
Mathematica

Revista eletrônica de divulgação matemática

Universidade Estadual do Oeste do Paraná - UNIOESTE
Centro de Ciências Exatas e Tecnológicas

Mathematica

Revista eletrônica de divulgação matemática

Conselho editorial:

Sandro Marcos Guzzo - UNIOESTE (Editor chefe)
Edson Carlos Licurgo Santos - UNIOESTE (Editor assistente)
Esdras Teixeira Costa - UFG (Editor assistente)
Fabiana Magda Garcia Papani - UNIOESTE (Editor assistente)
Karina Schiabel - UFSCar (Editor assistente)

Arte da Capa:

Diagramação:

Sandro Marcos Guzzo

Ficha Catalográfica...

Será solicitada assim que o ISSN estiver disponível...

Apresentação

A Revista Mathematica surgiu no final de 2019, com uma conversa informal entre alguns professores do curso de Matemática da Unioeste. Desejávamos inicialmente abrir um espaço para divulgação dos resultados obtidos em projetos de iniciação científica, monografias de conclusão de curso, dissertações de mestrado e pesquisas individuais de professores. Muitos destes trabalhos produzem belos textos de matemática que acabam por serem conhecidos por um número reduzido de pessoas. Surgiu então a ideia de criar uma revista que pudesse divulgar estes resultados na forma de textos curtos, didáticos e com conteúdo matemático que pudesse interessar aos alunos de graduação. Neste contexto, uma revista poderia também atrair a atenção de professores, pesquisadores e acadêmicos de instituições de todo o país.

A nossa intenção era publicar o primeiro número no ano de 2020. Contudo o ano de 2020 iniciou e com ele a pandemia causada pelo Coronavírus. Assim como todas as demais atividades, as atividades universitárias ficaram prejudicadas neste momento. As aulas foram suspensas e os projetos de pesquisa e de iniciação científica foram reestruturados. Não conseguimos no ano de 2020 divulgar a revista de forma satisfatória. Acreditamos ainda que mesmo chegando ao conhecimento de alguns professores e acadêmicos a existência da revista, havia outras preocupações maiores do que enviar um texto para uma revista nova e desconhecida.

Contudo, no ano de 2021, a Revista Mathematica recebeu quatro textos que foram avaliados e considerados aptos para publicação. Temos assim nosso primeiro volume. O conselho editorial agradece aos autores pelo envio dos trabalhos e também à comissão científica pelas contribuições feitas durante o processo de avaliação e correção dos trabalhos. Esperamos que em 2022 consigamos continuar o trabalho e publicar mais um volume.

O conselho editorial.

Índice de trabalhos

Uma demonstração da desigualdade isoperimétrica utilizando a desigualdade de Brunn-Minkowski	9
Métodos do Gradiente e de Newton em Otimização Quadrática Convexa	23
Solubilidade por meio de radicais de equações polinomiais de grau menor ou igual a quatro	39
Aproximação de funções por série de Fourier-Bessel	51

Uma demonstração da desigualdade isoperimétrica utilizando a desigualdade de Brunn-Minkowski

Felipe Costa - Universidade Estadual do Oeste do Paraná

(Recebido em 23/11/2021. Aceito em 03/12/2021. Publicado em 22/12/2021)

Resumo: Neste trabalho apresentamos uma demonstração da desigualdade isoperimétrica fazendo uso da desigualdade de Brunn-Minkowski. A vantagem desta abordagem é a sua extensão natural ao caso tridimensional, além de servir como um exemplo de aplicação do Cálculo Variacional no estudo de problemas geométricos.

Palavras-chave: Curvas planas; Desigualdade de Brunn-Minkowski; Desigualdade isoperimétrica.

1 Introdução

Para introduzir a desigualdade isoperimétrica vamos pensar em uma curva Γ como a trajetória contínua descrita por uma partícula no plano. Dizemos que Γ é uma curva fechada se o ponto inicial e o ponto final da trajetória são iguais. Neste caso, vamos denotar por Ω a região delimitada pela curva Γ .

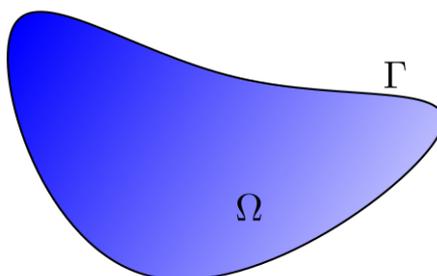


Figura 1: A curva Γ delimita a região Ω

A desigualdade isoperimétrica surge como uma consequência do

Problema isoperimétrico, que consiste em obter, dentre todas as curvas fechadas, sem autointerseções e de comprimento fixo, aquela que delimita a região de maior área.

Este é um dos problemas mais famosos e antigos da geometria, tendo como origem o mito da fundação de Cartago (veja VIRGÍLIO ou BLÅSJÖ).

Os Gregos já sabiam que a solução do problema isoperimétrico era uma circunferência. Mas foi somente em 1879, com K. Weierstrass, que uma demonstração suficientemente satisfatória foi apresentada. Na verdade, o resultado segue como uma consequência da teoria desenvolvida por Weierstrass denominada Cálculo das Variações, onde o problema isoperimétrico é um dos problemas típicos abordados por esta teoria.

Dada uma curva fechada Γ em \mathbb{R}^2 delimitando uma região Ω , vamos denotar por $L(\Gamma)$ o comprimento desta curva e por $A(\Omega)$ a área da região delimitada pela mesma. Se sabemos a priori que uma solução do problema isoperimétrico é uma circunferência \mathcal{C} , então qualquer outra curva fechada Γ com comprimento $L(\Gamma) = L(\mathcal{C})$ deve delimitar uma área $A(\Omega) \leq A(\mathcal{C})$. Suponhamos que o raio da circunferência \mathcal{C} seja igual a r , com isso podemos escrever

$$A(\Omega) \leq A(\mathcal{C}) = \pi r^2 = \frac{(2\pi r)^2}{4\pi} = \frac{L(\mathcal{C})^2}{4\pi} = \frac{L(\Gamma)^2}{4\pi},$$

ou seja,

$$A(\Omega) \leq \frac{L(\Gamma)^2}{4\pi}. \tag{1}$$

A desigualdade (1) é denominada desigualdade isoperimétrica. Observe que (1) foi obtida a partir da hipótese de que a circunferência é uma solução do problema isoperimétrico, o que, por enquanto, torna esta desigualdade apenas uma conjectura.

Por outro lado, caso provemos a desigualdade (1) para toda curva fechada Γ delimitando uma região Ω , então a circunferência seria uma solução do problema isoperimétrico, pois $L(\Gamma)^2/4\pi$ seria uma cota superior para todas as áreas de regiões delimitadas por curvas fechadas de comprimento $L(\Gamma)$ fixado, a qual é atingida quando Γ é uma circunferência de raio $L(\Gamma)/2\pi$. Isto responderia uma parte do problema isoperimétrico, restando apenas provar que a circunferência é a única curva plana com esta propriedade.

Teorema 1 (Desigualdade isoperimétrica). *Seja Γ uma curva fechada, sem autointerseções, de comprimento $L(\Gamma)$ e delimitando uma região Ω . Se $A(\Omega)$ é a área de Ω , então*

$$L(\Gamma)^2 - 4\pi A(\Omega) \geq 0.$$

Além disso, a igualdade ocorre se, e somente se, Γ é uma circunferência.

Para provar o Teorema 1, além das fórmulas de primeira variação do comprimento e da área, obtidas na Seção 3, vamos utilizar um resultado conhecido como Desigualdade de Brunn-Minkowski que trata sobre a área da soma de subconjuntos do espaço euclidiano \mathbb{R}^n . Evidentemente, estamos interessados no caso $n = 2$, mas vale ressaltar que a mesma desigualdade no caso $n = 3$ pode ser usada para estender a desigualdade isoperimétrica para o contexto de superfícies (veja MONTIEL e ROS).

Dados A e B dois subconjuntos quaisquer de \mathbb{R}^n , o conjunto soma $A + B$ é definido por

$$A + B = \{a + b \in \mathbb{R}^n : a \in A, b \in B\}.$$

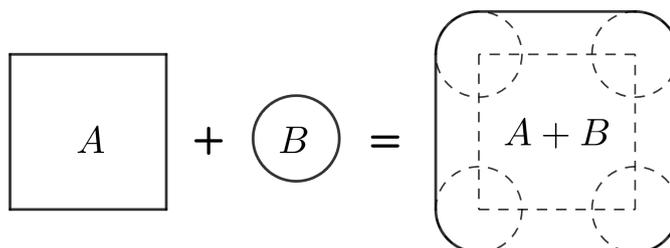


Figura 2: Soma do quadrado A com o círculo B

Queremos somar subconjuntos de \mathbb{R}^n que tenham áreas bem definidas. Estes conjuntos são chamados conjuntos Lebesgue mensuráveis, ou simplesmente conjuntos mensuráveis.

A Desigualdade de Brunn-Minkowski que será enunciada adiante é válida no contexto de conjuntos mensuráveis, mas, para os nossos propósitos, será suficiente considerar apenas subconjuntos abertos, visto que, todo aberto de \mathbb{R}^n é (Lebesgue) mensurável (veja STEIN e SHAKARCHI).

Teorema 2 (Desigualdade de Brunn-Minkowski). *Sejam Ω_1 e Ω_2 dois subconjuntos abertos e limitados de \mathbb{R}^n . Então*

$$A(\Omega_1)^{1/n} + A(\Omega_2)^{1/n} \leq A(\Omega_1 + \Omega_2)^{1/n}.$$

A demonstração do Teorema 2 pode ser encontrada em STEIN e SHAKARCHI.

Observe que está implícito na conclusão do Teorema 2 que soma de dois subconjuntos mensuráveis de \mathbb{R}^n é um conjunto mensurável. Na verdade, é possível mostrar se A ou B é um conjunto aberto em \mathbb{R}^n , então $A + B$ é um conjunto aberto. Assim, de acordo com o que foi destacado acima, $A + B$ é um conjunto mensurável.

Para utilizar o Teorema 2 em conjunto com as fórmulas de variação do comprimento e da área precisamos restringir a classe de curvas que estamos considerando, isto é, devemos considerar curvas de classe C^2 . Esta restrição deve-se ao fato de que a nossa abordagem passa pelo conceito de curvatura de uma curva. Além disso, queremos variar a curva Γ dentro de um conjunto aberto $N_\varepsilon(\Gamma)$ de modo que as curvas da variação ainda sejam fechadas, sem autointerseções e de classe C^2 . O conjunto aberto $N_\varepsilon(\Gamma)$ é chamado vizinhança tubular da curva Γ e pode ser pensado como uma faixa de largura 2ε ao longo de Γ . A construção de $N_\varepsilon(\Gamma)$ pode ser encontrada em ALENCAR, et al.

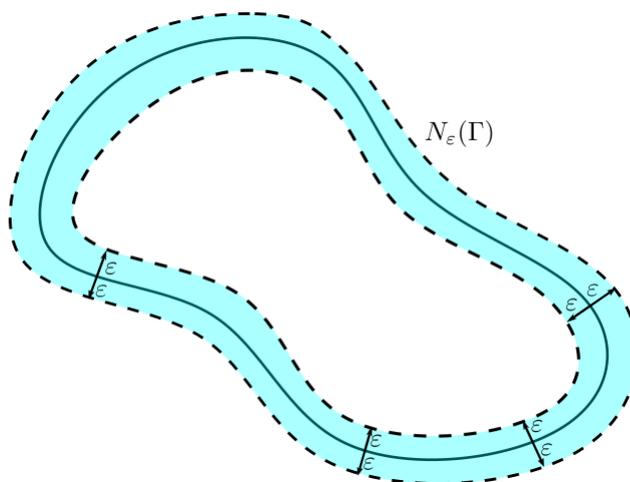


Figura 3: Vizinhança tubular

A parte da igualdade no Teorema 1 será demonstrada observando que as curvas planas fechadas de classe C^2 que satisfazem $L(\Gamma)^2 - 4\pi A(\Omega) = 0$ são aquelas que têm curvatura constante não nula. Esta informação caracteriza as circunferências.

2 Curvas em \mathbb{R}^2

Uma curva plana parametrizada é uma aplicação contínua $\alpha : I \rightarrow \mathbb{R}^2$, dada por $\alpha(t) = (x(t), y(t))$, onde $I \subset \mathbb{R}$ é um intervalo. A curva α será denominada de classe C^k se as funções coordenadas $x, y : I \rightarrow \mathbb{R}$ são de classe C^k .

O conjunto imagem da aplicação α , definido por

$$\Gamma = \{(x(t), y(t)) \in \mathbb{R}^2 : t \in I\}$$

é chamado traço da curva α . Neste caso, dizemos que α é uma parametrização de Γ .

Exemplo 1. Dado $r > 0$, a curva $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ dada por

$$\alpha(t) = (r \cos(t), r \sin(t))$$

tem como traço uma circunferência de raio r e centro na origem.

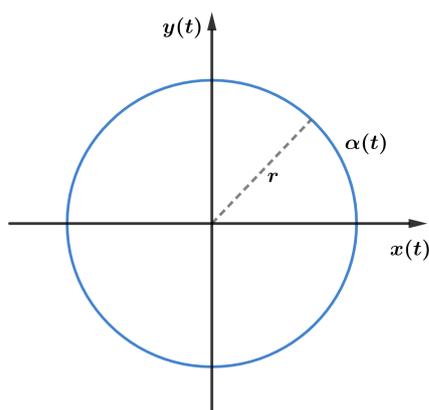


Figura 4: Circunferência de raio r e centro na origem

O próximo exemplo mostra que uma curva no plano pode ter autointerseções.

Exemplo 2. Considere a curva $\alpha : \mathbb{R} \rightarrow \mathbb{R}^2$ dada por

$$\alpha(t) = (t^3 - 4t, t^2 - 4).$$

O traço da curva α possui uma autointerseção na origem pois $\alpha(-2) = \alpha(2) = (0, 0)$.

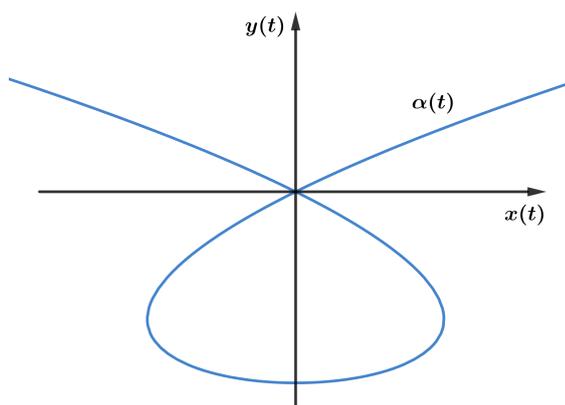


Figura 5: Traço da curva definida por $\alpha(t) = (t^3 - 4t, t^2 - 4)$

Se uma curva α está definida em um intervalo fechado $I = [a, b]$, os pontos $\alpha(a)$ e $\alpha(b)$ são chamados ponto inicial e ponto final de α , respectivamente. A curva $\alpha : [a, b] \rightarrow \mathbb{R}^2$ é dita fechada se $\alpha(a) = \alpha(b)$. Note a curva apresentada no exemplo 1 é fechada já que $\alpha(0) = \alpha(2\pi)$.

Dizemos que uma curva $\alpha : I \rightarrow \mathbb{R}^2$ é simples se a aplicação α é injetora. Observe que a curva apresentada no exemplo 2 não é simples pois $\alpha(-2) = \alpha(2)$.

Neste trabalho estamos particularmente interessados em curvas planas que reúnem as duas últimas propriedades acima. Precisamente, uma curva $\alpha : [a, b] \rightarrow \mathbb{R}^2$ é dita fechada e simples se $\alpha(a) = \alpha(b)$ e a restrição de α ao intervalo $[a, b]$ é injetora. Uma curva fechada e simples é também denominada curva de Jordan.

Seja $\alpha : I \rightarrow \mathbb{R}^2$ uma curva plana parametrizada dada por $\alpha(t) = (x(t), y(t))$. O vetor tangente de α em $t_0 \in I$ é definido por

$$\alpha'(t_0) = (x'(t_0), y'(t_0)).$$

Quando $\alpha'(t) \neq (0, 0)$ para todo $t \in I$, dizemos que a curva α é regular. Portanto, ao longo do traço de uma curva regular fica bem definida uma reta tangente em cada ponto $\alpha(t)$ na direção do vetor $\alpha'(t)$. Se $\alpha'(t_0) = (0, 0)$, então dizemos que α tem um ponto singular em t_0 .

De agora em diante, iremos considerar apenas curvas regulares de classe C^k , com $k \geq 2$, definidas em intervalos fechados e limitados, ficando assim assumida tacitamente estas hipóteses no decorrer deste texto.

O comprimento de uma curva $\alpha : [a, b] \rightarrow \mathbb{R}^2$ é, por definição,

$$L(\alpha) = \int_a^b \|\alpha'(t)\| dt,$$

onde

$$\|\alpha'(t)\| = \sqrt{(x'(t))^2 + (y'(t))^2}.$$

Observe que se $\|\alpha'(t)\| = 1$ para todo $t \in [a, b]$, então $L(\alpha) = b - a$, ou seja, o comprimento da

curva é igual ao comprimento do intervalo $[a, b]$. Por outro lado, se

$$\int_a^t \|\alpha'(s)\| ds = t - a, \quad \text{para todo } t \in [a, b],$$

então, pelo Teorema Fundamental do Cálculo, temos $\|\alpha'(t)\| = 1$ para todo $t \in [a, b]$.

Dizemos que a curva $\alpha : I \rightarrow \mathbb{R}^2$ está parametrizada pelo comprimento de arco se $\|\alpha'(t)\| = 1$ para todo $t \in I$. Uma propriedade muito importante envolvendo as curvas regulares e a parametrização pelo comprimento de arco é dada pela seguinte

Proposição 3. *Dada uma curva regular $\alpha : I \rightarrow \mathbb{R}^2$. Existe uma curva regular $\beta : [0, l] \rightarrow \mathbb{R}^2$ com o mesmo traço da curva α e que satisfaz $\|\beta'(t)\| = 1$ para todo $t \in [0, l]$.*

A proposição 3 garante que sempre é possível reparametrizar o traço de uma curva regular pelo comprimento de arco, conseqüentemente, o vetor tangente desta nova parametrização é unitário em todo ponto.

Exemplo 3. A curva $\alpha : [0, 2\pi] \rightarrow \mathbb{R}^2$ dada no exemplo 1 é regular, mas não está parametrizada pelo comprimento de arco para todo valor de $r \neq 1$ pois

$$\|\alpha'(t)\| = \sqrt{(-r \operatorname{sen}(t))^2 + (r \operatorname{cos}(t))^2} = r.$$

Por outro lado, é simples verificar que a curva $\beta : [0, 2\pi r] \rightarrow \mathbb{R}^2$ dada por

$$\beta(t) = (r \operatorname{cos}(t/r), r \operatorname{sen}(t/r))$$

possui o mesmo traço da curva α (é uma circunferência de raio r e centro na origem) e está parametrizada pelo comprimento de arco, independentemente do valor de $r > 0$.

Quando a curva $\alpha : I \rightarrow \mathbb{R}^2$ estiver parametrizada pelo comprimento de arco, denotaremos o vetor tangente (unitário) desta curva no ponto $t \in I$ por $T(t)$, ou seja,

$$T(t) = (x'(t), y'(t)).$$

A partir do vetor tangente $T(t)$, podemos definir o vetor normal (unitário) $N(t)$ da curva α no ponto $t \in I$ da seguinte maneira

$$N(t) = (-y'(t), x'(t)).$$

Observe que o vetor normal $N(t)$ é obtido através da rotação do vetor $T(t)$ por um ângulo de 90° no sentido anti-horário, assim, para cada $t \in I$, o conjunto $\{T(t), N(t)\}$ é uma base ortonormal positiva de \mathbb{R}^2 denominada referencial de Frenet de α .

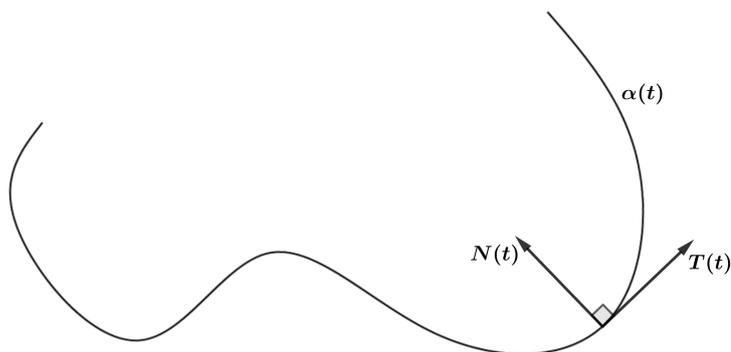


Figura 6: Referencial de Frenet de α

Estamos quase prontos para definir a função curvatura de uma curva no plano, restando apenas estabelecer o seguinte resultado:

Proposição 4. *Se uma função $X : I \rightarrow \mathbb{R}^2$ é diferenciável e $\|X\| : I \rightarrow \mathbb{R}$ é uma função constante, então $X'(t) \perp X(t)$ para todo $t \in I$.*

Prova. Derivando a equação $\|X(t)\|^2 = \text{constante}$, obtemos

$$0 = \frac{d}{dt}(X(t) \cdot X(t)) = X'(t) \cdot X(t) + X(t) \cdot X'(t) = 2X'(t) \cdot X(t).$$

Segue que $X'(t) \cdot X(t) = 0$ e, conseqüentemente, $X'(t) \perp X(t)$. □

Seja $\alpha : I \rightarrow \mathbb{R}^2$ uma curva de classe C^k , $k \geq 2$, parametrizada pelo comprimento de arco. Por definição, a função vetor tangente unitário $T : I \rightarrow \mathbb{R}^2$ satisfaz as hipóteses da Proposição 4. Assim, podemos concluir que $T'(t) \perp T(t)$ para todo $t \in I$. Este fato e a definição do vetor normal $N(t)$ garantem a existência de uma função $\kappa : I \rightarrow \mathbb{R}$, tal que

$$T'(t) = \kappa(t)N(t) \quad \text{para todo } t \in I. \tag{2}$$

Definição 5. A função $\kappa : I \rightarrow \mathbb{R}$, definida pela equação (2), é denominada função curvatura da curva α . Dizemos que $\kappa(t)$ é a curvatura de α no ponto $t \in I$.

Segue da equação (2) que a curvatura de α em $t \in I$ mede a variação do vetor tangente neste ponto, visto que

$$\|T'(t)\| = |\kappa(t)|.$$

Antes de prosseguirmos, deixemos registrado o seguinte par de equações, válidas para toda curva $\alpha : I \rightarrow \mathbb{R}^2$ parametrizada pelo comprimento de arco, conhecidas como Equações de Frenet de α :

$$\begin{cases} T'(t) = \kappa(t)N(t) \\ N'(t) = -\kappa(t)T(t) \end{cases} \quad \text{para todo } t \in I.$$

A primeira das Equações de Frenet é simplesmente a definição da função curvatura e a segunda equação segue da Proposição 4 aplicada à função vetor normal $N : I \rightarrow \mathbb{R}^2$ da curva α e do fato

de $\{T(t), N(t)\}$, $t \in I$, formar uma base ortonormal de \mathbb{R}^2 (a definição e os principais resultados sobre bases ortonormais em \mathbb{R}^n podem ser encontrados em ANTON e RORRES). Os detalhes sobre a dedução das Equações de Frenet podem ser encontrados em TENENBLAT.

Em relação ao traço da curva $\alpha : I \rightarrow \mathbb{R}^2$, podemos interpretar a curvatura como a medida do quanto uma curva deixa de ser uma reta. Esta interpretação pode ser justificada pelo seguinte resultado:

Proposição 6. *A função curvatura de uma curva $\alpha : I \rightarrow \mathbb{R}^2$ é identicamente nula se, e somente se, o traço de α está contido em uma reta.*

Prova. Se a função curvatura da curva α é identicamente nula, então $T'(t) = (0, 0)$ para todo $t \in I$. Como I é um intervalo, concluímos que a função vetor tangente $\alpha'(t)$ é constante, digamos, $\alpha'(t) = \alpha'(t_0) = v$ para todo $t \in I$, onde $t_0 \in I$ é um ponto fixado. Escolhendo arbitrariamente $t_1 \in I$, podemos escrever

$$\alpha(t) - \alpha(t_1) = \int_{t_1}^t \alpha'(t) dt = \int_{t_1}^t v dt = (t - t_1)v,$$

ou melhor,

$$\alpha(t) = \alpha(t_1) + (t - t_1)v.$$

Mostramos assim que o traço de α está contido na reta que passa pelo ponto $\alpha(t_1)$ na direção do vetor v .

Por outro lado, se o traço da curva α está contido em uma reta e α está parametrizada pelo comprimento de arco, então

$$\alpha(t) = P_0 + tv \quad \text{para todo } t \in I,$$

onde $P_0 \in \mathbb{R}^2$ é um ponto da reta e $v \in \mathbb{R}^2$ é um vetor diretor unitário desta reta. Daí, para cada $t \in I$, podemos escrever

$$\kappa(t)N(t) = T'(t) = \alpha''(t) = (0, 0).$$

Como o vetor $N(t)$ é unitário, concluímos que $\kappa(t) = 0$ para todo $t \in I$. □

Outro fato intuitivo que pode ser confirmado a partir do conceito de curvatura é a impressão de que uma circunferência parece “se curvar” de igual modo em cada um de seus pontos. A tradução matemática desta observação é que a função curvatura de uma circunferência é constante e não nula. Tanto esta afirmação quanto a sua recíproca (se a curvatura de uma curva de Jordan é constante e não nula, então esta curva é uma circunferência) podem ser provadas, e serão de grande importância neste texto, pois, como veremos, as curvas de Jordan de comprimento fixo e que delimitam a maior área possível são aquelas que têm curvatura constante. Portanto, faz-se necessário registrar a seguinte

Proposição 7. *Uma curva de Jordan $\alpha : I \rightarrow \mathbb{R}^2$ tem curvatura constante igual a $k_0 \neq 0$ se, e somente se, o traço de α é uma circunferência de raio $r = 1/k_0$.*

Utilizando a parametrização $\beta(t) = (r \cos(t/r), r \sin(t/r))$, dada no exemplo 3, podemos calcular facilmente a curvatura da circunferência de raio r já que, neste caso,

$$T'(t) = \beta''(t) = -\frac{1}{r}\beta(t) = \frac{1}{r}N(t).$$

Segue da equação (2) que $\kappa(t) = 1/r$ para todo $t \in [0, 2\pi r]$. A demonstração completa da proposição 7 pode ser encontrada em TENENBLAT.

Vamos finalizar esta seção enunciando um resultado que é uma consequência do Teorema de Green. Este resultado será usado para calcular a área delimitada por uma curva de Jordan. Mas antes disso, convém lembrar que uma curva de Jordan $\alpha : I \rightarrow \mathbb{R}^2$ está positivamente orientada se, para cada $t \in I$, o vetor normal $N(t)$ aponta para a região delimitada por α .

Lema 8. *Seja $\alpha : [a, b] \rightarrow \mathbb{R}^2$ uma curva de Jordan positivamente orientada e definida por $\alpha(t) = (x(t), y(t))$. Então a área da região Ω , delimitada pela curva α , é dada por*

$$A(\Omega) = \frac{1}{2} \int_a^b (x(t)y'(t) - y(t)x'(t))dt.$$

É importante observar que no lema 8 a curva α não precisa estar parametrizada pelo comprimento de arco. Além disso, denotando por $\alpha'(t)^\perp$ a rotação do vetor tangente de α no ponto $t \in I$ por um ângulo de 90° no sentido positivo (note que $T^\perp = N$), a fórmula que fornece a área da região Ω , delimitada por α , pode ser reescrita na forma

$$A(\Omega) = -\frac{1}{2} \int_a^b \alpha(t) \cdot \alpha'(t)^\perp dt.$$

3 Fórmulas de variação do comprimento e da área

Ao longo desta seção, $\alpha : [0, l] \rightarrow \mathbb{R}^2$ será uma curva de Jordan de classe C^k , $k \geq 2$, parametrizada pelo comprimento arco. Além disso, denotaremos por Γ o traço de α .

Dada uma função diferenciável $f : [0, l] \rightarrow \mathbb{R}$, podemos encontrar um número $\delta > 0$ tal que, se $|t| < \delta$, então $tf([0, l]) \subset (-\varepsilon, \varepsilon)$, onde $\varepsilon > 0$ é escolhido de modo que $N_\varepsilon(\Gamma)$ seja uma vizinhança tubular de Γ . Suponhamos que a função f satisfaz $f(0) = f(l)$.

Uma variação da curva $\alpha : [0, l] \rightarrow \mathbb{R}^2$ correspondente à função diferenciável $f : [0, l] \rightarrow \mathbb{R}$ é uma família de curvas dada por

$$\Gamma_t(f) = \{\alpha_f : [0, l] \rightarrow \mathbb{R}^2 : \alpha_f(s) = \alpha(s) - tf(s)N(s)\}.$$

Observe que $\Gamma_0(f) = \alpha : [0, l] \rightarrow \mathbb{R}^2$ para toda função f e que $\Gamma_t(1)$ é uma família de curvas paralelas à curva α para todo $t \in (-\delta, \delta)$.

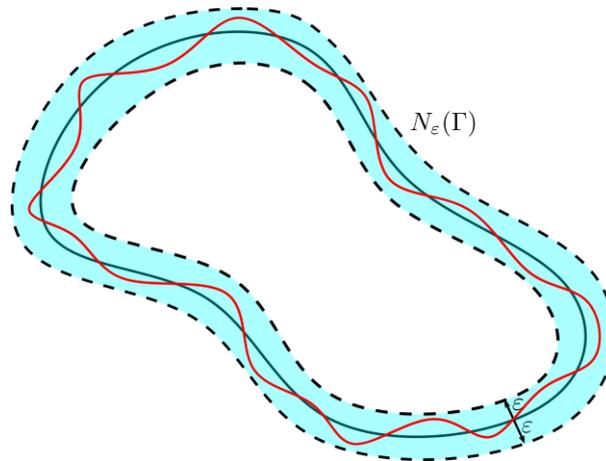


Figura 7: Variação da curva α

Lema 9 (Primeira variação do comprimento). *Seja $\Gamma_t(f)$, $|t| < \delta$, para algum $\delta > 0$, uma variação da curva de Jordan $\alpha : [0, l] \rightarrow \mathbb{R}^2$ correspondente à função $f : [0, l] \rightarrow \mathbb{R}$. Então a função $L : (-\delta, \delta) \rightarrow \mathbb{R}$ dada por*

$$t \in (-\delta, \delta) \mapsto L(t) = L(\Gamma_t(f))$$

é diferenciável e satisfaz

$$\left. \frac{d}{dt} \right|_{t=0} L(\Gamma_t(f)) = \int_0^l \kappa(s) f(s) ds.$$

Prova. Dada uma curva $\alpha_f \in \Gamma_t(f)$, temos

$$\alpha_f(s) = \alpha(s) - tf(s)N(s).$$

Derivando em relação a s encontramos

$$\alpha'_f(s) = T(s) - tf'(s)N(s) - tf(s)N'(s).$$

Usando as equações de Frenet ficamos com

$$\alpha'_f(s) = (1 + tf(s)\kappa(s))T(s) - tf'(s)N(s). \tag{3}$$

Como o conjunto $\{T(s), N(s)\}$ forma uma base ortonormal de \mathbb{R}^2 , temos

$$\|\alpha'_f(s)\|^2 = (1 + tf(s)\kappa(s))^2 + (tf'(s))^2.$$

Observe que, apesar de a curva $\alpha : [0, l] \rightarrow \mathbb{R}^2$ estar parametrizada pelo comprimento de arco, o mesmo não precisa ocorrer com as curvas α_f da variação de α . De todo modo, podemos escrever

$$L(\Gamma_t(f)) = \int_0^l \|\alpha'_f(s)\| ds = \int_0^l \sqrt{(1 + tf(s)\kappa(s))^2 + (tf'(s))^2} ds.$$

Finalmente, derivando sob o sinal de integração e fazendo $t = 0$, concluímos que

$$\left. \frac{d}{dt} \right|_{t=0} L(\Gamma_t(f)) = \int_0^l \left. \frac{d}{dt} \right|_{t=0} \sqrt{(1 + tf(s)\kappa(s))^2 + (tf'(s))^2} ds = \int_0^l \kappa(s) f(s) ds.$$

□

Lema 10 (Primeira variação da área). *Para cada $t \in (-\delta, \delta)$, seja $\Omega_t(f)$ a região delimitada pela curva de Jordan $\alpha_f \in \Gamma_t(f)$. Então a função $A : (-\delta, \delta) \rightarrow \mathbb{R}$ dada por*

$$t \in (-\delta, \delta) \mapsto A(t) = A(\Omega_t(f))$$

é diferenciável e satisfaz

$$\left. \frac{d}{dt} \right|_{t=0} A(\Omega_t(f)) = \int_0^l f(s) ds.$$

Prova. Dada uma curva $\alpha_f \in \Gamma_t(f)$, o lema 8 nos diz que

$$A(\Omega_t(f)) = -\frac{1}{2} \int_0^l \alpha_f(s) \cdot \alpha'_f(s)^\perp dt.$$

Utilizando a equação (3) e as definições de T e N encontramos

$$\alpha'_f(s)^\perp = (1 + tf(s)\kappa(s))N(s) + tf'(s)T(s).$$

Efetuada a operação $\alpha_f(s) \cdot \alpha'_f(s)^\perp$ obtemos

$$\alpha_f(s) \cdot \alpha'_f(s)^\perp = (1 + tf(s)\kappa(s))\alpha(s) \cdot N(s) - tf(s)(1 + tf(s)) + tf'(s)\alpha(s) \cdot T(s).$$

Derivando a expressão acima em relação a t , fazendo $t = 0$ e usando as equações de Frenet, obtemos

$$\left. \frac{d}{dt} \right|_{t=0} \alpha_f(s) \cdot \alpha'_f(s)^\perp = \alpha(s) \cdot (f(s)T(s))' - f(s).$$

Usando a identidade $[\alpha(s) \cdot (f(s)T(s))]' = f(s) + \alpha(s) \cdot (f(s)T(s))'$ podemos escrever

$$\left. \frac{d}{dt} \right|_{t=0} \alpha_f(s) \cdot \alpha'_f(s)^\perp = [\alpha(s) \cdot (f(s)T(s))]' - 2f(s).$$

Finalmente

$$\begin{aligned} \left. \frac{d}{dt} \right|_{t=0} A(\Omega_t(f)) &= -\frac{1}{2} \int_0^l \left. \frac{d}{dt} \right|_{t=0} \alpha_f(s) \cdot \alpha'_f(s)^\perp dt \\ &= -\frac{1}{2} \alpha(s) \cdot (f(s)T(s)) \Big|_0^l + \int_0^l f(s) ds \\ &= \int_0^l f(s) ds. \end{aligned}$$

Na última linha usamos a condição $f(0) = f(l)$ e o fato de α ser uma curva de Jordan regular. □

4 Demonstração da desigualdade isoperimétrica

Nesta seção vamos demonstrar a desigualdade isoperimétrica. Aqui faz-se necessário enunciar o Teorema isoperimétrico no contexto de curvas regulares de classe C^2 .

Teorema 11 (Desigualdade isoperimétrica). *Seja Γ o traço de uma curva de Jordan regular de classe C^2 , comprimento $L(\Gamma)$ e delimitando uma região $\Omega \subset \mathbb{R}^2$. Se $A(\Omega)$ é a área de Ω , então*

$$L(\Gamma)^2 - 4\pi A(\Omega) \geq 0.$$

Além disso, a igualdade ocorre se, e somente se, Γ é uma circunferência.

Prova. Suponhamos que a curva $\alpha : [0, l] \rightarrow \mathbb{R}^2$ cujo traço é igual a Γ está parametrizada pelo comprimento de arco. A demonstração será dividida em duas partes:

- (i) Para $t > 0$ suficientemente pequeno considere a variação $\Gamma_t(1)$ da curva α correspondente à função $f \equiv 1$. Por simplicidade, denotemos por Ω_t a região delimitada pelas curvas da variação $\Gamma_t(1)$. Deste modo, se D_t é um disco aberto de raio t e centrado na origem, temos

$$\Omega_t \supset \Omega + D_t.$$

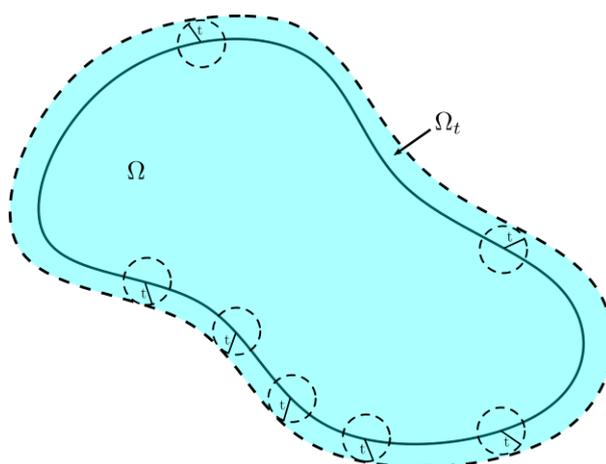


Figura 8: Conjunto $\Omega + D_t$

Da relação $A(\Omega_t) \geq A(\Omega + D_t)$ e da desigualdade de Brunn-Minkowski (Teorema 2) concluímos que

$$\begin{aligned} A(\Omega_t) &\geq A(\Omega + D_t) \\ &\geq [A(\Omega)^{1/2} + A(D_t)^{1/2}]^2 \\ &= A(\Omega) + 2A(\Omega)^{1/2}A(D_t)^{1/2} + A(D_t) \\ &= A(\Omega) + 2tA(\Omega)^{1/2}\sqrt{\pi} + \pi t^2 \\ &> A(\Omega) + 2tA(\Omega)^{1/2}\sqrt{\pi}, \end{aligned}$$

como $t > 0$,

$$\frac{A(\Omega_t) - A(\Omega)}{t} > 2A(\Omega)^{1/2}\sqrt{\pi}.$$

Fazendo $t \rightarrow 0^+$ encontramos

$$\left. \frac{d}{dt} \right|_{t=0} A(\Omega_t) \geq 2A(\Omega)^{1/2}\sqrt{\pi}.$$

Usando a fórmula de primeira variação da área (Lema 10) podemos escrever

$$\int_0^l f(s)ds \geq 2A(\Omega)^{1/2}\sqrt{\pi}.$$

Por outro lado, estamos supondo $f \equiv 1$ e $\alpha : [0, l] \rightarrow \mathbb{R}^2$ parametrizada pelo comprimento de arco, ou seja, $l = L(\Gamma)$. Segue que

$$L(\Gamma) \geq 2A(\Omega)^{1/2}\sqrt{\pi}$$

e assim,

$$L(\Gamma)^2 \geq 4\pi A(\Omega).$$

(ii) Se S_r é uma circunferência de raio r , então

$$L(S_r)^2 = 4\pi^2 r^2 = 4\pi A(D_r),$$

isto é,

$$L(S_r)^2 - 4\pi A(D_r) = 0.$$

Portanto, a igualdade na desigualdade isoperimétrica ocorre quando Γ é uma circunferência de raio arbitrário.

Suponhamos agora que Γ satisfaz

$$L(\Gamma)^2 - 4\pi A(\Omega) = 0.$$

Dada uma função diferenciável $f : [0, l] \rightarrow \mathbb{R}$ tal que $f(0) = f(l)$, considere a variação $\Gamma_t(f)$ da curva $\alpha : [0, l] \rightarrow \mathbb{R}^2$ (parametrização de Γ) correspondente à função f , onde $|t| < \delta$ para algum $\delta > 0$. Então, a função $h : (-\delta, \delta) \rightarrow \mathbb{R}$, dada por

$$h(t) = L(\Gamma_t(f))^2 - 4\pi A(\Omega_t(f)),$$

é diferenciável pelos Lemas 9 e 10 e atinge o seu valor mínimo no ponto $t = 0$ pois $h(t) \geq 0$ pela primeira parte da demonstração e $h(t) = 0$ por hipótese. Assim

$$0 = h'(0) = 2L(\Gamma) \left. \frac{d}{dt} \right|_{t=0} L(\Gamma_t(f)) - 4\pi \left. \frac{d}{dt} \right|_{t=0} A(\Omega_t(f)).$$

Usando as fórmulas de variação obtidas nos Lemas 9 e 10, podemos escrever

$$\int_0^l (2L(\Gamma)\kappa(s) - 4\pi)f(s)ds = 0$$

para toda função diferenciável $f : [0, l] \rightarrow \mathbb{R}$ satisfazendo $f(0) = f(l)$. Escolhendo

$$f(s) = 2L(\Gamma)\kappa(s) - 4\pi,$$

onde $\kappa : [0, l] \rightarrow \mathbb{R}$ é a função curvatura de α , concluímos que

$$2L(\Gamma)\kappa(s) - 4\pi = 0 \quad \text{para todo } s \in [0, l],$$

ou seja, a função curvatura $\kappa : [0, l] \rightarrow \mathbb{R}$ é tal que $\kappa \equiv 2\pi/L(\Gamma)$. Segue da proposição 7 que Γ é uma circunferência.

□

5 Sobre o caso tridimensional da desigualdade isoperimétrica

Conforme foi citado na introdução deste trabalho, os resultados aqui apresentados podem ser estendidos para superfícies em \mathbb{R}^3 . Neste caso, a desigualdade isoperimétrica é dada pelo

Teorema 12 (Desigualdade isoperimétrica espacial). *Seja $S \subset \mathbb{R}^3$ uma superfície compacta e conexa de área $A(S)$ e delimitando uma região Ω . Se $V(\Omega)$ é o volume de Ω , então*

$$A(S)^3 - 36\pi V(\Omega)^2 \geq 0.$$

Além disso, a igualdade ocorre se, e somente se, S é uma esfera.

A demonstração completa do Teorema 12 pode ser encontrada em MONTIEL e ROS, e segue a mesma linha apresentada neste trabalho com as devidas adaptações. Neste contexto, a desigualdade de Brunn-Minkowski é aplicada no caso $n = 3$ e as fórmulas de variação são estendidas para a variação da área da superfície S e para a variação do volume da região delimitada por S . Aqui, a curvatura média da superfície desempenha o papel da curvatura no caso de curvas planas.

Seguindo o mesmo caminho do caso bidimensional, é possível mostrar que as superfícies que verificam a igualdade no Teorema 12 são aquelas que têm curvatura média constante. Neste ponto, podemos aplicar o Teorema de Alexandrov que afirma que, se uma superfície compacta e conexa tem curvatura média constante, então esta superfície tem que ser a esfera.

Referências

- ALENCAR, H.; SANTOS, W.; NETO, G. Geometria diferencial das curvas no \mathbb{R}^2 . 1. ed. Rio de Janeiro: SBM, 2020. ISBN 978-65-990395-4-6.
- ANTON, H.; RORRES, C. Álgebra linear com aplicações. 8. ed. Porto Alegre: Bookman, 2001.
- BLÅSJÖ, V. The Isoperimetric Problem. The American Mathematical Monthly, v. 112, ed. 6, p. 526-566, 2005.
- MONTIEL, S.; ROS, A. Curves and surfaces. 2nd. ed. American Mathematical Soc., 2009. v. 69.
- STEIN, E. M.; SHAKARCHI, R. Real analysis: Measure Theory, Integration, and Hilbert Spaces. New Jersey: Princeton University Press, 2005. v. 3.
- TENENBLAT, K. Introdução à geometria diferencial. Editora Blucher, 2008.
- VIRGÍLIO, P. Eneida: Tradução de Manoel Odorico Mendes. Clássicos Jackson, 1854. v. 3.

Métodos do Gradiente e de Newton em Otimização Quadrática Convexa

Alex Augusto Nunes Machado
Simone Aparecida Miloca - Unioeste

(Recebido em 29/11/2021. Aceito em 03/12/2021. Publicado em 22/12/2021)

Resumo: Este trabalho tem por objetivos apresentar uma breve contextualização e formalização de conceitos envolvendo problemas pertencentes a uma área da Matemática denominada Otimização. O interesse está em uma classe particular nomeada Programação Quadrática Irrestrita o qual aparece em diversas situações práticas tanto na fase de modelagem como em processos de busca de soluções em outros métodos de otimização. Para tanto, descrevemos resultados teóricos envolvendo condições de otimalidade para problemas irrestritos bem como apresentamos dois algoritmos de busca direcionais, os métodos do Gradiente e de Newton, para então justificar matematicamente os motivos pelos quais o método de Newton resolve um problema dessa classe em apenas uma iteração, bem como demonstrar a fórmula utilizada para determinar o tamanho do passo no Método do Gradiente.

Palavras-chave: Programação Quadrática; Método de Newton; Método do Gradiente.

1 Introdução

Conforme apontam Ribeiro e Karas (2013), direta ou indiretamente a noção de otimização faz parte do dia-a-dia do ser humano moderno, dado que as pessoas responsáveis pela tomada de decisão (em qualquer campo do conhecimento) deparam-se com problemas nos quais necessitam obter resultados não somente precisos mas também eficientes.

Nesse sentido, conforme pontuam Hillier e Lieberman (2013, p. 1), desde o desenvolvimento da sociedade industrializada, e sobretudo com a eclosão da segunda guerra mundial na qual os militares britânicos e norte-americanos visando alocar os recursos da melhor maneira possível “(...) convocaram grande número de cientistas para aplicar uma abordagem científica para lidar com este e outros problemas táticos e estratégicos.”, a demanda por ferramentas que auxiliem nesse processo decisório por parte de organizações, públicas ou privadas, nos mais diversos campos, sejam no âmbito científico ou empresarial, cresceu substancialmente. É neste contexto que a Otimização Matemática se aplica.

Isto posto, o que vem a ser esse campo de estudo denominado *Otimização Matemática*? Luenberger e Ye (2016) em sua introdução escrevem que,

“O conceito de otimização agora está bem enraizado como um princípio subjacente à análise de muitos problemas complexos de decisão ou alocação. [...] aborda-se um problema de decisão complexo, envolvendo a seleção de valores para uma série de variáveis inter-relacionadas, focando a atenção em um único objetivo projetado para quantificar o desempenho e medir a qualidade da decisão. Este objetivo é maximizado (ou minimizado, dependendo da formulação) sujeito às restrições que podem limitar a seleção dos valores das variáveis de decisão.” (LUENBERGER; YE, 2016, p. 1, tradução do autor)

Já, Ribeiro e Karas (2013) apontam que a Otimização consiste em encontrar pontos de mínimo ou máximo de uma função $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ que está sujeita a determinadas restrições (outras funções). Tal função f é denominada *função objetivo* e emerge, quando pessoas responsáveis pela tomada de decisão deparam-se com problemas nos quais necessitam obter resultados precisos e, além disso,

“(...) um único aspecto adequado de um problema pode ser isolado e caracterizado por um objetivo, seja lucro ou perda em um ambiente de negócios, velocidade ou distância em um problema físico, retorno esperado no ambiente de investimentos de risco (...)” (LUENBERGER; YE, 2016, p. 1, tradução do autor).

Com isso, por meio de um processo de modelagem é possível traduzir a situação que demanda ser resolvida para o “mundo matemático”. Essa tradução, de modo geral, é representada por

$$\begin{array}{ll} \text{Otimizar} & f(x) \\ \text{sujeito a} & x \in \Omega \subset \mathbb{R}^n \end{array} \quad (1)$$

no qual Ω representa um conjunto de restrições que podem ser de igualdade ou desigualdade. Portanto, (1) formaliza em termos matemáticos o que vem a ser um problema de Otimização Matemática.

Assim, percebe-se então que o modelo consiste em uma espécie de representação do “mundo real” no “mundo matemático”. Esse processo de tradução como indicam Goldbarg e Luna (2005, p. 9) envolve “a percepção do elaborador do modelo (ou equipe de elaboração), uma faculdade cognitiva de alto nível. As *fórmulas* ou *equações* do modelo não existem prontas e acabadas na natureza, elas têm que ser identificadas ou criadas.”. Todavia, embora tal processo vise a extração da essência do problema, conforme salientam Hillier e Lieberman (2013), vale destacar que “a correspondência entre experiência e modelo formal está longe de ser perfeita: a tradução está sujeita a erros, simplificações e falhas de comunicação.” (MARTÍNEZ; SANTOS, 1995, p. 1).

Além disso, deve-se observar que o conjunto $\Omega \subset \mathbb{R}^n$ (e a função objetivo) pode assumir diversos formatos, uma vez que sua obtenção está diretamente ligada ao processo de formulação do modelo matemático. Nesse sentido conforme apontam Ribeiro e Karas (2013) e Goldbarg e Luna (2005) é possível agrupar os problemas de otimização em algumas categorias, cada qual possuindo técnicas específicas para sua resolução.

Algumas das categorias possíveis são: programação convexa, otimização linearmente restrita, programação separável, programação não convexa, programação geométrica, pro-

gramação fracionária e programação quadrática, sendo que esta última será o tema deste trabalho.

Mais especificamente, serão apresentados alguns conceitos e resultados básicos do campo de Otimização Matemática, bem como comentado brevemente a respeito dos algoritmos direcionais, em particular os Método do Gradiente e de Newton. Na sequência, as seções 4, 5 e 6 serão responsáveis por formalizar o problema de Programação Quadrática Irrestrita e provar matematicamente os motivos pelos quais o método de Newton necessita de apenas uma iteração para resolver problemas dessa classe além de demonstrar a fórmula utilizada para estabelecer o tamanho do passo no Método do Gradiente.

2 Condições de otimalidade

Visando estabelecer os critérios para verificar se algum determinado ponto $x \in \Omega$ é o que fornece o melhor valor para uma função f dada, isto é, formalizar matematicamente o que foi discutido na Introdução, esta seção apresenta algumas definições e resultados necessários para o avanço à Programação Quadrática Irrestrita.

Definição 1. Sejam $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x^* \in \Omega \subset \mathbb{R}^n$. Definimos x^* como sendo um *minimizador local de f em Ω* se, e só se, existe $\delta > 0$, tal que

$$f(x^*) \leq f(x),$$

para todo $x \in B(x^*, \delta) \cap \Omega$, com $B(x^*, \delta)$ sendo uma bola de centro x^* e raio δ .

Definição 2. Sejam $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e $x^* \in \Omega \subset \mathbb{R}^n$. Definimos x^* como sendo um *minimizador global de f em Ω* se, e só se,

$$f(x^*) \leq f(x),$$

para todo $x \in \Omega$.

Isto posto, uma pergunta natural é: “toda função $f: A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ possui um valor x^* que a otimiza?” A Proposição 3, conhecida como “*Teorema de Weierstrass*”, apresenta a resposta à esse questionamento, e sua demonstração pode ser encontrada em Ribeiro e Karas (2013, p. 36).

Proposição 3. *Se $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ é uma função contínua e Ω é um conjunto compacto não vazio, então existe minimizador global de f em Ω .*

Ou seja, sob certas circunstâncias, é possível garantir a existência de um minimizador global. Em outras palavras, isto significa que se dado um problema, seja ele de *qualquer natureza*, puder ser modelado matematicamente de tal modo que seu domínio seja um subconjunto compacto (limitado e fechado) de \mathbb{R}^n , então pode-se garantir a existência de um valor que minimiza f , independente da natureza do problema que o gerou.

Ademais, vale destacar que embora o Teorema de Weierstrass faça referência a minimização de uma função, tal proposição também pode ser aplicada no contexto de maximização de f , dado que maximizar f é o mesmo que minimizar $-f$. De fato, suponha que x^* seja maximizador global de $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, assim, por definição temos que para todo $x \in \Omega$,

$$f(x^*) \geq f(x)$$

logo, para todo $x \in \Omega$,

$$-f(x^*) \leq -f(x).$$

Portanto, segue por definição que x^* é minimizador global de $-f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$. A Figura 1 ilustra essas ideias.

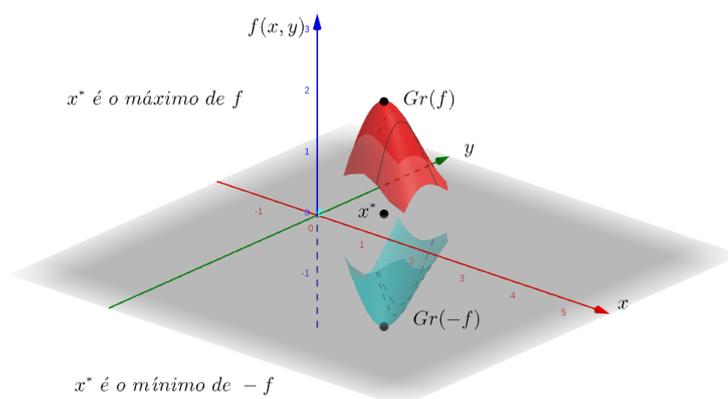


Figura 1: Minimizar f é equivalente a maximizar $-f$

Diante disso, a proposição 4 estabelece os critérios que permitem afirmar se algum x pertencente ao domínio de f pode ser um *candidato* a minimizador (ou maximizador) local.

Proposição 4. *Seja $f: \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável em $x^* \in \mathbb{R}^n$. Se x^* é minimizador local de f , então*

$$\nabla f(x^*) = 0.$$

Prova. Suponha $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável em $x^* \in \Omega$ e x^* minimizador local de f em Ω . Além disso, suponha $d \in \mathbb{R}^n \setminus \{0\}$.

Com isso, uma vez que x^* é minimizador local, segue por definição, que existe $\delta > 0$ tal que para todo $x \in B(x^*, \delta) \cap \Omega$,

$$f(x^*) \leq f(x),$$

e portanto, em particular, para $x = x^* + td$ com $t \in (0, \delta)$, isto é, para todo $t \in (0, \delta)$,

$$f(x^*) \leq f(x^* + td).$$

Assim,

$$0 \leq f(x^* + td) - f(x^*).$$

Além disso, fazendo uso da fórmula de Taylor de ordem 1, juntamente com a mudança de variável $td = x - x^*$, temos que

$$f(x^* + td) = f(x^*) + t\nabla f(x^*)^T d + r(t),$$

para $\lim_{t \rightarrow 0} \frac{r(t)}{t} = 0$.

Logo,

$$\begin{aligned} 0 \leq f(x^* + td) - f(x^*) &\implies 0 \leq f(x^*) + t\nabla f(x^*)^T d + r(t) - f(x^*) \\ &\implies 0 \leq t\nabla f(x^*)^T d + r(t) \implies \frac{0}{t} \leq \frac{t\nabla f(x^*)^T d}{t} + \frac{r(t)}{t} \\ &\implies 0 \leq \nabla f(x^*)^T d + \frac{r(t)}{t} \implies \lim_{t \rightarrow 0} 0 \leq \lim_{t \rightarrow 0} \left(\nabla f(x^*)^T d + \frac{r(t)}{t} \right) \\ &\implies 0 \leq \nabla f(x^*)^T d + \lim_{t \rightarrow 0} \frac{r(t)}{t}, \end{aligned}$$

e como $\lim_{t \rightarrow 0} \frac{r(t)}{t} = 0$, temos que

$$\nabla f(x^*)^T d \geq 0. \tag{D.6}$$

Diante disso, suponha por absurdo que $\nabla f(x^*) \neq 0$, logo tomando $d = -\nabla f(x^*)$ em (D.6), temos que

$$\nabla f(x^*)^T (-\nabla f(x^*)) \geq 0 \implies -\|\nabla f(x^*)\|^2 \geq 0 \implies \|\nabla f(x^*)\|^2 \leq 0,$$

o que é absurdo, portanto, não se pode supor $\nabla f(x^*) \neq 0$, logo $\nabla f(x^*) = 0$.

□

Posta em outras palavras, essa proposição informa que dado *qualquer* x no domínio de f se $\nabla f(x)$ não for nulo, então x não pode ser sequer um candidato a minimizador local. Sendo assim, elementos de \mathbb{R}^n que satisfazem essa proposição (também conhecida como *condição necessária de 1ª ordem*) recebem um nome especial, o qual será formalizado com a definição seguinte.

Definição 5. Sejam $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ diferenciável em Ω e $x^* \in \Omega$. Definimos x^* como sendo **ponto crítico** ou (*estacionário*) de f se, e só se, $\nabla f(x^*) = 0$.

Isto posto, convém enfatizar que as funções (objetivo e/ou restrições) envolvidas podem não ser tão elementares, ou mesmo que sejam, podem demandar muitas variáveis, e assim, a solução direta pode não ser uma boa opção.

Diante disso, entram em cena os *métodos de busca* que funcionam como um meio de, a partir de um palpite inicial $x_0 \in \mathbb{R}^n$ para a solução do problema, encontrar outros pontos, mais precisamente uma sequência de pontos de \mathbb{R}^n que, idealmente, conforme destacam Luenberger e Ye (2016) e Ribeiro e Karas (2013), geram uma sequência de pontos de \mathbb{R}^n que se aproximam da solução ótima. Diante disso, a próxima seção será destinada a apresentar alguns desses métodos de busca.

3 Algoritmos com buscas direcionais

Imaginemos a seguinte situação, encontrar o mínimo de uma dada função f que possui o seu gráfico ilustrado na Figura 2a.

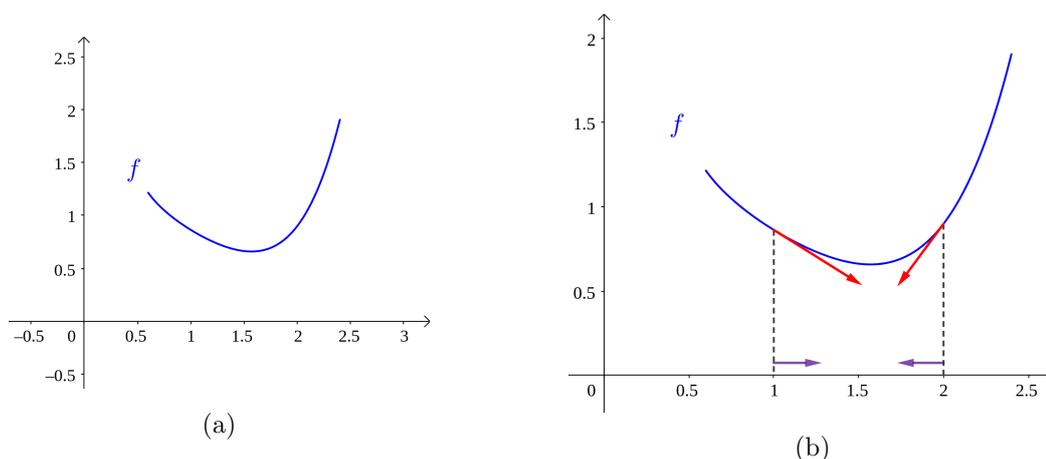


Figura 2: Minimização por bissecção

Intuitivamente podemos inferir que se escolhermos $x^* = 1$ então esse valor fará $\frac{d}{dx}f(1) < 0$ pois, visualmente, a função ainda pode assumir valores menores que $f(1)$ caso tomemos valores à direita de 1 para x^* . Assim, $f(1 + \varepsilon) < f(1)$, fazendo com que $f(1 + \varepsilon) - f(1) < 0$. Logo, $x^* = 1$ não poderá ser ponto crítico de f .

Porém, se tomarmos agora $x^* = 2$ sua derivada será positiva, pois novamente a intuição visual informa que f ainda pode assumir valores menores que $f(2)$ na condição de escolhermos algum valor à esquerda de 2. Deste modo, agora $f(2 + \varepsilon) > f(2)$, e assim $f(2 + \varepsilon) - f(2) > 0$, fazendo que $\frac{d}{dx}f(2) > 0$.

Diante disso, é razoável supor que embora ainda não saibamos qual é o valor que minimiza f , o intervalo $(1, 2)$ certamente conterá um valor que minimiza f , ao menos localmente. A Figura 2b ilustra essas ideias.

Ou seja, $\frac{d}{dx}f(x_0) < 0$ implica que o valor que minimiza f encontra-se à direita de x_0 , em contrapartida, se $\frac{d}{dx}f(x_0) > 0$ então x^* deverá estar à esquerda de x_0 .¹ Isto posto, uma estratégia para explorar essa característica consiste em

- 1° Tomar um intervalo $[a, b]$ de modo que $\frac{d}{dx}f(a) \cdot \frac{d}{dx}f(b) < 0$;
- 2° Calcular x_0 como sendo $\frac{b+a}{2}$;
- 3° Tomar o novo intervalo (a, b) de tal modo que ou $a = x_0$ se $\frac{d}{dx}f(x_0) < 0$ ou $b = x_0$ se $\frac{d}{dx}f(x_0) > 0$;

¹Porém, deve-se observar que isso é sempre verdade para o caso de a função a ser minimizada ser convexa. Para mais informações, consultar Hillier e Lieberman (2013, p. 953).

4° Calcular x_1 como sendo $\frac{b+a}{2}$, caso x_1 seja a solução ótima ou seja suficientemente bom pare, se não execute uma nova iteração.

Portanto, percebamos que tal estratégia consiste em ir reduzido o intervalo de busca pela metade até encontrar a solução desejada, ou ao menos obter uma solução com a precisão tão boa quanto se deseja. Essa abordagem é conhecida como método da bissecção e um exemplo de implementação é ilustrado pelo Algoritmo 1.

Tomar um intervalo $[a, b]$;

Escolher o palpite inicial como sendo $x_0 = \frac{b+a}{2}$;

$k \leftarrow 0$;

enquanto $\nabla f(x_k) \neq 0$ **faça**

se $\frac{d}{dx} f(x_k) < 0$ **então**

$a \leftarrow x_k$;

senão

$b \leftarrow x_k$;

fim

$x_k = \frac{b+a}{2}$;

fim

Algoritmo 1: Exemplo genérico de implementação do método da bissecção

Notemos que o método da bissecção de fato faz uso da ideia de, a partir de um valor inicial, obter outros de modo que $f(x_i) \leq f(x_{i-1})$. Todavia, como estender essas ideias ao \mathbb{R}^n ? Visando responder a esse questionamento temos a seguinte definição.

Definição 6. Sejam $f: \mathbb{R}^n \rightarrow \mathbb{R}$, $\bar{x} \in \mathbb{R}^n$ e $d \in \mathbb{R}^n \setminus \{0\}$ uma direção. Definimos d como sendo uma **direção de descida para f a partir de \bar{x}** se, e só se, existe $\delta > 0$ tal que $f(\bar{x}+td) < f(\bar{x})$ para todo $t \in (0, \delta)$.

Isto posto, embora a Definição 6 estabeleça como identificar *se dada direção d é ou não de descida*, a proposição seguinte vem para ajudar nesse processo de verificação.

Proposição 7. *Seja $f: \mathbb{R}^n \rightarrow \mathbb{R}$. Se $\nabla f(\bar{x})^T d < 0$, então d é uma direção de descida para f a partir de \bar{x} .*

Prova. Suponha $\nabla f(\bar{x})^T d < 0$, assim da definição de derivada direcional temos que

$$\frac{\partial}{\partial d} f(\bar{x}) < 0,$$

e como existe a derivada direcional na direção d , então f restrita à direção d é contínua neste ponto \bar{x} , e portanto, para $t \neq 0$,

$$\lim_{t \rightarrow 0} \frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0.$$

Logo, da definição de limite, existe $\delta > 0$ tal que

$$\frac{f(\bar{x} + td) - f(\bar{x})}{t} < 0,$$

para todo $t \in (-\delta, \delta)$. Assim, $f(\bar{x} + td) < f(\bar{x})$ para o caso particular em que $t \in (0, \delta)$.

□

Pondo em outras palavras a Proposição 7 está informando que caso a derivada na direção de d seja negativa, então existirão nessa mesma direção vetores de \mathbb{R}^n de modo que o valor de f aplicado em tais vetores sejam menores que $f(\bar{x})$.

Sendo assim, de posse dessas duas ferramentas o Algoritmo 2 estabelece uma estratégia genérica para minimizar funções de \mathbb{R}^n .

Escolher x_0 ;

$k \leftarrow 0$;

enquanto $\nabla f(x_k) \neq 0$ **faça**

Obter d_k tal que $\nabla f(x_k)^T d_k < 0$;

Obter $t_k > 0$ tal que $f(x_k + t_k d_k) < f(x_k)$;

$x_{k+1} \leftarrow x_k + t_k d_k$;

$k \leftarrow k + 1$;

fim

Algoritmo 2: Algoritmo de descida

Colocando em outros termos, o que o Algoritmo 2 faz em sua essência é, partindo de um ponto inicial, encontrar uma direção na qual a imagem de f diminui, isto é, obter uma direção de descida, para então encontrar o novo ponto inicial e recomeçar o processo levando em consideração essa atualização.

Na prática, a condição $\nabla f(x_k) = 0$ é substituída por algum critério de parada², seja pelo fato de que encontrar x_k que atenda a especificação pode requerer um alto custo computacional, ou simplesmente pela solução desejada não necessitar ser *a melhor*, mas sim uma *suficientemente boa*.

Além disso, conforme colocam Ribeiro e Karas (2013), as escolhas de d_k e t_k não podem ser arbitrárias pois, se assim o fossem, poderiam gerar uma sequência de pontos que, embora possuam pontos de acumulação estes podem não ser estacionários, isto é, podem não ser bons candidatos a minimizadores.

Diante disso, uma pergunta natural que emerge é: “*Como escolher d_k e t_k de modo conveniente ao propósito de minimizar f mais rapidamente?*”, a qual pode ser parafraseada em: “*Como escolher as direções e as amplitudes de cada passo em direção ao mínimo de modo eficiente?*”.

²Por exemplo, substituir $\nabla f(x_k) = 0$ por $\nabla f(x_k) < \varepsilon$ em que ε é escolhido de modo que x_k é suficientemente próximo da melhor solução possível.

Pois bem, no que se refere a escolha da direção em problemas de minimização, é natural a escolha de $-\nabla f(x_k)$ pois, da teoria do cálculo em \mathbb{R}^n , sabe-se que o vetor gradiente aponta no sentido de maior crescimento de f . Sendo assim, visando ilustrar essas ideias, consideremos a seguinte situação:

$$\text{minimizar } f(x) = -2x_1x_2 - 2x_2 + x_1^2 + 2x_2^2.$$

Para tanto, deve-se primeiramente estabelecer um critério de parada, nesse caso $\|\nabla f(x_k)\| < \varepsilon = 0,1$ e também um palpite inicial para a solução, nesse caso $x_0 = (0,0)$. Assim, dado que

$$\frac{d}{dx_1}f = -2x_2 + 2x_1 \text{ e } \frac{d}{dx_2}f = -2x_1 + 4x_2 - 2,$$

temos que

$$\|\nabla f(x_0)\| = \|\nabla f(0,0)\| = \|(-0 + 0, -0 + 0 - 2)\| = \|(0, -2)\| = \sqrt{0^2 + (-2)^2} = 2 > \varepsilon.$$

Portanto, iniciamos a **primeira iteração** com

$$-\nabla f(x_0) = -\nabla f(0,0) = -(0, -2) = (0, 2),$$

e assim, da definição de direção de descida (Definição 6), temos que

$$\begin{aligned} f(x_0 + t_0(-\nabla f(x_0))) &< f(x_0) \\ \iff f((0,0) + t_0(0,2)) &< f(0,0) \\ \iff f(0,2t_0) = -2 \cdot 0 \cdot (2t_0) - 2 \cdot (2t_0) + 0^2 + 2(2t_0)^2 \\ &< -2 \cdot 0 \cdot 0 - 2 \cdot 0 + 0^2 + 2 \cdot 0^2 \\ \iff -4t_0 + 8t_0^2 &< 0. \end{aligned}$$

Ou seja, qualquer t_0 tomado de modo que $-4t_0 + 8t_0^2 < 0$ já seria suficiente para “descermos” no sentido do minimizador. Todavia, é razoável optar pelo t_0 de tal modo que essa descida seja a mais rápida possível, sendo assim recaímos em outro problema de minimização, desta vez um de apenas uma variável.

Neste caso, fazendo uso da Proposição 4 para o caso particular de funções de uma variável, temos que

$$\frac{d}{dt_0}(-4t_0 + 8t_0^2) = 0 \iff -4 + 16t_0 = 0 \iff t_0 = \frac{1}{4}.$$

Portanto, o tamanho do passo no sentido de $-\nabla f(x_0)$ será $t_0 = \frac{1}{4}$, isto é,

$$x_1 = (0,0) + \frac{1}{4}(0,2) = \left(0, \frac{1}{2}\right).$$

Logo, encerra-se a primeira iteração e verifica-se novamente o critério de parada.

$$\begin{aligned} \|\nabla f(x_1)\| &= \left\| \nabla f\left(0, \frac{1}{2}\right) \right\| = \left\| \left(-2 \cdot \frac{1}{2} + 2 \cdot 0, -2 \cdot 0 + 4 \cdot \frac{1}{2} - 2\right) \right\| = \left\| \left(-1, -\frac{3}{2}\right) \right\| \\ &= \sqrt{(-1)^2 + \left(-\frac{3}{2}\right)^2} = \sqrt{\frac{4+9}{4}} = \frac{\sqrt{13}}{2} \simeq 1,803 > \varepsilon. \end{aligned}$$

Com isso, devido ao fato de que x_1 não satisfaz ao critério de parada, repete-se o processo aplicado na iteração 1 de modo a obter x_k que atenda ao critério de parada estabelecido.

Em resumo, esse processo fornece uma ideia de como fazer uso dessa estratégia de minimização envolvendo o vetor gradiente. Devido à sua importância essa estratégia recebe o nome de **Método do Gradiente**, também conhecida como **Método de Cauchy**. O Algoritmo 3 resume essa abordagem.

```
Escolher o critério de parada  $\varepsilon > 0$ ;  
Escolher  $x_0$ ;  
 $k \leftarrow 0$ ;  
enquanto  $\|\nabla f(x_k)\| > \varepsilon$  faça  
    Definir  $d_k = -\nabla f(x_k)$ ;  
    Obter  $t_k > 0$  tal que  $f(x_k + t_k d_k) < f(x_k)$ ;  
     $x_{k+1} \leftarrow x_k + t_k d_k$ ;  
     $k \leftarrow k + 1$ ;  
fim
```

Algoritmo 3: Método do Gradiente

Conforme destacam Ribeiro e Karas (2013, p. 91), este método converge globalmente, e sua demonstração pode ser encontrada na página 91 deste mesmo livro.

Isto posto, outra característica importante que o exemplo anterior ilustra é o fato de que a determinação do tamanho de passo k em sua respectiva iteração não é algo tão elementar, pois isso requer a resolução de outro problema de minimização, em particular minimizar f na direção d_k . E, conforme aponta Friedlander (1994, p.33) esse processo recebe o nome de **busca linear exata**.

Diante disso, dado que obter t_k tamanhos de passo exige a resolução de k problemas de minimização (cada qual associada a sua respectiva direção d_k), quando deseja-se utilizar a abordagem da busca linear exata para a obtenção do tamanho do passo na iteração k , a implementação do Método do Gradiente requer o uso de um segundo algoritmo de minimização. Porém, diferentemente do problema inicial, este “subproblema” consiste na minimização de funções de apenas uma variável, e assim podemos fazer uso de algoritmos já bem estabelecidos como o método da bissecção e o método da seção áurea.

Todavia, nem sempre a aplicação da busca linear exata para cada iteração é conveniente, seja por tal técnica demandar um alto custo computacional, ou mesmo pelo fato de que a minimização do subproblema associado a iteração k simplesmente não é possível. Sendo assim, entra em cena a **busca inexata** que, ao invés de obter o *melhor* tamanho de passo na direção d_k , visa encontrar um t_k *suficientemente* bom, isto é, não almeja minimizar o subproblema associado a iteração k , mas sim “[...] busca uma boa redução da função ao longo da direção, sem tentar minimizá-la.” Ribeiro e Karas (2013, p.77). Algumas das estratégias utilizadas na busca inexata são: Regra de Armijo, Regra de Goldstein, Regra de Wolfe e Regra de Wolfe

forte³.

Todavia, tendo em vista o que o foco deste trabalho, além de ser de caráter introdutório, abarca somente problemas de programação quadrática irrestrita, a estratégia escolhida para a determinação do tamanho de cada passo será a busca exata, pois como será discutido na seção seguinte, encontrar o tamanho do passo t_k nessa classe de problemas é sempre possível.

Voltemos ao problema

$$\begin{array}{ll} \text{Minimizar} & f(x) \\ \text{sujeito a} & x \in \mathbb{R}^n \end{array} \quad (2)$$

Como discutido na Seção 2, para resolvermos (2) devemos encontrar algum x^* de tal modo que $\nabla f(x^*) = 0$. Diante disso, visando alcançar esse objetivo, devemos então escolher não só uma direção de descida (d_k) mas também o tamanho do passo que será dado nessa direção (t_k). Uma das estratégias para isso consiste no Método do Gradiente, que apoia-se na ideia de “seguir” a direção de $-\nabla f(x_k)$ em cada iteração.

Todavia, uma outra abordagem consiste em tomar x_{k+1} como sendo o ponto que minimiza a aproximação de Taylor de segunda ordem⁴ na iteração k . Essa estratégia é conhecida como **Método de Newton**⁵ e é implementada pelo Algoritmo 4.

Escolher o critério de parada $\varepsilon > 0$;

Escolher x_0 ;

$k \leftarrow 0$;

enquanto $\|\nabla f(x_k)\| > \varepsilon$ **faça**

Definir $d_k = -(\nabla^2 f(x_k))^{-1} \nabla f(x_k)$;
 $x_{k+1} \leftarrow x_k + d_k$;
 $k \leftarrow k + 1$;

fim

Algoritmo 4: Método de Newton

Salientamos aqui que o cálculo da direção d_k do Algoritmo 4 é feito geralmente resolvendo um sistema de equações lineares que tem um custo computacional menor que o necessário para a inversão de matrizes (RIBEIRO; KARAS, 2013).

4 Programação quadrática irrestrita

De acordo com Hillier e Lieberman (2013), na classe de problemas de programação quadrática as restrições são lineares, porém sua função objetivo possui lei de formação contendo

³Para mais detalhes, ver Shi e Shen (2005).

⁴Vale destacar que também há a estratégia de utilizar a aproximação de Taylor de primeira ordem. Nesse caso utilizam-se propriedades da Jacobiana para chegar ao mesmo resultado. Ao leitor interessado em mais detalhes recomenda-se a leitura de Ribeiro e Karas (2013, p. 96)

⁵Mais especificamente Método de Newton Puro, pois $t_k = 1$ para todo k . Todavia, para uma discussão teórica mais detalhada desse método sugere-se a leitura de Ribeiro e Karas (2013) e Martins e Pereira (2010).

termos quadráticos ou produto entre duas variáveis. Em particular, consideremos o caso de programação quadrática irrestrita, caracterizada matricialmente por,

$$\begin{aligned} \text{Minimizar} \quad & f(x) = \frac{1}{2}x^T Ax + b^T x + c \\ \text{sujeito a} \quad & x \in \mathbb{R}^n, \end{aligned} \tag{3}$$

com $b \in \mathbb{R}^n$, $c \in \mathbb{R}$ e $A \in M_{n \times n}(\mathbb{R})$ simétrica definida positiva.

Notemos que

$$\begin{aligned} f(x) &= \frac{1}{2}x^T Ax + b^T x + c \\ &= \frac{1}{2} \left(\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \right) + \begin{bmatrix} b_1 & b_2 & \dots & b_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + c \\ &= \frac{1}{2} \left(\begin{bmatrix} x_1 & x_2 & \dots & x_n \end{bmatrix} \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n \end{bmatrix} \right) \\ &\quad + \begin{bmatrix} b_1x_1 & b_2x_2 & \dots & b_nx_n \end{bmatrix} + c \\ &= \frac{1}{2} \left(\begin{bmatrix} a_{11}x_1^2 + a_{12}x_1x_2 + \dots + a_{1n}x_1x_n + a_{21}x_1x_2 + a_{22}x_2^2 + \dots + a_{2n}x_2x_n \\ \vdots \\ a_{n1}x_1x_n + a_{n2}x_2x_n + \dots + a_{nn}x_n^2 \end{bmatrix} \right) \\ &\quad + \begin{bmatrix} b_1x_1 & b_2x_2 & \dots & b_nx_n \end{bmatrix} + c \\ &\equiv \frac{1}{2} (a_{11}x_1^2 + a_{12}x_1x_2 + \dots + a_{1n}x_1x_n, a_{21}x_1x_2 + a_{22}x_2^2 + \dots + a_{2n}x_2x_n \\ &\quad + \dots + a_{n1}x_1x_n + a_{n2}x_2x_n + \dots + a_{nn}x_n^2) \\ &\quad + b_1x_1 + b_2x_2 + \dots + b_nx_n + c. \end{aligned}$$

Deste modo, segue da Proposição 4 que se x^* minimiza f , então

$$\nabla f(x^*) = 0 \implies \begin{bmatrix} \frac{\partial}{\partial x_1} f(x^*) \\ \frac{\partial}{\partial x_2} f(x^*) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x^*) \end{bmatrix} = 0.$$

Com isso, derivando f analiticamente em relação a x_1, x_2, \dots, x_n , obtemos que

$$\begin{aligned} \frac{\partial}{\partial x_1} f &= \frac{1}{2} (2a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + a_{21}x_2 + 0 + \dots + 0 \\ &\quad + \dots + a_{n1}x_n + 0 + \dots + 0) + b_1 \\ &= \frac{1}{2} (2a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n + a_{21}x_2 + \dots + a_{n1}x_n) + b_1 \\ &= \frac{1}{2} (2a_{11}x_1 + (a_{12} + a_{21})x_2 + \dots + (a_{1n} + a_{n1})x_n) + b_1, \end{aligned}$$

e como A é simétrica, segue que

$$\begin{aligned} \frac{\partial}{\partial x_1} f &= \frac{1}{2}(2a_{11}x_1 + 2a_{21}x_2 + \cdots + 2a_{n1}x_n) + b_1 \\ &= \frac{1}{2} \cdot 2(a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n) + b_1 \\ &= a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n + b_1. \end{aligned}$$

Analogamente, mostra-se que $\frac{\partial}{\partial x_2} f = a_{21}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n + b_2, \dots, \frac{\partial}{\partial x_n} f = a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n$. Ou seja,

$$\begin{aligned} \begin{bmatrix} \frac{\partial}{\partial x_1} f(x_1) \\ \frac{\partial}{\partial x_2} f(x_2) \\ \vdots \\ \frac{\partial}{\partial x_n} f(x_n) \end{bmatrix} &= \begin{bmatrix} a_{11}x_1 + a_{21}x_2 + \cdots + a_{n1}x_n + b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{n2}x_n + b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n + b_n \end{bmatrix} = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{21} & a_{22} & \cdots & a_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad (\text{A ser simétrica}) \\ &= Ax + b. \end{aligned}$$

E portanto, em particular para x^* , temos que

$$\nabla f(x^*) = 0 \iff Ax^* + b = 0.$$

Isto é, uma forma de solucionar problemas de programação quadrática irrestrita consiste na resolução de um sistema linear, dado que nessa classe de problemas, obter x^* é teoricamente possível desde que A seja inversível.

Como poderá ser constado na Seção 5, o método de Newton explora essa característica para resolver problemas quadráticos (aqui sem restrições) em apenas uma iteração.

5 Método de Newton aplicado à programação quadrática irrestrita

Consideremos $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tal como definido em (3). Com isso, aplicar o Algoritmo 4 faz com que a direção de descida na iteração zero seja,

$$\begin{aligned} d_0 &= -(\nabla^2 f(x_0))^{-1} \nabla f(x_0) = -(\nabla(\nabla f(x_0)))^{-1} \nabla f(x_0) \\ &= -(\nabla(Ax_0 + b))^{-1} (Ax_0 + b) = -A^{-1}(Ax_0 + b) = -x_0 - A^{-1}b. \end{aligned}$$

E com isso,

$$x_1 = x_0 + (-x_0 - A^{-1}b) = -A^{-1}b.$$

Ou seja, a abordagem de Newton aplicada a problemas de programação quadrática irrestrita minimiza f em apenas uma iteração. Assim, reescrevendo o Algoritmo 4 obtemos:

Escolher o critério de parada ε ;

Escolher x_0 ;

$k \leftarrow 0$;

enquanto $\|\nabla f(x_k)\| > \varepsilon$ **faça**

 Definir $d_k = -x_k - A^{-1}b$;
 $x_{k+1} \leftarrow x_k + d_k$;
 $k \leftarrow k + 1$;

fim

Algoritmo 5: Método de Newton aplicado a programação quadrática irrestrita

6 Método do Gradiente aplicado à programação quadrática irrestrita

Novamente, consideremos $f: \mathbb{R}^n \rightarrow \mathbb{R}$ tal que $f(x) = \frac{1}{2}x^T Ax + b^T x + c$. Aplicando o 3 temos que a direção associada a iteração k será $d_k = -\nabla f(x_k)$, enquanto que, visando obter t_k de tal modo que a descida em direção ao minimizador seja a mais rápida possível devemos resolver o seguinte subproblema de otimização (a cada iteração):

$$\begin{aligned} \text{minimizar} \quad & g(t_k) = f(x_k + t_k d_k) \\ \text{sujeito a} \quad & t_k > 0. \end{aligned}$$

Diante disso, aplicando a proposição 4, temos que

$$g'(t_k) = 0.$$

Logo, obtemos que

$$\begin{aligned} \nabla f(x_k + t_k d_k)^T d_k = 0 &\iff \nabla f(x_{k+1})^T d_k = 0 \iff (Ax_{k+1} + b)^T d_k = 0 \\ &\iff (A(x_k + t_k d_k) + b)^T d_k = 0 \iff (Ax_k + t_k Ad_k + b)^T d_k = 0 \\ &\iff (Ax_k + b + t_k Ad_k)^T d_k = 0 \iff (\nabla f(x_k) + t_k Ad_k)^T d_k = 0 \\ &\iff (\nabla f(x_k)^T + t_k (Ad_k)^T) d_k = 0 \iff \nabla f(x_k)^T d_k + t_k (Ad_k)^T d_k = 0 \\ &\iff t_k (Ad_k)^T d_k = -\nabla f(x_k)^T d_k \iff t_k = -\frac{\nabla f(x_k)^T d_k}{(Ad_k)^T d_k}. \end{aligned}$$

Assim, das propriedades de transposto e do fato de que A é simétrica, segue que o tamanho do passo relativo a iteração k é dado por

$$t_k = -\frac{\nabla f(x_k)^T d_k}{d_k^T Ad_k} = \frac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T A \nabla f(x_k)}.$$

Ou seja, o Método do Gradiente aplicado a problemas de programação quadráticos é dado por:

Escolher o critério de parada ε ;

Escolher x_0 ;

$k \leftarrow 0$;

enquanto $\|\nabla f(x_k)\| > \varepsilon$ **faça**

Definir $d_k = -\nabla f(x_k)$;

Definir $t_k = \frac{\nabla f(x_k)^T \nabla f(x_k)}{\nabla f(x_k)^T A \nabla f(x_k)}$;

$x_{k+1} \leftarrow x_k + t_k d_k$;

$k \leftarrow k + 1$;

fim

Algoritmo 6: Método do Gradiente aplicado a programação quadrática irrestrita

7 Considerações finais

Este trabalho buscou apresentar uma breve contextualização e alguns conceitos envolvendo o campo da Otimização Matemática, além de definir uma classe particular de problemas da Otimização chamada de Programação Quadrática Irrestrita. Em especial, buscou-se mostrar que nessa classe o método de Newton resolve um problema em apenas uma iteração bem como provou-se a fórmula utilizada para determinar o tamanho do passo no Método do Gradiente.

Referências

FRIEDLANDER, A. *Elementos de programação não linear*. Campinas: Pontes, 1994.

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

HILLIER, F. S.; LIEBERMAN, G. J. *Introdução a Pesquisa Operacional*. Tradução Ariovaldo Griesi. Revisão técnica Pierre J. Ehrlich. 9. ed. Porto Alegre: AMGH, 2013.

LUENBERGER, D. G.; YE, Y. *Linear and Nonlinear Programming*. New York: Springer International Publishing Switzerland, 2016.

MARTÍNEZ, J. M.; SANTOS, S. A. Métodos computacionais de otimização. In: COLÓQUIO BRASILEIRO DE MATEMÁTICA, 20., 1995, Rio de Janeiro. *Anais...* Rio de Janeiro: IMPA, 1995. Disponível em: https://impa.br/wp-content/uploads/2017/04/20_CBM_95_03.pdf. Acesso em: 25 jan. 2021.

MARTINS, M. H. S.; PEREIRA, R. *Cálculo III e IV*. Florianópolis: UFSC/EAD/CED/CFM, 2010.

RIBEIRO, A. A.; KARAS, E. W. *Otimização contínua: aspectos teóricos e computacionais*. São Paulo: Cengage Learning, 2013.

SHI, Z.-J.; SHEN, J. New inexact line search method for unconstrained optimization. *Journal*

of optimization theory and applications, Springer, v. 127, n. 2, p. 425–446, 2005.

Solubilidade por meio de radicais de equações polinomiais de grau menor ou igual a quatro

*Elisangela Danielli de Lima - Universidade Tecnológica Federal do Paraná (UTFPR) - Medianeira
Sandro Marcos Guzzo - Universidade Estadual do Oeste do Paraná (UNIOESTE) - Cascavel*

(Recebido em 09/12/2021. Aceito em 13/12/2021. Publicado em 22/12/2021)

Resumo: Nosso objetivo neste trabalho é apresentar técnicas clássicas de resolução de equações polinomiais de grau menor ou igual a 4 por meio de radicais. Fazemos também um breve apanhado histórico sobre a busca de tais fórmulas. Em particular estamos interessados em comprovar que toda equação polinomial de grau menor ou igual a 4 é solúvel por meio de radicais, mesmo que as soluções (raízes) sejam números complexos.

Palavras-chave: Equações polinomiais; Solubilidade por radicais; Fórmula de Cardano.

1 Introdução

Nesta seção faremos um apanhado de definições e conceitos relevantes para este texto. Acreditamos que o leitor esteja familiarizado com a maioria destas definições e resultados. Mesmo assim vamos enunciar estes resultados por motivos de referência textual.

Denomina-se polinômio na variável x , qualquer expressão escrita na forma

$$a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_1 x + a_0, \quad (1)$$

sendo que a_0, a_1, \dots, a_n são coeficientes constantes (reais ou complexos) e $n \in \mathbb{N}$.

É comum designar um polinômio de forma mais econômica como uma função $p(x)$. Ao afirmar que $p(x)$ é um polinômio então deve ficar claro que

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \cdots + a_1 x + a_0.$$

Um polinômio $p(x)$ é chamado de polinômio nulo, se todos os coeficientes da expressão (1) forem iguais a zero. Neste caso escrevemos $p(x) = 0$. Dizer então que $p(x)$ é não nulo, e escrevemos $p(x) \neq 0$, significa que pelo menos um dos coeficientes da expressão (1) é não nulo.

Definição 1. Dado um polinômio $p(x)$ não nulo, então o grau de $p(x)$ é o maior número natural n de forma que $a_n \neq 0$. Representamos o grau de $p(x)$ por $gr(p)$.

Note que não está definido o grau do polinômio nulo. Mas o grau de um polinômio pode ser igual a 0. De acordo com a definição anterior, se $gr(p) = 0$ então $a_0 \neq 0$ e com isso $p(x) = a_0$ é obrigatoriamente não nulo.

Definição 2. Denomina-se equação polinomial na incógnita x a toda equação da forma

$$a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x + a_0 = 0, \quad (2)$$

sendo a_0, a_1, \dots, a_n são coeficientes constantes (reais ou complexos), com $a_n \neq 0$ e $n \in \mathbb{N}$. O número n é chamado de grau da equação polinomial.

Note que uma equação polinomial pode ser escrita na forma simplificada por $p(x) = 0$, sendo que $p(x)$ representa o polinômio do lado esquerdo da equação. Neste caso, o grau da equação polinomial é também o grau do polinômio $p(x)$.

Definição 3. Dado um número α , real ou complexo, o valor do polinômio $p(x)$ em α é o número denotado por $p(\alpha)$ dado por

$$p(\alpha) = a_n \alpha^n + a_{n-1} \alpha^{n-1} + a_{n-2} \alpha^{n-2} + \dots + a_1 \alpha + a_0,$$

isto é, $p(\alpha)$ é o número obtido ao substituir x por α . Se além disso $p(\alpha) = 0$, então α é dito uma raiz de $p(x)$.

Definição 4. Um número α , real ou complexo, é chamado de raiz ou de solução da equação polinomial,

$$a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x + a_0 = 0,$$

se

$$a_n \alpha^n + a_{n-1} \alpha^{n-1} + a_{n-2} \alpha^{n-2} + \dots + a_1 \alpha + a_0 = 0,$$

isto é, ao substituir a incógnita x por α , a igualdade torna-se verdadeira.

Note que se $p(x)$ é um polinômio, então uma raiz α de $p(x)$ é também uma raiz ou uma solução da equação $p(x) = 0$.

Embora o conceito de raiz de uma equação polinomial seja relativamente simples, em geral não é uma tarefa muito simples encontrar tais raízes (ou soluções). Algumas vezes as raízes são números irracionais ou mesmo números complexos. Resolver uma equação polinomial por meio de radicais significa encontrar as raízes da equação manipulando os coeficientes da equação. Desta forma, uma equação polinomial é solúvel por radicais quando suas soluções podem ser escritas como expressões que incluem apenas somas, subtrações, multiplicações, divisões, potências e raízes envolvendo seus coeficientes.

Os alunos do ensino fundamental já estão familiarizados com a obtenção de soluções de equações polinomiais de grau 1 ou 2. No caso de uma equação de grau 1, ou de primeiro grau,

$$ax + b = 0,$$

com $a \neq 0$, a solução $x = -\frac{b}{a}$ pode ser facilmente encontrada por manipulação dos coeficientes. No caso de uma equação de grau 2, ou de segundo grau,

$$ax^2 + bx + c = 0,$$

com $a \neq 0$, as soluções $x_0 = \frac{-b+\sqrt{b^2-4ac}}{2a}$ e $x_1 = \frac{-b-\sqrt{b^2-4ac}}{2a}$ também são bem conhecidas. Não importa se estas soluções são fornecidas com o que chamamos de “fórmulas prontas”, pois estas fórmulas prontas são obtidas manipulando-se os coeficientes da equação.

O Teorema Fundamental da Álgebra, demonstrado por Karl Friedrich Gauss em 1799 em sua tese de doutorado, garante que uma equação polinomial de grau n possui exatamente n raízes complexas. De acordo com o Teorema Fundamental da Álgebra, uma equação polinomial de grau n ,

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + a_{n-2} x^{n-2} + \dots + a_1 x + a_0 = 0, \quad (3)$$

com coeficientes reais ou complexos, possui exatamente n raízes r_1, r_2, \dots, r_n reais ou complexas (não necessariamente distintas) e além disso,

$$p(x) = a_n(x - r_1)(x - r_2) \dots (x - r_n).$$

O trabalho de encontrar estas n raízes pode ser feito por partes. Caso se possa determinar uma raiz r_1 da equação (3), então podemos escrever

$$p(x) = (x - r_1)q(x),$$

e o trabalho de encontrar as demais raízes de $p(x)$ pode continuar pela determinação das raízes de $q(x)$. A vantagem neste caso é que o grau de $q(x)$ é menor do que o grau de $p(x)$. Para ser mais preciso, se $gr(p) = n$, então teremos $gr(q) = n - 1$.

Nas próximas duas seções apresentaremos técnicas de solução por meio de radicais para equações polinomiais de graus 3 e 4. Equações polinomiais de grau maior ou igual a 5 não são solúveis por meio de radicais. Quem demonstrou isso foi o matemático Évariste Galois. O leitor interessado na Teoria de Galois ou em mais informações a respeito dos polinômios, como operações com polinômios, raízes de polinômios ou divisibilidade de polinômios, pode consultar Biazzi (2014).

2 Equações polinomiais do terceiro grau

Desde a descoberta de uma fórmula que resolvesse a equação do 2º grau, os matemáticos começaram a pensar em formas de resolver as equações de grau 3. Uma equação do terceiro grau é uma equação na forma

$$ax^3 + bx^2 + cx + d = 0, \quad (4)$$

com a, b, c e d coeficientes reais ou complexos e $a \neq 0$.

No início do século XVI, dois matemáticos italianos, Tartaglia e Cardano, conseguiram encontrar uma fórmula para a resolução das equações do terceiro grau. Girolamo Cardano (1501-1576) era um respeitado professor de Bolonha e Milão que frequentava a alta sociedade e além de matemático era médico, astrônomo, astrólogo e filósofo. Ele tinha grande vontade de aprender sobre os mais diversos assuntos, o que resultou na publicação de diversos livros.

Já Tartaglia, cujo nome oficial é Niccolò Fontana (1500-1557), era órfão de pai e muito pobre. Ainda na infância, quando sua cidade foi atacada pelos franceses, Niccolò sofreu um golpe que perfurou seu palato e o deixou gago para o resto da vida, e por isso recebeu o apelido Tartaglia, que significa gago. Apesar das dificuldades, Tartaglia era autodidata e muito inteligente e se tornou professor de matemática.

A obtenção de uma fórmula resolutive para as equações do terceiro grau teve início por volta de 1510 quando Scipione del Ferro, professor de Matemática da Universidade de Bolonha encontrou uma fórmula para resolver as equações do tipo $x^3 + px + q = 0$. Ele não publicou essa fórmula, mas compartilhou-a com seu aluno Antonio Maria Fior que posteriormente iniciou um desafio com Tartaglia que consistia em um passar para o outro uma lista de equações de terceiro grau para serem resolvidas. Tartaglia aceitou o desafio e em 1535 conseguiu encontrar uma forma geral para resolver as equações do tipo $x^3 + px^2 + q = 0$, além daquelas já conhecidas por Fior. Cardano, que na época estava escrevendo uma obra englobando Álgebra, Aritmética e Geometria, sabendo da resolução obtida por Tartaglia pediu a ele que a revelasse para que fosse colocada em seu livro, mas Tartaglia não concordou alegando que tinha intenção de publicá-la numa obra própria.

Cardano implorou, sob juramento ao Evangelho, que não publicaria a fórmula e Tartaglia decidiu confiar-lhe o segredo. Em 1545, em sua obra *Ars Magna*, Cardano quebrou a promessa feita a Tartaglia e publicou a fórmula, dizendo ainda que 30 anos antes Scipione del Ferro já havia obtido os mesmos resultados. Tartaglia foi a público a fim de esclarecer os fatos e a traição de Cardano, mas no fim das contas a fórmula obtida por Tartaglia é conhecida até hoje como Fórmula de Cardano.

Contudo, o método desenvolvido por Tartaglia só é aplicável em equações de terceiro grau que não possuem o termo quadrático e portanto não podem ser aplicados na equação geral (4). Ludovico Ferrari, professor da Universidade de Bolonha e discípulo de Cardano, desenvolveu um método que permite eliminar o termo quadrático da equação original (4) transformando-a em uma equação do tipo $x^3 + px + q = 0$.

A técnica de Ferrari consiste em aplicar a mudança de variáveis $x = y - \frac{b}{3a}$ na equação (4), obtendo

$$a \left(y - \frac{b}{3a} \right)^3 + b \left(y - \frac{b}{3a} \right)^2 + c \left(y - \frac{b}{3a} \right) + d = 0,$$

e portanto

$$ay^3 + \left(c - \frac{b^2}{3a} \right) y + \left(\frac{2b^3}{27a^2} - \frac{bc}{3a} + d \right) = 0.$$

Designando agora $p = c - \frac{b^2}{3a}$ e $q = \frac{2b^3}{27a^2} - \frac{bc}{3a} + d$, recaímos em uma equação de grau 3 sem o termo quadrático.

O que precisamos para obter as raízes da equação geral (4) é encontrar pelo menos uma raiz. A partir desta raiz encontrada podemos reduzir o grau do polinômio e usar a fórmula de Bháskara para determinar as outras duas raízes. Iniciamos, sem perda de generalidade,

considerando a equação

$$x^3 + bx^2 + cx + d = 0,$$

pois caso o coeficiente dominante a seja diferente de 1, podemos dividir toda a equação por a e obter a equação com coeficiente dominante 1. De acordo com a técnica de Ferrari, fazendo a mudança de variáveis $x = y - \frac{b}{3}$, obtemos então

$$y^3 + py + q = 0, \tag{5}$$

com $p = c - \frac{b^2}{3}$ e $q = \frac{2b^3}{27} - \frac{bc}{3} + d$.

A ideia central agora é comparar a expressão (5) com a expressão do cubo de uma soma. Se u e v são dois números reais ou complexos, então

$$(u + v)^3 = u^3 + 3u^2v + 3uv^2 + v^3,$$

ou equivalentemente

$$(u + v)^3 - 3uv(u + v) - (u^3 + v^3) = 0. \tag{6}$$

Note que esta última expressão é uma equação cúbica, no termo $(u + v)$, que não apresenta o termo quadrático. Comparando então a equação (6) com a equação (5), vemos que a soma $(u + v)$ é uma raiz de (5), desde que

$$\begin{cases} p = -3uv \\ q = -(u^3 + v^3). \end{cases}$$

Basta agora determinarmos u e v solução do sistema (não linear) acima e a raiz procurada para a equação (5) será $(u + v)$. Dentre as várias maneiras para resolver este sistema, vamos utilizar uma técnica não convencional que pode ser aplicada neste caso particular. Reescrevemos o sistema como

$$\begin{cases} u^3v^3 = -\frac{p^3}{27} \\ u^3 + v^3 = -q, \end{cases}$$

e queremos encontrar u^3 e v^3 conhecendo sua soma e seu produto. Por conta disso, sabemos então que u^3 e v^3 são as raízes da equação quadrática

$$w^2 - (u^3 + v^3)w + u^3v^3 = 0,$$

isto é,

$$w^2 + qw - \frac{p^3}{27} = 0.$$

Desta forma

$$u^3 = \frac{-q + \sqrt{q^2 + 4\frac{p^3}{27}}}{2} = \frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}$$

e

$$v^3 = \frac{-q - \sqrt{q^2 + 4\frac{p^3}{27}}}{2} = \frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}.$$

Portanto

$$u = \sqrt[3]{\frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad \text{e} \quad v = \sqrt[3]{\frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}},$$

e uma raiz procurada da equação (5) é

$$y = u + v = \sqrt[3]{\frac{-q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{\frac{-q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

Finalmente para voltar na equação original (4), basta ver que $x = y + \frac{b}{3}$.

O leitor atento verá que u^3 e v^3 , obtidos por este método, podem ser números complexos. Não queremos aqui entrar em muitos detalhes sobre os números complexos, mas caso u^3 e v^3 sejam complexos, eles serão complexos conjugados. As raízes cúbicas de dois números complexos conjugados continuam complexos conjugados uma da outra, e a soma de dois números complexos conjugados é um número real. Então, mesmo que tenhamos que determinar as raízes cúbicas de dois números complexos, a raiz y encontrada por esta técnica será um número real. Para mais detalhes sobre números complexos, números complexos conjugados e raízes de números complexos sugerimos Zill (2011).

Como ilustração da técnica obtida, vamos considerar dois exemplos nos quais queremos determinar as raízes (reais ou complexas) de polinômios de grau 3.

Exemplo 1. Considerando a equação $x^3 - 6x^2 + 6x - 5 = 0$, que chamaremos de equação original. Sabemos que a substituição de Ferrari $x = y - \frac{b}{3} = y + 2$ eliminará o termo quadrático. Fazendo então esta substituição, obtemos

$$(y + 2)^3 - 6(y + 2)^2 + 6(y + 2) - 5 = 0,$$

e após a reorganização dos termos,

$$y^3 - 6y - 9 = 0.$$

Claro que para chegarmos nesta última equação, sem o termo quadrático, poderíamos usar as fórmulas obtidas anteriormente. Como $b = -6$, $c = 6$ e $d = -5$, podemos calcular $p = c - \frac{b^2}{3} = -6$ e $q = \frac{2b^3}{27} - \frac{bc}{3} + d = -9$, obtendo o polinômio em y , sem o termo quadrático.

Agora, para encontrar u e v , temos que

$$u^3 = \frac{9}{2} + \sqrt{\frac{(-9)^2}{4} + \frac{(-6)^3}{27}} = 8 \quad \text{e} \quad v^3 = \frac{9}{2} - \sqrt{\frac{(-9)^2}{4} + \frac{(-6)^3}{27}} = 1,$$

e portanto

$$u = \sqrt[3]{8} = 2 \quad \text{e} \quad v = \sqrt[3]{1} = 1.$$

Segue que $y = u + v = 3$, donde $x = y - \frac{b}{3} = 3 - \frac{-6}{3} = 5$ é uma raiz procurada da equação de grau 3. De posse dessa raiz, vamos dividir o polinômio original por $(x - 5)$. Temos então

$$x^3 - 6x^2 + 6x - 5 = (x - 5)(x^2 - x + 1),$$

e agora basta encontrar as duas raízes restantes com a equação quadrática $x^2 - x + 1 = 0$. Estas raízes são precisamente

$$x_1 = \frac{1 + \sqrt{1 - 4}}{2} = \frac{1 + \sqrt{3}i}{2},$$

e

$$x_2 = \frac{1 - \sqrt{1 - 4}}{2} = \frac{1 - \sqrt{3}i}{2}.$$

Segue que as 3 raízes da equação original, obtidas por meio de radicais, são $x_0 = 5$, $x_1 = \frac{1}{2} + \frac{\sqrt{3}}{2}i$, e $x_2 = \frac{1}{2} - \frac{\sqrt{3}}{2}i$.

O leitor poderia agora dizer que o exemplo anterior poderia ser resolvido de forma mais rápida. É conhecido que uma equação polinomial geral da forma (2) possui uma solução racional na forma $\frac{p}{q}$ se, e somente se, $p|a_0$ e $q|a_n$. Neste caso, os únicos números racionais que poderiam ser solução da equação polinomial do exemplo anterior são ± 5 e ± 1 . Poderíamos rapidamente testar estes quatro valores e encontrar a primeira raiz $x_0 = 5$. Este exemplo permite isso, porém há casos em que as raízes podem não ser racionais ou mesmo casos em que a_0 possui muitos divisores e o trabalho de procurar raízes por substituição pode se tornar tão exaustivo quando o método que aplicamos, como no exemplo seguinte.

Exemplo 2. Consideremos a equação $x^3 - 24x^2 + 160x - 256 = 0$, que chamaremos de equação original. Fazendo a substituição de Ferrari $x = y - \frac{b}{3} = y + 8$, temos

$$(y + 8)^3 - 24(y + 8)^2 + 160(y + 8) - 256 = 0,$$

e reorganizando os termos

$$y^3 - 32y = 0.$$

Embora seja uma equação de grau 3 ainda, já detectamos que esta equação agora pode ser escrita na forma $y(y^2 - 32) = 0$. Desta forma as raízes procuradas agora são mais simples de serem obtidas e são

$$y_0 = 0, \quad y_1 = \sqrt{32} = 4\sqrt{2} \quad \text{e} \quad y_2 = -\sqrt{32} = -4\sqrt{2}.$$

Voltando para a variável original x , como $x = y + 8$, temos as 3 raízes da equação original $x_0 = 8$, $x_1 = 8 + 4\sqrt{2}$ e $x_2 = 8 - 4\sqrt{2}$.

3 Equações polinomiais do quarto grau

Uma equação do quarto grau é uma equação na forma

$$ax^4 + bx^3 + cx^2 + dx + e = 0, \tag{7}$$

com a, b, c, d e e coeficientes reais ou complexos, e $a \neq 0$. No mesmo documento onde publicou a resolução da equação do terceiro grau, na Ars Magna, Cardano também publicou um método

para a resolução de uma equação do quarto grau, desenvolvida pelo seu discípulo Ludovico Ferrari.

A resolução da equação de quarto grau pelo Método de Ferrari, consiste primeiro em transformar a equação original (7) em uma equação na forma

$$x^4 + px^2 + r = qx, \quad (8)$$

ou seja uma equação do quarto grau sem o termo cúbico. Isto é conseguido fazendo a mudança de variável $x = y - \frac{b}{4a}$ na equação original (7), para chegar na forma (8). Vamos aos detalhes.

Inicialmente podemos considerar, sem perda de generalidade, a equação

$$x^4 + bx^3 + cx^2 + dx + e = 0,$$

e utilizando a mudança de variável sugerida por Ferrari, $x = y - \frac{b}{4}$, obtemos

$$\left(y - \frac{b}{4}\right)^4 + b\left(y - \frac{b}{4}\right)^3 + c\left(y - \frac{b}{4}\right)^2 + d\left(y - \frac{b}{4}\right) + e = 0,$$

e portanto

$$y^4 + \left(c - \frac{3b^2}{8}\right)y^2 + \left(d + \frac{b^3}{8} - \frac{bc}{2}\right)y + \left(\frac{b^2c}{16} - \frac{3b^4}{256} - \frac{bd}{4} + e\right) = 0.$$

Designando $p = c - \frac{3b^2}{8}$, $q = d + \frac{b^3}{8} - \frac{bc}{2}$ e $r = \frac{b^2c}{16} - \frac{3b^4}{256} - \frac{bd}{4} + e$, e chegamos na forma desejada

$$y^4 + py^2 + qy + r = 0.$$

Feito isso, a ideia agora é escrever

$$y^4 + py^2 + r = -qy,$$

e tornar os dois membros quadrados perfeitos. Queremos acrescentar então em ambos os membros desta equação termos que tornem ambos os membros quadrados perfeitos. Para qualquer valor de $\alpha \in \mathbb{R}^*$, acrescentando αy^2 e $\frac{q^2}{4\alpha}$ em ambos os membros, obtemos

$$y^4 + (\alpha + p)y^2 + \left(r + \frac{q^2}{4\alpha}\right) = \alpha y^2 - qy + \frac{q^2}{4\alpha} = \left(\sqrt{\alpha}y - \frac{q}{2\sqrt{\alpha}}\right)^2.$$

Então queremos encontrar α para que o lado esquerdo da equação também seja um quadrado perfeito. Notemos que uma expressão quadrática $z^2 + mz + n$ é um quadrado perfeito se e somente se $\Delta = m^2 - 4n = 0$. Desta forma procuramos α de forma que

$$(\alpha + p)^2 - 4\left(r + \frac{q^2}{4\alpha}\right) = 0,$$

ou ainda

$$\alpha^3 + 2p\alpha^2 + (p^2 - 4r)\alpha - q^2 = 0.$$

Resolvendo agora esta equação do terceiro grau, determinamos α que torna ambos os membros da equação

$$y^4 + (\alpha + p)y^2 + \left(r + \frac{q^2}{4\alpha}\right) = \alpha y^2 - qy + \frac{q^2}{4\alpha}$$

quadrados perfeitos. Feito isso, já que ambos os membros são quadrados perfeitos, basta extrair raiz quadrada de ambos os membros obtendo

$$\left(y^2 + \frac{\alpha + p}{2}\right) = \pm \left(\sqrt{\alpha}y - \frac{q}{2\sqrt{\alpha}}\right).$$

Esta última expressão nos fornece duas equações quadráticas em y e portanto quatro valores para y . De posse dos quatro valores obtidos para y , usamos $x = y - \frac{b}{4}$ para obter as quatro raízes procuradas para a equação original.

Notemos que o valor de α foi obtido pela solução de uma equação do terceiro grau. Conforme visto anteriormente α poderá assumir três valores distintos e então poderíamos perguntar se isto nos conduziria a 12 raízes da equação original, sendo 4 raízes para cada valor de α considerado. O que acontece nesse caso é que temos quatro raízes que serão combinadas duas a duas de acordo com a escolha de α . Lembremos que há 3 formas distintas de agrupar 4 elementos dois a dois. Além disso, este α pode ser complexo, ou mesmo um número real negativo, o que conduziria a soluções complexas da equação original.

Vamos utilizar as ideias desenvolvidas em um exemplo.

Exemplo 3. Consideremos a equação $x^4 + 4x^3 - 6x^2 - 36x - 31 = 0$. Primeiramente faremos a eliminação do termo cúbico. Usando a substituição de Ferrari $x = y - \frac{b}{4} = y - 1$, obtemos

$$(y - 1)^4 + 4(y - 1)^3 - 6(y - 1)^2 - 36(y - 1) - 31 = 0$$

e após a reorganização dos termos,

$$y^4 - 12y^2 - 16y - 4 = 0.$$

Claro que para chegarmos nesta última equação, sem o termo cúbico, poderíamos utilizar as fórmulas obtidas nesta seção. Como $b = 4$, $c = -6$, $d = -36$ e $e = -31$, calculamos $p = c - \frac{3b^2}{8} = -12$, $q = d + \frac{b^3}{8} - \frac{bc}{2} = -16$ e $r = \frac{b^2c}{16} - \frac{3b^4}{256} - \frac{bd}{4} + e = -4$, chegando na equação em y sem o termo cúbico.

Agora reescrevemos

$$y^4 - 12y^2 - 4 = 16y,$$

e acrescentamos em ambos os membros os termos αy^2 e $\frac{q^2}{4\alpha} = \frac{64}{\alpha}$ obtendo

$$y^4 + (\alpha - 12)y^2 + \left(\frac{64}{\alpha} - 4\right) = \alpha y^2 + 16y + \frac{64}{\alpha}, \quad (9)$$

sendo que α deve ser uma solução da equação cúbica

$$\alpha^3 + 2p\alpha^2 + (p^2 - 4r)\alpha - q^2 = 0,$$

isto é,

$$\alpha^3 - 24\alpha^2 + 160\alpha - 256 = 0.$$

Usando os métodos da seção anterior para equação de grau 3, obtemos três valores para α que são $\alpha_0 = 8$, $\alpha_1 = 8 + 4\sqrt{2}$ e $\alpha_2 = 8 - 4\sqrt{2}$ (ver segundo exemplo da seção anterior). Lembremos que como comentado anteriormente, a escolha do valor de α não alterará as raízes procuradas, apenas a ordem de obtenção delas. Naturalmente o valor $\alpha = 8$ é o mais simples de ser utilizado agora. Com $\alpha = 8$ em (9) obtemos

$$y^4 - 4y^2 + 4 = 8y^2 + 16y + 8,$$

donde reescrevemos ambos os membros como quadrados perfeitos

$$(y^2 - 2)^2 = (\sqrt{8}y + \sqrt{8})^2.$$

Extraindo raiz quadrada em ambos os membros, obtemos duas equações quadráticas em y ,

$$y^2 - 2 = \sqrt{8}y + \sqrt{8} \quad \text{e} \quad y^2 - 2 = -(\sqrt{8}y + \sqrt{8}),$$

ou ainda

$$y^2 - \sqrt{8}y - (2 + \sqrt{8}) = 0 \quad \text{e} \quad y^2 + \sqrt{8}y - (2 - \sqrt{8}) = 0,$$

Resolvendo estas duas equações quadráticas obtemos as quatro raízes da equação de quarto grau modificada,

$$\begin{aligned} y_0 &= \sqrt{2} + \sqrt{4 + \sqrt{8}}, \\ y_1 &= \sqrt{2} - \sqrt{4 + \sqrt{8}}, \\ y_2 &= -\sqrt{2} + \sqrt{4 - \sqrt{8}}, \\ y_3 &= -\sqrt{2} - \sqrt{4 - \sqrt{8}}. \end{aligned}$$

Para a equação original lembramos que $x = y - 1$ e portanto as raízes da equação original são

$$\begin{aligned} x_0 &= -1 + \sqrt{2} + \sqrt{4 + \sqrt{8}}, \\ x_1 &= -1 + \sqrt{2} - \sqrt{4 + \sqrt{8}}, \\ x_2 &= -1 - \sqrt{2} + \sqrt{4 - \sqrt{8}}, \\ x_3 &= -1 - \sqrt{2} - \sqrt{4 - \sqrt{8}}. \end{aligned}$$

Observe que nesse caso foram encontradas quatro soluções irracionais.

Conclusões

Nos estudos que originaram este trabalho tivemos a oportunidade de observar o quanto a busca por soluções de equações polinomiais fez com que a matemática evoluísse ao longo dos anos. A própria linguagem matemática sofreu diversas adaptações e, podemos dizer até, melhorias para que fosse mais prático resolver equações. Algumas descobertas da área de resoluções de equações alteraram de forma definitiva os rumos da matemática.

Muitos problemas de diversas áreas são modelados por equações polinomiais, portanto sabemos que elas são de grande importância. Porém, como existem diversos métodos computacionais capazes de encontrar soluções com bastante aproximação, a resolução por meio de radicais hoje em dia acaba sendo deixada de lado.

Um fato que chama a atenção nessa temática é que apesar de as últimas atualizações nas resoluções por radicais terem mais de 300 anos, a maior parte dos alunos de ensino médio não conhece essas fórmulas ou mesmo a história tão rica por trás delas. Um trabalho extremamente rico, histórica e matematicamente, seria apresentar para essas novas gerações o quanto a matemática evoluiu através da busca pelas soluções por radicais.

Referências

- Biazzi, Ricardo N. *Polinômios Irredutíveis - Critérios e Aplicações*. Dissertação, UNESP. Rio Claro, 2014.
- Costa, Ueslei F. *Equações Algébricas: do Papiro de Ahmes até Évariste Galois*. Dissertação, UFTM. Uberaba, 2020.
- Garbi, Gilberto G. *O Romance das Equações Algébricas*. 3ª edição revista e ampliada. Editora Livraria da Física. São Paulo, 2009.
- Zill, Dennis G., Shanahan, Patrick D. *Curso introdutório à análise complexa com aplicações*. Rio de Janeiro, LTC, 2011.

Aproximação de funções por série de Fourier-Bessel

Cintya Akemi Okawa - Universidade Estadual do Oeste do Paraná (UNIOESTE)
Sandro Marcos Guzzo - Universidade Estadual do Oeste do Paraná (UNIOESTE)

(Recebido em 13/12/2021. Aceito em 20/12/2021. Publicado em 22/12/2021)

Resumo: Algumas séries fornecem maneiras de representar funções relativamente complicadas em termos de funções mais elementares ou familiares. As séries de potência são os exemplos mais clássicos quando se fala de expansão de uma função em série. No entanto, existem outras séries importantes na Matemática, como a série de Fourier, a série de Laurent e a série de Fourier-Bessel que é o foco do nosso estudo. Dessa forma, este trabalho tem como objetivo apresentar as Funções de Bessel e mostrar como representar algumas funções como uma série envolvendo estas funções, que é conhecida como série de Fourier-Bessel. Para isso, apresentaremos alguns aspectos da Equação de Bessel, das Funções de Bessel e das propriedades que permitem expressar funções como uma série das Funções de Bessel. Por fim, com o intuito de ilustrar o procedimento, apresentaremos os gráficos de algumas aproximações utilizando o software *Scilab*.

Palavras-chave: Equação de Bessel. Funções de Bessel. Série de Fourier-Bessel.

1 Introdução

Algumas séries fornecem maneiras de representar funções relativamente complicadas em termos de funções elementares e/ou familiares. Em livros de cálculo diferencial e integral são encontradas formas de expandir uma função f como série de potência. Em especial, são apresentadas as expressões

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!} (x-a)^n \quad \text{e} \quad f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

conhecidas, respectivamente, como série de Taylor e de Maclaurin. A ideia neste caso é escrever uma função f em termos das funções mais simples x^n . Claro que isto só é possível com certas hipóteses sobre a função f .

Além das séries de Taylor e de Maclaurin, outras séries são conhecidas na literatura matemática. Uma dessas é a série de Fourier-Bessel. Como o nome sugere, a série de Fourier-Bessel está relacionada com as Funções de Bessel. A ideia central é, com certas hipóteses, escrever uma dada função f como uma soma infinita das Funções de Bessel. Neste caso, as Funções de Bessel não são necessariamente “mais simples”, mas podem ser mais adequadas para certas aplicações.

A Função de Bessel de ordem ν , denotada por J_ν , é uma função que satisfaz a equação

$$t^2 y'' + ty' + (t^2 - \nu^2)y = 0,$$

com $t > 0$, $y = y(t)$ e $\nu \in \mathbb{R}$, chamada de equação diferencial de Bessel de ordem ν . É importante notarmos que ν não é a ordem da equação diferencial. A “ordem ν ” é uma referência específica para a equação de Bessel. A ordem da equação diferencial no sentido da teoria geral das equações diferenciais é 2.

A cada $\nu \in \mathbb{R}$ está associada uma equação de Bessel e sua respectiva solução denotada por J_ν . Isto nos fornece uma família infinita de funções a serem estudadas. Estas funções possuem várias propriedades interessantes. Em particular destacamos a ortogonalidade, que possibilita o estudo da série que envolve as Funções de Bessel.

2 Equação de Bessel e Funções de Bessel

A equação de Bessel está relacionada com vários problemas físicos, dentre eles os que envolvem ondas eletromagnéticas, condução de calor, vibração, difusão, processamento de sinais (filtro de Bessel) e dinâmica de corpos flutuantes. Além disso, recentemente, as Funções de Bessel aparecem no problema inverso da propagação de ondas, com aplicações em medicina, astronomia e imagem acústica. Embora existam todas estas aplicações, neste trabalho estamos interessados apenas no aspecto teórico.

Nesta seção, partiremos da equação de Bessel, buscando sua solução, e definiremos as Funções de Bessel. Procederemos desta forma, pois partir da definição de uma equação e procurar por suas soluções parece ser um caminho mais natural. No entanto, os conceitos poderiam ser definidos, de modo equivalente, partindo das Funções de Bessel e encontrando uma equação para a qual estas funções são soluções. Este processo pode ser encontrado em Bowman (2010).

Para o estudo da equação de Bessel e das Funções de Bessel que faremos neste texto, será suficiente considerar $\nu = n \in \mathbb{N}$. Para o estudo do caso geral, sugerimos ao leitor consultar Okawa (2021) ou Bowman (2010).

A equação diferencial de segunda ordem

$$t^2 y'' + ty' + (t^2 - n^2)y = 0, \tag{1}$$

com $t > 0$, $y = y(t)$ e $n \in \mathbb{N}$ é a equação diferencial de Bessel de ordem n . Como esta é uma equação de segunda ordem, sabemos da teoria das equações diferenciais que ela deve possuir duas soluções linearmente independentes. Ao leitor não familiarizado com a teoria geral das equações diferenciais recomendamos Zill (2016).

A equação diferencial (1) não é uma equação com coeficientes constantes. Vamos encontrar uma solução em forma de série de potências em torno do ponto $t_0 = 0$. Queremos então encontrar uma solução da forma

$$y(t) = \sum_{k=0}^{\infty} c_k t^{k+r},$$

em que r é uma constante a ser determinada. Sabemos que no intervalo de convergência da série, a série é derivável e podemos obter $y'(t)$ derivando a série termo a termo, mantendo o intervalo de convergência, exceto possivelmente pelos extremos deste intervalo. Deste modo, temos

$$y'(t) = \sum_{k=0}^{\infty} (k+r)c_k t^{k+r-1},$$

e

$$y''(t) = \sum_{k=0}^{\infty} (k+r)(k+r-1)c_k t^{k+r-2}.$$

Substituindo agora na equação (1)

$$\begin{aligned} & t^2 y'' + t y' + (t^2 - n^2) y \\ &= t^2 \sum_{k=0}^{\infty} (k+r)(k+r-1)c_k t^{k+r-2} + t \sum_{k=0}^{\infty} (k+r)c_k t^{k+r-1} + (t^2 - n^2) \sum_{k=0}^{\infty} c_k t^{k+r} \\ &= \sum_{k=0}^{\infty} (k+r)(k+r-1)c_k t^{k+r} + \sum_{k=0}^{\infty} (k+r)c_k t^{k+r} + \sum_{k=0}^{\infty} c_k t^{k+r+2} - \sum_{k=0}^{\infty} n^2 c_k t^{k+r} \\ &= \sum_{k=0}^{\infty} [(k+r)(k+r) - n^2] c_k t^{k+r} + \sum_{k=0}^{\infty} c_k t^{k+r+2} \\ &= t^r \left[(r^2 - n^2)c_0 + ((1+r)^2 - n^2)c_1 t + \sum_{k=0}^{\infty} c_k t^{k+2} + \sum_{k=2}^{\infty} ((k+r)^2 - n^2)c_k t^k \right] \\ &= t^r \left[(r^2 - n^2)c_0 + ((1+r)^2 - n^2)c_1 t + \sum_{k=2}^{\infty} ((k+r)^2 - n^2)c_k t^k + \sum_{k=2}^{\infty} c_{k-2} t^k \right] \\ &= t^r \left[(r^2 - n^2)c_0 + ((1+r)^2 - n^2)c_1 t + \sum_{k=2}^{\infty} [((k+r)^2 - n^2)c_k + c_{k-2}] t^k \right] = 0. \end{aligned}$$

Dessa forma, temos que

$$\begin{aligned} (r+n)(r-n)c_0 &= 0 \\ (1+r-n)(1+r+n)c_1 &= 0 \\ (k+r-n)(k+r+n)c_k + c_{k-2} &= 0. \end{aligned}$$

Vamos pedir na primeira equação que $c_0 \neq 0$ pois caso contrário não conseguiremos uma restrição para r . Então, com $c_0 \neq 0$ na primeira equação, temos

$$(r+n)(r-n) = 0, \tag{2}$$

que é chamada de equação indicial. Suas raízes são $r_1 = n$ e $r_2 = -n$. Com a segunda equação, obtemos

$$(1+r-n)(1+r+n)c_1 = 0,$$

que nos conduz a $c_1 = 0$. A terceira equação nos dá uma fórmula recursiva para o cálculo dos coeficientes

$$c_k = -\frac{c_{k-2}}{(k+r-n)(k+r+n)} \quad k = 2, 3, 4, \dots \tag{3}$$

Agora, se $n \geq 0$, então $r_1 = n$ em (3) nos dá a fórmula de recorrência

$$c_k = -\frac{c_{k-2}}{k(2n+k)} \quad k = 2, 3, 4, \dots$$

e como $c_1 = 0$, temos que $c_3 = 0, c_5 = 0, \dots$, isto é

$$c_{2k+1} = 0,$$

para todo $k \in \mathbb{N}$. Assim, nos resta saber o que acontece com os c_{2k} . Calculando os coeficientes a partir da fórmula, obtemos

$$\begin{aligned} c_2 &= -\frac{c_0}{2(2n+2)} \\ c_4 &= -\frac{c_2}{4(2n+4)} = \frac{c_0}{4 \cdot 2(2n+2)(2n+4)} \\ c_6 &= -\frac{c_4}{6(2n+6)} = -\frac{c_0}{6 \cdot 4 \cdot 2(2n+2)(2n+4)(2n+6)} \\ &\vdots \end{aligned}$$

donde

$$\begin{aligned} c_2 &= -\frac{c_0}{1(n+1)2^2} \\ c_4 &= \frac{c_0}{2 \cdot 1(n+1)(n+2)2^4} \\ c_6 &= -\frac{c_0}{3 \cdot 2 \cdot 1(n+1)(n+2)(n+3)2^6} \\ &\vdots \end{aligned}$$

e em geral

$$c_{2k} = \frac{(-1)^k c_0}{k!(n+1)(n+2) \cdots (n+k)2^{2k}}. \quad (4)$$

Para facilitar a escrita dos coeficientes, vamos escolher

$$c_0 = \frac{1}{2^n n!}, \quad (5)$$

e utilizando a propriedade do fatorial, notamos que

$$(n+k)(n+(k-1)) \cdots (n+2)(n+1)n! = (n+k)!.$$

Note que esta escolha de c_0 é por pura conveniência, pois de acordo com (4), c_0 é uma constante de liberdade da solução da equação. Esta constante de liberdade ficará ajustada quando for imposta uma condição inicial da equação diferencial.

Assim, substituindo (5) em (4), obtemos

$$\begin{aligned} c_{2k} &= \frac{(-1)^k}{k!(n+k) \cdots (n+2)(n+1)n!2^{2k}2^n} \\ &= \frac{(-1)^k}{k!(n+k)!2^{2k+n}}. \end{aligned}$$

Obtemos assim a primeira solução para a equação de Bessel

$$y_1(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{t}{2}\right)^{2k+n}.$$

A função y_1 definida desta forma é comumente denotada por J_n e conhecida como Função de Bessel de primeira espécie de ordem n . Desta forma,

$$y_1(t) = J_n(t) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(n+k)!} \left(\frac{t}{2}\right)^{2k+n}. \quad (6)$$

Em particular,

$$J_0(t) = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{2^{2k}(k!)^2}, \quad (7)$$

é a Função de Bessel de primeira espécie de ordem zero. Note que esta é uma série de potências de t e como em toda série de potências, por simplicidade na escrita da série, consideramos $t^0 = 1$ mesmo para $t = 0$. Nestes termos, $J_0(0) = 1$ e também $J_n(0) = 0$ para todo $n \in \mathbb{N}^*$.

Como o nosso interesse principal envolve as funções J_n obtidas, não vamos nos preocupar em determinar a segunda solução da equação (1), linearmente independente com J_n . O leitor interessado nesta segunda solução pode consultar Bowman (2010).

A seguir veremos uma propriedade das Funções de Bessel de primeira espécie J_n . Embora estas funções possuam muitas propriedades operatórias importantes, só enunciaremos a propriedade de interesse deste texto. O leitor interessado em mais propriedades das Funções de Bessel de primeira espécie, pode consultar Bowmann (2010).

Proposição 1. *Sejam $n \in \mathbb{N}$ e J_n a Função de Bessel de primeira espécie de ordem n . Então para todo $t \in [0, \infty)$, temos*

- i) $\frac{d}{dt}(t^n J_n(t)) = t^n J_{n-1}(t);$
- ii) $t \frac{d}{dt} J_n(t) = n J_n(t) - t J_{n+1}(t).$

Prova. De acordo com a definição da função J_n ,

$$\begin{aligned} \frac{d}{dt}(t^n J_n(t)) &= \frac{d}{dt} t^n \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+n}}{2^{2k+n} k!(n+k)!} \\ &= \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+2n}}{2^{2k+n} k!(n+k)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k (2k+2n) t^{2k+2n-1}}{2^{2k+n} k!(n+k)(n+k-1)!} \\ &= t^n \sum_{k=0}^{\infty} \frac{(-1)^k 2(k+n) t^{2k+n-1}}{2^{2k+n} k!(n+k)(n+k-1)!} \end{aligned}$$

$$= t^n \sum_{k=0}^{\infty} \frac{(-1)^k}{k!((n-1)+k)!} \left(\frac{t}{2}\right)^{2k+n-1} = t^n J_{n-1}(t).$$

Para a segunda expressão, temos que a igualdade é trivialmente satisfeita para $t = 0$, já que $J_n(0) = 0$ para todo $n \in \mathbb{N}^*$. Agora, para $t \neq 0$, temos que

$$\frac{d}{dt}(t^{-n} J_n(t)) = \frac{d}{dt} \left(t^{-n} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+n}}{2^{2k+n} k! (n+k)!} \right) = \frac{d}{dt} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k}}{2^{2k+n} k! (n+k)!}$$

e observe que agora o caso $k = 0$ é constante. Então

$$\begin{aligned} \frac{d}{dt}(t^{-n} J_n(t)) &= \sum_{k=1}^{\infty} \frac{(-1)^k (2k) t^{2k-1}}{2^{2k+n} k! (n+k)(n+k)!} \\ &= \sum_{k=1}^{\infty} \frac{(-1)^k t^{2k-1}}{2^{2k+n-1} (k-1)! (n+k)(n+k)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^{(k+1)} t^{2(k+1)-1}}{2^{2(k+1)+n-1} (k+1-1)! (n+k+1)(n+k+1)!} \\ &= - \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+1}}{2^{2k+n+1} k! (n+k+1)(n+k+1)!} \\ &= -t^{-n} \sum_{k=0}^{\infty} \frac{(-1)^k t^{2k+n+1}}{2^{2k+n+1} k! (n+1+k)(n+1+k)!} = -t^{-n} J_{n+1}(t). \end{aligned}$$

Desta forma,

$$-t^{-n} J_{n+1}(t) = \frac{d}{dt}(t^{-n} J_n(t)) = -nt^{-n-1} J_n(t) + t^{-n} J'_n(t),$$

e multiplicando ambos os membros por t^{n+1} vem

$$-t J_{n+1}(t) = -n J_n(t) + t J'_n(t),$$

e a igualdade desejada. □

Podemos generalizar um pouco mais os resultados da Proposição 1. Note que, nas hipóteses da Proposição, para qualquer $\alpha \in (0, \infty)$, temos

$$\frac{d}{dt}(t^n J_n(\alpha t)) = \alpha t^n J_{n-1}(\alpha t), \tag{8}$$

e também

$$t \frac{d}{dt} J_n(\alpha t) = n J_n(\alpha t) - \alpha t J_{n+1}(\alpha t). \tag{9}$$

A ortogonalidade das funções de Bessel é um dos resultados mais importantes e que permite definirmos a série de Fourier-Bessel. Vamos agora definir o produto interno que utilizaremos no espaço vetorial $\mathcal{C}((0, 1); \mathbb{R})$.

Definição 2. Seja V um espaço vetorial real. Um produto interno sobre V é uma aplicação

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

que satisfaz

- i) $\langle u, u \rangle \geq 0$ e além disso, $\langle u, u \rangle = 0$ se e somente se $u = 0$;
- ii) $\langle u, v \rangle = \langle v, u \rangle$;
- iii) $\langle \alpha u, v \rangle = \alpha \langle u, v \rangle$;
- iv) $\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$,

para todos $u, v \in V$ e $\alpha \in \mathbb{R}$.

Proposição 3. *Sejam $\mathcal{C} = C^0((0, 1); \mathbb{R})$ o espaço vetorial das funções contínuas de $(0, 1)$ em \mathbb{R} . A aplicação*

$$\begin{aligned} \langle \cdot, \cdot \rangle : \mathcal{C} \times \mathcal{C} &\rightarrow \mathbb{R} \\ (f, g) &\mapsto \langle f, g \rangle = \int_0^1 tf(t)g(t)dt \end{aligned} \quad (10)$$

é um produto interno sobre o espaço \mathcal{C} .

Não é difícil provar que esta função satisfaz as propriedades (i)-(iv) da definição de produto interno. Basta usar as propriedades das integrais e da multiplicação de números reais.

Agora que já temos um produto interno, podemos falar da ortogonalidade das Funções de Bessel. Vamos ver que essa ortogonalidade está relacionada com as raízes da equação $J_n(t) = 0$. Em Bowman (2010), encontramos que, se ν é um número real qualquer, a equação $J_\nu(t) = 0$ possui infinitas raízes reais não negativas.

Considerando uma Função de Bessel específica J_n e as suas infinitas raízes reais positivas α_i , para $i \in \mathbb{N}^*$, com $\alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_n < \dots$, construímos o conjunto

$$\{J_n(\alpha_i t); \quad i \in \mathbb{N}^*\} \subset \mathcal{C}.$$

Vamos provar que este conjunto é um conjunto de funções ortogonais em relação ao produto interno considerado em (10). Para as próximas proposições observemos que uma função $J_n(\alpha t)$ satisfaz a equação diferencial

$$t^2 y'' + ty' + (\alpha^2 t^2 - n^2)y = 0. \quad (11)$$

De fato, como $J_n(t)$ é solução da equação

$$t^2 y'' + ty' + (t^2 - n^2)y = 0,$$

fazendo $x = \alpha t$, obtemos

$$\frac{dy}{dt} = \frac{dy}{dx} \frac{dx}{dt} = \alpha \frac{dy}{dx}$$

e

$$\frac{d^2 y}{dt^2} = \frac{d}{dt} \left(\frac{dy}{dt} \right) = \frac{d}{dx} \left(\frac{dy}{dt} \right) \frac{dx}{dt} = \alpha^2 \frac{d^2 y}{dx^2}.$$

Substituindo em (11), temos

$$\left(\frac{x}{\alpha}\right)^2 \alpha^2 \frac{d^2 y}{dx^2} + \left(\frac{x}{\alpha}\right) \alpha \frac{dy}{dx} + \left(\alpha^2 \left(\frac{x}{\alpha}\right)^2 - n^2\right) y = 0,$$

e então

$$x^2 \frac{d^2 y}{dx^2} + x \frac{dy}{dx} + (x^2 - n^2) y = 0,$$

que é a equação de Bessel de ordem n na variável x , cuja solução é $J_n(x)$. E como $x = \alpha t$, temos que $J_n(\alpha t)$ é solução de (11).

Proposição 4. *Se J_n é uma Função de Bessel de ordem n e α_i para $i \in \mathbb{N}^*$ são as raízes positivas da função J_n , então o conjunto*

$$\{J_n(\alpha_i t); \quad i \in \mathbb{N}^*\},$$

é ortogonal em relação ao produto interno em (10).

Prova. Precisamos provar que dadas duas funções distintas deste conjunto, o produto interno entre estas duas funções se anula. De outra forma, se α_1 e α_2 são duas raízes positivas distintas de J_n , então

$$\langle J_n(\alpha_1 t), J_n(\alpha_2 t) \rangle = 0.$$

Notemos que $u = J_n(\alpha_1 t)$ é uma solução da equação

$$tu'' + u' + \left(\alpha_1^2 - \frac{n^2}{t^2}\right) tu = 0, \tag{12}$$

e $v = J_n(\alpha_2 t)$ é solução de

$$tv'' + v' + \left(\alpha_2^2 - \frac{n^2}{t^2}\right) tv = 0. \tag{13}$$

Multiplicando (12) por v e (13) por u , encontramos

$$tvu'' + vu' + \left(\alpha_1^2 - \frac{n^2}{t^2}\right) tuv = 0,$$

e

$$tuv'' + uv' + \left(\alpha_2^2 - \frac{n^2}{t^2}\right) tuv = 0,$$

e subtraindo a segunda da primeira, vem

$$t(vu'' - uv'') + (vu' - uv') + (\alpha_1^2 - \alpha_2^2) tuv = 0,$$

que pode ainda ser escrito como

$$\frac{d}{dt}(t(u'v - uv')) = (\alpha_2^2 - \alpha_1^2) tuv.$$

Integrando ambos os membros em t , no intervalo $(0, 1)$, obtemos

$$(\alpha_2^2 - \alpha_1^2) \int_0^1 tuv dt = t(u'v - uv') \Big|_{t=0}^1,$$

e substituindo u e v ,

$$\begin{aligned} (\alpha_2^2 - \alpha_1^2) \int_0^1 t J_n(\alpha_1 t) J_n(\alpha_2 t) dt &= t(J_n'(\alpha_1 t) J_n(\alpha_2 t) - J_n(\alpha_1 t) J_n'(\alpha_2 t)) \Big|_{t=0}^1 \\ &= J_n'(\alpha_1) J_n(\alpha_2) - J_n(\alpha_1) J_n'(\alpha_2). \end{aligned}$$

Como α_1 e α_2 são raízes de J_n e $\alpha_1^2 \neq \alpha_2^2$, então

$$\int_0^1 t J_n(\alpha_1 t) J_n(\alpha_2 t) dt = 0,$$

donde $\langle J_n(\alpha_1 t), J_n(\alpha_2 t) \rangle = 0$ e então $J_n(\alpha_1 t)$ e $J_n(\alpha_2 t)$ são ortogonais em relação ao produto interno em (10). Como α_1 e α_2 são duas raízes arbitrárias, isso vale para todo par de raízes diferentes da função J_n . \square

Proposição 5. *Se J_n é uma Função de Bessel de ordem n e α é qualquer uma das raízes da função J_n , então*

$$\langle J_n(\alpha t), J_n(\alpha t) \rangle = \frac{1}{2} J_{n+1}^2(\alpha) > 0. \quad (14)$$

Prova. Primeiramente a equação (11) pode ser reescrita como

$$(ty')' + \left(\alpha^2 t - \frac{n^2}{t} \right) y = 0, \quad (15)$$

e $y = J_n(\alpha t)$ é ainda uma solução desta equação.

Dessa forma, temos que

$$(ty')' = \left(\frac{n^2}{t} - \alpha^2 t \right) y,$$

e multiplicando a equação por $2tJ_n'(\alpha t)$ e substituindo $y = J_n(\alpha t)$, temos

$$2tJ_n'(\alpha t)(tJ_n'(\alpha t))' = 2(n^2 - \alpha^2 t^2)J_n(\alpha t)J_n'(\alpha t),$$

e então

$$\frac{d}{dt} ((tJ_n'(\alpha t))^2) = (n^2 - \alpha^2 t^2) ((J_n(\alpha t))^2)'$$

Vamos integrar os dois lados da igualdade no intervalo $(0, 1)$. Para o lado esquerdo da igualdade, usamos (9) e obtemos

$$\begin{aligned} (tJ_n'(\alpha t))^2 \Big|_{t=0}^1 &= (nJ_n(\alpha t) - \alpha t J_{n+1}(\alpha t))^2 \Big|_{t=0}^1 \\ &= (nJ_n(\alpha) - \alpha J_{n+1}(\alpha))^2 - (nJ_n(0))^2 = \alpha^2 J_{n+1}^2(\alpha), \end{aligned}$$

já que $nJ_n(\alpha) = nJ_n(0) = 0$ para todo $n \in \mathbb{N}$.

Já para o lado direito, encontramos a integral

$$\int_0^1 (n^2 - \alpha^2 t^2) (J_n^2(\alpha t))' dt.$$

Vamos aplicar uma integral por partes fazendo $u = n^2 - \alpha^2 t^2$ e $\frac{dv}{dt} = (J_n^2(\alpha t))'$, e assim $\frac{du}{dt} = -2\alpha^2 t$ e $v = J_n^2(\alpha t)$. Logo

$$\begin{aligned} \int_0^1 (n^2 - \alpha^2 t^2)(J_n^2(\alpha t))' dt &= [(n^2 - \alpha^2 t^2)J_n^2(\alpha t)]_{t=0}^1 + 2\alpha^2 \int_0^1 t J_n^2(\alpha t) dt \\ &= (n^2 - \alpha^2)J_n^2(\alpha) - n^2 J_n^2(0) + 2\alpha^2 \int_0^1 t J_n^2(\alpha t) dt. \end{aligned}$$

Notemos que, $(n^2 - \alpha^2)J_n^2(\alpha) - n^2 J_n^2(0) = 0$, pois α é raiz de $J_n(t)$. Da série de potências (6) vemos que $J_n(0) = 0$ para todo $n \neq 0$, e para o caso em que $n = 0$, temos $n^2 = 0$, donde $n^2 J_n^2(0) = 0$ para todo $n \in \mathbb{N}$. Segue que,

$$\int_0^1 (n^2 - \alpha^2 t^2)(J_n^2(\alpha t))' dt = 2\alpha^2 \langle J_n(\alpha t), J_n(\alpha t) \rangle.$$

Juntando então os dois lados da igualdade, obtemos

$$2\alpha^2 \langle J_n(\alpha t), J_n(\alpha t) \rangle = \alpha^2 J_{n+1}^2(\alpha),$$

e assim,

$$\langle J_n(\alpha t), J_n(\alpha t) \rangle = \frac{1}{2} J_{n+1}^2(\alpha).$$

Bowman (2010) afirma que as funções J_n e J_{n+1} não possuem raízes em comum. Assim, como α é raiz de J_n , então α não é raiz de J_{n+1} , donde o último termo da última equação é estritamente positivo. \square

Agora que provamos a ortogonalidade das funções de Bessel, podemos tratar da série de Fourier-Bessel.

3 A série de Fourier-Bessel

Como já vimos anteriormente as funções $J_n(\alpha_i t)$, em que α_i é raiz de $J_n(t) = 0$, são funções duas a duas ortogonais em relação a função t e vimos também que a equação $J_n(t) = 0$ possui infinitas raízes. Logo, a família das funções $J_n(\alpha_i t)$ forma uma base de Hilbert para o espaço vetorial $\mathcal{C} = C^0((0, 1); \mathbb{R})$ das funções contínuas com domínio $(0, 1)$. Dessa forma, enunciamos o próximo teorema.

Teorema 6. *Sejam $\alpha_1 < \alpha_2 < \alpha_3 < \dots < \alpha_k < \dots$ as raízes positivas da equação $J_n(t) = 0$, com $n \geq 0$, e $f \in \mathcal{C}$. Então f pode ser escrita na forma*

$$f(t) = \sum_{k=1}^{\infty} C_k J_n(\alpha_k t), \tag{16}$$

que é chamada de série de Fourier-Bessel de ordem n da função f . Além disso, a equação $J_n(t) = 0$ é chamada de condição de contorno.

Neste trabalho não argumentaremos sobre a demonstração deste resultado. No entanto, Watson (1922) faz uma discussão sobre este tipo de expansão em série provando sua veracidade. A questão que daremos atenção é como determinar os coeficientes C_k admitindo que uma função f possa ser escrita como uma série da forma (16).

Admitindo que uma função $f \in \mathcal{C}$ possa ser escrita na forma (16), temos que

$$f(t) = C_1 J_n(\alpha_1 t) + C_2 J_n(\alpha_2 t) + C_3 J_n(\alpha_3 t) + \dots$$

Para determinar o coeficiente C_k , tomamos o produto interno desta igualdade com a função $J_n(\alpha_k t)$, e como $\langle C_i J_n(\alpha_i t), C_k J_n(\alpha_k t) \rangle = 0$ para todo $i \neq k$, resta

$$\langle f(t), J_n(\alpha_k t) \rangle = \langle C_k J_n(\alpha_k t), J_n(\alpha_k t) \rangle = C_k \langle J_n(\alpha_k t), J_n(\alpha_k t) \rangle,$$

donde

$$C_k = \frac{\langle f(t), J_n(\alpha_k t) \rangle}{\langle J_n(\alpha_k t), J_n(\alpha_k t) \rangle}.$$

Utilizando a definição do produto interno e a igualdade (14), obtemos

$$C_k = \frac{2}{(J_{n+1}(\alpha_k))^2} \int_0^1 t f(t) J_n(\alpha_k t) dt,$$

para todo $k \in \mathbb{N}^*$.

4 Exemplos de funções escritas como série de Fourier-Bessel

Nesta seção, faremos alguns exemplos de funções escritas como séries de Fourier-Bessel e faremos alguns testes utilizando o software *Scilab*¹ para analisarmos a aproximação das séries com relação a função original. Neste contexto, utilizaremos funções simples que nos permitam uma fácil comparação de resultados. Além disso, o uso de funções mais simples facilitará o trabalho analítico com o produto interno definido em (10). Para um grande conjunto de funções o produto interno necessitará de métodos numéricos.

Como primeiro exemplo, vamos escrever a função $f(t) = 1$ no intervalo $(0, 1)$ como uma série de Fourier-Bessel de ordem 0. Queremos então escrever

$$f(t) = \sum_{k=1}^{\infty} C_k J_0(\alpha_k t),$$

sendo α_k as infinitas raízes da função J_0 . Como já vimos,

$$C_k = \frac{2}{(J_{n+1}(\alpha_k))^2} \int_0^1 t f(t) J_n(\alpha_k t) dt,$$

isto é,

$$C_k = \frac{2}{(J_1(\alpha_k))^2} \int_0^1 t J_0(\alpha_k t) dt.$$

¹Software livre, versão , que pode ser obtido em <https://www.scilab.org/>.

Usando a identidade (8) temos

$$\begin{aligned}
 C_k &= \frac{2}{(J_1(\alpha_k))^2} \frac{1}{\alpha_k} \int_0^1 \alpha_k t J_0(\alpha_k t) dt \\
 &= \frac{2}{(J_1(\alpha_k))^2} \frac{1}{\alpha_k} \int_0^1 \frac{d}{dt} (t J_1(\alpha_k t)) dt \\
 &= \frac{2}{(J_1(\alpha_k))^2} \frac{1}{\alpha_k} (t J_1(\alpha_k t)) \Big|_{t=0}^1 \\
 &= \frac{2}{(J_1(\alpha_k))^2} \frac{1}{\alpha_k} J_1(\alpha_k) = \frac{2}{\alpha_k J_1(\alpha_k)},
 \end{aligned}$$

donde obtemos

$$f(t) = 2 \sum_{k=1}^{\infty} \frac{J_0(\alpha_k t)}{\alpha_k J_1(\alpha_k)}. \tag{17}$$

Podemos considerar alguns termos da série em (17) para gerar um gráfico ilustrativo do comportamento da série. A Figura 1 mostra exemplos com 10, 50 e 200 termos.

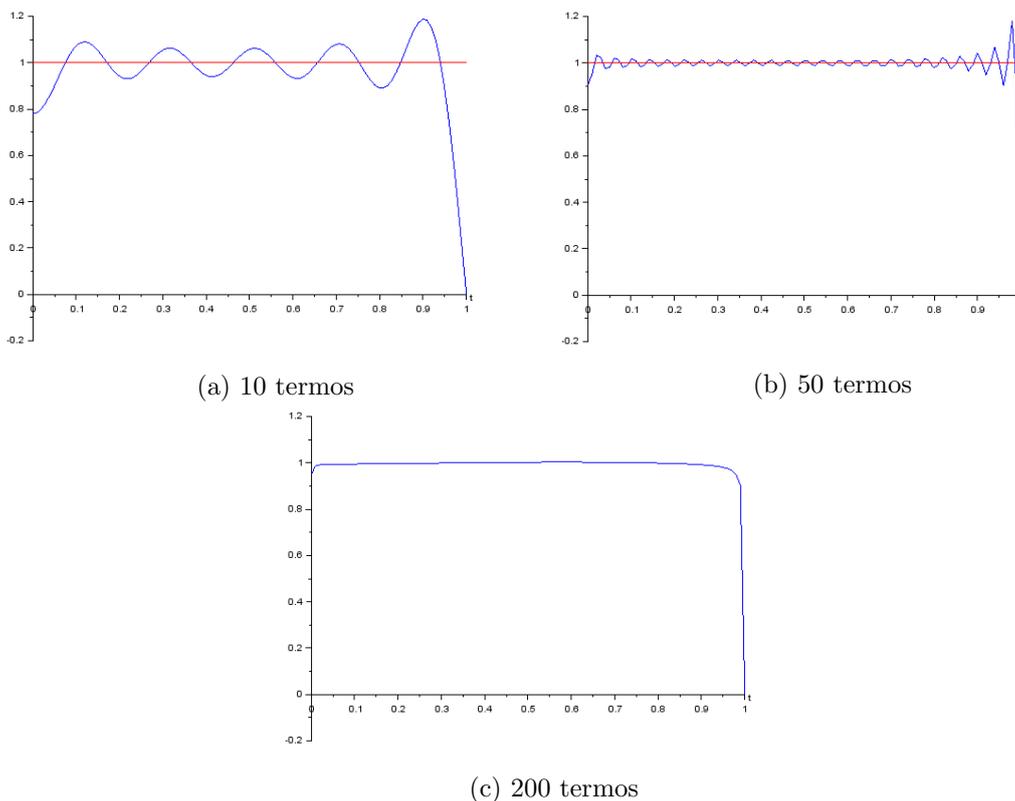


Figura 1: Representações gráficas para a expansão de $f(t) = 1$.

Observemos que para calcular os k_0 primeiros termos da série, precisamos das raízes α_k da função J_n para $k = 1, 2, \dots, k_0$.

Como segundo exemplo, vamos escrever a função $f(t) = t$ no intervalo $(0, 1)$ como uma expansão em série de Fourier-Bessel de ordem 1. Queremos então escrever

$$f(t) = \sum_{k=1}^{\infty} C_k J_1(\alpha_k t).$$

Calculando os coeficientes, temos

$$C_k = \frac{2}{(J_2(\alpha_k))^2} \int_0^1 t^2 J_1(\alpha_k t) dt,$$

e usando novamente a identidade (8) obtemos

$$\begin{aligned} C_k &= \frac{2}{(J_2(\alpha_k))^2} \frac{1}{\alpha_k} \int_0^1 \alpha_k t^2 J_1(\alpha_k t) dt \\ &= \frac{2}{(J_2(\alpha_k))^2} \frac{1}{\alpha_k} \int_0^1 \frac{d}{dt} (t^2 J_2(\alpha_k t)) dt \\ &= \frac{2}{(J_2(\alpha_k))^2} \frac{1}{\alpha_k} (t^2 J_2(\alpha_k t)) \Big|_{t=0}^1 \\ &= \frac{2}{(J_2(\alpha_k))^2} \frac{1}{\alpha_k} J_2(\alpha_k) = \frac{2}{\alpha_k J_2(\alpha_k)}, \end{aligned}$$

donde obtemos

$$f(t) = 2 \sum_{k=1}^{\infty} \frac{J_1(\alpha_k t)}{\alpha_k J_2(\alpha_k)}. \quad (18)$$

A Figura 2 mostra aproximações da série com 10, 50 e 100 termos.

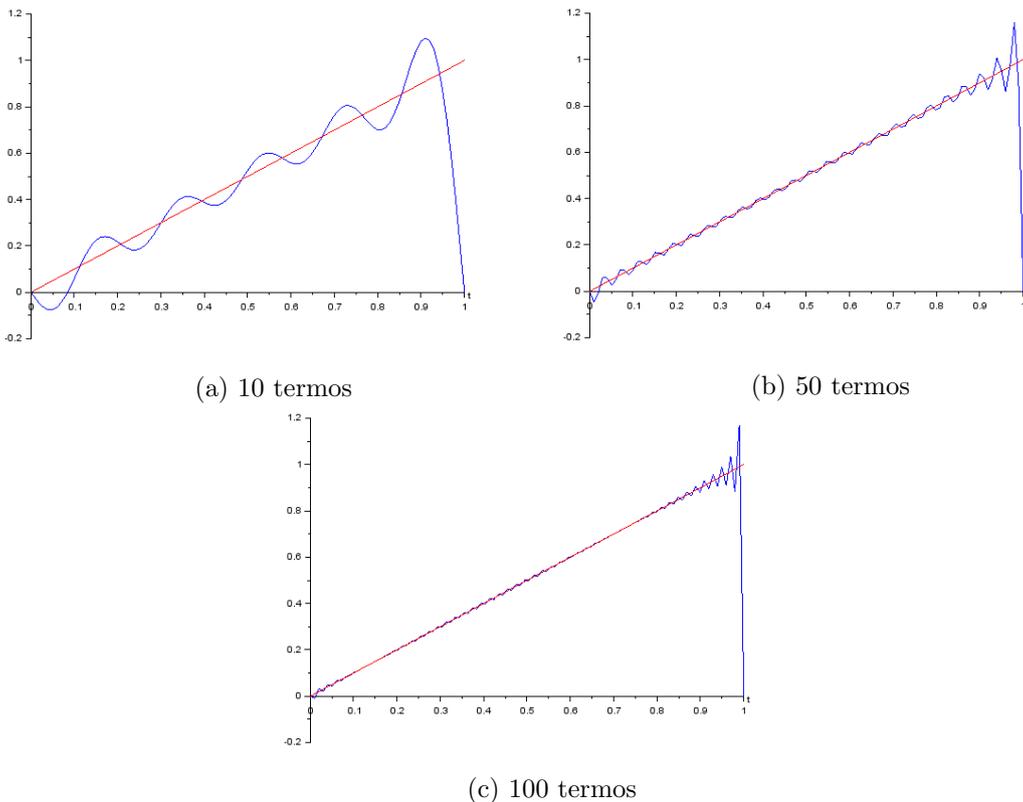


Figura 2: Representações gráficas para a expansão de $f(t) = t$.

Notemos que nestes dois exemplos apresentados, os gráficos da série de Fourier-Bessel convergem para o gráfico da função original. Mas próximo de $t = 1$ elas decrescem até atingir a imagem igual a zero em $t = 1$. Esse comportamento se assemelha ao fenômeno de Gibbs que acontece próximo de pontos de descontinuidade nas aproximações de funções por série de

Fourier. No caso das séries de Fourier-Bessel, esse comportamento ocorre pois em $t = 1$ todos os termos da série se anulam já que α_k são as raízes de J_n .

Assim, a série de Fourier-Bessel converge para a função original, mas em $t = 1$ a série é sempre igual a zero devido à condição de contorno com a qual estamos trabalhando. Dessa forma, próxima de $t = 1$ a função apresenta uma grande oscilação e depois um rápido decrescimento pois as funções J_n são contínuas e por isso essa transição é feita de modo contínuo.

De maneira geral, estes dois exemplos apresentados podem ser generalizados. Consideremos $f(t) = t^m$ para $m \in \mathbb{N}$ fixado, e queremos escrever

$$f(t) = t^m = \sum_{k=1}^{\infty} C_k J_m(\alpha_k t),$$

em que α_k são as (infinitas) raízes positivas da função J_m .

Procedendo de modo análogo aos dois exemplos anteriores, temos os coeficientes C_k dados por

$$C_k = \frac{2}{(J_{m+1}(\alpha_k))^2} \int_0^1 t^{m+1} J_m(\alpha_k t) dt,$$

isto é,

$$C_k = \frac{2}{\alpha_k J_{m+1}(\alpha_k)},$$

e assim, temos que

$$t^m = f(t) = 2 \sum_{k=1}^{\infty} \frac{J_m(\alpha_k t)}{\alpha_k J_{m+1}(\alpha_k)}.$$

Observe que para a função $f(t) = t^m$ escolhemos a Função de Bessel J_m . Isto porque os coeficientes podem ser determinados de forma analítica. Nada impediria escolhermos outra função J_n com índice n diferente da potência m , mas teríamos daí que recorrer a métodos numéricos para calcular os coeficientes.

Considerações finais

Com base no exposto neste trabalho, conseguimos boas aproximações para as funções do tipo t^n no intervalo $(0, 1)$ utilizando séries de Fourier-Bessel de ordem n . Ao estudarmos as propriedades das funções de Bessel de primeira espécie, em especial a ortogonalidade existente na família dessas funções, vimos que essa ortogonalidade se dá em relação à função t e que depende das raízes da equação $J_n(t) = 0$. Esta última restrição, que chamamos de condição de contorno, nos permite encontrar aproximações, ou expansões, de funções em série de Fourier-Bessel no intervalo $(0, 1)$. No entanto, é possível trabalhar com intervalos maiores mudando a condição de contorno envolvida. Sendo assim, um estudo futuro pode envolver expansões de outras funções em série de Fourier-Bessel e também em intervalos maiores.

Agradecimentos

Os autores agradecem às professoras Sandra Maria Tieppo e Simone Aparecida Miloca, pelas contribuições feitas em um trabalho anterior que originou este texto.

Referências

Bowman, Frank. *Introduction to Bessel Functions*. New York: Dover Publications, 2010.

Okawa, Cintya A. *Um estudo sobre a série de Fourier-Bessel*. Monografia (Graduação em Matemática) - Universidade Estadual do Oeste do Paraná - UNIOESTE. Cascavel - Pr, p. 63. 2021.

Watson, G. N. *Theory of Bessel function*. London: Cambridge University Press, 1922.

Zill, Dennis G. *Equações diferenciais com aplicações em modelagem*. 3ª edição. São Paulo: Cengage Learning, 2016.

Créditos

Este volume foi gerado em código \LaTeX , editado no TeXstudio² versão 4.1.2 (64 bit, Qt5), e compilado pelo MiKTeX³ versão 2.9.6000 (32 bit). Para esta edição foram utilizados os seguintes pacotes \LaTeX :

This issue was created in \LaTeX code, edited on TeXstudio² version 4.1.2 (64 bit, Qt5), and compiled by MiKTeX³ version 2.9.6000 (32 bit). This issue uses the following \LaTeX packages:

<code>amssymb</code>	<code>babel</code>	<code>indentfirst</code>
<code>amsfonts</code>	<code>inputenc</code>	<code>enumerate</code>
<code>amsmath</code>	<code>multicol</code>	<code>fancyhdr</code>
<code>latexsym</code>	<code>color</code>	<code>subcaption</code>
<code>amsthm</code>	<code>float</code>	<code>algorithm2e</code>
<code>makeidx</code>	<code>graphicx</code>	

²<https://www.texstudio.org/>

³<https://miktex.org/>

Índice de autores

Alex Augusto Nunes Machado, 23

Cintya Akemi Okawa, 51

Elisangela Danielli de Lima, 39

Felipe Costa, 9

Sandro Marcos Guzzo, 39, 51

Simone Aparecida Miloca, 23