável por $\epsilon = \frac{1}{kn}$:

$$\left| \lambda - \frac{m}{n} \right| < \frac{1}{nk}.$$

□

Um exemplo da aplicação desse teorema é apresentado a seguir.

Determinar uma aproximação racional para $\sqrt{3}$ considerando $k = 8$.

Para aplicar o teorema acima, é necessário trabalhar com todos os 8 múltiplos de $\sqrt{3}$.

Para uma primeira verificação, foi trabalhado com $3\sqrt{3} = \sqrt{27}$. Veja que $\alpha_3 = 5$, pois

$$25 < 27 < 36 \Rightarrow 5 < 3\sqrt{3} < 6.$$

Supondo

$$\begin{aligned} 3\sqrt{3} - 5 &< \frac{1}{8} \\ \Rightarrow 24\sqrt{3} - 40 &< 1 \\ \Rightarrow 24\sqrt{3} &< 41 \\ \Rightarrow 1728 &< (41)^2 \\ \Rightarrow 1728 &< 1681, \end{aligned}$$

chegamos em um absurdo, o que implica que $\beta_3 = 3\sqrt{3} - 5$ não está em $I_1 = [0, \frac{1}{8}]$.

Considerando $n = 7$, temos que $7\sqrt{3} = \sqrt{147}$ satisfaz

$$144 < 147 < 169 \Rightarrow 12 < 7\sqrt{3} < 13.$$

Suponha, por absurdo, que

$$7\sqrt{3} - 12 \geq \frac{1}{8}.$$

Daí,

$$\begin{aligned} 56\sqrt{3} &\geq 97 \\ \Rightarrow (56)^2 3 &\geq 97^2 \\ \Rightarrow 9408 = 3136 \cdot 3 &\geq 9409. \end{aligned}$$

Com isso, verifica-se que

$$\begin{aligned}\beta_7 &= 7\sqrt{3} - 12 < \frac{1}{8} \quad (\beta_7 \in I_1) \\ \Rightarrow -\frac{1}{8} &< 0 < 7\sqrt{3} - 12 < \frac{1}{8} \\ \Rightarrow -\frac{1}{7.8} &< \sqrt{3} - \frac{12}{7} < \frac{1}{7.8} \\ \Rightarrow -\frac{1}{56} &< \sqrt{3} - \frac{12}{7} < \frac{1}{56} \\ \Rightarrow \left| \sqrt{3} - \frac{12}{7} \right| &< \frac{1}{56}.\end{aligned}$$

Observe que chegamos na mesma aproximação do exemplo do Teorema 2.

Conclusões

Os Teoremas 2 e 3 forneceram aproximações racionais interessantes para um dado irracional. De maneira comparativa, podemos pontuar que o segundo teorema requer o uso de um n suficientemente grande para gerar o resultado. Já o terceiro teorema 3, apesar de possuir passos e análise mais elaborada, chega no mesmo erro do Teorema 2 com uma partição (k) menor. Sendo assim, ambos os resultados tem suas vantagens e desvantagens. Do teorema 3 ainda podemos encontrar uma aproximação com erro delimitado por $\frac{1}{n^2}$:

$$\left| \alpha - \frac{m}{n} \right| < \frac{1}{nk} < \frac{1}{n^2},$$

uma vez que $n < k$. Essa aproximação parece ser mais simples de se obter, porém tem um erro maior e ainda assim é necessário encontrar o n correto que faz com que β_n esteja no intervalo I_1 . Por fim, pontuamos que o Caso 2 do Teorema 3 demanda uma análise maior ao se obter a aproximação, sendo necessário trabalhar com dois múltiplos do irracional α (um para cada β).

Referências

NIVEN, I. **Números: Racionais e Irracionais**. 1^o ed. Rio de Janeiro: Sociedade Brasileira de Matemática, 2012.

Um estudo sobre os fatores que promovem a evasão e a permanência dos discentes no curso de licenciatura em matemática da Unioeste – campus Cascavel no período de 2016 – 2022.

Nadya Beatriz Antunes Barroso
Universidade Estadual do Oeste do Paraná - Unioeste
nadya.barroso@unioeste.br

Pamela Gonçalves
Universidade Estadual do Oeste do Paraná - Unioeste
pamela.goncalves@unioeste.br

Resumo: Esse trabalho tem o objetivo de analisar os fatores de evasão e permanência dos alunos no curso de Licenciatura em Matemática da Unioeste — Campus Cascavel no período de 2016 – 2022. Será inicialmente apresentado o histórico da instituição e do curso, as legislações vigentes, pesquisas relacionadas com a evasão e a permanência de alunos. O objetivo é estudar os fatores da evasão e permanência dos alunos, com o propósito de tentar perceber padrões na tomada de decisões dos alunos em relação ao curso. A metodologia será baseada na pesquisa qualitativa com abordagem descritiva. Para a obtenção de dados, serão aplicados dois questionários, um para os alunos que evadiram e outro para os alunos que permaneceram no curso, com o objetivo de identificar as causas da evasão e permanência.

Palavras-chave: Ensino Superior; Evasão; Permanência; Unioeste.

1 Introdução

Ingressar no ensino superior é uma oportunidade que para muitos se julga necessária para ter uma boa qualificação profissional e trabalhar com o que gosta. Com o constate crescimento dos cursos de graduação e áreas de atuação o ensino superior se torna cada vez mais atrativo e gerando bastante concorrência. No Brasil, as formas de ingresso mais conhecidas para as instituições de graduação é pelos vestibulares ou pelo Exame nacional do Ensino médio – Enem, aonde a nota obtida pode ser usada em programas como Programa Universidade para todos – PROUNI que oferta bolsas de estudos de 50% e 100% nos cursos de graduação em instituições privadas, Sistema de seleção unificado – SISU que disponibiliza vagas dos cursos de ensino superior nas instituições públicas e o Fundo do financiamento

estudantil – FIES que possibilita o financiamento de cursos de graduação das instituições privadas.

Segundo o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - (INEP) no ano de 2017 e 2020 os números de instituições e cursos de graduação apresentaram os seguintes resultados:

Tabela 1: Dados do Ensino Superior Inep - 2017 e 2020.

2017		
Instituições privadas	Total de instituições	Total de cursos de
	2.152	24.955
Instituições públicas	Total de instituições	Total de cursos de
	296	10.425
Total	2.448	35.380
2020		
Instituições privadas	Total de instituições	Total de cursos de
	2.143	31.147
Instituições públicas	Total de instituições	Total de cursos de
	304	10.806
Total	2.457	41.953

Fonte: Inep 2017 e 2020

Como demonstra a tabela, de 2017 a 2020 houve um aumento de aproximadamente 0,36% no total de instituições e 18,57% em no total de cursos de graduações. Já em relação as matrículas no ensino superior, no ano de 2020 o Brasil possuía um total de 8.680.354 alunos matriculados e em 2017, 8.286.663, também apresentando aumento. Analisando o número total de matrículas por modalidade acadêmica, no ano de 2017 segundo o INEP, os cursos de bacharelado, licenciatura e tecnológico obtiveram respectivamente 5.662.351 matrículas, 1.589.449 matrículas e 999.289 matrículas. Foi observado que (INEP 2017.p. 29):

Há uma certa estabilidade na distribuição da matrícula entre os graus acadêmicos, nos últimos três anos: 69% bacharelado, 19% licenciatura e 12% em cursos tecnológicos. Nos últimos 10 anos, entretanto, o número de alunos em cursos de licenciatura cresceu apenas 49,7%, enquanto o número de alunos de cursos tecnológicos cresceu 141%. No mesmo período, os cursos de bacharelado cresceram 65,6%.

Em um vasto campo de opções, os alunos com o ensino médio completo que tem o intuito de fazer uma graduação dispõem de uma grande decisão, que faz surgir alguns questionamentos: Qual área escolher? Em que área irei me adaptar melhor? Qual a melhor

escolha para meu futuro? Segundo Montmarquette, Cannings E Mahseredjian, (2002 apud. MARTINS E MACHADO, 2018), o processo de tomada de decisão do indivíduo ocorre considerando dois fatores: a existência de sucesso no curso escolhido e o rendimento esperado pela formação recebida. Tendo em vista a ansiedade do ingresso no ensino superior, os alunos com frequência, acabam escolhendo o curso desejado sem pesquisar referente a área, muitas vezes por vontade de familiares ou até mesmo pelo simples fato de ter conseguido uma vaga, o que poderá ser uma possível causa para uma futura evasão.

A adaptação no Ensino Superior - (ES) é difícil para muito dos ingressantes, uma vez que exige mais tempo, determinação e comprometimento aos estudos e atividades acadêmicas do que estão acostumados. Analisar os motivos de evasão se tornam um campo de estudo muito amplo, visto que existem muitos fatores que podem influenciar o discente a evadir do curso. Segundo Silva Filho et al. (2007, p.642):

a evasão estudantil no ensino superior é um problema internacional que afeta o resultado dos sistemas educacionais. As perdas de estudantes que iniciam, mas não terminam seus cursos são desperdícios sociais, acadêmicos e econômicos. No setor público, são recursos públicos investidos sem o devido retorno. No setor privado, é uma importante perda de receitas. Em ambos os casos, a evasão é uma fonte de ociosidade de professores, funcionários, equipamentos e espaço físico).

Por outro lado, podemos refletir sobre a permanência, pois para muitos alunos a conclusão do ES se torna um sonho em construção. Fatores como qualidade de ensino, oportunidades acadêmicas e planos de carreira, são cruciais para a permanência do aluno nos cursos de graduação. Segundo Cislighi (2010, p.20) “a aquisição de conhecimentos assim como a percepção dos progressos e da qualidade na sua formação são fortes motivadores para que os estudantes permaneçam frequentando os cursos e concluam com sucesso a sua passagem pela IES”. Há muitas pesquisas sobre a evasão, cuja preocupação é imensa, porém há de se pensar nas causas da permanência, gerando alguns questionamentos: O que faz o aluno evadir ou permanecer no curso? Quais as expectativas do aluno em relação a sua formação? Quais fatores ajudam ou influenciam na decisão de permanência no curso?

Portanto, esse trabalho tem como objetivo estudar os fatores da evasão e permanência dos alunos no curso de Licenciatura em Matemática da Unioeste – Campus Cascavel nos anos de 2016 a 2022 com o propósito de analisar padrões na tomada de decisões dos alunos em relação ao curso. O recorte, se concentrará no curso de licenciatura em matemática (Unioeste/Campus Cascavel), uma vez que faz parte da formação da pesquisadora, e o período escolhido é relacionado com o tempo de imersão da autora no curso.

2 Objetivo de pesquisa

Este trabalho tem como objetivo identificar os fatores de evasão e permanência no curso de Licenciatura em Matemática na Unioeste – campus Cascavel no período de 2016 a 2022, possibilitando mapear os motivos da escolha de concluir e desistir do curso. Durante a pesquisa queremos possibilitar que os alunos pontuem sua(s) justificativas em relação ao curso e que as informações obtidas ajudem novos alunos que estão passando pela tomada a decisão.

3 Desenvolvimento da pesquisa

A Unioeste é uma universidade pública e gratuita, tendo seu reconhecimento como universidade em 23 de dezembro de 1994, é uma instituição de ensino *multicampi*, formado por cinco campus distribuído em cinco municípios, Cascavel, Francisco Beltrão, Toledo, Foz do Iguaçu e Marechal Cândido do Rondon.

A Unioeste foi criada a partir da junção de cinco faculdades municipais isoladas, em Cascavel (FECIVEL criada em 1972), em Foz do Iguaçu (FACISA criada em 1979), em Marechal Candido Rondon (FACIMAR criada em 1980) em Toledo (FACITOL criada em 1980) e em Francisco Beltrão (FACIBEL criada em 1974). A Unioeste campus – Cascavel, possui 19 cursos de graduação presencial, onde vamos abordar durante a pesquisa especialmente o curso de Licenciatura em Matemática Unioeste - campus Cascavel.

O curso de Licenciatura em Matemática na Unioeste – Campus Cascavel, foi implantado em 1987, está vinculado com o Centro de Ciências Exatas e Tecnológicas (CCET), tendo como duração mínima quatro anos e máxima de sete anos no período noturno. O ingresso no curso pode ser através do processo seletivo do vestibular que oferta 20 vagas ou pelo Sistema de Seleção Unificada - SISU ofertando também 20 vagas. O curso possui nota cinco no Exame Nacional de Desempenho dos Estudantes – ENADE, sendo aplicado a última vez em 2017. O Projeto Político Pedagógico –(PPP) vigente no curso é a Resolução 256/2016 – CEPE aprovado no dia 8 de dezembro de 2016. O curso prevê formar profissionais para atuar no campo da docência em matemática e visa:

Desenvolver em seus alunos a capacidade social de comunicar-se matematicamente, favorecendo a utilização dos conhecimentos matemáticos para a compreensão consciente do mundo circundante, para que eles percebam que esta é uma ciência resultante de uma ação social, que é dinâmica e criativa (MATEMÁTICA – Unioeste, 2022.)¹

A alta evasão no curso de matemática e a pouca formação de discentes é um assunto muito discutido, possuindo também como fator preocupante a permanência prolongada dos alunos. Segundo Appio (2013) o maior número de abandonos no curso de matemática da Unioeste está no concentrado no primeiro ano e através da sua pesquisa aponta as reprovações constantes em algumas matérias contribuem para a evasão do curso.

¹ Disponível em: <https://dmat-unioeste.mat.br/>. Acesso em 10 de novembro 2022.

No curso de Licenciatura em Matemática da Unioeste/Campus Cascavel, existe uma grande problemática em relação a desistência no curso relacionada com o número de reprovações em determinadas disciplinas como mostra a pesquisa de Appio (2013), resultados apontaram que o maior número de reprovações está concentrado no primeiro ano do curso. Além disso, segundo Moraes, (2019, p. 13).

O estudante que coleciona aprovações nas disciplinas do seu curso e ensaja a construção de uma base pertinente de conhecimentos dispõe de maior chance em persistir no seu objetivo de completar a graduação. Em contraste, o estudante que acumula insucessos nas disciplinas, como tradicionalmente ocorre no princípio dos cursos de exatas, acaba se desmotivando em dar continuidade aos seus estudos.

Analisar os fatores de permanência dos alunos não só tem grande importância para os discentes, mas também para a instituição, ter uma visão do que gera continuação do discente no curso. Segundo Cislaghi (2008) o que faz um estudante ficar na dúvida e reconsiderar o objetivo de obter uma formação, ou o que pode aumentar sua motivação para concluir o curso que frequenta, demanda de conhecimentos e o uso de técnicas específicas. Como aponta Maciel, Lima e Gimenez (2016. apud. SALAS et al., 2020):

permanecer na educação superior e concluí-la são ações determinadas por vários elementos e envolvem recursos humanos e econômicos, conjunto que representa, cada vez mais, esforços institucionais e o desenvolvimento de políticas específicas para favorecer o sucesso dos estudantes da Educação Superior pública (p. 761).

Transparecer os fatores de permanência dos alunos no curso de licenciatura em matemática pode ser um grande aliado na tomada de decisões dos discentes, ressaltar as muitas oportunidades que o curso oferece e reacender os motivos de escolha do curso.

4 Metodologia

O trabalho apresentado terá como fundamentação a pesquisa qualitativa, que segundo Vieira (1996 apud Zanella 2011 p. 35) é caracterizada por não usar instrumentos estatísticos para a obtenção e análise de dados. Já para Richardson (1999) podemos afirmar que as investigações que se voltam para uma análise qualitativa têm como objeto de estudo situações complexas ou bastante particulares.

Será abordado também a metodologia de pesquisa descritiva que Gil (2007 apud. Zanella, 2011 p. 34) descreve como, “conhecer a realidade estudada, suas características e seus problemas”. Após a submissão ao processo de análise do comitê de ética², a pesquisa será feita através de aplicação de dois questionários: um para alunos que evadiram de curso e outro para os que permaneceram, ambos no período de 2016 a 2022.

² <https://www.unioeste.br/portal/comites/cep/inicial> . Acesso em nov, 2022.

Em relação a coleta de dados dos alunos que evadiram, será feito um levantamento junto a secretaria acadêmica do campus e no colegiado do curso de licenciatura em matemática. Será realizada também uma pesquisa nos sites disponíveis da instituição com dados dos alunos disponibilizados durante os anos de 2016-2022.

5 Considerações finais

A realidade do curso de licenciatura em matemática da Unioeste é um alto índice de alunos que evadiram e poucos formados, como relatam as pesquisas de Appio (2013), Silva (2020) e Andreta (2013) que serão utilizadas como base e motivação para a conclusão desse trabalho. Explorar e transparecer os motivos dos alunos para sua tomada de decisão é de extrema importância, para que fique claro as realidades enfrentadas pelos discentes durante a conclusão do curso. Por fim, espera-se que esse trabalho seja relevante e motivador para pesquisas futuras, visto que o campo da evasão e permanência de alunos no ensino superior é ainda um campo de pesquisa pouco explorado.

Referências

- APPIO, Andressa. **Classificando dados de evasão do curso de licenciatura em matemática da unioeste – campus cascavel no período de 2003-2010**. 2012. 50 f. TCC (Graduação) - Curso de Matemática, Universidade Estadual do Oeste do Paraná, Cascavel, 2012.
- ANDRETA, Aldioni Adaiani. **INFLUÊNCIA DE FATORES SOCIOECONÔMICOS NA EVASÃO DE ACADÊMICOS DO CURSO DE MATEMÁTICA DA UNIOESTE – CAMPUS DE CASCAVEL**. 2011. 49 f. TCC (Graduação) - Curso de Matemática, Universidade Estadual do Oeste do Paraná, Cascavel, 2011
- CISLAGHI, Renato. **Um modelo de sistema de gestão do conhecimento em um framework para a promoção da permanência discente no ensino de graduação**. 2008. 273 f. Tese (Doutorado) - Curso de Engenharia e Gestão do Conhecimento, Universidade Federal de Santa Catarina, Florianópolis, 2008.
- CURSO de Matemática. **Matemática**. 2022 – Universidade Estadual do Oeste do Paraná. Disponível em: <https://dmat-unioeste.mat.br/>. Acesso em 10 nov. 2022.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Censo da Educação Superior, 2020. Brasília: MEC, 2021.
- INSTITUTO NACIONAL DE ESTUDOS E PESQUISAS EDUCACIONAIS ANÍSIO TEIXEIRA (INEP). Censo da Educação Superior, 2017. Brasília: MEC, 2018.

MARTINS, Felipe dos Santos; MACHADO, Danielle Carusi. **Uma análise da escolha do curso superior no Brasil**. 2018. Disponível em:

<https://www.scielo.br/j/rbepop/a/PGXB5BznhrHkXQX5ktZGdkv/?format=pdf&lang=pt>.
Acesso em: 16 out. 2022.

MORAES, Kaluti; HEIDEMANN, Leonardo; ESPINOSA, Tobias. Métodos ativos de ensino podem ser entendidos como recursos para o combate à evasão em cursos de Ciências Exatas? Uma análise pautada nas ideias de Vincent Tinto. **Caderno Brasileiro de Ensino de Física**, Rio Grande do Sul, v. 37, n. 2, p. 369-405, ago. 2020.

RICHARDSON, Roberto J. **Pesquisa social: métodos e técnicas**. São Paulo: Atlas, 1999. 334 p

SALAS, Shila Antuanett Neciosup; VELASQUE, Luciane de Souza; GIRALDO, Victor Augusto; QUINTANEIRO, Wellerson. **Permanência na Educação Superior Pública: um olhar de licenciandos em Matemática de duas universidades**. 2020. 25 f. Tese (Doutorado) - Curso de Matemática, Universidade Estadual de Montes Claros, Montes Claros, 2020

SILVA, Gabriela Artini da. **Classificação de Dados em Busca de Padrões que Evidenciem a Evasão no Curso de Licenciatura em Matemática**. 2020. 48 f. TCC (Graduação) - Curso de Matemática, Universidade Estadual do Oeste do Paraná, Cascavel, 2020.

SILVA FILHO, Roberto Leal Lobo e; MOTEJUNAS, Paulo Roberto; HIPÓLITO, Oscar; LOBO, Maria Beatriz de Carvalho Melo. **A evasão no ensino superior brasileiro**. 2007. Disponível em: <https://www.scielo.br/j/cp/a/x44X6CZfd7hqF5vFNnHhVWg/?lang=pt&format=pdf>. Acesso em: 15 out. 2022.

ZANELLA, Liane Carly Hermes. **Metodologia de Pesquisa**. 2011. Disponível em: <https://www.atfcursosjuridicos.com.br/repositorio/material/3-leitura-extra-02.pdf>. Acesso em: 17 out. 2022.

PNLD: Dados históricos, seu funcionamento e algumas considerações relevantes acerca da aquisição

Ricardo Mondini Ferrazza¹
Unioeste
ricardoferrazza7@gmail.com

Thais de Souza²
Unioeste
thaissouza38@hotmail.com

Jean Sebastian Toillier³
Unioeste
jean.toillier@unioeste.com

Resumo: Este trabalho tem como objetivo fazer uma análise acerca do Plano Nacional do Livro Didático, com ênfase nas mudanças ocorridas em sua linha do tempo, seu funcionamento e levantamento de alguns pontos como: quem são as editoras que são escolhidas? Os professores são contemplados com as coleções escolhidas? E ao final, nossas conclusões, embasadas em textos científicos.

Palavras-chave: PNLD; Educação Matemática; FNDE.

1 Introdução

Este texto é formulado a partir de dois projetos de iniciação científica voluntária, que tinham como objetivo inicial o estudo dos paradidáticos, principalmente aqueles que fazem parte do acervo de livros didáticos antigos do Laboratório de Ensino de Matemática (LEM), da Universidade Estadual do Oeste do Paraná (Unioeste), campus de Cascavel. Assim, foi feita a catalogação desses livros, uma breve análise de seus conteúdos, além do estudo do Plano Nacional do Livro e do Material Didático (PNLD). Ainda durante a catalogação, percebemos que os paradidáticos tinham ligação direta com os livros didáticos, pois faziam um complemento do conteúdo, utilizando de histórias que auxiliam no entendimento do aluno.

Desse modo, para o presente texto, abordaremos de modo exclusivo a história do PNLD, pois é um projeto que está em vigor desde 1929, que sofreu mudanças ao longo dos

¹ Acadêmico do curso de licenciatura em matemática; voluntário no PIC.

² Acadêmica do curso de licenciatura em matemática; voluntária no PIC.

³ Professor do curso de Licenciatura em Matemática da Unioeste, campus de Cascavel.

anos, além de ter nomes diferentes e diversas formas de execuções. Falaremos sobre o longo histórico, seu modo de funcionamento e faremos algumas críticas a escolha da editora contemplada.

2 Histórico do Plano Nacional do Livro Didático - PNLD

O atual PNLD teve início em 1929, com a criação do Instituto Nacional do Livro (INL), um órgão para legislar sobre as políticas do livro didático, auxiliando na legitimidade do livro didático e que, conseqüentemente, contribuiu no aumento da sua produção. Na sequência, em 1938, é instituída a Comissão Nacional do Livro Didático (CNLD), que foi consolidada pelo Decreto de Lei nº 1.006, de 30/12/1938, que estabelecia sua primeira política de legislação e controle de produção e circulação do livro didático no país. Após sete anos, por meio do Decreto-Lei nº 8.460, de 26/12/1945, se torna mais sólida a legislação sobre as condições de produção, importação e utilização do livro didático, limitando ao professor a escolha do livro a ser utilizado pelos alunos, conforme definido no art. 5º (BRASIL,2022).

Um acordo em 1966 entre o ministério da educação (MEC) e a Agência Norte-Americana para o Desenvolvimento Internacional (Usaid) libera a criação da Comissão do Livro Técnico e Livro Didático (Colted), tendo em vista coordenar as ações referentes à produção, edição e distribuição do livro didático. Segundo o Fundo Nacional de Desenvolvimento da Educação (FNDE), foi assegurado ao MEC recursos suficientes para a distribuição gratuita do 51 milhões de exemplares no período de três anos, nos próximos anos garantiu financiamento do governo a partir de verbas públicas, dando continuidade ao programa.

Em 1968 é criado o FNDE que é uma autarquia federal vinculada ao Ministério da Educação (MEC). Isso quer dizer que ele é uma entidade pública, fiscalizada pelo Estado, mas que possui autonomia para sua gestão. Sendo responsável também pelo PNLD. Assim, em 1970, implementou-se o sistema de coedição de livros com as editoras nacionais, com recursos do INL e, no ano seguinte, o INL passou a desenvolver o Programa do Livro Didático para o Ensino fundamental (Pliddef), quando assumiu as atribuições administrativas e de gerenciamento dos recursos financeiros que estavam até então sob responsabilidade da Colted. Nesse período foi necessária a contrapartida das unidades federativas após o término do convênio MEC/Usaid, o que se efetivou com a implantação do sistema de contribuição financeira das unidades federadas para o fundo do livro didático (BRASIL, 2022).

Após o decreto nº 77.107 de 4 de fevereiro de 1976, o governo federal assumiu a compra de boa parcela dos livros para distribuir para parte das escolas e unidades federadas. Com isso, foi extinto o INL, e foi criada a Fundação do Material Escolar (Fename), que

passou a ser responsável pela execução do programa do livro didático tendo auxílio das unidades federativas as quais eram responsáveis por promover os recursos, porém a falta de recursos acabava excluindo maioria das escolas municipais do programa.

Segundo o portal do FNDE (BRASIL 2022), sete anos após a Fename assumir a execução do programa do livro didático, é criada a Fundação de Assistência ao Estudante (FAE) que incorpora o Plidef que a substitui. Nesse mesmo período, o grupo de trabalho encarregado do exame dos problemas relativos aos livros didáticos propõe a participação dos professores na escolha dos livros e ampliação do programa incluindo as demais séries do ensino fundamental. Finalmente, em 1985, com a edição do decreto nº 91.542, de 19/08/1985, o Plidef dá lugar ao Programa Nacional do Livro Didático, implantando diversas mudanças, como:

- Indicação do livro didático pelos professores;
- Reutilização do livro;
- Extensão da oferta aos alunos de 1^a e 2^a série das escolas públicas e comunitárias;
- Controle do processo decisório pela FAE;
- Garantia do critério de escolha do livro pelos professores.

No ano de 1992 devido às limitações orçamentarias, a distribuição do livro didático é comprometida, tendo um recuo na abrangência da distribuição, restringindo-se o atendimento até a 4^a série do ensino fundamental. Já em 1993, a resolução do FNDE vincula recursos para a aquisição dos livros didáticos destinados aos alunos das redes públicas de ensino, estabelecendo-se, assim, um fluxo regular de verbas para a aquisição e distribuição do livro didático (BRASIL, 2022).

Com participação do MEC/FAE/UNESCO entre os anos de 1993 e 1994 são definidos critérios para a avaliação dos livros didáticos, sendo feita uma publicação de um guia. De forma progressiva, se tem a volta da universalização da distribuição do livro didático no ensino fundamental e uma expansão das disciplinas envolvidas na produção dos livros didáticos (BRASIL, 2022).

Em 1996 é iniciado o processo de avaliação pedagógica dos livros inscritos no PNLD, tendo neste mesmo ano o primeiro “Guia do livro Didático” de 1^a a 4^a série. A avaliação foi feita pelo MEC conforme critérios previamente discutidos. Esse procedimento foi aperfeiçoado e é aplicado até os dias atuais. Assim, livros que apresentem erros conceituais, indução a erros, desatualização, preconceito ou discriminação de qualquer natureza são excluídos do Guia do Livro Didático⁴.

⁴ O Guia de livros didáticos é uma peça fundamental do PNLD e tem, a princípio, três funções. A primeira delas é de orientação aos docentes da Educação Básica para que possam melhor realizar o processo de escolha das obras que serão utilizadas nas escolas do Brasil. O Guia enuncia, também, os pressupostos da avaliação pedagógica, efetivados em observância ao Edital do Programa e em conformidade com afirmações acadêmicas atualizadas. Apresenta reflexões fundamentais à formação docente no tocante aos processos de mediação pedagógica. (BRASIL,2022).

Em 1997, a FAE é extinta e passa a responsabilidade pela política de execução do PNLD para o Fundo Nacional de Desenvolvimento da Educação (FNDE). Neste mesmo ano, o programa é ampliado e o Ministério da Educação passa a adquirir, de forma contínua livros didáticos para de alfabetização, língua portuguesa, matemática, ciências, história e geografia para todos os alunos de 1^a a 8^a série do ensino Fundamental das escolas públicas.

No início dos anos 2000, é inserida no PNLD a distribuição de dicionários da língua portuguesa para uso dos alunos de 1^a a 4^a série, e pela primeira vez na história do programa, a entrega dos livros didáticos é feita no ano anterior do seu uso. Além disso, a partir de 2001, de forma gradativa, foi ampliado o atendimento aos alunos com deficiência visual que frequentam as salas de aula do ensino regular das escolas públicas, utilizando de livros didáticos em braile. Atualmente, esses alunos são atendidos também com livros em libras, caractere ampliado e na versão *MecDaisy* (versão digital de livros com suporte de leitura de texto e descrição de gráficos) (BRASIL,2022).

Em 2002, os anos finais tiveram distribuição integral do livro didático. Já em 2004, foi estipulada a meta de que os alunos matriculados no ensino fundamental tivessem dicionários de língua portuguesa para utilizarem durante sua vida escolar (BRASIL, 2022).

Segundo o FNDE em 2003 é publicada a Resolução do FNDE n^o 38, de 15/10/2003, que institui o Programa Nacional do Livro Didático para o Ensino Médio (PNLEM). No ano de 2004 os anos iniciais recebem distribuição integral. O atendimento ao Ensino médio foi instituído progressivamente e no ano de 2004 é criada uma ferramenta importante para a execução do programa, o *Siscort*, sistema direcionado a registrar e controlar o remanejamento de livros e distribuição da Reserva Técnica referente ao ensino fundamental, o sistema foi implantado em todos os estados (BRASIL, 2022).

No ano de 2005 é feita a distribuição do PNLD e do PNLEM com distribuição parcial ao ensino médio, ainda foram distribuídos três tipos de dicionários para serem adaptados ao nível de ensino do aluno. Para os alunos que têm surdez e utilizam a Língua Brasileira de Sinais (Libras), houve distribuição do dicionário enciclopédico ilustrado trilíngue - Língua Brasileira de Sinais/Língua Portuguesa/Língua Inglesa (BRASIL,2022).

O FNDE em 2007 adquiriu 110,2 milhões de livros para reposição e complementação dos livros. Nesse mesmo ano, para o ensino médio houve distribuição parcial, seguindo a meta progressiva de universalização do material para esse segmento (BRASIL, 2022).

Com a publicação da **resolução do FNDE 18, de 24/04/2007**, é regulamentado o Programa Nacional do Livro Didático para a Alfabetização de Jovens e Adultos (PNLA) para distribuição, a título de doação, de obras didáticas às entidades parceiras do Programa Brasil Alfabetizado (PBA), com vistas à alfabetização e à escolarização de pessoas com idade de 15 anos ou mais (BRASIL, 2022).

Para utilização em 2009, houve distribuição integral nos anos finais e distribuição parcial no ensino médio. Houve aquisição de 114,8 milhões de livros didáticos para 36,6 milhões de alunos da educação pública, para utilização a partir de 2010. Ainda em 2009, foram investidos R\$18,8 milhões na compra de 2,8 milhões de obras do PNLA.

Além disso, são publicadas duas importantes resoluções. A primeira, a **resolução FNDE nº. 51, de 16/09/2009**, regulamentando o Programa Nacional do Livro Didático para a Educação de Jovens e Adultos (PNLD EJA). O programa abrange o PNLA, visto que atende estudantes jovens e adultos também em sua fase de alfabetização. A segunda, **resolução do FNDE nº. 60, de 20/11/2009**, estabelece novas regras para participação no PNLD: a partir de 2010, as redes públicas de ensino e as escolas federais devem aderir ao programa para receber os livros didáticos.

Em 2010, os anos iniciais têm distribuição integral e o atendimento ao EJA é ampliado. Foram adquiridas, pela primeira vez, obras complementares para os alunos de 1º e 2º anos do ensino fundamental, além disso é publicado o **Decreto nº. 7.084, de 27/01/2010**, que dispõe sobre os procedimentos para execução dos programas de material didático: o Programa Nacional do Livro Didático (PNLD) e o Programa Nacional Biblioteca da Escola (PNBE) (BRASIL, 2022).

Segundo o FNDE (BRASIL, 2022), em 2011, foi distribuído integralmente livros para o ensino médio, inclusive na modalidade Educação de Jovens e Adultos. Foram distribuídos ainda livros para os alunos do ensino fundamental da Educação de Jovens e Adultos, conforme previa a **Resolução do FNDE nº. 51 de 2009**. Com o investimento de R\$ 140,6 mil reais, foram distribuídos 14,1 milhões de livros, atendendo 5 milhões de alunos.

Importantes e significativos avanço nos programas do livro ocorreram na área de tecnologia. Em 2012, foi publicado edital para formação de parcerias para estruturação e operação de serviço público e gratuito de disponibilização de materiais digitais a usuários da educação nacional. O edital tinha por objetivo a constituição de acordos de cooperação entre o FNDE e instituições interessadas para a estruturação e a operação de serviço virtual para disponibilização de obras digitais e outros conteúdos educacionais digitais para professores, estudantes e outros usuários da rede pública de ensino brasileira, com ênfase nos títulos do PNLD, do PNBE e de outras ações governamentais na área de material escolar, por meio de tecnologia que assegure o atendimento em escala nacional e proteja os direitos autorais digitais e a propriedade intelectual dos acervos (BRASIL, 2022).

Além disso, também em 2012, pela primeira vez, os editores puderam inscrever no âmbito do PNLD 2014, objetos educacionais digitais complementares aos livros impressos. Esse novo material multimídia, que inclui jogos educativos, simuladores e infográficos animados, foi enviado para as escolas em DVD para utilização pelos alunos dos anos finais do ensino fundamental no ano letivo de 2014. O DVD é um recurso adicional para as escolas que ainda não tinham internet. Os novos livros didáticos trouxeram também endereços on-line

para que os estudantes tivessem acesso ao material multimídia, complementando o assunto estudado, com o intuito de tornar as aulas mais modernas e interessantes (BRASIL, 2022).

Além disso, também em 2012, pela primeira vez, os editores puderam inscrever no âmbito do PNLD 2014, objetos educacionais digitais complementares aos livros impressos. Esse novo material multimídia, que inclui jogos educativos, simuladores e infográficos animados, foi enviado para as escolas em DVD para utilização pelos alunos dos anos finais do ensino fundamental no ano letivo de 2014. O DVD é um recurso adicional para as escolas que ainda não tinham internet. Os novos livros didáticos trouxeram também endereços on-line para que os estudantes tivessem acesso ao material multimídia, complementando o assunto estudado (BRASIL, 2022).

Já para o ano letivo de 2015, foi lançado em 2012 o edital que prevê que as editoras podem apresentar obras multimídias, reunindo livro impresso e livro digital (BRASIL, 2022). A versão digital deve trazer o mesmo conteúdo do material impresso mais os objetos educacionais digitais, como vídeos, animações, simuladores, imagens, jogos, textos, entre outros itens para auxiliar na aprendizagem. O edital também permite a apresentação de obras somente na versão impressa, para viabilizar a participação das editoras que ainda não dominam as novas tecnologias. Esse material foi destinado aos alunos e professores do ensino médio da rede pública.

O Decreto nº 9.099, de 18 de julho de 2017, unificou as ações de aquisição e distribuição de livros didáticos e literários, anteriormente contempladas pelo Programa Nacional do Livro Didático (PNLD) e pelo Programa Nacional Biblioteca da Escola (PNBE). Com nova nomenclatura, o Programa Nacional do Livro e do Material Didático – PNLD também teve seu alvo ampliado com a possibilidade de inclusão de outros materiais de apoio à prática educativa para além das obras didáticas e literárias: obras pedagógicas, *softwares* e jogos educacionais, materiais de reforço e correção de fluxo, materiais de formação e materiais destinados à gestão escolar, entre outros (BRASIL,2022).

3 Funcionamento

A adesão das escolas e os sistemas de ensino estaduais, municipais e federais segundo o FNDE (BRASIL, 2022) deve ser feita formalmente, seguindo editais publicados no site do programa observando as normas, prazos e procedimento estabelecidos pelo Ministério da Educação. Os beneficiários que não desejarem mais fazer parte do projeto devem fazer a solicitação de suspensão das remessas ou realizar a exclusão do programa. Aos que fazem parte do programa a adesão deve estar sempre atualizada, sendo feita até o final do mês de maio do ano anterior àquele em que a entidade deseja ser atendida.

O mesmo ocorre para as editoras que desejam enviar coleções, estas também devem atender aos editais presentes no site. As coleções que são inscritas passam por um processo de

triagem realizado pelo Instituto de Pesquisas Tecnológicas do Estado de São Paulo (IPT). Os livros que forem selecionados são encaminhados ao MEC, órgão responsável pela avaliação pedagógica. Os avaliadores dos livros são escolhidos pela SEB conforme os critérios divulgados no edital, alguns desses critérios incluem se as obras seguem a ABNT e se possuem resistência para que durem por pelo menos 3 anos. Assim são elaboradas resenhas dos livros aprovados que passam a compor o guia e livros didáticos.

A e corpo docente analisam as obras e escolhem os livros que serão utilizados em sua escola. A escolha é realizada via internet. Segundo o FNDE (BRASIL, 2022), após a compilação dos dados referentes aos pedidos realizados pela internet inicia o processo de negociação com as editoras. A aquisição é realizada por ilegitimidade de licitação tendo em vista que as escolhas dos livros são efetivadas pelas escolas e que são editoras específicas que detém o direito de produção de cada livro. Concluída a negociação é firmado contrato e informa-se as quantidades de livros a serem produzidos e as localidades de entrega para as editoras. Esses processos podem ser organizados conforme a Figura 1.

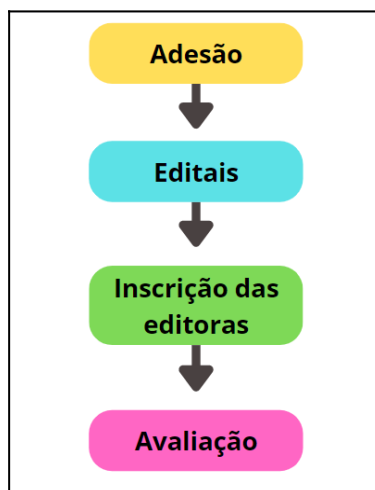


Figura 1: Primeira fase FNDE

Fonte: Acervo dos autores

A distribuição dos livros é realizada pela Empresa Brasileira de Correios e Telégrafos, que leva os livros diretamente das editoras para as escolas. Os livros chegam às escolas entre outubro do ano anterior ao atendimento e o início do ano letivo, vale ressaltar que nas zonas rurais, as obras são entregues nas sedes das prefeituras ou nas secretarias municipais de educação, estas devem efetivar a entrega dos livros.

Segundo FNDE os livros possuem a seguinte distribuição: Para o fundamental cada aluno tem o direito a um exemplar dos livros de língua portuguesa, matemática, ciências, história, geografia e língua estrangeira. Esses livros seguem o ciclo de três anos, ou seja, ao

final do ano letivo os alunos devem realizar a devolução dos livros para a escola para que possam ser usados por outros alunos. Os únicos livros que não seguem esse critério são os de língua estrangeira que podem ser de inglês ou espanhol são os únicos consumíveis. O critério de reutilização é mantido para o ensino médio, este que são entregues os livros de língua portuguesa, matemática, geografia, história, biologia, física e química, mantêm-se também o critério de utilização dos livros de língua estrangeira. No caso da sociologia e da filosofia, os alunos receberão um livro em volume único, ou seja, este será utilizado durante os três anos do ensino médio.

O ciclo de uso dos livros é de três anos, ocorrendo de forma alternada, são contemplados alunos do 1º ao 5º ano do ensino fundamental, 6º ano ao 9º ano do ensino fundamental e 1º ao 3º ano do ensino médio. Em cada ano um dos três anos do ensino básico é contemplado com a entrega de livros até se fechar o ciclo de três anos assim se repetindo. Entre os ciclos são realizadas reposições caso ocorram extravios ou perdas e acréscimos por matrícula. Segundo o FNDE a distribuição é realizada de acordo projeções do censo escolar referente a dois anos anteriores ao ano do programa, que é o censo disponível no momento do processamento da escolha feita pelas escolas. Dessa maneira, poderá haver pequenas oscilações entre o número de livros e o de alunos. Para realizar o ajuste, garantindo o acesso de todos os alunos aos materiais, é necessário fazer o seu remanejamento, daquelas escolas onde estejam excedendo para aquelas onde ocorra falta de livros. As escolas podem recorrer ainda à reserva técnica, percentual de livros disponibilizado às Secretarias Estaduais de Educação para atender a novas turmas e matrículas.

4 Críticas ao Programa

Apesar dos pontos positivos presente no projeto, vale ressaltar algumas críticas quem vem historicamente ligadas a publicação de livros didáticos no Brasil, que com o tempo se estenderam apenas no mercado de publicações e chegaram até o PNLD.

A partir do desenvolvimento do mercado financeiro, grandes corporações se tornam protagonistas em algumas áreas, o que não é diferente quando pensamos na produção de livros didáticos no território nacional. Segundo Célia Cristina de Figueiredo (2007), no Brasil há grandes editoras desde os anos de 1970 até os dias de hoje, dentre elas destacamos o Grupo Saraiva, Abril, Moderna, FTD, IBEP e Scipione, as quais têm monopolizado o mercado de publicação dos livros didáticos, tendo apenas poucos títulos sendo publicados por outros grupos, esses não tão populares. Mas qual é o critério para essa escolha contínua de grandes corporações?

Em uma crítica publicada no site do Senado são levantadas por Tatiana Feitosa Britto as mesmas críticas expressas por Cecília.

Na prática, a avaliação vem causando bastante desconforto junto ao mercado editorial. Nas primeiras experiências, grandes editoras tiveram livros consagrados pelo uso reprovados pelos avaliadores. Para

se ter uma ideia, na avaliação de 1996, de um total de 466 livros inscritos, 281 (60%) foram classificados como não recomendados. Mais recentemente, coleções que haviam sido aprovadas com louvor em avaliações anteriores foram excluídas do Guia do Livro Didático (BRITTO,2011, p. 10).

Segundo Célia Cristina de Figueiredo Cassiano (2007) na década de 1980 essas editoras se caracterizavam por sua grande maioria terem gestão familiar e serem brasileiras. A presença no mercado até hoje deve-se ao fato que com o passar dos anos as editoras aprenderam como realizar o *marketing* para seus livros o que os tornou mais atrativos para serem adquiridos pelas escolas. Porém, segundo Tatiane Feitosa de Britto (2011) como não há alternância das obras aprovadas, com os mesmos autores constando nas publicações de livro, a variedade fica escassa, principalmente para os anos finais da educação básica e isso torna mais complexa a escolha dos livros a serem utilizados pelos docentes em sala de aula.

Com as coleções publicadas pelas editoras, é de caráter dos docentes fazer a escolha das coleções que será utilizado para auxílio. Em cada região há critérios para as escolhas dos livros, seja para cada escola, grupo de escola ou único para toda rede de ensino. Em casos de escolha por rede de ensino, são contados os votos para cada coleção e a que tiver mais votos será então escolhido para o ciclo do próximo ano, em alguns casos mais de uma coleção é escolhida para o uso, porém, normalmente, apenas uma coleção é escolhida de toda rede de escolas. Segundo Tatiane Feitosa de Britto (2011) esse tipo de escolha é mais barato, facilita para a impressão dos livros, o que é viável para o programa e para as editoras, porém, essa escolha pode não ter sido a mais proveitosa em relação ao conteúdo.

O também segundo Tatiane Feitosa de Britto (2011) problema referente a escolha de livros por rede de escolhas se torna um dos problemas que deixam o programa menos atrativo. Tendo em vista o tempo que leva para a análise de cada livro, que é feita a partir da observação dos conteúdos presentes em cada coleção, atividades, pesquisar, textos e até mesmo o uso de imagens, isso pode ser visto como uma falta de consideração pelo trabalho de análise realizado pelos professores das escolas públicas de ensino, isso acaba gerando uma frustração.

Conclusões

Com base no que foi apresentado concluímos que o PNLD teve e tem um longo trajeto, com inúmeras mudanças, sendo consolidado, expandido e aprimorado. Em análise ao histórico concluímos que na maioria das vezes as grandes editoras foram as contempladas, levantamos algumas críticas, mas, por mais que essas críticas existam o programa é essencial para o ensino público no Brasil. Entretanto, sabemos que a utilização de livros de forma gratuita pelos alunos nas escolas brasileiras é um processo importante para uma democratização de acesso aos estudos.

Agradecimentos

A Professora Dulcyene Maria Ribeiro e ao professor Jean Sebastian Toillier pela orientação, apoio, confiança, pelo suporte no seu pouco tempo livre, pelas suas correções e incentivos. Em homenagem a Juliano Canute.

Referências

BRASIL. **PNLD- O que é e como funciona**. Brasil escola, s/a. Disponível em: <https://educador.brasilecola.uol.com.br/sugestoes-pais-professores/pnld-que-e-como-funciona.htm>. Acesso em: 02 nov. 2022

BRITTO, Tatiane Feitosa de. **O Livro Didático, o Mercado Editorial e os Sistemas de Ensino Apostilados**. Centro de Estudos da Consultoria do Senado, 2011. Disponível em: <https://www12.senado.leg.br/publicacoes/estudos-legislativos/tipos-de-estudos/textos-para-discussao/td-92-o-livro-didatico-o-mercado-editorial-e-os-sistemas-de-ensino-apostilados>. Acesso em: 03 nov. 2022.

CASSIANO, Célia Cristina de Figueiredo. **O mercado do livro didático no Brasil: da criação do programa nacional do livro didático (PNLD) à entrada do capital internacional espanhol (1985-2007)**. 2007. 252 f. Tese (Doutorado) - Curso de História, Pontifícia Universidade Católica de São Paulo, São Paulo, 2007. Disponível em: <https://tede2.pucsp.br/bitstream/handle/10614/1/Celia%20Cristina%20de%20Figueiredo%20Cassiano.pd>. Acesso em: 02 nov. 2022.

PNLD. **Ministério da Educação**, s/a. Disponível em: <http://portal.mec.gov.br/busca-geral/318-programas-e-aco-es-1921564125/pnld-439702797/12391-pnld>. Acesso em: 02 nov. 2022.

BRASIL. Decreto nº 1.006, de 30 de dezembro de 1938. Estabelece as condições de produção, importação e utilização do livro didático. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, v. 4, p. 350, 1938.

BRASIL. Decreto nº 8.460, de 26 de dezembro de 1945. Consolida a legislação sobre as condições de produção, importação e utilização do livro didático. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, p. 19208, 1945.

BRASIL. Decreto nº 77.107, de 4 de fevereiro de 1976. Dispõe sobre a edição e distribuição de livros textos e dá outras providências. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, p. 1681, 1976.

BRASIL. Decreto nº 91.542, de 19 de agosto de 1985. Institui o Programa Nacional do Livro Didático, dispõe sobre sua execução e dá outras providências. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, p. 12178, 1985.

BRASIL. Decreto nº 7.084, de 27 de janeiro de 2010. Dispõe sobre os programas de material didático e dá outras providências. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, 2010.

BRASIL. Decreto nº 9.099, de 18 de julho de 2017. Dispõe sobre o Programa Nacional do Livro e do Material Didático. Lex: Coleção de Leis do Brasil, São Paulo: Diário Oficial da União, 2017.