UNIVERSIDADE ESTADUAL PAULISTA
Instituto de Biociências, Letras e Ciências Exatas
DEPARTAMENTO DE CIÊNCIAS DE COMPUTAÇÃO E ESTATÍSTICA

Interpretação eletrostática e zeros de polinômios

Alessandro Santana Martins

Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada
MAP - 092

Rua Cristovão Colombo, 2265
15054-000 - São José do Rio Preto - SP - Brasil
Telefone: (017) 221-2444
Fax: (017) 221-2445

Interpretação eletrostática e zeros de polinômios

Alessandro Santana Martins

Dissertação apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de São José do Rio Preto, São Paulo, para a obtenção do título de Mestre em Matemática Aplicada.

Orientadora: Prof.ª Dr.ª Eliana Xavier Linhares de Andrade

São José do Rio Preto

25 de fevereiro de 2005

Martins, Alessandro Santana.
Interpretação eletrostática e zeros de polinômios /
Alessandro Santana Martins. - São José do Rio Preto : [s.n.], 2005
78 f. : 30 cm.

Orientador: Eliana Xavier Linhares de Andrade
Dissertação (mestrado) - Universidade Estadual Paulista.
Instituto de Biociências, Letras e Ciências Exatas

1. Polinômios ortogonais. 2. Zeros de Polinômios.
3. Interpretação eletrostática.
I. Andrade, Eliana Xavier Linhares de. II. Universidade
Estadual Paulista. Instituto de Biociências, Letras e
Ciências Exatas. III. Título.

CDU - 517.587

Aos meus pais, Geraldo e Candida
e aos meus irmãos, Edemilson e Mario Marcos,
ofereço.

À minha esposa, Ana Luiza
e às minhas filhas Ingrid e Emily,
dedico.

Agradecimentos

Primeiramente, devo agradecer a Deus por ter me dado a vida e oportunidade de poder conviver ao lado de pessoas maravilhosas.

Aos meus pais Geraldo e Candida, que sempre me apoiaram e me incentivaram em tudo que decidi realizar e que enchem minha vida de muito amor, carinho e bons exemplos.

À minha esposa Ana Luiza, querida companheira de todas as horas, a quem pude confiar todas as minhas preocupações, pelo carinho, atenção, paciência, compreensão e que foi uma das principais fontes de incentivo.

Às minhas filhas Ingrid e Emily, que são a cara do pai e me trazem muitas felicidades.

Agradecimentos especiais à minha madrinha Marilda, padrinho Mario Lucio, Tia Rosa e primo Maykeel por terem me recepcionado tão bem em São José do Rio Preto me fazendo sentir-me em casa por todo meu mestrado.

A todos os meus familiares, padrinhos e madrinas por compreenderem a minha ausência constante. Em especial, aos meus irmãos Edemilson e Mario Marcos, Estela, Victor, Augusto, Sonia, Fatima, Luiz, Karina, Ilza, Euripedes e Alessandra.

Aos amigos da Credicofrul que souberam entender a minha saída, mas que sempre estarão comigo onde eu estiver.

Um agradecimento especial à Profa. Dra. Eliana Xavier Linhares de Andrade que tornou possível a realização deste e de vários outros trabalhos, pela dedicação e paciência que sempre teve comigo no mestrado e, acima de tudo, pela amizade e companheirismo.

À Profa. Dra. Cleonice Fátima Bracciali, pelo incentivo e confiança na minha pessoa.

Ao Prof. Dr. Alagacone Sri Ranga pelo apoio, companheirismo e por ter contribuído em minha formação acadêmica.

Ao Prof. Dr. Dimitar Kolev Dimitrov, pela atenção, apoio e preocupação que sempre teve comigo.

A todos os professores e funcionários que de alguma forma contribuíram para a realização deste trabalho. Em especial aos funcionários do DCCE, Getúlio, Sergio, Luiza, Olga e Sandra.

Aos meus amigos Celso, Alcides, Roberto, Delcimar, Ricardo, Flavier e outros que sempre me acompanharam nos momentos de alegria e dificuldades.

A todos os meus amigos de Pós-Graduação, em especial, Daniel, Feltrim, Fernando, José Renato, Claudia, Karina, André, Fabiano, Paulo, Orestes e outros.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico - CNPq - Brasil, pelo auxílio financeiro.

“ Tentar e falhar é, pelo menos, aprender. Não chegar
a tentar é sofrer a inestimável perda
do que poderia ter sido. ”

Geraldo Eustáquio

Resumo

O principal objetivo deste trabalho é estudar um problema de eletrostática geral que envolve ambos, um campo externo e restrições sobre cargas livres. Foram fornecidas condições necessárias e suficientes para o mínimo da energia em termos de soluções polinomiais de uma equação diferencial de Lamé modificada. Além disso, foram dadas novas demonstrações, mais simples, de resultados clássicos de Stieltjes e Szegő. Finalmente, foi obtida uma interpretação eletrostática para os zeros dos polinômios comumente chamados de Hermite-Laurent.

Palavras-chave: Polinômios ortogonais, zeros de polinômios, interpretação eletrostática.

Abstract

A general electrostatic problem which involves both an external field and restrictions on the free charges is studied. Necessary and sufficient conditions for the minimum of the energy are furnished in terms of polynomial solutions of a modified Lamé differential equation. New simplified proofs of classical results of Sitieltjes and Szegő are given. An electrostatic interpretation of the so-called Hermite-Laurent polynomials is obtained.

Keywords: Orthogonal polynomials, zeros of polynomials, electrostatic interpretation.

Sumário

Capítulo 1

Introdução

A teoria de polinômios ortogonais tem vasta aplicação em todos os tipos de problemas da Matemática Pura e das Ciências Aplicadas. Esses polinômios são ferramentas essenciais para a solução de muitos problemas e vêm contribuindo nos estudos relacionados a Equações Diferenciais, Frações Contínuas, Estabilidade Numérica, Algoritmos Rápidos e Super-rápidos, com aplicações que abrangem da Teoria dos Números à Teoria da Aproximação, da Combinatória à Representação de Grupos, da Mecânica Quântica à Física Estatística e da Teoria de Sistemas a Processamento de Sinais.

Historicamente, a primeira contribuição que revelou a importância dos polinômios ortogonais foi o resultado de 1812 de Gauss, que afirma que a única fórmula de quadratura da forma

$$\int_a^b f(x)\mu(x)dx \approx \sum_{k=1}^{n} A_k f(x_k),$$

que é exata para polinômios de grau $2n-1$, tem, como nós $x_1, x_2, \ldots, x_n$, os zeros do polinômio de grau n que é ortogonal em (a,b) com relação à função peso $\mu(x)$. Portanto, nesses últimos dois séculos, a localização precisa dos zeros desses polinômios tem desafiado a curiosidade de vários célebres matemáticos.

Um outro motivo para o interesse nos zeros dos polinômios ortogonais clássicos é que eles admitem uma bela interpretação eletrostática. Descreveremos brevemente esta interpretação para os zeros do polinômio de Jacobi, $P_n^{(\alpha,\beta)}(x)$. Lembremos que os polinômios de

Jacobi são ortogonais em $[-1,1]$ com relação à função peso $(1-x)^{\alpha}(1+x)^{\beta}$.

Consideremos o seguinte campo elétrico no intervalo $[-1,1]$. Sejam $\alpha, \beta > -1$ e tomemos duas cargas fixas com forças $(\alpha+1)/2$ e $(\beta+1)/2$ localizadas em 1 e -1, respectivamente. Suponhamos que existam n cargas unitárias livres localizadas no intervalo $(-1,1)$. Consideremos o campo elétrico que obedece à lei do potencial logarítmico. Do ponto de vista da eletrostática, isto significa que as cargas estão distribuídas ao longo de fios infinitos, perpendiculares à reta real. Portanto, se as cargas livres estão localizadas em $x_1, \ldots, x_n$, a energia do campo é dada por

$$L_1(x_1,\ldots,x_n) = \sum_{k=1}^{n} \left((\alpha+1)/2 \log \frac{1}{|1-x_k|} + (\beta+1)/2 \log \frac{1}{|1+x_k|} \right) + \sum_{1 \le i < k \le n} \log \frac{1}{|x_k - x_i|}.$$

A única posição das cargas para a qual a energia atinge o mínimo global é quando $x_1, \ldots, x_n$ coincidem com os zeros de $P_n^{(\alpha,\beta)}(x)$. Esta fascinante interpretação deve-se a Stieltjes [21, 22, 23] que demonstrou que a energia do campo tem mínimo local nos zeros do polinômio de Jacobi de grau n. Szegő [24, Seção 6.7] demonstrou que, de fato, a energia tem um único mínimo, estabelecendo desta forma a estabilidade do equilíbrio.

A beleza deste resultado inspirou Stieltjes a estudar uma generalização natural: dadas $m+1$ cargas positivas r_j fixas em a_j, $0 \le j \le m$, com $a_0 < \ldots < a_m$, determinar todas as posições de equilíbrio possíveis de n cargas livres pertencentes ao intervalo (a_0, a_m). Nesse caso, a energia do campo é dada por

$$L(x_1,\ldots,x_n) = \sum_{j=0}^{m} r_j \sum_{k=1}^{n} \log \frac{1}{|x_k - a_j|} + \sum_{1 \le i < k \le n} \log \frac{1}{|x_k - x_i|}. \tag{1.1}$$

Surpreendentemente, este problema está fortemente relacionado à questão da caracterização das soluções polinomiais da equação diferencial

$$A(x)y''(x) + 2B(x)y'(x) + C(x)y(x) = 0, \tag{1.2}$$

onde $A(x) = (x-a_0)\cdots(x-a_m)$, $B(x)$ e $C(x)$ são polinômios de grau m e $m-1$, respectivamente, e

$$\frac{B(x)}{A(x)} = \sum_{j=0}^{m} \frac{r_j}{x - a_j}. \tag{1.3}$$

A equação (1.2) é conhecida como equação de Lamé na forma algébrica. Heine [9] provou que, dados $A(x)$ e $B(x)$, existem, no máximo, $(n+m-1)!/(n!(m-1)!)$ polinômios $C(x)$ tais que, para cada um deles, (1.2) tem solução polinomial $y(x)$ de grau n. Stieltjes mostrou que há exatamente $(n+m-1)!/(n!(m-1)!)$ polinômios $C(x)$ tais que, para cada um deles, existe uma solução polinomial $y(x)$ de grau n para a equação de Lamé que tem somente zeros reais. Além disso, Stieltjes demonstrou que cada caso corresponde a uma distribuição das n cargas livres nos m intervalos (a_{j-1}, a_j), $j = 1, \ldots, m$. Este resultado é conhecido como Teorema de Heine-Stieltjes [24, Teorema 6.8]. O polinômio $C(x)$ é chamado de Van Vleck e o correspondente polinômio $y(x)$, que é solução de (1.2), é chamado polinômio de Stieltjes. Van Vleck [27] foi o primeiro a demonstrar que os zeros de $C(x)$ pertencem ao intervalo (a_0, a_m). Devido a sua beleza, o problema da interpretação eletrostática de zeros de polinômios continuou a desafiar a atenção de matemáticos famosos. Klein [13], Bôcher [4] e Pólya [17] provaram que, se $a_j \in \mathbb{C}$, $j = 0, \ldots, m$, os zeros dos polinômios de Stieltjes pertencem ao fecho convexo de $a_0, \ldots, a_m$ desde que os coeficientes r_j sejam positivos. Marden [15] mostrou este resultado sob exigências mais fracas. Ele permitiu que os coeficientes r_j fossem números complexos com partes reais positivas. Alam [1] estendeu o resultado de Marden. O Capítulo 2.9 da monografia de Marden [16] contém informações mais detalhadas sobre a equação de Lamé e nos referimos aos artigos [26, 25] para a sua conexão com a eletrostática dos zeros de polinômios ortogonais. Em um artigo recente, Grünbaum [7] forneceu uma interpretação eletrostática para os zeros dos polinômios de Koornwinder-Krall e Ismail [10] mostrou que os zeros de uma classe geral de polinômios ortogonais são pontos de equilíbrio único em um campo eletrostático e calculou a energia deste campo. Grünbaum [8] estabeleceu um resultado sobre a eletrostática dos zeros de polinômios obtidos dos de Jacobi através de aplicações repetidas da transformada de Darboux.

É muito interessante observar que, apesar deste grande número de contribuições de famosos matemáticos, o problema eletrostático que envolve, ao mesmo tempo, cargas fixas negativas e positivas não foi tratado na literatura. Em particular, a unicidade do polinômio de Van Vleck (ou a unicidade da posição de equilíbrio) nunca foi obtida neste caso. Cargas negativas apareceram, pela primeira vez, no artigo de Grünbaum [7], artigo este que trata de um problema que envolve duas cargas negativas entre duas positivas fixas. Infelizmente, ele não demonstra a unicidade e, de fato, o resultado até pode ser considerado errado visto que as posições finais das cargas livres dependem de suas posições iniciais.

Dimitrov e Van Assche [6] foram os primeiros a demonstrar a unicidade dos polinômios de Van Vleck que envolvem, ao mesmo tempo, coeficientes r_j negativos e positivos, o que implica na existência e unicidade da solução $y(x)$ da equação diferencial (1.2), ou seja, do polinômio de Stieltjes. Este resultado é equivalente à existência e unicidade do mínimo da energia na presença de cargas fixas negativas e positivas simultaneamente.

O resultado de Dimitrov e Van Assche [6] pode ser formulado como segue:

Teorema 1.1 *Sejam* $A(x) = (x - a_0)(x - a_1)(x - a_2)(x - a_3), \quad a_0 < \ldots < a_3$, *um polinômio de grau quatro e* $B(x)$ *um polinômio cúbico para as quais os coeficientes* r_{j-1} *e* r_j, *na decomposição em frações parciais* (1.3), *são positivos e os dois coeficientes restantes são negativos. Se*

a) a seqüência r_0, r_1, r_2, r_3 *admite somente uma mudança de sinal, isto é, se* $j = 1$ *ou* $j = 3$,

ou

b) para $n > 1 - (r_0 + r_1 + r_2 + r_3)$, *a seqüência* r_0, r_1, r_2, r_3 *admite duas mudanças de sinal, isto é, se* $r_0 < 0$, $r_1 > 0$, $r_2 > 0$ *e* $r_3 < 0$,

então existe um único par $(C(x), y(x))$, *com* $C(x)$ *um polinômio de grau dois e* $y(x) = (x - x_1) \cdots (x - x_n)$ *uma solução de* (1.2), *tal que* $a_{j-1} < x_1 < \ldots < x_n < a_j$.

Um outro resultado demonstrado por eles, e que segue imediatamente do teorema acima, é que os zeros dos polinômios de Gegenbauer-Laurent são os únicos pontos de equilíbrio do campo gerado por duas cargas positivas de mesmo valor $(2\lambda + 1)/4$ nos pontos a e b, $0 < a < b$, e cargas negativas $-1/2$ no ponto $-\sqrt{ab}$ e $-(n + \lambda - 1)/2$ na origem.

Estimulados pelo trabalho de Dimitrov e Van Assche [6], nesta dissertação discutiremos o problema da interpretação eletrostática do campo gerado, tanto por cargas positivas como negativas, porém na presença de restrições.

Uma exigência natural para a localização das cargas livres é que elas pertençam ao simplex

$$\Xi := \bigcap_{k=1}^{s} \{a_{j_k} < x_{\mu_k} < \ldots x_{\mu_{k+1}-1} < a_{j_k+1} : r_{j_k}, r_{j_k+1} > 0\},$$

onde $1 = \mu_1 \leq \mu_2 \leq \ldots \mu_s \leq \mu_{s+1} - 1 = n$ e s é o número de intervalos $[a_{j_k}, a_{j_k+1}]$ com cargas positivas r_{j_k} e r_{j_k+1} nos pontos extremos.

Como observaremos mais adiante, existem casos em que o campo eletrostático não é gerado por energia externa, isto é, não há cargas fixas. Nestes casos, por Ξ entenderemos o conjunto

$$-\infty < x_1 < x_2 < \ldots < x_n < \infty.$$

Em várias situações, as cargas livres estão sujeitas a certas restrições da forma $R(\boldsymbol{x}) = R(x_1, \ldots, x_n) = 0$, que definem um compacto Ω em $\mathbb{R}^n$. Uma exigência natural para Ω e que iremos adotar daqui em diante, é que $\Xi \cap \Omega \neq$. Neste trabalho, consideraremos funções $R(\boldsymbol{x})$ que satisfazem às seguintes propriedades:

1. $R(\boldsymbol{x})$ é uma função simétrica com relação às variáveis $x_1, \ldots, x_n$, isto é, $R(x_1, \ldots, x_n) = R(x_{j_1}, \ldots, x_{j_n})$, onde $(j_1, \ldots, j_n)$ é uma permutação qualquer de $\{1, \ldots, n\}$;
2. $\dfrac{\partial R(\boldsymbol{x})}{\partial x_k}$ não depende de $(x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n)$, ou seja, $\dfrac{\partial^2 R(\boldsymbol{x})}{\partial x_i \partial x_j} = 0$, $\forall i \neq j$;
3. $D(x_k) = A(x_k)\dfrac{\partial R(\boldsymbol{x})}{\partial x_k}$ é um polinômio de grau p em x_k.

Consideremos, agora, o campo eletrostático, que obedece à lei do potencial logarítmico, descrito a seguir. Cargas r_j são distribuídas uniformemente ao longo de retas perpendiculares ao eixo real e que a interceptam nos pontos a_j, $j = 0, 1, \ldots, m$. Além disso, n cargas unitárias móveis são, também, uniformemente distribuídas ao longo de retas perpendiculares ao eixo real que passam pelos pontos $x_1, \ldots, x_n$ sujeitos à restrição $R(x_1, \ldots, x_n) = 0$.

Nossa contribuição pode ser dada pelo seguinte resultado:

Teorema 1.2 *Considere o campo elétrico descrito acima com a energia definida por* (1.1). *Seja* $\boldsymbol{x} = (x_1, \ldots, x_n)$, *tal que* $R(\boldsymbol{x}) = 0$, *um ponto de mínimo local de* $L(\boldsymbol{x})$ *em* Ξ. *Então,* *existem um polinômio* $C(x)$, *cujo grau não excede* $\max\{m-1, p-1\}$, *e uma constante* ρ, *tais que o polinômio* $y(x) = (x-x_1)\cdots(x-x_n)$ *é solução da equação de Lamé modificada*

$$A(x)y''(x) + 2\{B(x) - \rho D(x)\}y'(x) + C(x)y(x) = 0. \tag{1.4}$$

Reciprocamente: a) Sejam um polinômio $C(x)$ *de grau menor ou igual a* $\max\{m-1, p-1\}$ *e uma constante* ρ *tais que* (1.4) *possui uma solução polinomial* $y(x) = (x-x_1)\cdots(x-x_n)$, *cujos zeros pertencem a* Ξ *e satisfazem* $R(\boldsymbol{x}) = 0$. *Então,*

$$\frac{\partial L(\boldsymbol{x})}{\partial x_k} - \rho \frac{\partial R(\boldsymbol{x})}{\partial x_k} = 0, \qquad k = 1, 2, \ldots, n,$$

onde $\boldsymbol{x} = (x_1, \ldots, x_n)$ *é o vetor formado pelos zeros de* $y(x)$.
b) Existem um polinômio $C(x)$ *e uma constante* ρ *para os quais* (1.4) *possui uma solução polinomial* $y(x) = (x-\eta_1)\cdots(x-\eta_n)$, *cujos zeros são as coordenadas de* $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$, *que é um ponto de mínimo global da energia em* $\Xi \cap \Omega$.

Como conseqüência deste resultado, mostramos que os zeros do chamado polinômio de Hermite-Laurent de grau n são as coordenadas do ponto extremo da energia do campo gerado por duas cargas negativas fixas, $-(2n-3)/4$ na origem e $-1/2$ no ponto $-\beta$, $\beta > 0$, e n cargas unitárias livres nos pontos $x_1, \ldots, x_n$ sujeito à restrição $\dfrac{1}{4\alpha n}\displaystyle\sum_{k=1}^{n}\left(x_k + \frac{\beta^2}{x_k}\right) = N$, $\alpha > 0$. Este estudo será mostrado em detalhes no Capítulo 4. Acreditamos que esta posição de mínimo seja, também, a posição de mínimo global, estabelecendo, com isso, a unicidade do problema.

Organizamos, então, esta dissertação da seguinte forma.

Capítulo 2 - *Resultados preliminares* - Contém pré-requisitos matemáticos que necessitaremos no decorrer do trabalho.

Capítulo 3 - *Eletrostática e zeros de polinômios* - É o principal capítulo de nosso trabalho, onde estudamos a interpretação eletrostática do campo gerado por cargas fixas negativas e positivas. Neste capítulo, também mostramos nossa contribuição para o estudo da interpretação eletrostática na presença de restrições e apresentamos, na última seção, essa interpretação para os zeros dos polinômios ortogonais clássicos.

Capítulo 4 - *Interpretação eletrostática dos zeros dos polinômios de Laurent ortogonais* - Estudamos, neste capítulo, a relação entre os polinômios ortogonais simétricos e os L-polinômios ortogonais. Com isso, chegamos a uma interpretação eletrostática para os zeros desses polinômios.

Finalmente, relacionamos, nas *Referências Bibliográficas*, os livros e artigos por nós consultados e/ou citados.

Capítulo 2

Resultados preliminares

Neste capítulo, apresentamos alguns resultados básicos da Análise Matemática, pré-requisitos essenciais ao desenvolvimento do trabalho. Muitos resultados serão considerados sem demonstração e podem ser encontrados nos textos clássicos sobre o assunto.

2.1 Teoria dos multiplicadores de Lagrange

Nesta seção, abordaremos a teoria dos multiplicadores de Lagrange por ser de grande importância na análise da interpretação eletrostática dos zeros de polinômios. A maioria dos resultados aqui apresentados pode ser encontrada em [3], assim como suas demonstrações.

2.1.1 Derivadas

Definição 2.1 *Seja $f : \mathbb{R}^n \to \mathbb{R}$ uma função. Considere $\boldsymbol{x} \in \mathbb{R}^n$ fixo e a expressão*

$$\lim_{\alpha \to 0} \frac{f(\boldsymbol{x} + \alpha \boldsymbol{e}_i) - f(\boldsymbol{x})}{\alpha},$$

onde $\boldsymbol{e}_i$ é o i-ésimo vetor unitário (todas as coordenadas são 0*, exceto a i-ésima, que é* 1*). Se o limite acima existe, ele é chamado de* ***derivada parcial*** *de f no ponto $\boldsymbol{x}$ com relação à i-ésima coordenada e é denotado por $\dfrac{\partial f(\boldsymbol{x})}{\partial x_i}$ ou $\nabla_i f(\boldsymbol{x})$ (x_i denota a i-ésima coordenada do vetor $\boldsymbol{x}$).*

Se todas as derivadas parciais existem, o gradiente de f em $\boldsymbol{x}$ é definido como

$$\nabla f(\boldsymbol{x}) = \begin{pmatrix} \dfrac{\partial f(\boldsymbol{x})}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x})}{\partial x_n} \end{pmatrix}.$$

Definição 2.2 *Seja $\boldsymbol{y} \in \mathbb{R}^n$. Definimos a **derivada direcional** de f na direção $\boldsymbol{y}$ como*

$$f'(\boldsymbol{x}; \boldsymbol{y}) = \lim_{\alpha \to 0^+} \frac{f(\boldsymbol{x} + \alpha \boldsymbol{y}) - f(\boldsymbol{x})}{\alpha},$$

se esse limite existe.

Das definições acima, podemos concluir se $f'(\boldsymbol{x}; \boldsymbol{e}_i) = -f'(\boldsymbol{x}; -\boldsymbol{e}_i)$, então $f'(\boldsymbol{x}; \boldsymbol{e}_i) = \dfrac{\partial f(\boldsymbol{x})}{\partial x_i}$.

Se todas as derivadas direcionais de f em um vetor $\boldsymbol{x}$ existem e $f'(\boldsymbol{x}; \boldsymbol{y})$ é uma função linear de $\boldsymbol{y}$, dizemos que f é diferenciável em $\boldsymbol{x}$. Este tipo de derivada é também chamada **derivada de Gateaux**. Observe que f é diferenciável em $\boldsymbol{x}$ se, e somente se, o gradiente $\nabla f(\boldsymbol{x})$ existe e satisfaz $\boldsymbol{y}^T \nabla f(\boldsymbol{x}) = f'(\boldsymbol{x}; \boldsymbol{y})$ para algum $\boldsymbol{y} \in \mathbb{R}^n$.

Definição 2.3 *Se f é diferenciável em um conjunto aberto S e o gradiente $\nabla f(\boldsymbol{x})$ é uma função contínua em $\boldsymbol{x}$, então f é continuamente diferenciável em S.*

Tomemos, agora, uma função $f : \mathbb{R}^n \to \mathbb{R}^m$ de valores vetoriais, ou seja, se $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$,

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), \ldots, f_m(\boldsymbol{x})).$$

Definição 2.4 *Uma função $f : \mathbb{R}^n \to \mathbb{R}^m$ é diferenciável (continuamente diferenciável) se cada componente f_i de f é diferenciável (continuamente diferenciável).*

A **matriz gradiente** de f, que denotaremos por $\nabla f(\boldsymbol{x})$, é a matriz $n \times m$ cuja i-ésima coluna é o gradiente de f_i, $\nabla f_i(\boldsymbol{x})$. Assim,

$$\nabla f(\boldsymbol{x}) = (\nabla f_1(\boldsymbol{x}) \cdots \nabla f_m(\boldsymbol{x})) = \begin{pmatrix} \dfrac{\partial f_1(\boldsymbol{x})}{\partial x_1} & \cdots & \dfrac{\partial f_m(\boldsymbol{x})}{\partial x_1} \\ \dfrac{\partial f_1(\boldsymbol{x})}{\partial x_2} & \cdots & \dfrac{\partial f_m(\boldsymbol{x})}{\partial x_2} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_1(\boldsymbol{x})}{\partial x_n} & \cdots & \dfrac{\partial f_m(\boldsymbol{x})}{\partial x_n} \end{pmatrix}.$$

Definição 2.5 *A **matriz Jacobiana** de f é a transposta da matriz gradiente. Assim, o (i,j)-ésimo elemento da matriz Jacobiana é igual à derivada parcial* $\dfrac{\partial f_i(\boldsymbol{x})}{\partial x_j}$.

Suponhamos, agora, que cada uma das derivadas parciais de uma função $f : \mathbb{R}^n \to \mathbb{R}$ é uma função continuamente diferenciável de $\boldsymbol{x} \in \mathbb{R}^n$. Usaremos a notação $\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}$ para indicar a i-ésima derivada parcial de $\dfrac{\partial f(\boldsymbol{x})}{\partial x_j}$ no ponto $\boldsymbol{x} \in \mathbb{R}^n$.

Definição 2.6 *Definimos a **matriz Hessiana** de f, $\nabla^2 f(\boldsymbol{x})$, como a matriz cujo (i,j)-ésimo elemento é igual a* $\dfrac{\partial^2 f(\boldsymbol{x})}{\partial x_i \partial x_j}$, *que denotaremos por* $\nabla^2_{ij} f(\boldsymbol{x})$.

Sejam $f : \mathbb{R}^n \to \mathbb{R}^k$ e $g : \mathbb{R}^k \to \mathbb{R}^m$ funções continuamente diferenciáveis e seja h a sua composta, isto é, $h(\boldsymbol{x}) = g(f(\boldsymbol{x}))$, $\boldsymbol{x} \in \mathbb{R}^n$. Então, derivando pela regra da cadeia, obtemos

$$\nabla h(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) \nabla g(f(\boldsymbol{x})), \qquad \forall\ \boldsymbol{x} \in \mathbb{R}^n.$$

Considere, agora, $f : \mathbb{R}^{m+n} \to \mathbb{R}$ uma função de $(\boldsymbol{x}, \boldsymbol{y})$, onde $\boldsymbol{x} = (x_1, \ldots, x_m) \in \mathbb{R}^m$ e $\boldsymbol{y} = (y_1, \ldots, y_n) \in \mathbb{R}^n$. Escrevemos

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \dfrac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial x_1} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial x_m} \end{pmatrix}, \qquad \nabla_{\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y}) = \begin{pmatrix} \dfrac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial y_1} \\ \vdots \\ \dfrac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial y_n} \end{pmatrix},$$

$$\nabla_{\boldsymbol{xx}} f(\boldsymbol{x}, \boldsymbol{y}) = \left(\frac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial x_i \partial x_j} \right), \qquad \nabla_{\boldsymbol{xy}} f(\boldsymbol{x}, \boldsymbol{y}) = \left(\frac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial x_i \partial y_j} \right) \quad \text{e} \quad \nabla_{\boldsymbol{yy}} f(\boldsymbol{x}, \boldsymbol{y}) = \left(\frac{\partial f(\boldsymbol{x}, \boldsymbol{y})}{\partial y_i \partial y_j} \right).$$

Seja $f : \mathbb{R}^{m+n} \to \mathbb{R}^k$, isto é, $f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_k(\boldsymbol{x}))$, $\boldsymbol{x} \in \mathbb{R}^{m+n}$. Então,

$$\nabla_{\boldsymbol{x}} f(\boldsymbol{x}, \boldsymbol{y}) = (\nabla_{\boldsymbol{x}} f_1(\boldsymbol{x}, \boldsymbol{y}) \ldots \nabla_{\boldsymbol{x}} f_k(\boldsymbol{x}, \boldsymbol{y})) \quad \text{e} \quad \nabla_{\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y}) = (\nabla_{\boldsymbol{y}} f_1(\boldsymbol{x}, \boldsymbol{y}) \ldots \nabla_{\boldsymbol{y}} f_k(\boldsymbol{x}, \boldsymbol{y})).$$

A seguir, daremos alguns exemplos de relações úteis que podem ser demonstradas usando-se a regra da cadeia.

- $\nabla \left(f(A\boldsymbol{x}) \right) = A^T \nabla f(A\boldsymbol{x})$, onde A é uma matriz;

- $\nabla_{\boldsymbol{x}}(f(h(\boldsymbol{x}), \boldsymbol{y})) = \nabla h(\boldsymbol{x}) \nabla_h f(h(\boldsymbol{x}), \boldsymbol{y})$;
- $\nabla_{\boldsymbol{x}}(f(h(\boldsymbol{x}), g(\boldsymbol{x}))) = \nabla h(\boldsymbol{x}) \nabla_h f(h(\boldsymbol{x}), g(\boldsymbol{x})) + \nabla g(\boldsymbol{x}) \nabla_g f(h(\boldsymbol{x}), g(\boldsymbol{x}))$.

Vamos enunciar, agora, alguns teoremas sobre funções diferenciáveis. A demonstração do teorema a seguir pode ser encontrada em [14, pag. 89].

Teorema 2.1 (Teorema do Valor Médio) *Seja $f : \mathbb{R} \to \mathbb{R}$ continuamente diferenciável em um intervalo aberto I. Então, para cada $x, y \in I$, $x < y$, existe um $\xi \in (x, y)$, tal que*

$$f(y) - f(x) = \nabla f(\xi)(y - x).$$

Para enunciarmos o próximo teorema, nececitaremos da seguinte definição.

Definição 2.7 *Para toda norma $||\cdot|| \in \mathbb{R}^n$, $\epsilon > 0$ e $\boldsymbol{x}^* \in \mathbb{R}^n$, considere o conjunto $S = \{\boldsymbol{x} \mid ||\boldsymbol{x} - \boldsymbol{x}^*|| < \epsilon\}$. O conjunto aberto S é chamado* ***esfera aberta centrada em $\boldsymbol{x}^*$****.*

Teorema 2.2 (Expansões de 2ª ordem) *Seja $f : \mathbb{R}^n \to \mathbb{R}$ duas vezes continuamente diferenciável em uma esfera aberta S centrada em $\boldsymbol{x} \in \mathbb{R}^n$. Então, para todo $\boldsymbol{y} \in \mathbb{R}^n$, tal que $\boldsymbol{x} + \boldsymbol{y} \in S$, temos*

i) $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + \boldsymbol{y}^T \nabla f(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{y}^T \left(\int_0^1 \left(\int_0^t \nabla^2 f(\boldsymbol{x} + \tau \boldsymbol{y}) d\tau \right) dt \right) \boldsymbol{y}$;

ii) existe um $\alpha \in [0, 1]$ tal que $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + \boldsymbol{y}^T \nabla f(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{y}^T \nabla^2 f(\boldsymbol{x} + \alpha \boldsymbol{y}) \boldsymbol{y}$;

iii) $f(\boldsymbol{x}+\boldsymbol{y}) = f(\boldsymbol{x}) + \boldsymbol{y}^T \nabla f(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{y}^T \nabla^2 f(\boldsymbol{x}) \boldsymbol{y} + o(||\boldsymbol{y}||^2)$.

Teorema 2.3 (Teorema da Função Implícita) *Seja $f : \mathbb{R}^{m+n} \to \mathbb{R}^n$ uma função de $\boldsymbol{x} \in \mathbb{R}^n$ e $\boldsymbol{y} \in \mathbb{R}^m$ tal que:*

i) $f(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}}) = 0$;

ii) f é contínua e sua matriz gradiente $\nabla_{\boldsymbol{y}} f(\boldsymbol{x}, \boldsymbol{y})$ é não singular e contínua em um conjunto aberto que contém $(\overline{\boldsymbol{x}}, \overline{\boldsymbol{y}})$.

Então, existem conjuntos abertos $S_{\overline{\boldsymbol{x}}} \subset \mathbb{R}^n$ e $S_{\overline{\boldsymbol{y}}} \subset \mathbb{R}^m$ contendo $\overline{\boldsymbol{x}}$ e $\overline{\boldsymbol{y}}$, respectivamente, e uma função contínua $\phi : S_{\overline{\boldsymbol{x}}} \to S_{\overline{\boldsymbol{y}}}$, tal que $\overline{\boldsymbol{y}} = \phi(\overline{\boldsymbol{x}})$ e $f(\boldsymbol{x}, \phi(\boldsymbol{x})) = 0$ para todo $\boldsymbol{x} \in S_{\overline{\boldsymbol{x}}}$. A função ϕ é única no seguinte sentido: se $\boldsymbol{x} \in S_{\overline{\boldsymbol{x}}}$, $\boldsymbol{y} \in S_{\overline{\boldsymbol{y}}}$ e $f(\boldsymbol{x}, \boldsymbol{y}) = 0$, então $\boldsymbol{y} = \phi(\boldsymbol{x})$. Além disso, se, para algum $p > 0$, f é p vezes continuamente diferenciável, o mesmo é verdade para ϕ e

$$\nabla\phi(\boldsymbol{x}) = -\nabla_{\boldsymbol{x}} f(\boldsymbol{x}, \phi(\boldsymbol{x})) \left(\nabla_{\boldsymbol{y}} f(\boldsymbol{x}, \phi(\boldsymbol{x}))\right)^{-1}, \qquad \forall\ \boldsymbol{x} \in S_{\overline{\boldsymbol{x}}}.$$

Demonstração: Ver [14, pag. 160]. ■

2.1.2 Conjuntos e funções convexas

Definição 2.8 *Seja C um subconjunto do $\mathbb{R}^n$. Dizemos que C é convexo se*

$$\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y} \in C, \qquad \forall\ \boldsymbol{x}, \boldsymbol{y} \in C, \quad \forall\ \alpha \in [0,1].$$

Definição 2.9 *Seja C um subconjunto convexo do $\mathbb{R}^n$. Uma função $f : C \to \mathbb{R}$ é chamada convexa se*

$$f\big(\alpha\boldsymbol{x} + (1-\alpha)\boldsymbol{y}\big) \leq \alpha f(\boldsymbol{x}) + (1-\alpha) f(\boldsymbol{y}), \qquad \forall\ \boldsymbol{x}, \boldsymbol{y} \in C, \quad \forall\ \alpha \in [0,1]. \tag{2.1}$$

A função f é chamada côncava se $-f$ é convexa. A função f é chamada estritamente convexa se a desigualdade acima é estrita para todo $\boldsymbol{x}, \boldsymbol{y} \in C$ com $\boldsymbol{x} \neq \boldsymbol{y}$ e todo $\alpha \in (0,1)$. Para uma função $f : \mathbb{R}^n \to \mathbb{R}$, dizemos também que f é convexa sobre o conjunto convexo C se (2.1) vale.

A seguir, enunciaremos um resultado que nos fornece meios para verificar a convexidade de um conjunto.

Teorema 2.4 *i) Para toda coleção $\{C_i \mid i \in I\}$ de conjuntos convexos, o conjunto intersecção $\bigcap_{i \in I} C_i$ é convexo.*

ii) Se C_1 e C_2 são conjuntos convexos, então o conjunto $\{\boldsymbol{x}_1 + \boldsymbol{x}_2 \mid \boldsymbol{x}_1 \in C_1,\ \boldsymbol{x}_2 \in C_2\}$ é também convexo.

iii) A imagem de um conjunto convexo através de uma transformação linear é convexa.

iv) Se C é um conjunto convexo e $f : C \to \mathbb{R}$ é uma função convexa, os conjuntos $\{\boldsymbol{x} \in C \mid f(\boldsymbol{x}) \leq \alpha\}$ e $\{\boldsymbol{x} \in C \mid f(\boldsymbol{x}) < \alpha\}$ são convexos para todo escalar α.

Veremos, agora, um resultado para reconhecermos funções convexas.

Teorema 2.5 *i) Uma função linear é convexa;*

ii) Qualquer norma de vetor é convexa;

iii) A soma ponderada de funções convexas, com pesos positivos, é convexa;

iv) Se I é um conjunto de índices, $C \subset \mathbb{R}^n$ é um conjunto convexo e $f_i : C \to \mathbb{R}$ é convexa para cada $i \in I$, então a função $h : C \to \mathbb{R} \cup \{\infty\}$, definida por

$$h(\boldsymbol{x}) = \sup_{i \in I} f_i(\boldsymbol{x}),$$

é também convexa.

Para funções diferenciáveis, existe uma caracterização alternativa de convexidade dada pelo seguinte teorema.

Teorema 2.6 *Sejam $C \subset \mathbb{R}^n$ um conjunto convexo e $f : C \to \mathbb{R}$ diferenciável em C.*

i) A função f é convexa se, e somente se,

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + (\boldsymbol{y} - \boldsymbol{x})^T \nabla f(\boldsymbol{x}), \qquad \forall\ \boldsymbol{x}, \boldsymbol{y} \in C. \tag{2.2}$$

ii) Se a desigualdade (2.2) é estrita para $\boldsymbol{x} \neq \boldsymbol{y}$, então f é estritamente convexa.

Para funções convexas duas vezes diferenciáveis, existe a seguinte caracterização de convexidade.

Teorema 2.7 *Sejam $C \subset \mathbb{R}^n$ um conjunto convexo, $f : C \to \mathbb{R}$ duas vezes diferenciável em C e Q uma matriz real simétrica $n \times n$.*

i) *Se $\nabla^2 f(\boldsymbol{x})$ é semi-definida positiva, ou seja, $\boldsymbol{x}^T \nabla^2 f(\boldsymbol{x})\boldsymbol{x} \geq 0$, $\forall\ \boldsymbol{x} \in C$, então f é convexa.*

ii) *Se $\nabla^2 f(\boldsymbol{x})$ é definida positiva, isto é, $\boldsymbol{x}^T \nabla^2 f(\boldsymbol{x})\boldsymbol{x} > 0$, $\forall\ \boldsymbol{x} \in C$, então f é estritamente convexa.*

iii) *Se $C = \mathbb{R}^n$ e f é convexa, então $\nabla^2 f(\boldsymbol{x})$ é semi-definida positiva para todo $\boldsymbol{x} \in C$.*

iv) *A função $f(\boldsymbol{x}) = \boldsymbol{x}^T Q \boldsymbol{x}$ é convexa se, e somente se, Q é semi-definida positiva.*

v) *A função $f(\boldsymbol{x}) = \boldsymbol{x}^T Q \boldsymbol{x}$ é estritamente convexa se, e somente se, Q é definida positiva.*

2.1.3 Mínimos local e global

Sejam $X \subset \mathbb{R}^n$ e $f : X \to \mathbb{R}$ uma função.

Definição 2.10 *Um vetor $\boldsymbol{x}^* \in X$ é um mínimo local de f se existe $\epsilon > 0$ tal que*

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}),$$

para todo $\boldsymbol{x} \in X$ satisfazendo $||\boldsymbol{x} - \boldsymbol{x}^*|| \leq \epsilon$, onde $||\cdot||$ é uma norma de vetor.

Definição 2.11 *Um vetor $\boldsymbol{x}^* \in X$ é um mínimo global de f se*

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}), \qquad \forall\ \boldsymbol{x} \in X.$$

Máximos local e global são definidos analogamente. Em particular, $\boldsymbol{x}^*$ é um máximo local (global) de f se $\boldsymbol{x}^*$ é um mínimo local (global) de $-f$.

Sob hipóteses de convexidade, a distinção entre mínimos local e global é desnecessária como mostra o teorema a seguir.

Teorema 2.8 *Se $C \subset \mathbb{R}^n$ é um conjunto convexo e $f : C \to \mathbb{R}$ é uma função convexa, então o mínimo local de f é também um mínimo global. Se f é estritamente convexa, então existe no máximo um mínimo global.*

Demonstração: Suponha que $\boldsymbol{x}^*$ é um mínimo local de f, mas não um mínimo global. Então, existe algum $\boldsymbol{x} \neq \boldsymbol{x}^*$, tal que $f(\boldsymbol{x}) < f(\boldsymbol{x}^*)$. Usando a desigualdade (2.1), concluímos que $f(\alpha\boldsymbol{x}^* + (1-\alpha)\boldsymbol{x}) < f(\boldsymbol{x}^*)$ para todo $\alpha \in [0,1)$. Isto contradiz à hipótese de que $\boldsymbol{x}^*$ é um mínimo local.

Suponha, agora, que f é estritamente convexa e que existam $\boldsymbol{x}^*$ e $\boldsymbol{x}^{**}$, com $\boldsymbol{x}^* \neq \boldsymbol{x}^{**}$, dois mínimos globais de f. Então, $\boldsymbol{x} = (\boldsymbol{x}^* + \boldsymbol{x}^{**})/2$ deve pertencer a C, já que C é convexo. Além disso, o valor de f deve ser menor em $\boldsymbol{x}$ do que em $\boldsymbol{x}^*$ e $\boldsymbol{x}^{**}$, pois f é estritamente convexa. Como $\boldsymbol{x}^*$ e $\boldsymbol{x}^{**}$ são mínimos globais, obtemos uma contradição. ∎

2.1.4 Condições necessárias para mínimo local

Modelos matemáticos de otimização podem, em geral, ser representados por um conjunto de restrições X e uma função custo $f : X \to \mathbb{R}$. O conjunto X consiste das decisões admissíveis X e o custo $f(\boldsymbol{x})$ é uma medida de qualidade da decisão x.

Queremos encontrar uma decisão ótima, isto é, um $\boldsymbol{x}^* \in X$, tal que

$$f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}), \qquad \forall\ \boldsymbol{x} \in X.$$

Vamos nos restringir ao caso em que cada decisão $\boldsymbol{x}$ é um vetor n-dimensional, isto é, $\boldsymbol{x}$ é uma n-upla de números reais $(x_1, \ldots, x_n)$. Assim, $X \subset \mathbb{R}^n$.

Primeiramente, vamos considerar o problema de otimização sem restrições onde $X = \mathbb{R}^n$, isto é,

$$\begin{aligned} &minimizar \;\; f(\boldsymbol{x}) \\ &sujeito \;\; a \;\; \boldsymbol{x} \in \mathbb{R}^n. \end{aligned}$$

Se a função custo é diferenciável, podemos usar gradientes e expansões em séries de Taylor para comparar o custo de um vetor com o custo de vetores próximos.

Em muitos casos, é importante saber que existe pelo menos um mínimo global de uma função f em um conjunto X. A existência de pelo menos um mínimo global é garantida se f é uma função contínua e X é um subconjunto compacto do $\mathbb{R}^n$. Este é o teorema de Weierstrass que enunciaremos a seguir e cuja demonstração pode ser encontrada em [14, pag. 44].

Teorema 2.9 (Teorema de Weierstrass) *Toda função real contínua $f : X \to \mathbb{R}$ definida num compacto $X \subset \mathbb{R}^n$ atinge seu máximo e seu mínimo em X, isto é, existem pontos $\boldsymbol{x}^*, \boldsymbol{x}^{**} \in X$, tais que $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x}) \leq f(\boldsymbol{x}^{**})$ para todo $\boldsymbol{x} \in X$.*

Forneceremos, agora, condições necessárias para que um ponto seja ponto de mínimo.

Teorema 2.10 *Seja $\boldsymbol{x}^*$ um mínimo local de $f : \mathbb{R}^n \to \mathbb{R}$.*

i) **(Condição necessária de 1ª ordem)** *Se f é continuamente diferenciável em um conjunto aberto S que contém $\boldsymbol{x}^*$, então*

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

ii) **(Condição necessária de 2ª ordem)** Se f é duas vezes continuamente diferenciável em S, então

$$\nabla^2 f(\boldsymbol{x}^*)$$

é semi-definida positiva, $\forall\ \boldsymbol{x} \in \mathbb{R}^n$.

Demonstração: Sejam $\boldsymbol{d} \in \mathbb{R}^n$ fixo e $g : \mathbb{R} \to \mathbb{R}$ uma função definida por: $g(\alpha) = f(\boldsymbol{x}^* + \alpha \boldsymbol{d})$. Usando a regra da cadeia, obtemos

$$0 \leq \lim_{\alpha \to 0^+} \frac{f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) - f(\boldsymbol{x}^*)}{\alpha} = \left. \frac{dg(\alpha)}{d\alpha} \right|_{\alpha=0} = \boldsymbol{d}^T \nabla f(\boldsymbol{x}^*),$$

onde a desigualdade segue da hipótese de $\boldsymbol{x}^*$ ser um mínimo local. Como $\boldsymbol{d}$ é arbitrário, a mesma desigualdade vale quando substituímos $\boldsymbol{d}$ por $-\boldsymbol{d}$. Portanto, $\boldsymbol{d}^T \nabla f(\boldsymbol{x}^*) = 0$ para todo $\boldsymbol{d} \in \mathbb{R}^n$, o que mostra que $\nabla f(\boldsymbol{x}^*) = 0$.

Suponhamos, agora, que f é duas vezes continuamente diferenciável e seja $\boldsymbol{d}$ um vetor qualquer do $\mathbb{R}^n$. Logo, para todo $\alpha \in \mathbb{R}$, a expansão em série de Taylor até 2ª ordem fornece

$$f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) - f(\boldsymbol{x}^*) = \alpha \nabla f(\boldsymbol{x}^*)^T \boldsymbol{d} + \frac{\alpha^2}{2} \boldsymbol{d}^T \nabla f^2(\boldsymbol{x}^*) \boldsymbol{d} + o(\alpha^2).$$

Usando a condição $\nabla f(\boldsymbol{x}^*) = 0$ e a otimalidade local de $\boldsymbol{x}^*$, concluímos que existe um $\epsilon > 0$ suficientemente pequeno, tal que, para todo $\alpha \in (0, \epsilon)$,

$$0 \leq \frac{f(\boldsymbol{x}^* + \alpha \boldsymbol{d}) - f(\boldsymbol{x}^*)}{\alpha^2} = \frac{1}{2} \boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}^*) \boldsymbol{d} + \frac{o(\alpha^2)}{\alpha^2}.$$

Como $\lim_{\alpha \to 0} \dfrac{o(\alpha^2)}{\alpha^2} = 0$, obtemos $\boldsymbol{d}^T \nabla^2 f(\boldsymbol{x}^*) \boldsymbol{d} \geq 0$, mostrando que $\nabla^2 f(\boldsymbol{x}^*)$ é semi-definida positiva. ■

O teorema a seguir mostra o caso em que a função custo é convexa.

Teorema 2.11 *Seja $f : C \to \mathbb{R}$ uma função convexa sobre um conjunto convexo C.*

i) Um mínimo local de f em C é também um mínimo global em C. Se f é estritamente convexa, então existe no máximo um mínimo global de f.

ii) Se f é convexa e o conjunto C é aberto, então $\nabla f(\boldsymbol{x}^*) = 0$ é condição necessária e suficiente para um vetor $\boldsymbol{x}^* \in C$ ser um mínimo global de f em C.

Demonstração: A parte *i)* foi demonstrada no Teorema 2.8. Para mostrar a parte *ii)*, note que, pelo Teorema 2.6,

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \nabla f(\boldsymbol{x}^*)^T (\boldsymbol{x} - \boldsymbol{x}^*), \qquad \forall\ \boldsymbol{x} \in C.$$

Se $\nabla f(\boldsymbol{x}^*) = 0$, obtemos $f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*)$ para todo $\boldsymbol{x} \in C$ e, então, $\boldsymbol{x}^*$ é um mínimo global. ■

2.1.5 Multiplicadores de Lagrange

O conjunto de restrições de um problema de otimização é, normalmente, especificado em termos de restrições com igualdades ou desigualdades. Se levarmos em conta esta estrutura, obtemos uma sofisticada coleção de condições de otimalidade envolvendo algumas variáveis auxiliares chamadas **multiplicadores de Lagrange**. Essas variáveis facilitam a caracterização de soluções ótimas.

Consideraremos, aqui, problemas de otimização que envolvem restrições com igualdades da forma

$$\begin{aligned} &minimizar \quad f(\boldsymbol{x}) \\ &sujeito \ \ a \quad h(\boldsymbol{x}) = 0, \end{aligned}$$

onde as funções $f : \mathbb{R}^n \to \mathbb{R}$ e $h : \mathbb{R}^n \to \mathbb{R}^m$ são continuamente diferenciáveis.

Vamos escrever $h(\boldsymbol{x})$ da forma

$$h(\boldsymbol{x}) = (h_1(\boldsymbol{x}), \dots, h_m(\boldsymbol{x})), \qquad \boldsymbol{x} \in \mathbb{R}^n.$$

O principal resultado sobre multiplicadores de Lagrange é o seguinte.

Teorema 2.12 *Seja $\boldsymbol{x}^*$ um mínimo local de f sujeito a $h(\boldsymbol{x}) = 0$ e suponhamos que os gradientes das restrições, $\nabla h_1(\boldsymbol{x}^*), \dots, \nabla h_m(\boldsymbol{x}^*)$, sejam linearmente independentes. Então, existem únicos escalares $\rho_1^*, \dots, \rho_m^*$, chamados **multiplicadores de Lagrange**, tais que*

$$\nabla f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \rho_i^* \nabla h_i(\boldsymbol{x}^*) = 0. \tag{2.3}$$

Se, além disso, f e h são duas vezes continuamente diferenciáveis, então

$$\boldsymbol{y}^T \left(\nabla^2 f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \rho_i^* \nabla^2 h_i(\boldsymbol{x}^*) \right) \boldsymbol{y} \geq 0, \qquad \forall\ \boldsymbol{y} \in V(\boldsymbol{x}^*), \tag{2.4}$$

onde $V(\boldsymbol{x}^*)$ é o subespaço das variações viáveis de 1ª ordem

$$V(\boldsymbol{x}^*) = \{\boldsymbol{y} \mid \nabla h_i(\boldsymbol{x}^*)^T \boldsymbol{y} = 0, \quad i = 1, \dots, m\}. \tag{2.5}$$

Existem duas formas de interpretar a equação (2.3):

i) O gradiente da função custo, $\nabla f(\boldsymbol{x}^*)$, pertence ao subespaço gerado pelos gradientes das restrições, ou

ii) o gradiente da função custo, $\nabla f(\boldsymbol{x}^*)$, é ortogonal ao subespaço das variações viáveis de 1ª ordem $V(\boldsymbol{x}^*)$. Este é o subespaço das variações $\Delta\boldsymbol{x}$ para os quais o vetor $\boldsymbol{x} = \boldsymbol{x}^* + \Delta\boldsymbol{x}$ satisfaz à restrição $h(\boldsymbol{x}) = 0$ até a 1ª ordem. Assim, de acordo com a condição (2.3) dos multiplicadores de Lagrange, no mínimo local $\boldsymbol{x}^*$, a variação de 1ª ordem do custo $\nabla f(\boldsymbol{x}^*)^T \Delta\boldsymbol{x}$ é zero para todas as variações $\Delta\boldsymbol{x}$ neste subespaço. Isto é análogo à condição $\nabla f(\boldsymbol{x}^*) = 0$ da otimização sem restrições.

Antes de demonstrarmos o Teorema 2.12, daremos um exemplo para o caso de problemas com restrições lineares.

Exemplo 2.1 *Considere o problema*

$$\begin{aligned} &minimizar \quad f(\boldsymbol{x}) \\ &sujeito \;\; a \quad A\boldsymbol{x} = \boldsymbol{b}, \end{aligned} \tag{2.6}$$

onde A *é uma matriz* $m \times n$ *com linhas linearmente independentes e* $\boldsymbol{b} \in \mathbb{R}^m$ *é um vetor dado.*

Reordenando as coordenadas de $\boldsymbol{x}$ se necessário (podemos supor que as primeiras m colunas de A são linearmente independentes), A pode ser particionada da seguinte forma:

$$A = (B \quad R),$$

onde B é uma matriz inversível $m \times m$ e R é uma matriz $m \times (n-m)$. Do mesmo modo, particionamos $\boldsymbol{x}$ como

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_R \end{pmatrix},$$

onde $\boldsymbol{x}_B \in \mathbb{R}^m$ e $\boldsymbol{x}_R \in \mathbb{R}^{n-m}$. Podemos escrever, então, o problema (2.6) como

$$\begin{aligned} &minimizar \quad f(\boldsymbol{x}_B, \boldsymbol{x}_R) \\ &sujeito \;\; a \quad B\boldsymbol{x}_B + R\boldsymbol{x}_R = \boldsymbol{b}. \end{aligned} \tag{2.7}$$

Usando a equação da restrição para expressar $\boldsymbol{x}_B$ em termos de $\boldsymbol{x}_R$, obtemos

$$\boldsymbol{x}_B = B^{-1}(\boldsymbol{b} - R\boldsymbol{x}_R).$$

Substituindo na função f, podemos converter o problema (2.7) no problema de otimização sem restrições

$$\begin{aligned} &minimizar \quad F(\boldsymbol{x}_B) \equiv f(B^{-1}(\boldsymbol{b} - R\boldsymbol{x}_R), \boldsymbol{x}_R) \\ &sujeito \;\; a \quad \boldsymbol{x}_R \in \mathbb{R}^{n-m}. \end{aligned} \tag{2.8}$$

Se $(\boldsymbol{x}_B^*, \boldsymbol{x}_R^*)$ é um mínimo local do problema com restrições (2.6), então $\boldsymbol{x}_R^*$ é um mínimo local da função custo "reduzida" F, ou seja,

$$0 = \nabla F(\boldsymbol{x}_R^*) = -R^T(B^T)^{-1}\nabla_B f(\boldsymbol{x}^*) + \nabla_R f(\boldsymbol{x}^*), \tag{2.9}$$

onde $\nabla_B f$ e $\nabla_R f$ denotam os gradientes de f com relação a $\boldsymbol{x}_B$ e $\boldsymbol{x}_R$, respectivamente. Definindo

$$\boldsymbol{\rho}^* = -(B^T)^{-1}\nabla_B f(\boldsymbol{x}^*), \tag{2.10}$$

as equações (2.9) e (2.10) podem ser dadas, respectivamente, por

$$\nabla_R f(\boldsymbol{x}^*) + R^T\boldsymbol{\rho}^* = 0 \tag{2.11}$$

e

$$\nabla_B f(\boldsymbol{x}^*) + B^T\boldsymbol{\rho}^* = 0. \tag{2.12}$$

De (2.11) e (2.12), obtemos

$$\nabla f(\boldsymbol{x}^*) + A^T\boldsymbol{\rho}^* = 0, \tag{2.13}$$

que é a condição (2.3) dos multiplicadores de Lagrange para o caso do problema com restrições lineares (2.6). Note que o vetor $\boldsymbol{\rho}^*$ que satisfaz esta condição é único, pois as colunas de A^T são linearmente independentes.

Para demonstrar a condição (2.4), vamos mostrar que ela é equivalente à condição necessária de 2ª ordem sem restrições dada por

$$0 \leq \boldsymbol{d}^T\nabla^2 F(\boldsymbol{x}_R^*)\boldsymbol{d}, \qquad \forall\ \boldsymbol{d} \in \mathbb{R}^{n-m}. \tag{2.14}$$

Usando as equações (2.8) e (2.9), obtemos

$$\nabla^2 F(\boldsymbol{x}_R) = \nabla\Big(- R^T(B^T)^{-1}\nabla_B f(B^{-1}(\boldsymbol{b} - R\boldsymbol{x}_R), \boldsymbol{x}_R) + \nabla_R f(B^{-1}(\boldsymbol{b} - R\boldsymbol{x}_R), \boldsymbol{x}_R)\Big). \tag{2.15}$$

Considere a seguinte partição da matriz Hessiana $\nabla^2 f(\boldsymbol{x}^*)$:

$$\nabla^2 f(\boldsymbol{x}^*) = \begin{pmatrix} \nabla^2_{BB} f(\boldsymbol{x}^*) & \nabla^2_{BR} f(\boldsymbol{x}^*) \\ \nabla^2_{RB} f(\boldsymbol{x}^*) & \nabla^2_{RR} f(\boldsymbol{x}^*) \end{pmatrix}.$$

Avaliando (2.15) em $\boldsymbol{x}_R = \boldsymbol{x}_R^*$, obtemos

$$\nabla^2 F(\boldsymbol{x}_R^*) = R^T(B^T)^{-1}\nabla^2_{BB} f(\boldsymbol{x}^*)B^{-1}R - R^T(B^T)^{-1}\nabla^2_{BR} f(\boldsymbol{x}^*) - \nabla^2_{RB} f(\boldsymbol{x}^*)B^{-1}R + \nabla^2_{RR} f(\boldsymbol{x}^*),$$

que, por (2.14), é semi-definida positiva. Como as restrições são lineares, $\nabla^2 h_i(\boldsymbol{x}^*) = 0$.

Logo, para todo $\boldsymbol{d} \in \mathbb{R}^{n-m}$, concluímos que

$$0 \leq \boldsymbol{d}^T \nabla^2 F(\boldsymbol{x}_R^*) \boldsymbol{d} = \boldsymbol{y}^T \nabla^2 f(\boldsymbol{x}^*) \boldsymbol{y} = \boldsymbol{y}^T \left(\nabla^2 f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \rho_i^* \nabla^2 h_i(\boldsymbol{x}^*) \right) \boldsymbol{y}, \tag{2.16}$$

onde $\boldsymbol{y}$ é o vetor

$$\boldsymbol{y} = \begin{pmatrix} -B^{-1} R \boldsymbol{d} \\ \boldsymbol{d} \end{pmatrix}. \tag{2.17}$$

Podemos ver que o subespaço $V(\boldsymbol{x}^*)$ de variações viáveis (2.5) é dado por

$$\begin{aligned} V(\boldsymbol{x}^*) &= \{(\boldsymbol{y}_B, \boldsymbol{y}_R) \mid B\boldsymbol{y}_B + R\boldsymbol{y}_R = 0\} \\ &= \{(\boldsymbol{y}_B, \boldsymbol{y}_R) \mid \boldsymbol{y}_B = -B^{-1} R \boldsymbol{d},\ \boldsymbol{y}_R = \boldsymbol{d},\ \boldsymbol{d} \in \mathbb{R}^{n-m}\} \end{aligned}$$

e, então, a equação (2.16) é equivalente à condição de 2ª ordem (2.4) dos multiplicadores de Lagrange.

Vamos, agora, demonstrar o Teorema dos Multiplicadores de Lagrange generalizando a análise do exemplo anterior.

Demonstração do Teorema 2.12: Vamos supor $m < n$. Se $m = n$, qualquer vetor, inclusive $-\nabla f(\boldsymbol{x}^*)$, pode ser expresso como combinação linear dos vetores linearmente independentes $\nabla h_1(\boldsymbol{x}^*), \ldots, \nabla h_m(\boldsymbol{x}^*)$, demonstrando, deste modo, o teorema.

Reordenando as coordenadas de $\boldsymbol{x}$, se necessário, podemos particionar o vetor $\boldsymbol{x}$ em $\boldsymbol{x} = (\boldsymbol{x}_B, \boldsymbol{x}_R)$, onde a submatriz quadrada $\nabla_B h(\boldsymbol{x}^*)$ (a matriz gradiente de h com relação a $\boldsymbol{x}_B$) é inversível. A restrição

$$h(\boldsymbol{x}_B, \boldsymbol{x}_R) = 0$$

tem solução $(\boldsymbol{x}_B^*, \boldsymbol{x}_R^*)$ e podemos, então, usar o Teorema da Função Implícita para expressar $\boldsymbol{x}_B$ em termos de $\boldsymbol{x}_R$ através de uma única função continuamente diferenciável $\phi : S \to \mathbb{R}^m$, onde S é uma esfera centrada em $\boldsymbol{x}_R^*$ (ϕ é duas vezes continuamente diferenciável se h também é). Em particular, temos $\boldsymbol{x}_B^* = \phi(\boldsymbol{x}_R^*)$, $h(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R) = 0$ para todo $\boldsymbol{x}_R \in S$ e

$$\nabla \phi(\boldsymbol{x}_R) = -\nabla_R h(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R) \left(\nabla_B h(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R)\right)^{-1}, \qquad \forall\ \boldsymbol{x}_R \in S,$$

onde $\nabla_R h$ é a matriz gradiente de h com relação a $\boldsymbol{x}_R$.

Procedemos, agora, como no caso anterior de restrições lineares. Observemos que $\boldsymbol{x}_R^*$ é um mínimo (sem restrição) da função custo "reduzida"

$$F(\boldsymbol{x}_R) = f(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R).$$

Aplicamos, então, as correspondentes condições necessárias de 1ª e 2ª ordens (sem restrições). A condição (2.3) segue repetindo-se os cálculos feitos para a obtenção das equações (2.10) a (2.13), tomando-se

$$B^T = \nabla_B h(\boldsymbol{x}^*), \qquad R^T = \nabla_R h(\boldsymbol{x}^*), \qquad A^T = \nabla h(\boldsymbol{x}^*) \quad \text{e} \quad \boldsymbol{\rho}^* = -(B^T)^{-1}\nabla_B f(\boldsymbol{x}^*).$$

A demonstração da condição (2.4) requer cálculos demorados e lembram o que foi usado na equação (2.16) do exemplo anterior. Em particular, para $\boldsymbol{d} \in \mathbb{R}^{n-m}$, seja $\boldsymbol{y}$ dado por (2.17).

Consideremos, ainda,

$$H_i(\boldsymbol{x}_R) = h_i(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R), \qquad i = 1, \dots, m \tag{2.18}$$

e seja $\phi_i(\boldsymbol{x}_R)$ a i-ésima coordenada de ϕ, onde

$$\phi(\boldsymbol{x}_R) = \begin{pmatrix} \phi_1(\boldsymbol{x}_R) \\ \vdots \\ \phi_m(\boldsymbol{x}_R) \end{pmatrix}.$$

Então, derivando duas vezes a relação $F(\boldsymbol{x}_R) = f(\phi(\boldsymbol{x}_R), \boldsymbol{x}_R)$, obtemos, através de cálculos diretos, que

$$\boldsymbol{d}^T \nabla^2 F(\boldsymbol{x}_R)\boldsymbol{d} = \boldsymbol{y}^T \nabla^2 f(\boldsymbol{x}^*)\boldsymbol{y} + \boldsymbol{d}^T \left(\sum_{j=1}^{m} \nabla^2 \phi_j(\boldsymbol{x}_R^*) \frac{\partial f(\boldsymbol{x}^*)}{\partial x_j} \right) \boldsymbol{d}. \tag{2.19}$$

Analogamente, derivando duas vezes a equação (2.18) e igualando a zero, temos

$$0 = \boldsymbol{d}^T \nabla^2 H_i(\boldsymbol{x}_R)\boldsymbol{d} = \boldsymbol{y}^T \nabla^2 h_i(\boldsymbol{x}^*)\boldsymbol{y} + d^T \left(\sum_{j=1}^{m} \nabla^2 \phi_j(\boldsymbol{x}_R^*) \frac{\partial h_i(\boldsymbol{x}^*)}{\partial x_j} \right) \boldsymbol{d}.$$

Multiplicando a equação acima por ρ_i^* e somando para $i = 1, \ldots, m$, obtemos

$$0 = \sum_{i=1}^{m} \rho_i^* \boldsymbol{y}^T \nabla^2 h_i(\boldsymbol{x}^*) \boldsymbol{y} + \boldsymbol{d}^T \left(\sum_{j=1}^{m} \nabla^2 \phi_j(\boldsymbol{x}_R^*) \sum_{i=1}^{m} \rho_i^* \frac{\partial h_i(\boldsymbol{x}^*)}{\partial x_j} \right) \boldsymbol{d}. \tag{2.20}$$

Somando as equações (2.19) e (2.20) e usando as relações $\boldsymbol{d}^T \nabla^2 F(\boldsymbol{x}_R) \boldsymbol{d} \geq 0$ e

$$\frac{\partial f(\boldsymbol{x}^*)}{\partial x_j} + \sum_{i=1}^{m} \rho_i^* \frac{\partial h_i(\boldsymbol{x}^*)}{\partial x_j} = 0, \qquad j = 1, \ldots, m,$$

concluímos que

$$0 \leq \boldsymbol{y}^T \left(\nabla^2 f(\boldsymbol{x}^*) + \sum_{i=1}^{m} \rho_i^* \nabla^2 h_i(\boldsymbol{x}^*) \right) \boldsymbol{y},$$

para todo $\boldsymbol{y}$ da forma (2.17). Como mostramos no exemplo anterior, $\boldsymbol{y}$ pertence ao subespaço $V(\boldsymbol{x}^*)$ se, e somente se, $\boldsymbol{y}$ é dado por (2.17). Assim, demonstramos a condição de 2ª ordem (2.4) dos multiplicadores de Lagrange. ■

2.1.6 Restrições tipo desigualdades

Neste trabalho, não estudaremos detalhadamente as considerações de primeira ordem para todas as restrições do tipo desigualdade. Entretanto, existem alguns problemas extremos específicos onde, pela estrutura do domínio e a simplicidade da própria função, é possível reduzir um problema extremo com restrições tipo desigualdade a um correspondente que só envolve igualdade. Um desses casos é quando a região considerada é estrelada e a função investigada é homogênea.

Definição 2.12 *O conjunto $U \subset \mathbb{R}$, que contém a origem, é chamado* ***estrelado*** *se, para todo $x \in U$, temos que $tx \in U$ para todo $t \in [0, 1]$.*

Observe que a última exigência é equivalente a $\{y = tx \ : \ 0 \leq t \leq 1\} \subset U$. Vale ainda mencionar que todo conjunto convexo que contém a origem é estrelado.

Definição 2.13 *A função $f : \mathbb{R}^n \to \mathbb{R}$, $f(\boldsymbol{x}) \geq 0$, é homogênea de ordem $\alpha > 0$ se, para todo $\gamma \in \mathbb{R}$, $f(\gamma \boldsymbol{x}) = |\gamma|^\alpha f(\boldsymbol{x})$.*

Lema 2.1 *Seja* $f : U \to \mathbb{R}$ *uma função contínua e homogênea de ordem* α, *onde* U *é um conjunto estrelado e compacto. Então,* $\max_{x \in U} f(x)$ *é atingido em um ponto* $x^* \in \partial U$.

Demonstração: Suponhamos que $\max_{x \in U} f(x)$ seja atingido em um ponto $x^* \notin \partial U$. Como U é um conjunto compacto, ele é limitado. Logo, existe uma constante c, tal que $\bar{x} = cx^* \in \partial U$. Usando o fato de U ser estrelado, concluimos que a constante c é única e, além disso, $c > 1$. Como f é homogênea, então

$$f(\bar{x}) = f(cx^*) = |c|^\alpha f(x^*) > f(x^*).$$

Chegamos, assim, a uma contradição. ∎

2.2 Polinômios ortogonais

Faremos, agora, um breve estudo sobre os polinômios ortogonais. Somente as definições e propriedades mais importantes para o desenvolvimento deste trabalho serão apresentadas.

As aplicações dos polinômios ortogonais associados às chamadas medidas clássicas, como as de Jacobi, Laguerre e Hermite, têm, particularmente, papel fundamental em muitos problemas das ciências e das engenharias. Ao contrário do que gostaríamos, os polinômios ortogonais associados às medidas não clássicas ainda não gozam de semelhante privilégio. Isto se deve, em parte, às restrições encontradas para gerá-los.

Sejam (a, b) um intervalo real, $-\infty \leq a < b \leq \infty$, e $\phi(x)$ uma função real limitada, não decrescente e com infinitos pontos de aumento em (a, b).

Definição 2.14 *Se os momentos definidos por*

$$\mu_r = \int_a^b x^r d\phi(x)$$

existem para $r = 0, 1, 2, \ldots,$ *então* $d\phi(x)$ *é chamada distribuição (medida positiva) em* (a, b).

Se $d\phi(x) = \omega(x)dx$, então $\omega(x) \geq 0$ em (a, b), mas não identicamente nula, é chamada função peso.

Se $d\phi(x)$ é interpretada como uma distribuição de massa sobre a reta real positiva, então os momentos μ_1 e μ_2 correspondem, respectivamente, aos primeiro e segundo momentos da distribuição de massa. Essa nomenclatura tem sua origem na Mecânica.

Definimos o suporte de $d\phi$ (*supp* $(d\phi)$) como sendo o conjunto dos pontos de aumento de ϕ, ou seja,

$$\textit{supp}\ (d\phi) := \Big\{ x \in \mathbb{R} : \phi(x+\varepsilon) - \phi(x-\varepsilon) > 0, \ \ \text{para todo} \ \ \varepsilon > 0 \Big\}.$$

Podemos, agora, definir polinômios ortogonais. Seja $\mathbb{P}_n$ o espaço de todos os polinômios algébricos de grau menor ou igual a n.

Definição 2.15 *Dizemos que uma seqüência de polinômios $\{P_n(x)\}_{n=0}^{\infty}$ é uma seqüência de polinômios ortogonais com relação à distribuição $d\phi(x)$ no intervalo (a,b) se*

$$(i) \ P_n(x) \textit{ é de grau exatamente } \ n, \quad n \neq 0.$$

$$(ii) \ \langle P_n, P_m \rangle = \int_a^b P_n(x) P_m(x) d\phi(x) = \begin{cases} 0, & se \quad n \neq m, \\ \rho_n \neq 0, & se \quad n = m. \end{cases}$$

Definição 2.16 *Dizemos que uma seqüência de polinômios ortogonais é uma seqüência de polinômios ortonormais, denotada por $\{P_n^*(x)\}_{n=0}^{\infty}$, se $\rho_n = 1$.*

Notação: Representaremos os polinômios ortogonais $P_n(x)$ por

$$P_n(x) = \sum_{i=0}^{n} a_{n,i} x^i, \qquad a_{n,n} \neq 0.$$

Se $a_{n,n} = 1$, os polinômios ortogonais são chamados de polinômios ortogonais mônicos e iremos denotá-los por $\widehat{P}_n(x)$.

2.2.1 Algumas propriedades de polinômios ortogonais

Veremos, agora, algumas propriedades importantes dos polinômios ortogonais $P_n(x)$. Um estudo detalhado pode ser encontrado em [5] e [24].

Teorema 2.13 *Sejam $P_0(x), P_1(x), \ldots, P_m(x)$ pertencentes a uma seqüência de polinômios ortogonais. Então, eles são linearmente independentes.*

Demonstração: Sejam c_j, $j = 0, 1, \ldots, m$, constantes tais que $\sum_{j=0}^{m} c_j P_j(x) = 0$. Logo, para cada polinômio $P_k(x)$, $0 \leq k \leq m$, obtemos

$$\left\langle \sum_{j=0}^{m} c_j P_j, P_k \right\rangle = \langle 0, P_k \rangle = 0,$$

ou seja,

$$\sum_{j=0}^{m} c_j \langle P_j, P_k \rangle = 0.$$

Mas, como

$$\sum_{j=0}^{m} c_j \langle P_j, P_k \rangle = c_k \underbrace{\langle P_k, P_k \rangle}_{>0} = 0,$$

obtemos, $c_k = 0$, $k = 0, 1, \ldots, m$. ∎

O teorema acima nos diz que os polinômios ortogonais $P_k(x)$, $k = 0, 1, 2, \ldots, n$, formam uma base para $\mathbb{P}_n$. Como conseqüência, obtemos o seguinte resultado.

Corolário 2.1 *Sejam $\{P_n(x)\}_{n=0}^{\infty}$ e $\{Q_n(x)\}_{n=0}^{\infty}$ duas seqüências de polinômios ortogonais no intervalo (a, b) com relação à distribuição $d\phi(x)$. Então,*

$$P_j(x) = c_j Q_j(x), \qquad j = 0, 1, \ldots,$$

onde c_j é uma constante que depende apenas de j.

Teorema 2.14 *Seja $\{P_n(x)\}_{n=0}^{\infty}$ uma seqüência de polinômios e $d\phi(x)$ uma distribuição em (a, b). Então, as seguintes afirmações são equivalentes:*

(a) $\{P_n(x)\}_{n=0}^{\infty}$ *é uma seqüência de polinômios ortogonais com relação à distribuição $d\phi(x)$ em (a, b);*

(b) $\langle P_n, \pi \rangle = 0$ *para todo polinômio $\pi(x)$ de grau $\leq n - 1$;*

(c) $\langle x^s, P_n \rangle = \int_a^b x^s P_n(x) d\phi(x) \begin{cases} = 0, & se \quad 0 \leq s < n, \\ \neq 0, & se \quad s = n. \end{cases}$

Demonstração: $(a) \Rightarrow (b)$ Seja $\pi(x)$ um polinômio de grau $\leq n-1$. Como $\{P_n(x)\}_{n=0}^{\infty}$ é uma seqüência de polinômios ortogonais, pelo Teorema 2.13, $P_0, P_1, \ldots, P_{n-1}$ formam uma base para $\mathbb{P}_{n-1}$. Assim,

$$\pi(x) = \sum_{k=0}^{n-1} \alpha_k P_k(x)$$

e, então,

$$\langle P_n, \pi \rangle = \sum_{k=0}^{n-1} \alpha_k \langle P_n, P_k \rangle = 0.$$

$(b) \Rightarrow (c)$ Temos, por hipótese, que $\langle P_n, \pi \rangle = 0$ para todo $\pi(x)$ de grau $\leq n-1$. Assim, $\langle P_n, x^s \rangle = 0$ se $s < n$.

Consideremos, agora, $s = n$. Então, $\pi(x) = x^n \in \mathbb{P}_n$. Logo,

$$x^n = \sum_{j=0}^{n} \alpha_j P_j(x), \qquad \text{com} \quad \alpha_n = \frac{1}{a_{n,n}} \neq 0.$$

Assim,

$$\langle P_n, x^n \rangle = \sum_{j=0}^{n} \alpha_j \langle P_n, P_j \rangle = \alpha_n \langle P_n, P_n \rangle \neq 0.$$

$(c) \Rightarrow (a)$ (i) Consideremos $m < n$. Seja $P_m(x) = \displaystyle\sum_{k=0}^{n} a_{m,k} x^k, \qquad a_{m,m} \neq 0.$

Por hipótese,

$$\langle P_m, P_n \rangle = \sum_{k=0}^{m} a_{m,k} \left\langle x^k, P_n \right\rangle = 0.$$

Para $m > n$, a demonstração é análoga.

(ii) Seja $m = n$. Então,

$$\langle P_n, P_n \rangle = \sum_{k=0}^{n} a_{n,k} \left\langle x^k, P_n \right\rangle = a_{n,n} \langle x^n, P_n \rangle \neq 0. \qquad \blacksquare$$

Definição 2.17 *O determinante definido por*

$$H_n = \begin{vmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{vmatrix}, \quad n \geq 0, \tag{2.21}$$

é chamado determinante de Hankel de ordem $n+1$.

Teorema 2.15 *Se os momentos* μ_r, $r = 0, 1, \ldots, 2n$, *existem, o determinante de Hankel* (2.21) *é diferente de zero.*

Demonstração: Consideremos o sistema linear homogêneo

$$\begin{pmatrix} \mu_0 & \mu_1 & \ldots & \mu_n \\ \mu_1 & \mu_2 & \ldots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \ldots & \mu_{2n} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Observe que o determinante da matriz dos coeficientes é H_n. Mostremos que a única solução do sistema linear acima é $b_0 = b_1 = \ldots = b_n = 0$. Dessa forma, temos $H_n \neq 0$.

Substituindo os momentos do sistema linear acima por sua definição, obtemos

$$\left\{ \begin{array}{ccccccccc} b_0 \int_a^b d\phi(x) & + & b_1 \int_a^b x d\phi(x) & + & \cdots & + & b_n \int_a^b x^n d\phi(x) & = & 0 \\ b_0 \int_a^b x d\phi(x) & + & b_1 \int_a^b x^2 d\phi(x) & + & \cdots & + & b_n \int_a^b x^{n+1} d\phi(x) & = & 0 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ b_0 \int_a^b x^n d\phi(x) & + & b_1 \int_a^b x^{n+1} d\phi(x) & + & \cdots & + & b_n \int_a^b x^{2n} d\phi(x) & = & 0 \end{array} \right. .$$

Multiplicando as equações, respectivamente, por $b_0, b_1, \ldots, b_n$ e somando-as, encontramos

$$\begin{aligned} b_0^2 \int_a^b d\phi(x) + b_1^2 \int_a^b x^2 d\phi(x) + \cdots + b_n^2 \int_a^b x^{2n} d\phi(x) + 2b_0 b_1 \int_a^b x d\phi(x) + \cdots \\ + \quad 2b_0 b_n \int_a^b x^n d\phi(x) + 2b_1 b_n \int_a^b x^{n+1} d\phi(x) + \cdots = 0, \end{aligned}$$

ou seja,

$$\int_a^b (b_0 + b_1 x + \cdots + b_n x^n)^2 d\phi(x) = 0.$$

Logo, $Q(x) = \sum_{i=0}^{n} b_i x^i \equiv 0$ para todo $x \in \mathbb{R}$ e, portanto, $b_0 = b_1 = \cdots = b_n = 0$. ■

Teorema 2.16 *Os determinantes de Hankel,* H_n, *são diferentes de zero para* $n = 0, 1, \ldots$ *se, e somente se, existe uma única seqüência de polinômios ortogonais* $\{P_n(x)\}_{n=0}^{\infty}$ *no intervalo* (a, b) *relativamente à distribuição* $d\phi(x)$.

Demonstração: Seja $P_n(x) = \sum_{k=0}^{n} a_{n,k}x^k$, $a_{n,n} \neq 0$. Se $P_n(x)$ pertence a alguma seqüência de polinômios ortogonais, temos que

$$\langle x^m, P_n \rangle = \begin{cases} 0, & \text{se } m < n, \\ \rho_n \neq 0, & \text{se } m = n. \end{cases}$$

Observe que

$$\begin{aligned} \langle x^m, P_n \rangle &= \int_a^b x^m (a_{n,n}x^n + a_{n,n-1}x^{n-1} + \cdots + a_{n,1}x + a_{n,0}) d\phi(x) \\ &= a_{n,n} \int_a^b x^{m+n} d\phi(x) + a_{n,n-1} \int_a^b x^{m+n-1} d\phi(x) + \cdots + a_{n,1} \int_a^b x^{m+1} d\phi(x) \\ &\quad + a_{n,0} \int_a^b x^m d\phi(x) \\ &= a_{n,n}\mu_{m+n} + a_{n,n-1}\mu_{m+n-1} + \cdots + a_{n,1}\mu_{m+1} + a_{n,0}\mu_m. \end{aligned}$$

Logo, fazendo $m = 0, 1, \ldots, n$, obtemos, respectivamente,

$$\begin{cases} a_{n,n}\mu_n & + & \cdots & + & a_{n,1}\mu_1 & + & a_{n,0}\mu_0 & = & 0 \\ a_{n,n}\mu_{n+1} & + & \cdots & + & a_{n,1}\mu_2 & + & a_{n,0}\mu_1 & = & 0 \\ \vdots & & & & \vdots & & \vdots & & \vdots \\ a_{n,n}\mu_{2n} & + & \cdots & + & a_{n,1}\mu_{n+1} & + & a_{n,0}\mu_n & = & \rho_n \neq 0 \end{cases}.$$

Na forma matricial,

$$\begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{pmatrix} \begin{pmatrix} a_{n,0} \\ a_{n,1} \\ \vdots \\ a_{n,n} \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \rho_n \neq 0 \end{pmatrix}.$$

Supondo que existe uma única seqüência de polinômios ortogonais, então existem únicos $a_{n,0}, a_{n,1}, \ldots, a_{n,n}$, com $a_{n,n} \neq 0$, que satisfazem ao sistema linear acima. Portanto, o determinante de Hankel é diferente de zero.

Reciprocamente, se os determinantes de Hankel são diferentes de zero, a solução do sistema acima é única e, além disso, $a_{n,n} \neq 0$ pois $a_{n,n} = \dfrac{\rho_n H_{n-1}}{H_n}$. ■

Uma propriedade bastante importante e útil dos polinômios ortogonais é o seguinte teorema.

Teorema 2.17 (Relação de Recorrência de Três Termos) *Seja* $\{P_n(x)\}_{n=0}^{\infty}$ *uma seqüencia de polinômios ortogonais em* (a,b) *relativamente à distribuição* $d\phi(x)$*. Então,*

$$P_{n+1}(x) = (\gamma_{n+1}x - \beta_{n+1})P_n(x) - \alpha_{n+1}P_{n-1}(x), \quad n \geq 0, \tag{2.22}$$

com $P_{-1}(x) = 0,\ P_0(x) = 1,\ \alpha_{n+1},\ \beta_n,\ \gamma_n \in \mathbb{R},\ n \geq 1$ *e*

$$\gamma_{n+1} = \frac{a_{n+1,n+1}}{a_{n,n}} \neq 0, \quad \beta_{n+1} = \gamma_{n+1}\frac{\langle xP_n, P_n\rangle}{\langle P_n, P_n\rangle}, \quad \alpha_{n+1} = \frac{\gamma_{n+1}}{\gamma_n}\frac{\langle P_n, P_n\rangle}{\langle P_{n-1}, P_{n-1}\rangle} \neq 0.$$

Demonstração: Temos que $P_n(x) = \displaystyle\sum_{i=0}^{n} a_{n,i}x^i$. Como $xP_n(x)$ é um polinômio de grau $n+1$, então

$$xP_n(x) = \sum_{i=0}^{n+1} b_iP_i(x).$$

Igualando os coeficientes dos termos de maior grau em ambos os membros da igualdade anterior, encontramos $b_{n+1} = \dfrac{a_{n,n}}{a_{n+1,n+1}}$. Porém, do Teorema 2.14 (b),

$$\langle xP_n, P_j\rangle = \int_a^b P_n(x)xP_j(x)d\phi(x) = 0, \quad \text{para} \quad j \leq n-2.$$

Mas, para $j \leq n-2$, temos que

$$\langle xP_n, P_j\rangle = \sum_{i=0}^{n+1} b_i \langle P_i, P_j\rangle = b_j \langle P_j, P_j\rangle .$$

Logo, $b_j = 0$ para $j \leq n-2$. Assim,

$$P_{n+1}(x) = \frac{1}{b_{n+1}}xP_n(x) - \frac{b_{n-1}}{b_{n+1}}P_{n-1}(x) - \frac{b_n}{b_{n+1}}P_n(x),$$

ou seja,

$$P_{n+1}(x) = (\gamma_{n+1}x - \beta_{n+1})P_n(x) - \alpha_{n+1}P_{n-1}(x),$$

com $\gamma_{n+1} = \dfrac{1}{b_{n+1}}$, $\quad \beta_{n+1} = \dfrac{b_n}{b_{n+1}} \quad$ e $\quad \alpha_{n+1} = \dfrac{b_{n-1}}{b_{n+1}}$.

Como $b_{n+1} = \dfrac{a_{n,n}}{a_{n+1,n+1}}$, temos que $\gamma_{n+1} = \dfrac{a_{n+1,n+1}}{a_{n,n}} \neq 0$.

De (2.22), calculando o produto interno $\langle P_{n+1}(x), P_n(x)\rangle$, obtemos que

$$\beta_{n+1} = \gamma_{n+1}\frac{\langle xP_n, P_n\rangle}{\langle P_n, P_n\rangle}.$$

De maneira análoga, chegamos que $\alpha_{n+1} = \dfrac{\gamma_{n+1}}{\gamma_n}\dfrac{\langle P_n, P_n\rangle}{\langle P_{n-1}, P_{n-1}\rangle}$. ■

Observe que, para a seqüência de polinômios ortogonais mônicos $\{\widehat{P}_n(x)\}_{n=0}^{\infty}$, $\gamma_{n+1} = 1$ para $n \geq 0$.

Um resultado bem importante sobre os zeros dos polinômios ortogonais é o seguinte.

Teorema 2.18 *Os zeros dos polinômios ortogonais $P_n(x)$, $n \geq 1$, associados à distribuição $d\phi(x)$ no intervalo (a, b) são reais, distintos e pertencem ao intervalo (a, b).*

Demonstração: Suponhamos que $P_n(x)$ não tenha raízes em (a, b). Logo, $P_n(x) > 0$ ou $P_n(x) < 0$ para todo $x \in (a, b)$. Então,

$$\int_a^b P_n(x)d\phi(x) \neq 0.$$

Mas, $\displaystyle\int_a^b P_n(x)d\phi(x) = 0$, que é uma contradição. Assim, existe pelo menos uma raiz real de $P_n(x)$ em (a, b). Suponhamos que $x_{n,1}, x_{n,2}, \ldots, x_{n,r}$ $(r < n)$ são raízes de multiplicidade ímpar de $P_n(x)$ em (a, b). Então,

$$P_n(x) = (x - x_{n,1})(x - x_{n,2}) \cdots (x - x_{n,r})q(x),$$

onde $q(x)$ é um polinômio de grau $(n - r)$ que tem somente raízes complexas ou de multiplicidade par em (a, b). Logo, $q(x)$ não muda de sinal em (a, b). Assim,

$$I = \int_a^b (x - x_{n,1})(x - x_{n,2}) \cdots (x - x_{n,r})P_n(x)d\phi(x) = 0.$$

Porém,

$$I = \int_a^b (x - x_{n,1})^2(x - x_{n,2})^2 \cdots (x - x_{n,r})^2 q(x)d\phi(x) \neq 0,$$

que é uma contradição. Logo, $P_n(x)$ tem $r \geq n$ raízes de multiplicidade ímpar em (a, b). Portanto, $P_n(x)$ tem n raízes simples em (a, b). ■

Vamos denotar por $x_{n,1} < x_{n,2} < \ldots < x_{n,n}$ os zeros de $P_n(x)$ em ordem crescente.

Teorema 2.19 (Teorema da Separação dos Zeros) *Seja $\{P_j(x)\}_{j=0}^{\infty}$ uma seqüência de polinômios ortogonais. Então, entre dois zeros consecutivos de $P_{n+1}(x)$ existe somente um zero de $P_n(x)$, ou seja, $x_{n+1,i} < x_{n,i} < x_{n+1,i+1}, \quad i = 1, 2, \cdots, n$.*

A demonstração deste resultado pode ser encontrada em [2].

Definição 2.18 *Uma função distribuição $\phi(x)$ definida em um intervalo $[-b, b]$ é chamada simétrica se $d\phi(x) = -d\phi(-x)$.*

Uma importante propriedade satisfeita pelos polinômios ortogonais simétricos é o seguinte resultado.

Teorema 2.20 *Seja $\{P_n(x)\}_{n=0}^{\infty}$ uma seqüência de polinômios ortogonais com relação a uma distribuição $d\phi(x)$ em um intervalo simétrico com relação à origem $(-b, b)$. Então, as seguintes afirmações são equivalentes:*

(a) *$d\phi(x)$ é simétrica, isto é, $d\phi(x) = -d\phi(-x)$;*

(b) *$P_n(-x) = (-1)^n P_n(x)$, $n \geq 0$;*

(c) *na fórmula de recorrência, $\beta_{n+1} = 0$, $n \geq 1$.*

Demonstração: $(a) \Rightarrow (b)$ Temos, por hipótese, que $d\phi(x)$ é simétrica. Logo, fazendo $y = -x$, para $m \neq n$ obtemos

$$0 = \int_{-b}^{b} P_n(y)P_m(y)d\phi(y) = \int_{b}^{-b} P_n(-x)P_m(-x)d\phi(-x) = \int_{-b}^{b} P_n(-x)P_m(-x)d\phi(x).$$

Assim, $\{P_n(-x)\}_{n=0}^{\infty}$ é uma seqüência de polinômios ortogonais em $(-b, b)$ com relação à distribuição $d\phi(x)$. Pelo Corolário 2.1, $P_n(-x) = c_n P_n(x)$, onde c_n é uma constante. Comparando os coeficientes dos termos de maior grau, chegamos que $c_n = (-1)^n$.
$(b) \Rightarrow (c)$ Da relação de recorrência (2.22), temos que

$$\beta_{n+1} = \gamma_{n+1} \frac{\langle xP_n, P_n \rangle}{\langle P_n, P_n \rangle}.$$

Mas,

$$\begin{aligned}\langle xP_n, P_n\rangle &= \int_{-b}^{b} x[P_n(x)]^2 d\phi(x) = \int_{b}^{0} (-x)[P_n(-x)]^2 d\phi(-x) + \int_{0}^{b} x[P_n(x)]^2 d\phi(x)\\ &= -\int_{0}^{b} x(-1)^{2n}[P_n(x)]^2 d\phi(x) + \int_{0}^{b} x[P_n(x)]^2 d\phi(x) = 0.\end{aligned}$$

$(c) \Rightarrow (a)$ Por hipótese, temos que $\beta_{n+1} = 0$. Logo,

$$\begin{aligned}0 = \langle xP_n, P_n\rangle &= \int_{-b}^{b} x[P_n(x)]^2 d\phi(x)\\ &= \int_{0}^{b} x[P_n(x)]^2 d\phi(-x) + \int_{0}^{b} x[P_n(x)]^2 d\phi(x).\end{aligned}$$

Portanto,

$$\int_{0}^{b} x[P_n(x)]^2 d\phi(x) = \int_{0}^{b} x[P_n(x)]^2 (-d\phi(-x)),$$

de onde segue o resultado. ■

2.2.2 Polinômios ortogonais clássicos

Nesta subseção, apresentaremos algumas das principais propriedades dos polinômios de Jacobi, incluindo, como caso especial, os polinômios Ultrasféricos ou de Gegenbauer. Abordaremos, também, os polinômios de Laguerre e Hermite.

Polinômios de Jacobi - $P_n^{(\alpha,\beta)}(x)$

São ortogonais no intervalo $[-1, 1]$ com relação à função peso $w(x) = (1-x)^{\alpha}(1+x)^{\beta}$, $\alpha, \beta > -1$.

Os polinômios de Jacobi podem ser definidos, através da Formula de Rodrigues (Szegő [24], p.67), por

$$P_n^{(\alpha,\beta)}(x) = \frac{(-1)^n}{2^n n!}(1-x)^{-\alpha}(1+x)^{-\beta}\frac{d^n}{dx^n}\Big\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\Big\}. \qquad (2.23)$$

Aplicando a conhecida regra de Leibnitz para calcularmos $\dfrac{d^n}{dx^n}\Big\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\Big\}$, obtemos

$$\begin{aligned}\frac{d^n}{dx^n}\Big\{(1-x)^{n+\alpha}(1+x)^{n+\beta}\Big\} &= \sum_{k=0}^{n}\binom{n}{k}\frac{d^{n-k}}{dx^{n-k}}(1-x)^{n+\alpha}\frac{d^k}{dx^k}(1+x)^{n+\beta}\\ &= \sum_{k=0}^{n}(-1)^{n-k}\binom{n}{k}(\alpha+n)(\alpha+n-1)\cdots(\alpha+k+1)(\beta+n)\\ &\quad\times(\beta+n-1)\cdots(\beta+n-k+1)(1-x)^{\alpha+k}(1+x)^{\beta+n-k}.\end{aligned}$$

Substituindo em (2.23), concluímos que

$$\begin{aligned}P_n^{(\alpha,\beta)}(x) &= \frac{1}{2^n}\sum_{k=0}^{n}\frac{(-1)^k}{k!(n-k)!}(\alpha+n)(\alpha+n-1)\cdots(\alpha+k+1)(\beta+n)(\beta+n-1)\cdots\\ &\quad\times(\beta+n-k+1)(1-x)^k(1+x)^{n-k}.\end{aligned}$$

Como $(1-x)^k=\displaystyle\sum_{j=0}^{k}(-1)^j\binom{k}{j}x^j$ e $(1+x)^{n-k}=\displaystyle\sum_{i=0}^{n-k}\binom{n-k}{i}x^i$, aplicando o produto de Cauchy em $(1-x)^k(1+x)^{n-k}$, obtemos

$$\begin{aligned}P_n^{(\alpha,\beta)}(x) &= \frac{1}{2^n}\sum_{k=0}^{n}\frac{(-1)^k}{k!(n-k)!}(\alpha+n)\cdots(\alpha+k+1)(\beta+n)\cdots(\beta+n-k+1)\\ &\quad\times\left\{(-1)^k x^n+\left[(-1)^{k-1}\binom{k}{k-1}+(-1)^k\binom{n-k}{n-k-1}\right]x^{n-1}+\cdots\right\}\\ &= \frac{1}{2^n}\sum_{k=0}^{n}\frac{(-1)^k}{k!(n-k)!}(\alpha+n)\cdots(\alpha+k+1)(\beta+n)\cdots(\beta+n-k+1)x^n+\cdots\quad.\end{aligned}$$

Daí,

$$a_{n,n}=\frac{1}{2^n}\sum_{k=0}^{n}\frac{(-1)^k}{k!(n-k)!}\frac{\Gamma(\alpha+n+1)}{\Gamma(\alpha+k+1)}\frac{\Gamma(\beta+n+1)}{\Gamma(\beta+n-k+1)},$$

onde $\Gamma(x)$ é a função gama, que pode ser definida como a integral de Euler de segunda espécie

$$\Gamma(x)=\int_0^\infty e^{-t}t^{x-1}dt,\quad Re(x)>0 \;\text{ ou }\; x>0.$$

É bem conhecido que a função gama satisfaz à propriedade $\Gamma(x+1)=x\Gamma(x)$, para $x>0$.

A partir da fórmula de Rodrigues, obtemos a seguinte relação de ortogonalidade satisfeita pelos polinômios de Jacobi:

$$\left\langle P_n^{(\alpha,\beta)}, P_m^{(\alpha,\beta)} \right\rangle = \begin{cases} 0, & \text{se} \quad n \neq m, \\ \dfrac{2^{\alpha+\beta+1}\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)n!\Gamma(\alpha+\beta+n+1)}, & \text{se} \quad n = m. \end{cases}$$

A relação de recorrência para os polinômios de Jacobi é da seguinte forma

$$P_n^{(\alpha,\beta)}(x) = (\gamma_{n+1}^{(\alpha,\beta)} x - \beta_{n+1}^{(\alpha,\beta)}) P_{n-1}^{(\alpha,\beta)}(x) - \alpha_{n+1}^{(\alpha,\beta)} P_{n-2}^{(\alpha,\beta)}(x), \qquad n \geq 2,$$

onde $P_0^{(\alpha,\beta)}(x) = 1$, $P_1^{(\alpha,\beta)}(x) = \frac{1}{2}(\alpha+\beta+2)x + \frac{1}{2}(\alpha-\beta)$,

$$\gamma_{n+1}^{(\alpha,\beta)} = \frac{(2n+\alpha+\beta-1)(2n+\alpha+\beta)}{2n(n+\alpha+\beta)}, \qquad \beta_{n+1}^{(\alpha,\beta)} = \frac{(2n+\alpha+\beta-1)(\beta^2-\alpha^2)}{2n(n+\alpha+\beta)(2n+\alpha+\beta-2)}$$

e

$$\alpha_{n+1}^{(\alpha,\beta)} = \frac{2(n+\alpha-1)(n+\beta-1)(2n+\alpha+\beta)}{2n(n+\alpha+\beta)(2n+\alpha+\beta-2)},$$

O seguinte teorema será de grande utilidade no decorrer deste trabalho.

Teorema 2.21 *Os polinômios de Jacobi, $P_n^{(\alpha,\beta)}(x)$, satisfazem à seguinte equação diferencial homogênea de 2ª ordem:*

$$(1-x^2)y''(x) + [\beta - \alpha - (\alpha+\beta+2)x]y'(x) + n(n+\alpha+\beta+1)y(x) = 0 \tag{2.24}$$

ou

$$\frac{d}{dx}\Big\{(1-x)^{\alpha+1}(1+x)^{\beta+1}y'(x)\Big\} + n(n+\alpha+\beta+1)(1-x)^{\alpha}(1+x)^{\beta}y(x) = 0. \tag{2.25}$$

Demonstração: Seja $y(x) = P_n^{(\alpha,\beta)}(x)$. Calculando a derivada de $(1-x)^{\alpha+1}(1+x)^{\beta+1}y'(x)$, obtemos

$$\begin{aligned} \frac{d}{dx}\Big\{(1-x)^{\alpha+1}(1+x)^{\beta+1}y'(x)\Big\} &= -(\alpha+1)(1-x)^{\alpha}(1+x)^{\beta+1}y'(x) + (\beta+1)(1-x)^{\alpha+1} \\ &\quad \times (1+x)^{\beta}y'(x) + (1-x)^{\alpha+1}(1+x)^{\beta+1}y''(x) \\ &= (1-x)^{\alpha}(1+x)^{\beta}z(x), \end{aligned} \tag{2.26}$$

onde $z(x) = [-(\alpha+1)(1+x) + (\beta+1)(1-x)]y'(x) + (1-x^2)y''(x)$. Como $y(x)$ é um polinômio de grau n, então $z(x)$ é também de grau n. Mostremos que $z(x) = cy(x)$, onde c é uma constante. Consideremos, então, a integral

$$I = \int_{-1}^{1} \frac{d}{dx}\Big\{(1-x)^{\alpha+1}(1+x)^{\beta+1}y'(x)\Big\}q(x)dx, \tag{2.27}$$

onde $q(x)$ é um polinômio de grau $n-1$. Integrando por partes, obtemos

$$I = -\int_{-1}^{1} (1-x)^{\alpha+1}(1+x)^{\beta+1}q'(x)y'(x)dx$$

já que $\alpha+1$ e $\beta+1$ são positivos. Integrando por partes novamente

$$I = \int_{-1}^{1} y(x)\frac{d}{dx}\Big\{(1-x)^{\alpha+1}(1+x)^{\beta+1}q'(x)\Big\}dx = \int_{-1}^{1} y(x)r(x)(1-x)^{\alpha}(1+x)^{\beta}dx,$$

onde $r(x) = -(\alpha+1)(1+x)q'(x) + (\beta+1)(1-x)q'(x) + (1-x^2)q''(x)$ é um polinômio de grau $n-1$.

Portanto, $I = 0$. Logo, de (2.26) e (2.27),

$$\int_{-1}^{1} z(x)q(x)(1-x)^{\alpha}(1+x)^{\beta}dx = 0.$$

Como $q(x) \in \mathbb{P}_{n-1}$, pelo Corolário 2.1, $z(x) = cy(x)$, onde c é uma constante. Comparando, então, os termos de maior grau de $z(x)$ e $cy(x)$, obtemos $c = -n(n+\alpha+\beta+1)$. ■

Teorema 2.22 *Sejam $\alpha > -1$, $\beta > -1$. A equação diferencial*

$$(1-x^2)y''(x) + [\beta - \alpha - (\alpha+\beta+2)x]y'(x) + \gamma y(x) = 0, \tag{2.28}$$

onde γ é um parâmetro, tem solução polinomial não identicamente nula se, e somente se, γ tem a forma $n(n+\alpha+\beta+1)$, $n = 0, 1, 2, \ldots$. Esta solução é $cteP_n^{(\alpha,\beta)}(x)$ e nenhuma solução linearmente independente de $P_n^{(\alpha,\beta)}(x)$ pode ser um polinômio.

Demonstração: Seja $y(x) = \displaystyle\sum_{i=0}^{\infty} a_i(x-1)^i$. Então,

$$y'(x) = \sum_{i=1}^{\infty} ia_i(x-1)^{i-1} \quad \text{e} \quad y''(x) = \sum_{i=2}^{\infty} i(i-1)a_i(x-1)^{i-2}.$$

Substituindo na equação (2.28), encontramos

$$(1-x^2)\sum_{i=2}^{\infty} i(i-1)a_i(x-1)^{i-2}+[\beta-\alpha-(\alpha+\beta+2)x]\sum_{i=1}^{\infty} ia_i(x-1)^{i-1}+\gamma\sum_{i=0}^{\infty} a_i(x-1)^i=0.$$

Desde que $(1-x^2)=-(1+x)(x-1)$, a equação acima pode ser escrita como

$$-(1+x)\sum_{i=2}^{\infty} i(i-1)a_i(x-1)^{i-1}+[\beta-\alpha-(\alpha+\beta+2)x]\sum_{i=1}^{\infty} ia_i(x-1)^{i-1} \\ +\gamma\sum_{i=0}^{\infty} a_i(x-1)^i=0. \tag{2.29}$$

Somando-se e subtraindo-se $\alpha+\beta+2$ ao termo entre colchetes de (2.29) e, como $(1+x)=(2+x-1)$, obtemos

$$-(2+x-1)\sum_{i=2}^{\infty} i(i-1)a_i(x-1)^{i-1}-[2(\alpha+1)+(\alpha+\beta+2)(x-1)]\sum_{i=1}^{\infty} ia_i(x-1)^{i-1} \\ +\gamma\sum_{i=0}^{\infty} a_i(x-1)^i=0.$$

Mas, a equação acima pode ser escrita da forma

$$-\sum_{i=0}^{\infty} i(i-1)a_i(x-1)^i-(\alpha+\beta+2)\sum_{i=0}^{\infty} ia_i(x-1)^i+\gamma\sum_{i=0}^{\infty} a_i(x-1)^i \\ -\sum_{i=0}^{\infty} 2i(i+1)a_{i+1}(x-1)^i-\sum_{i=0}^{\infty} 2(\alpha+1)(i+1)a_{i+1}(x-1)^i=0.$$

Logo,

$$\sum_{i=0}^{\infty}\Big\{[\gamma-i(i-1)-i(\alpha+\beta+2)]a_i-[2i(i+1)+2(\alpha+1)(i+1)]a_{i+1}\Big\}(x-1)^i=0$$

e, portanto,

$$[\gamma-i(i+\alpha+\beta+1)]a_i-[2(i+1)(i+\alpha+1)]a_{i+1}=0, \qquad i=0,1,\dots \tag{2.30}$$

Se $y(x)$ é um polinômio de grau n, $a_i=0$ para $i=n+1,n+2,\dots$. Logo, o coeficiente de a_n deve se anular, isto é, $\gamma=n(n+\alpha+\beta+1)$.

Reciprocamente, se $\gamma=n(n+\alpha+\beta+1)$, então, em (2.30),

$$2(i+1)(i+\alpha+1)a_{i+1}=0, \qquad i=n,n+1,\dots .$$

Logo, como $\alpha > -1$, obtemos $a_{i+1} = 0$, $i = n, n+1, \ldots$. Agora, sejam $\gamma = n(n+\alpha+\beta+1)$ e $z(x)$ uma segunda solução de (2.24) ou (2.25). Então,

$$\left\{\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}y'(x)] + n(n+\alpha+\beta+1)(1-x)^{\alpha}(1+x)^{\beta}y(x)\right\} z(x) = 0 \quad (2.31)$$

e

$$\left\{\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}z'(x)] + n(n+\alpha+\beta+1)(1-x)^{\alpha}(1+x)^{\beta}z(x)\right\} y(x) = 0. \quad (2.32)$$

Subtraindo (2.32) de (2.31), obtemos

$$\left\{\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}]y'(x)\right\} z(x) - \left\{\frac{d}{dx}[(1-x)^{\alpha+1}(1+x)^{\beta+1}]z'(x)\right\} y(x) = 0.$$

Logo,

$$\begin{aligned}(1-x)^{\alpha}(1+x)^{\beta}\Big\{&[-(\alpha+1)(1+x)+(\beta+1)(1-x)]\,[y'(x)z(x)-y(x)z'(x)]\\ &+(1-x^2)[y'(x)z(x)-y(x)z'(x)]'\Big\} = 0.\end{aligned}$$

Observe que podemos escrever esta última equação da seguinte forma:

$$\frac{d}{dx}\Big\{(1-x)^{\alpha+1}(1+x)^{\beta+1}[y'(x)z(x)-y(x)z'(x)]\Big\} = 0.$$

Portanto, para todo x,

$$(1-x)^{\alpha+1}(1+x)^{\beta+1}[y'(x)z(x)-y(x)z'(x)] = \text{cte}.$$

Se fizermos $x \to \pm 1$ na equação acima, vemos que $y(x)$ e $z(x)$ não podem ambos ser polinômios, a menos que a constante do lado direito seja zero, isto é, a menos que $y(x)$ e $z(x)$ sejam linearmente dependentes. ■

Polinômios de Gegenbauer - $C_n^{(\lambda)}(x)$

São múltiplos dos polinômios de Jacobi quando $\alpha = \beta = \lambda - 1/2 \neq 0$. São também conhecidos por **Polinômios Ultrasféricos** e são ortogonais no intervalo $(-1, 1)$ com relação à função peso $\omega(x) = (1-x^2)^{\lambda-1/2}$, $\lambda > -\dfrac{1}{2}$.

Os polinômios de Gegenbauer satisfazem

$$C_n^{(\lambda)}(x) = \frac{\Gamma(\lambda+1/2)\Gamma(n+2\lambda)}{\Gamma(2\lambda)\Gamma(n+\lambda+1/2)} P_n^{(\lambda-1/2,\lambda-1/2)}(x).$$

A relação de recorrência para os polinômios de Gegenbauer é da seguinte forma:

$$C_{n+1}^{(\lambda)}(x) = \frac{2(n+\lambda)}{n+1} x C_n^{(\lambda)}(x) - \frac{n+2\lambda-1}{n+1} C_{n-1}^{(\lambda)}(x), \qquad n \geq 0,$$

onde $C_{-1}^{(\lambda)}(x) = 0$ e $C_0^{(\lambda)}(x) = 1$.

A relação de ortogonalidade é dada por

$$\left\langle C_n^{(\lambda)}, C_m^{(\lambda)} \right\rangle = \begin{cases} 0, & \text{se} \quad n \neq m, \\ \dfrac{2^\lambda \Gamma(\lambda+n+1/2)\Gamma(\lambda+n+1/2)}{(\lambda+n)n!\Gamma(2\lambda+n)}, & \text{se} \quad n = m. \end{cases}$$

Além disso, $C_n^{(\lambda)}$, satisfaz à seguinte equação diferencial de 2ª ordem

$$(1-x^2)y''(x) - (2\lambda+1)xy'(x) + n(n+2\lambda)y(x) = 0. \tag{2.33}$$

Polinômios de Laguerre - $L_n{}^{(\alpha)}(x)$

Os polinômios de Laguerre são ortogonais no intervalo $[0,\infty)$ com relação à função peso $w(x) = x^\alpha e^{-x}$, $\alpha > -1$. Podem ser definidos, pela Fórmula de Rodrigues, por

$$L_n^{(\alpha)}(x) = \frac{(-1)^n}{n!} x^{-\alpha} e^x \frac{d^n}{dx^n}[x^{\alpha+n} e^{-x}].$$

Aplicando a regra de Leibnitz na relação acima, obtemos sua representação explícita:

$$L_n^{(\alpha)}(x) = \sum_{k=0}^{n} \frac{(-1)^k}{k!} \binom{n+\alpha}{n-k} x^k. \tag{2.34}$$

Se $x_{n,1}, x_{n,2}, \ldots, x_{n,n}$ são os zeros de $L_n^{(\alpha)}(x)$, da equação anterior obtemos que

$$x_{n,1} + x_{n,2} + \cdots + x_{n,n} = n(n+\alpha). \tag{2.35}$$

A relação de recorrência de três termos é da forma

$$nL_n^{(\alpha)}(x) = (-x+2n+\alpha-1)L_{n-1}^{(\alpha)}(x) - (n+\alpha-1)L_{n-2}^{(\alpha)}(x), \qquad n \geq 2,$$

com as condições iniciais $L_0^{(\alpha)}(x) = 1$ e $L_1^{(\alpha)}(x) = -x + \alpha + 1$.

De (2.34), obtemos facilmente

$$a_{n,0} = L_n^{(\alpha)}(0) = \binom{n+\alpha}{n} \qquad \text{e} \qquad a_{n,n} = \frac{(-1)^n}{n!}.$$

Os polinômios de Laguerre satisfazem à seguinte equação diferencial de 2ª ordem

$$xy''(x) + (\alpha + 1 - x)y'(x) + ny(x) = 0,$$

cuja demonstração é análoga à do Teorema 2.21 para os polinômios de Jacobi. Um argumento similar àquele usado no Teorema 2.22 mostra que uma condição necessária e suficiente para que a equação diferencial

$$xy''(x) + (\alpha + 1 - x)y'(x) + \gamma y(x) = 0$$

tenha solução polinomial é que $\gamma \in \mathbb{Z}^+$, o conjunto dos inteiros não-negativos. Além disso, se $\gamma = n$, as únicas soluções polinomiais de (2.2.2) são $y(x) = cL_n^{(\alpha)}(x)$, onde c é uma constante arbitrária não-negativa.

A relação de ortogonalidade é dada por:

$$\left\langle L_n^{(\alpha)}, L_m^{(\alpha)} \right\rangle = \int_0^\infty L_n^{(\alpha)}(x) L_m^{(\alpha)}(x) x^\alpha e^{-x} dx = \begin{cases} 0, & \text{se} \quad n \neq m, \\ \dfrac{\Gamma(n+\alpha+1)}{n!}, & \text{se} \quad n = m, \end{cases}$$

que é facilmente obtida da fórmula de Rodrigues.

Polinômios de Hermite - $H_n(x)$

Os polinômios de Hermite são ortogonais no intervalo $(-\infty, \infty)$ com relação à função peso $w(x) = e^{-x^2}$. Pela Fórmula de Rodrigues, podem ser definidos por

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n}[e^{-x^2}].$$

Sua representação explícita é dada por

$$\frac{H_n(x)}{n!} = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{(-1)^k}{k!} \frac{(2x)^{n-2k}}{(n-2k)!}, \tag{2.36}$$

onde $\lfloor a \rfloor$ = significa o maior inteiro menor ou igual a a.

De (2.36), obtemos, $H_n(x) = (-1)^n H_n(-x)$. Logo, se $x_{n,1} < x_{n,2} < \cdots < x_{n,n}$ são os zeros de $H_n(x)$, então $x_{n,n-i+1} = -x_{n,i}$, $i = 1, 2, \ldots, \lfloor n/2 \rfloor$.

De (2.36), obtemos, também, que

$$x_{n,1}^2 + x_{n,2}^2 + \cdots + x_{n,n}^2 = \frac{n(n-1)}{2}. \tag{2.37}$$

Além disso, $a_{n,n} = 2^n$, $n \geq 0$ e, para os polinômios pares,

$$a_{2m,0} = H_{2m}(0) = (-1)^m \frac{(2m)!}{m!}, \qquad m = 1, 2, \ldots\ .$$

Os polinômios de Hermite, $H_n(x)$, $n \geq 0$, satisfazem à seguinte equação diferencial de 2ª ordem

$$y''(x) - 2xy'(x) + 2ny(x) = 0 \tag{2.38}$$

ou

$$\frac{d}{dx}\left\{e^{-x^2} y'(x)\right\} + 2ne^{-x^2} y(x) = 0 \tag{2.39}$$

e sua demonstração também é análoga ao resultado para os polinômios de Jacobi (Teorema 2.21). Além disso, a equação diferencial

$$y''(x) - 2xy'(x) + \gamma y(x) = 0$$

possui solução polinomial se, e somente se, γ é um número par não-negativo. Em particular, se $\gamma = 2n$, a única solução polinomial de (2.38) é $y(x) = cH_n(x)$, sendo c uma constante arbitrária não-negativa. A demonstração deste fato também é análoga àquela sobre os polinômios de Jacobi.

Da fórmula de Rodrigues, é fácil obter a relação de ortogonalidade

$$\langle H_n, H_m \rangle \;=\; \int_{-\infty}^{\infty} H_n(x) H_m(x) e^{-x^2} dx = \begin{cases} 0, & \text{se} \quad n \neq m, \\ 2^n n! \sqrt{\pi}, & \text{se} \quad n = m. \end{cases}$$

Além disso, satisfazem à seguinte fórmula de recorrência de três termos

$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \qquad n \geq 0,$$

com $H_{-1}(x) = 0$ e $H_0(x) = 1$.

2.3 Polinômios similares aos ortogonais

A introdução do problema de momento forte pelos pesquisadores Jones, Thron e Waadeland [11] abriu caminho para o estudo de polinômios que apresentam propriedades semelhantes às dos polinômios ortogonais. O problema de momento forte pode ser expresso da seguinte forma:

"Dado uma seqüência $\{\mu_m\}_{m=-\infty}^{\infty}$ de números reais, em que condições existe uma medida não-negativa $d\psi(t)$, tal que

$$\mu_m = \int_a^b t^m d\psi(t), \qquad m = \ldots, -2, -1, 0, 1, 2 \ldots ? \quad " \tag{2.40}$$

Se os momentos μ_m, $m \in \mathbb{Z}$, definidos por (2.40) existem e são finitos, $d\psi(t)$ é chamada uma distribuição forte. Se $d\psi(t) = \nu(t)dt$, então $\nu(t)$ é uma função peso forte.

Jones, Thron e Waadeland resolveram o problema (2.40) acima quando $d\psi(t)$ é uma medida forte de Stieltjes, isto é, $(a, b) \subseteq (0, \infty)$. Com o objetivo de estudar o problema de momento forte para o caso em que $(a, b) \subseteq (-\infty, \infty)$, conhecido como problema de momento forte de Hamburger, Sri Ranga [19] introduziu uma seqüência de polinômios mônicos $B_n^{(k)}(t)$, similares aos polinômios ortogonais, definidos por:

$$\int_a^b t^{k-n+s} B_n^{(k)}(t) d\psi(t) \;=\; \begin{cases} 0, & \text{se} \quad 0 \leq s \leq n-1, \\ \breve{\rho}_n^{(k)} > 0, & \text{se} \quad s = n, \end{cases}$$

para $n = 1, 2, \ldots$ e $k \in \mathbb{Z}$, onde $-\infty \leq a < b \leq \infty$.

Vamos considerar, aqui, somente os polinômios $B_n^{(k)}(t)$ para $k = 0$ e intervalos $[a, b]$ que estejam totalmente contidos do lado positivo do eixo real, isto é, $0 \leq a < b \leq \infty$. Assim, $d\psi(t)$ é uma distribuição forte de Stieltjes. Denotaremos tais polinômios por $B_n(t)$ e os chamaremos, simplesmente, de polinômios similares.

Portanto, os polinômios $B_n(t)$, para $n = 0, 1, 2, \ldots$, são dados por:

$$\int_a^b t^{-n+s} B_n(t) d\psi(t) \;=\; \begin{cases} 0, & \text{se} \quad 0 \leq s \leq n-1, \\ \breve{\rho}_n > 0, & \text{se} \quad s = n. \end{cases} \tag{2.41}$$

O resultado a seguir é de primordial importância para a prova da existência dos polinômios $B_n(t)$.

Teorema 2.23 *Para as distribuições fortes de Stieltjes, $d\psi(t)$, os determinantes de Hankel, $H_n^{(m)}$, satisfazem:*

$$H_n^{(m)} = \begin{vmatrix} \mu_m & \mu_{m+1} & \cdots & \mu_{m+n-1} \\ \mu_{m+1} & \mu_{m+2} & \cdots & \mu_{m+n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{m+n-1} & \mu_{m+n} & \cdots & \mu_{m+2n-2} \end{vmatrix} > 0,$$

para m, n inteiros, $n \geq 1$.

Demonstração: Similar à do Teorema 2.15. ∎

Como vamos considerar os polinômios $B_n(t)$ mônicos, se denotarmos

$$B_n(t) = \sum_{k=0}^{n} b_{n,n-k} t^{n-k}, \tag{2.42}$$

então $b_{n,n} = 1$ para todo $n \geq 0$.

Usando (2.40), podemos escrever (2.41) como um sistema de equações lineares algébricas nos coeficientes de $B_n(t)$:

$$\left\{ \begin{array}{ccccccccc} \mu_{-n}b_{n,0} & + & \mu_{-n+1}b_{n,1} & + & \cdots & + & \mu_0 b_{n,n} & = & 0 \\ \mu_{-n+1}b_{n,0} & + & \mu_{-n+2}b_{n,1} & + & \cdots & + & \mu_1 b_{n,n} & = & 0 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ \mu_{-1}b_{n,0} & + & \mu_0 b_{n,1} & + & \cdots & + & \mu_{n-1}b_{n,n} & = & 0 \\ \mu_0 b_{n,0} & + & \mu_1 b_{n,1} & + & \cdots & + & \mu_n b_{n,n} & = & \check{\rho}_n \end{array} \right. \tag{2.43}$$

Usando a regra de Cramer, o coeficiente $b_{n,n}$ pode ser dado por:

$$b_{n,n} = \frac{\check{\rho}_n H_n^{(-n)}}{H_{n+1}^{(-n)}} = 1.$$

Assim,

$$\check{\rho}_n = \frac{H_{n+1}^{(-n)}}{H_n^{(-n)}} > 0, \qquad n \geq 0.$$

Se substituirmos a última equação de (2.43) por (2.42) e, novamente, aplicarmos a regra de Cramer, obtemos:

$$B_n(t) = \frac{1}{H_n^{(-n)}} \begin{vmatrix} \mu_{-n} & \mu_{-n+1} & \cdots & \mu_0 \\ \mu_{-n+1} & \mu_{-n+2} & \cdots & \mu_1 \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{-1} & \mu_0 & \cdots & \mu_{n-1} \\ 1 & t & \cdots & t^n \end{vmatrix}, \quad n \geq 0. \tag{2.44}$$

Assim, vemos que uma condição necessária e suficiente para a existência de $B_n(t)$ é que $H_n^{(-n)} \neq 0$. Agora, se substituirmos $t = 0$ em (2.44), obtemos

$$B_n(0) = (-1)^n \frac{H_n^{(-n+1)}}{H_n^{(-n)}} = b_{n,0}, \qquad n \geq 0.$$

Logo, uma condição necessária e suficiente para a existência dos polinômios $B_n(t)$ com $B_n(0) \neq 0$ é que $H_n^{(-n)} \neq 0$ e $H_n^{(-n+1)} \neq 0$.

Usando a relação

$$\int_a^b t^{-(n+1)} B_n(t) d\psi(t) = \mu_{-n-1} b_{n,0} + \mu_{-n} b_{n,1} + \cdots + \mu_{-1} b_{n,n}$$

em (2.43), encontramos

$$\eta_n = \int_a^b t^{-(n+1)} B_n(t) d\psi(t) = (-1)^n \frac{H_{n+1}^{(-n-1)}}{H_n^{(-n)}}, \qquad n \geq 0,$$

que é diferente de zero.

2.3.1 Algumas propriedades dos polinômios similares

Os polinômios similares, $B_n(t)$, exibem muitas propriedades semelhantes às dos polinômios ortogonais. A mais interessante e útil delas é a relação de recorrência de três termos que daremos a seguir.

Teorema 2.24 *Seja $B_n(t)$ o polinômio definido em* (2.41). *Então, a seguinte relação de recorrência é verdadeira para $n \geq 1$:*

$$B_{n+1}(t) \;=\; (t - \breve{\beta}_{n+1}) B_n(t) - \breve{\alpha}_{n+1} t B_{n-1}(t), \tag{2.45}$$

onde $B_0(t) = 1,\ B_1(t) = z - \breve{\beta}_1$ *e os coeficientes* $\breve{\alpha}_{n+1}$ *e* $\breve{\beta}_{n+1}$ *satisfazem:*

$$\breve{\alpha}_{n+1} = \frac{\breve{\rho}_n}{\breve{\rho}_{n-1}} > 0,$$

$$\breve{\beta}_1 = \frac{\mu_0}{\mu_{-1}}, \qquad \breve{\beta}_{n+1} = -\alpha_{n+1}\frac{\eta_{n-1}}{\eta_n}, \qquad n \geq 1. \tag{2.46}$$

Demonstração: Como $B_n(t)$ é um polinômio de grau n, o polinômio $B_{n+1}(t) - tB_n(t)$ é de grau, no máximo, n. Portanto, pode ser escrito da forma

$$B_{n+1}(t) - tB_n(t) = -\breve{\beta}_{n+1}B_n(t) - \breve{\alpha}_{n+1}tB_{n-1}(t) + p_{n-1}(t), \tag{2.47}$$

onde $\breve{\alpha}_{n+1}$ e $\breve{\beta}_{n+1}$ são tais que $p_{n-1}(t)$ é de grau $n-1$.

Tomemos $p_{n-1}(t) = \displaystyle\sum_{j=0}^{n-1} c_j t^j$. Multiplicando-se ambos os membros da equação (2.47) por t^{-n+s} e integrando, obtemos:

$$\begin{aligned} \int_a^b t^{-n+s}B_{n+1}(t)d\psi(t) \ &- \ \int_a^b t^{-n+s+1}B_n(t)d\psi(t) \\ &= \ -\breve{\beta}_{n+1}\int_a^b t^{-n+s}B_n(t)d\psi(t) - \breve{\alpha}_{n+1}\int_a^b t^{-(n-1)+s}B_{n-1}(t)d\psi(t) \\ &\quad + \sum_{j=0}^{n-1} c_j \int_a^b t^{-n+j+s}d\psi(t). \end{aligned} \tag{2.48}$$

Fazendo, respectivamente, $s = 0, 1, 2, \ldots, n-1$, na equação acima, obtemos o sistema linear:

$$\begin{cases} \displaystyle\sum_{j=0}^{n-1} c_j\mu_{-n+j+i} = 0, & i = 0, 1, \ldots, n-2, \\ \displaystyle\sum_{j=0}^{n-1} c_j\mu_{-1+j} = \breve{\alpha}_{n+1}\breve{\rho}_{n-1} - \breve{\rho}_n, \end{cases}$$

cujas incógnitas são os coeficientes do polinômio $p_{n-1}(t)$ e cujo determinante da matriz dos coeficientes é $H_n^{(-n)}$ que, como já sabemos, é maior do que zero. Assim, se tomarmos $\breve{\alpha}_{n+1} = \dfrac{\breve{\rho}_n}{\breve{\rho}_{n-1}}$, o sistema será homogêneo e sua única solução é $c_j = 0$, $j = 0, 1, \ldots, n-1$, o que significa dizer que $p_{n-1}(t) \equiv 0$. Portanto, (2.45) está demonstrado.

Para determinarmos o coeficiente $\breve{\beta}_{n+1}$, tomemos $s = -1$ em (2.48). Logo, $0 = -\breve{\beta}_{n+1}\eta_n - \breve{\alpha}_{n+1}\eta_{n-1}$ e, daí, segue imediatamente (2.46). ■

Sobre os zeros dos polinômios similares $B_n(t)$ vale, ainda, o seguinte resultado, cuja demonstração é análoga ao resultado para polinômios ortogonais.

Teorema 2.25 *Os zeros dos polinômios similares $B_n(t)$ $n \geq 1$, são reais, distintos e pertencem ao intervalo (a, b).*

Os polinômios $B_n(t)$ estão relacionados aos polinômios de Laurent ortogonais ou L-polinômios ortogonais, $R_n(t)$, com relação a uma distribuição forte, tratados por Jones e outros em [11, 12], da seguinte forma:

$$R_{2m}(t) = t^{-m} B_{2m}(t), \qquad R_{2m+1}(t) = t^{-m-1} B_{2m+1}(t), \qquad m \geq 0.$$

Esta relação pode ser facilmente verificada através de (2.41) e da definição de polinômios de Laurent dados em [12].

Lembremos que um polinômio de Laurent é uma função da forma $\displaystyle\sum_{n=p}^{q} c_n t^n$, onde $p \leq q$, $p, q \in \mathbb{Z}$ e $c_n \in \mathbb{R}$.

Um estudo detalhado sobre os polinômios similares aos ortogonais pode ser encontrado em [2, 19].

Capítulo 3

Eletrostática e zeros de polinômios

O estudo apresentado nas seções 3.1 e 3.3 é baseado, principalmente, no livro de Szegő [24] e no artigo de Dimitrov e Van Assche [6].

Na seção 3.2, apresentamos nossa contribuição à interpretação eletrostática dos zeros de polinômios para o caso em que há restrições.

3.1 Interpretação eletrostática e equação de Lamé

Consideremos, então, o campo eletrostático que obedece à lei do potencial logarítmico descrito na Introdução, ou seja, cargas r_j são distribuídas uniformemente ao longo de retas perpendiculares ao eixo real e que a interceptam nos pontos a_j, $j = 0, 1, \ldots, m$. Além disso, n cargas unitárias positivas móveis estão distribuídas uniformemente ao longo de retas perpendiculares ao eixo real e que passam pelos pontos $x_1, x_2, \ldots, x_n$ pertencentes ao intervalo (a_0, a_m). Queremos que as cargas livres se movam apenas entre cargas fixas positivas consecutivas para evitar aglutinação. Lembremos que o ponto $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ pertence ao simplex

$$\Xi := \bigcap_{k=1}^{s} \{a_{j_k} < x_{\mu_k} < \ldots x_{\mu_{k+1}-1} < a_{j_k+1} : r_{j_k}, r_{j_k+1} > 0\},$$

onde $1 = \mu_1 \leq \mu_2 \leq \ldots \mu_s \leq \mu_{s+1} - 1 = n$ e s é o número de intervalos $[a_{j_k}, a_{j_k+1}]$ com cargas positivas r_{j_k} e r_{j_k+1} nos pontos extremos.

A energia do campo é, então, dada por

$$\begin{aligned}
L(\boldsymbol{x}) &= \sum_{j=0}^{m} r_j \sum_{k=1}^{n} \log \frac{1}{|x_k - a_j|} + \sum_{1\le i<k\le n} \log \frac{1}{|x_k - x_i|} \\
&= \sum_{j=0}^{m} \sum_{k=1}^{n} \log |x_k - a_j|^{-r_j} + \sum_{1\le i<k\le n} \log |x_k - x_i|^{-1} \\
&= -\sum_{j=0}^{m} \sum_{k=1}^{n} \log |x_k - a_j|^{r_j} - \sum_{1\le i<k\le n} \log |x_k - x_i|.
\end{aligned}$$

A energia é mínima se, e somente se, a função

$$T(\boldsymbol{x}) = e^{-L(\boldsymbol{x})} = \prod_{j=0}^{m} \prod_{k=1}^{n} |x_k - a_j|^{r_j} \prod_{1\le i<k\le n} |x_k - x_i| \tag{3.1}$$

é máxima.

Observe que $T(\boldsymbol{x})$ se anula na fronteira de Ξ e é estritamente positiva em seu interior. Portanto, o máximo de $T(\boldsymbol{x})$ é atingido num ponto $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)$ em Ξ onde $\dfrac{\partial T(\boldsymbol{x})}{\partial x_k} = 0$ para $k = 1, 2, \ldots, n$. Mas,

$$\begin{aligned}
T(\boldsymbol{x}) &= \prod_{j=0}^{m} \Big\{ |x_1 - a_j|^{r_j} |x_2 - a_j|^{r_j} |x_2 - x_1| |x_3 - a_j|^{r_j} |x_3 - x_1| |x_3 - x_2| \ldots \\
&\quad \times |x_n - a_j|^{r_j} |x_n - x_1| |x_n - x_2| \ldots |x_n - x_{n-1}| \Big\} \\
&= \prod_{j=0}^{m} \Big\{ |x_1 - a_j|^{r_j} |x_2 - a_j|^{r_j} |x_2 - x_1| \ldots |x_{k-1} - a_j|^{r_j} |x_{k-1} - x_1| \ldots \\
&\quad \times |x_{k-1} - x_{k-2}| |x_{k+1} - a_j|^{r_j} |x_{k+1} - x_1| \ldots |x_{k+1} - x_{k-1}| \ldots |x_n - a_j|^{r_j} |x_n - x_1| \ldots \\
&\quad \times |x_n - x_{k-1}| |x_n - x_{k+1}| \ldots |x_n - x_{n-1}| |x_k - a_j|^{r_j} \Big\} \prod_{\substack{i=1 \\ i\ne k}}^{n} |x_k - x_i|.
\end{aligned}$$

Portanto,

$$T(\boldsymbol{x}) = \tau(\bar{\boldsymbol{x}} \backslash x_k) |\omega_k(x_k)| \prod_{j=0}^{m} |x_k - a_j|^{r_j}, \tag{3.2}$$

onde $\tau(\bar{\boldsymbol{x}} \backslash x_k)$ é uma função que não depende de x_k, $\omega(x) := \prod_{i=1}^{n} (x - x_i)$ e $\omega_k(x) := \dfrac{\omega(x)}{x - x_k}$.

Como $[\omega(x)]^{(l+1)} = [(x - x_k)\omega_k(x)]^{(l+1)}$, aplicando a regra de Leibnitz, obtemos

$$[\omega(x)]^{(l+1)} = (x - x_k)\omega_k^{(l+1)}(x) + (l+1)\omega_k^{(l)}(x).$$

Logo, para $x = x_k$,

$$(l+1)\omega_k^{(l)}(x_k) = \omega^{(l+1)}(x_k), \qquad l \geq 0. \tag{3.3}$$

Observe que o *sinal* $\omega_k(x_k) = (-1)^{n-k}$. Logo, podemos escrever $|\omega_k(x_k)| = (-1)^{n-k}\omega_k(x_k)$. Portanto, (3.2) pode ser escrito como

$$T(\boldsymbol{x}) = (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\omega_k(x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}.$$

Derivando a equação acima com relação a x_k, obtemos

$$\begin{aligned}
\frac{\partial T(\boldsymbol{x})}{\partial x_k} &= (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\frac{\partial}{\partial x_k}\left(\omega_k(x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\right)\\
&= (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\left\{\frac{r_0}{|x_k - a_0|}\omega_k(x_k) + \cdots + \frac{r_m}{|x_k - a_m|}\omega_k(x_k)\right.\\
&\quad \left. + \omega_k'(x_k)\right\}\\
&= (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\left\{\left[\sum_{j=0}^{m}\frac{r_j}{|x_k - a_j|}\right]\omega_k(x_k) + \omega_k'(x_k)\right\}\\
&= (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\left\{\frac{B(x_k)}{A(x_k)}\omega_k(x_k) + \omega_k'(x_k)\right\}\\
&= (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\omega_k(x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\left\{\frac{B(x_k)}{A(x_k)} + \frac{\omega_k'(x_k)}{\omega_k(x_k)}\right\},
\end{aligned}$$

para $k = 1, 2, \ldots, n$.

De (3.3), concluímos que

$$\frac{\partial T(\boldsymbol{x})}{\partial x_k} = (-1)^{n-k}\tau(\bar{\boldsymbol{x}}\backslash x_k)\omega_k(x_k)\prod_{j=0}^{m}|x_k - a_j|^{r_j}\left\{\frac{B(x_k)}{A(x_k)} + \frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)}\right\}. \tag{3.4}$$

O máximo de $T(\boldsymbol{x})$ é atingido onde $\dfrac{\partial T(\boldsymbol{x})}{\partial x_k} = 0$, ou seja, onde

$$A(x_k)\omega''(x_k) + 2B(x_k)\omega'(x_k) = 0,$$

por (3.4).

Como $A(x)$ é de grau $m+1$, $B(x)$ é de grau m e $\omega(x)$ é de grau n, a expressão $A(x)\omega''(x) + 2B(x)\omega'(x)$ é um polinômio de grau $m+n-1$. No ponto de máximo de $T(\boldsymbol{x})$ este último polinômio anula-se nos zeros de $\omega(x)$. Então, pelo Teorema Fundamental da Álgebra,

$$A(x)\omega''(x) + 2B(x)\omega'(x) + C(x)\omega(x) = 0$$

para algum polinômio $C(x)$ de grau $m-1$. Mostramos, então, o seguinte resultado.

Lema 3.1 *Seja $T : \mathbb{R}^n \to \mathbb{R}$ definida em* (3.1). *Então, $\partial T(\boldsymbol{x})/\partial x_k = 0$, $k = 1, \ldots, n$, se, e somente se, as coordenadas do vetor $\boldsymbol{x}$ são os zeros do polinômio de grau n que é solução da equação de Lamé* (1.2).

O resultado a seguir nos fornece condições para que a solução seja única.

Teorema 3.1 *Se a energia do campo eletrostático em questão tem um único ponto de mínimo em Ξ, então existe um único par $(C(x), y(x))$, com $C(x)$ um polinômio de Van Vleck e $y(x)$ um polinômio de Stieltjes, para a equação de Lamé* (1.2) *(Ver pag.* 3*). Reciprocamente, se existe um único par $(C(x), y(x))$ de polinômios de Van Vleck e Stieltjes para* (1.2), *tal que os zeros do polinômio de Stieltjes pertencem a Ξ, então a energia do campo descrito tem um único ponto de mínimo.*

Demonstração: Suponhamos que exista um polinômio $\widetilde{C}(x)$ de grau $m-1$, $\widetilde{C}(x) \not\equiv C(x)$, tal que a equação diferencial

$$A(x)z''(x) + 2B(x)z'(x) + \widetilde{C}(x)z(x) = 0 \tag{3.5}$$

também admite uma solução $z(x) = (x-\xi_1)(x-\xi_2)\cdots(x-\xi_n)$ com zeros distintos em (a_{j-1}, a_j).

Seguindo Szegő [24, §6.83 pag. 154], vamos introduzir a função $H(x) = \displaystyle\prod_{j=0}^{m} |x-a_j|^{2r_j}$.

Logo, para $x \neq a_j$, $j = 0, \ldots, m$, temos que

$$\begin{aligned} H'(x) &= 2r_0(x-a_0)^{2r_0-1}(x-a_1)^{2r_1}\ldots(x-a_m)^{2r_m} \\ &\quad +2r_1(x-a_0)^{2r_0}(x-a_1)^{2r_1-1}(x-a_2)^{2r_2}\ldots(x-a_m)^{2r_m}+\ldots \\ &\quad +2r_m(x-a_0)^{2r_0}(x-a_1)^{2r_1}\ldots(x-a_{m-1})^{2r_{m-1}}(x-a_m)^{2r_m-1} \\ &= 2\prod_{j=0}^{m}(x-a_j)^{2r_j}\sum_{j=0}^{m}\frac{r_j}{x-a_j} = 2H(x)\frac{B(x)}{A(x)}. \end{aligned}$$

Multiplicando por $z(x)$ a equação (1.2), por $y(x)$ a equação (3.5) e subtraindo as duas equações resultantes, encontramos

$$A(x)[y'(x)z(x)-y(x)z'(x)]' + 2B(x)[y'(x)z(x)-y(x)z'(x)] = [\widetilde{C}(x)-C(x)]y(x)z(x). \quad (3.6)$$

Multiplicando ambos os membros de (3.6) por $\dfrac{H(x)}{A(x)}$, obtemos

$$\frac{d}{dx}\Big\{H(x)[y'(x)z(x)-y(x)z'(x)]\Big\} = \frac{\widetilde{C}(x)-C(x)}{A(x)}y(x)z(x)H(x). \quad (3.7)$$

Tomemos $x_0 = \xi_0 = a_{j-1}$, $x_{n+1} = \xi_{n+1} = a_j$ e consideremos $\dfrac{\widetilde{C}(x)-C(x)}{A(x)} > 0$ em (a_{j-1}, a_j). É claro que $y(x) = (x-x_1)(x-x_2)\cdots(x-x_n)$ não muda de sinal em (x_i, x_{i+1}), $i = 1, \ldots, n-1$. Suponhamos que $z(x)$ não muda de sinal em (x_i, x_{i+1}), $i = 1, \ldots, n-1$, isto é, entre dois zeros consecutivos de $y(x)$. Então, de (3.7),

$$\frac{d}{dx}\Big\{H(x)[y'(x)z(x)-y(x)z'(x)]\Big\}$$

também não muda de sinal em (x_i, x_{i+1}), $i = 1, \ldots, n-1$.

Considere os pontos $\tilde{x} = x_i + \epsilon$ e $\hat{x} = x_{i+1} - \epsilon$, $\epsilon > 0$.

1) Se $y(x)z(x) > 0$, então $H(x)[y'(x)z(x)-y(x)z'(x)]$ é crescente. Assim,

   - se $y(x) > 0 \Rightarrow z(x) > 0 \Rightarrow y'(\tilde{x})z(\tilde{x}) > 0$ e $y'(\hat{x})z(\hat{x}) < 0$;
   - se $y(x) < 0 \Rightarrow z(x) < 0 \Rightarrow y'(\tilde{x})z(\tilde{x}) > 0$ e $y'(\hat{x})z(\hat{x}) < 0$.

2) Se $y(x)z(x) < 0$, então $H(x)[y'(x)z(x)-y(x)z'(x)]$ é decrescente. Logo,

   - se $y(x) > 0 \Rightarrow z(x) < 0 \Rightarrow y'(\tilde{x})z(\tilde{x}) < 0$ e $y'(\hat{x})z(\hat{x}) > 0$;
   - se $y(x) < 0 \Rightarrow z(x) > 0 \Rightarrow y'(\tilde{x})z(\tilde{x}) < 0$ e $y'(\hat{x})z(\hat{x}) > 0$.

Portanto, $sinal(yz) = sinal(y'z)$ em $\tilde{x}$ e $sinal(yz) = -sinal(y'z)$ em $\hat{x}$. Observe que $sinal\{H(y'z - yz')\} = sinal(y'z)$ em $\tilde{x}$ e em $\hat{x}$, isto é,

- se $y(x)z(x) > 0$, então $sinal\{H(y'z - yz')\} > 0$ em $\tilde{x}$ e $sinal\{H(y'z - yz')\} < 0$ em $\hat{x}$;
- se $y(x)z(x) < 0$, então $sinal\{H(y'z - yz')\} < 0$ em $\tilde{x}$ e $sinal\{H(y'z - yz')\} > 0$ em $\hat{x}$.

Logo, das duas observações acima, $H(x)[y'(x)z(x) - y(x)z'(x)]$ muda de sinal em (x_i, x_{i+1}) da seguinte forma:

- Se $y(x)z(x) > 0$, $H(\tilde{x})[y'(\tilde{x})z(\tilde{x}) - y(\tilde{x})z'(\tilde{x})] > 0$ e $H(\hat{x})[y'(\hat{x})z(\hat{x}) - y(\hat{x})z'(\hat{x})] < 0$. Então, $H(x)[y'(x)z(x) - y(x)z'(x)]$ não pode ser crescente, o que é uma contradição por 1).
- Se $y(x)z(x) < 0$, $H(\tilde{x})[y'(\tilde{x})z(\tilde{x}) - y(\tilde{x})z'(\tilde{x})] < 0$ e $H(\hat{x})[y'(\hat{x})z(\hat{x}) - y(\hat{x})z'(\hat{x})] > 0$. Logo, $H(x)[y'(x)z(x) - y(x)z'(x)]$ não pode ser decrescente. Por 2), chegamos novamente a uma contradição.

Portanto, $y(x)z(x)$ deve mudar de sinal pelo menos uma vez em (x_i, x_{i+1}). Concluímos, então, que $z(x)$ tem pelo menos uma raiz em (x_i, x_{i+1}).

Sejam x_1 e x_n o primeiro e último zeros de $y(x)$ em $[a_{j-1}, a_j]$, respectivamente. Mostremos, agora, que $z(x)$ deve mudar de sinal também em $[a_{j-1}, x_1]$ e $[x_n, a_j]$. Lembremos que $\dfrac{\widetilde{C}(x) - C(x)}{A(x)} > 0$.

Suponhamos que $z(x)$ não muda de sinal nesses intervalos. Como os zeros de $z(x)$ pertencem a (a_{j-1}, a_j), então $y(x)$ e $z(x)$ têm ambos sinais constantes em $[a_{j-1}, x_1]$ e $[x_n, a_j]$. Mas, $y(x)$ e $z(x)$ são polinômios de mesmo grau e com coeficientes dos termos de maior grau positivos. Isto significa que ambos têm o mesmo sinal em $[a_{j-1}, x_1]$ e em $[x_n, a_j]$. Então, de (3.7), $H(x)[y'(x)z(x) - y(x)z'(x)]$ é uma função crescente nesses dois intervalos. Além disso,

$$\lim_{x \to a_{j-1}^+} H(x)[y'(x)z(x) - y(x)z'(x)] = 0 \quad \text{e} \quad \lim_{x \to a_j^-} H(x)[y'(x)z(x) - y(x)z'(x)] = 0.$$

Assim, a expressão $[y'(x)z(x) - y(x)z'(x)]$ deve ser positiva em $[a_{j-1}, x_1]$ e negativa em $[x_n, a_j]$. Mas, $y(x_1) = 0$. Logo, para $x = x_1$, temos que $sinal(y'z) = sinal(y'z - yz') > 0$.

Como x_1 é o menor zero de $y(x)$, então $sinal(y'y) < 0$ para $x = x_1 - \epsilon$, $\epsilon > 0$. Como $y(x)$ e $z(x)$ são do mesmo tipo e, além disso, $z(x)$ também não tem zeros em $[a_{j-1}, x_1)$, então $sinal(y'z) = sinal(y'y) < 0$. Isto é uma contradição.

Podemos usar o mesmo argumento para $x = x_n$. Mostramos, assim, que $z(x)$ tem, pelo menos, $n+1$ zeros se $\dfrac{\widetilde{C}(x) - C(x)}{A(x)} > 0$. Absurdo! Portanto, $\dfrac{\widetilde{C}(x) - C(x)}{A(x)}$ deve ser negativa em algum ponto de $a_{j-1} < x < a_j$.

Repetindo a mesma demonstração, mas trocando os papéis de $(C(x), y(x))$ e $(\widetilde{C}(x), z(x))$, concluímos que $\dfrac{\widetilde{C}(x) - C(x)}{A(x)}$ deve ser positiva em algum ponto de $a_{j-1} < x < a_j$. Isto significa que $\widetilde{C}(x) - C(x)$ deve mudar de sinal pelo menos uma vez em $[a_{j-1}, a_j]$, $j = 1, 2, \ldots, m$. Logo, $\widetilde{C}(x) - C(x)$ tem, pelo menos, m zeros em $[a_0, a_m]$. Impossível, pois $\widetilde{C}(x) - C(x)$ é de grau $m - 1$. Portanto, $\widetilde{C}(x) - C(x) \equiv 0$. ■

Estamos, agora, em condições de demonstrar o Teorema 1.1. Consideremos, então, o problema eletrostático descrito no Capítulo 1 com cargas fixas negativas e positivas.

Demonstração do Teorema 1.1: Já vimos que a existência do polinômio de Stieltjes com as propriedades desejadas é equivalente à existência de um ponto de máximo $\boldsymbol{x}^* = (x_1^*, x_2^*, \ldots, x_n^*)$ de $T(\boldsymbol{x})$ no simplex $\Xi := \{a_{j-1} < x_1 < \ldots < x_n < a_j\}$. Como já mencionamos, $T(\boldsymbol{x})$ é estritamente positiva em Ξ e se anula na fronteira de Ξ. Assim, o ponto $\boldsymbol{x}^*$ existe.

Vamos, agora, estabelecer a unicidade. Seja, então, $\widetilde{C}(x)$ um polinômio de grau dois, $\widetilde{C}(x) \not\equiv C(x)$, tal que a equação diferencial

$$A(x)z''(x) + 2B(x)z'(x) + \widetilde{C}(x)z(x) = 0$$

também admite uma solução $z(x) = (x - \xi_1) \cdots (x - \xi_n)$, com zeros distintos $\xi_1 < \ldots < \xi_n$ em (a_{j-1}, a_j). Pelo Teorema 3.1, temos que:

1) se $(\widetilde{C}(x) - C(x))/A(x)$ é positivo entre dois zeros x_i e x_{i+1}, consecutivos, com $0 \leq i \leq n$, então $z(x)$ deve ter um zero em (x_i, x_{i+1}). Por outro lado, se $(\widetilde{C}(x) - C(x))/A(x)$ é negativo entre ξ_i e ξ_{i+1}, com $0 \leq i \leq n$, então $y(x)$ deve ter um zero em (ξ_i, ξ_{i+1});

2) $\widetilde{C}(x) - C(x)$ deve mudar de sinal pelo menos uma vez em (a_{j-1}, a_j).

Já que podemos nos restringir às soluções mônicas da equação de Lamé, os coeficientes dos termos de maior grau de $\widetilde{C}(x)$ e $C(x)$ são iguais e, então, $\widetilde{C}(x) - C(x)$ é uma função linear. Para demonstrar a unicidade do polinômio de Van Vleck, mostraremos que $\widetilde{C}(x) - C(x)$ tem pelo menos dois zeros, o que é possível somente se $\widetilde{C}(x) \equiv C(x)$.

Suponhamos que $\widetilde{C}(x) \not\equiv C(x)$. Então, por **2)**, $\widetilde{C}(x) - C(x)$ tem pelo menos um zero em (a_{j-1}, a_j) e, por **1)**, concluímos que os zeros de $y(x)$ e $z(x)$ se entrelaçam.

Observe que podemos escrever $y(x) = z(x) + q_{n-1}(x)$, onde $q_{n-1}(x) \in \mathbb{P}_{n-1}$. Então,

$$\begin{aligned}\frac{y(x)}{z(x)} &= 1 + \frac{q_{n-1}(x)}{z(x)} = 1 + \frac{1}{z(x)} \sum_{k=1}^{n} \frac{z(x)}{(x-\xi_k)z'(\xi_k)} q_{n-1}(\xi_k) \\ &= 1 + \sum_{k=1}^{n} \frac{q_{n-1}(\xi_k)/z'(\xi_k)}{x-\xi_k} = 1 + \sum_{k=1}^{n} \frac{\delta_k}{x-\xi_k},\end{aligned}$$

onde $\delta_k = \dfrac{q_{n-1}(\xi_k)}{z'(\xi_k)}$.

Portanto, os coeficientes δ_k da decomposição acima têm todos o mesmo sinal. Como

$$\left[\frac{y(x)}{z(x)}\right]' = \left[1 + \sum_{k=1}^{n} \frac{\delta_k}{x-\xi_k}\right]' = -\sum_{k=1}^{n} \frac{\delta_k}{(x-\xi_k)^2},$$

então

$$\begin{aligned}H(x)\left[y'(x)z(x) - y(x)z'(x)\right] &= H(x)z^2(x)\left[\frac{y'(x)z(x) - y(x)z'(x)}{z^2(x)}\right] \\ &= H(x)z^2(x)\left[\frac{y(x)}{z(x)}\right]' \\ &= -H(x)\sum_{k=1}^{n} \delta_k z_k^2(x), \qquad (3.8)\end{aligned}$$

onde $z_k(x) = \dfrac{z(x)}{x-\xi_k}$. Suponhamos que os δ_k sejam todos positivos (o caso em que todos os δ_k são negativos é análogo). Então, a função

$$G(x) = -H(x)\sum_{k=1}^{n} \delta_k z_k^2(x)$$

é não positiva, se anula somente nos zeros de $H(x)$ e converge para $-\infty$ nos pólos de $H(x)$.

Caso *a)* Observe que, neste caso, $H(x)$ tem dois zeros consecutivos e dois pólos consecutivos, isto é,

$$i) \quad r_0 > 0,\ r_1 > 0,\ r_2 < 0,\ r_3 < 0 \quad ou \quad ii) \quad r_0 < 0,\ r_1 < 0,\ r_2 > 0,\ r_3 > 0.$$

Logo, $H(x)$ terá uma das seguintes formas:

$$H(x) = (x-a_0)^{2r_0}(x-a_1)^{2r_1}\frac{1}{(x-a_2)^{|2r_2|}}\frac{1}{(x-a_3)^{|2r_3|}} \tag{3.9}$$

ou

$$H(x) = \frac{1}{(x-a_0)^{|2r_0|}}\frac{1}{(x-a_1)^{|2r_1|}}(x-a_2)^{2r_2}(x-a_3)^{2r_3}. \tag{3.10}$$

Como $G(x)$ se anula em a_0 e a_1, se $H(x)$ é dada por (3.9) (se anula em a_2 e a_3, se $H(x)$ tem a forma (3.10)), pelo Teorema de Rolle concluímos que $G'(x)$ tem um zero no intervalo (a_0, a_1) $((a_2, a_3))$.

Temos, também, que $G(x)$ converge para $-\infty$ nos pólos a_2 e a_3 (a_0 e a_1). Então, ela tem um ponto de máximo local no intervalo (a_2, a_3) $((a_0, a_1))$. Logo, $G'(x)$ deve mudar de sinal neste intervalo. Mas, de (3.7) e (3.8),

$$G'(x) = \frac{\widetilde{C}(x) - C(x)}{A(x)}y(x)z(x)H(x).$$

Assim, $\widetilde{C}(x) - C(x)$ muda de sinal em (a_0, a_1) e em (a_2, a_3), isto é, tem um zero em cada intervalo. Isto é uma contradição, pois $\widetilde{C}(x) - C(x)$ é linear. Portanto, a unicidade esta demonstrada.

Caso *b)* Neste caso, devemos ter $r_0 < 0$, $r_1 > 0$, $r_2 > 0$ e $r_3 < 0$. Logo,

$$H(x) = \frac{1}{(x-a_0)^{|2r_0|}}(x-a_1)^{2r_1}(x-a_2)^{2r_2}\frac{1}{(x-a_3)^{|2r_3|}}.$$

Como $n > 1-(r_0+r_1+r_2+r_3)$, temos que $G(x)$ tende a $-\infty$ quando x tende a $-\infty$ ou $+\infty$. Mas, a_0 e a_3 são pólos de $H(x)$ e isto significa que $G(x)$ terá um ponto de máximo local entre $-\infty$ e a_0 e, também, entre a_3 e $+\infty$. Assim, temos dois zeros adicionais da derivada de $G(x)$, o que, em vista de (3.7), fornece mais dois zeros de $\widetilde{C}(x) - C(x)$. ∎

3.2 Campos eletrostáticos na presença de restrições e a equação de Lamé modificada

A principal contribuição deste trabalho será tratada nesta seção. Demonstraremos, aqui, o Teorema 1.2, que trata do caso em que as n cargas unitárias livres estão sujeitas a restrições.

Para este propósito, consideremos primeiramente as condições necessárias de 1ª ordem para o correspondente problema extremo.

Sabemos, pela seção anterior, que a energia do campo é dada por $L(\boldsymbol{x})$. Essa energia é mínima se, e somente se, $T(\boldsymbol{x})$, dado em (3.1), é máxima. Como o problema em questão envolve restrições, ele pode ser formulado da seguinte forma:

$$\begin{aligned} &maximizar \;\; T(\boldsymbol{x}) \\ &sujeito \;\; a \;\; R(\boldsymbol{x}) = 0. \end{aligned} \tag{3.11}$$

Sabemos que $T(\boldsymbol{x})$ é uma função contínua e estritamente positiva em Ξ e, portanto, (3.11) é equivalente a

$$\begin{aligned} &maximizar \;\; T(\boldsymbol{x}) \\ &sujeito \;\; a \;\; T(\boldsymbol{x})R(\boldsymbol{x}) = 0. \end{aligned}$$

O método dos multiplicadores de Lagrange pode ser usado para resolver o problema acima. Para isso, vamos definir a função $F(\boldsymbol{x}) = T(\boldsymbol{x}) - \rho T(\boldsymbol{x})R(\boldsymbol{x})$, onde ρ é uma constante. Precisamos, então, encontrar o ponto que satisfaz

$$\frac{\partial F(\boldsymbol{x})}{\partial x_k} = 0, \qquad k = 1, 2, \ldots, n,$$

ou seja,

$$\frac{\partial T(\boldsymbol{x})}{\partial x_k} - \rho \frac{\partial}{\partial x_k}\left[T(\boldsymbol{x})R(\boldsymbol{x})\right] = 0, \qquad k = 1, 2, \ldots, n.$$

Aplicando a regra da cadeia ao segundo termo da equação anterior, obtemos

$$\frac{\partial T(\boldsymbol{x})}{\partial x_k} - \rho\left\{\frac{\partial T(\boldsymbol{x})}{\partial x_k}R(\boldsymbol{x}) + T(\boldsymbol{x})\frac{\partial R(\boldsymbol{x})}{\partial x_k}\right\} = 0.$$

Como, no ponto de máximo, $R(\boldsymbol{x}) = 0$, a equação acima pode ser escrita da forma

$$\frac{\partial T(\boldsymbol{x})}{\partial x_k} - \rho T(\boldsymbol{x})\frac{\partial R(\boldsymbol{x})}{\partial x_k} = 0.$$

De (3.2) e (3.4), a equação anterior torna-se

$$\frac{B(x_k)}{A(x_k)} + \frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)} - \rho\frac{\partial R(\boldsymbol{x})}{\partial x_k} = 0, \tag{3.12}$$

ou seja,

$$A(x_k)\omega''(x_k) + 2\{B(x_k) - \rho D(x_k)\}\omega'(x_k) = 0,$$

lembrando que $D(x_k) = A(x_k)\dfrac{\partial R(\boldsymbol{x})}{\partial x_k}$ (Ver página 5).

O primeiro membro da equação acima é um polinômio de grau $\max\{m+n-1, n+p-1\}$ que se anula nos zeros de $\omega(x)$. Para destacarmos o fato de que os zeros de $\omega(x)$ são $x_1, \ldots, x_n$, usaremos a notação

$$\omega(t; \boldsymbol{x}) := (t - x_1)\cdots(t - x_n).$$

O Teorema Fundamental da Álgebra garante a existência de um polinômio $C(t)$ com grau menor ou igual a $\max\{m-1, p-1\}$, tal que

$$A(t)\omega''(t;\boldsymbol{x}) + 2\{B(t) - \rho D(t)\}\omega'(t;\boldsymbol{x}) + C(t)\omega(t;\boldsymbol{x}) = 0. \tag{3.13}$$

Obviamente, a implicação recíproca é também verdadeira, isto é, se $(C(t), \rho, \omega(t;\boldsymbol{x}))$ é uma tripla para (3.13), onde $\boldsymbol{x} \in \Xi \cap \Omega$, então as condições de 1ª ordem

$$\begin{cases} \dfrac{\partial F(\boldsymbol{x})}{\partial x_k} = 0, \quad k = 1, 2, \ldots, n \\ R(\boldsymbol{x}) = 0 \end{cases} \tag{3.14}$$

são satisfeitas. Portanto, demonstramos o resultado enunciado a seguir.

Lema 3.2 *O ponto $\boldsymbol{x} \in \Xi \cap \Omega$ satisfaz às condições necessárias de 1ª ordem* (3.14) *para o problema extremo* (3.11) *se, e somente se, existe uma tripla* $(C(t), \rho, \omega(t;\boldsymbol{x}))$ *para a equação de Lamé modificada* (3.13).

Estamos, agora, em condições de demonstrar o Teorema 1.2.

Demonstração do Teorema 1.2: Seja $\boldsymbol{x}^*$ um ponto de mínimo da energia $L(\boldsymbol{x})$ sujeito à restrição $R(\boldsymbol{x}) = 0$. Portanto, $\boldsymbol{x}^*$ é uma solução de (3.11). Conseqüentemente, pelo

Teorema 2.10, $\boldsymbol{x}^*$ necessariamente satisfaz (3.14). Então, o Lema 3.2 implica na existência de $(C(t), \rho, \omega(t; \boldsymbol{x}))$ para (3.13).

Para estabelecermos a afirmação *a)*, se $(C(t), \rho, \omega(t; \boldsymbol{x}))$, com $\boldsymbol{x} \in \Xi$ e $R(\boldsymbol{x}) = 0$, satisfaz (3.13), pelo lema anterior, o ponto $\boldsymbol{x}$ é solução de (3.14).

Finalmente, demonstraremos *b)*. Como já observamos, o problema de minimizar $L(\boldsymbol{x})$ em $\Xi \cap \Omega$ é equivalente ao de maximizar $T(\boldsymbol{x})$ na mesma região. Seja $\overline{\Xi}$ o fecho de Ξ. Como $\overline{\Xi}$ é fechado, Ω é um compacto e $\Xi \cap \Omega \neq$, então o conjunto $\overline{\Xi} \cap \Omega$ é um subconjunto compacto e não-vazio do $\mathbb{R}^n$. Pelo Teorema de Weiertrass, a função contínua $T(\boldsymbol{x})$ atinge seu maior valor em um ponto de $\overline{\Xi} \cap \Omega$, digamos $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)$.

Por outro lado, $T(\boldsymbol{x})$ é estritamente positiva em Ξ e anula-se na fronteira $\partial\Xi$. Portanto, $\boldsymbol{\eta}$ é um ponto interior a $\overline{\Xi}$, isto é, $\boldsymbol{\eta} \in \Xi$. Logo, $\boldsymbol{\eta}$ é um ponto de máximo local de $T(\boldsymbol{x})$ em Ξ, sujeito à restrição $R(\boldsymbol{x}) = 0$. Novamente, o Teorema 2.10 mostra que $\boldsymbol{\eta}$ satisfaz (3.14). Finalmente, o Lema 3.2 implica que as coordenadas $\eta_1, \ldots, \eta_n$ do ponto extremo $\boldsymbol{\eta}$ são os zeros de um polinômio de Stieltjes $\omega(t; \boldsymbol{x})$ para (3.13). ■

3.3 Interpretação eletrostática dos zeros de polinômios ortogonais clássicos

Nesta seção, estudaremos o problema da interpretação eletrostática dos zeros dos polinômios de Jacobi, Laguerre e Hermite. Os resultados envolvendo os polinômios de Laguerre e Hermite foram obtidos de maneira diferente daquela feita no texto clássico de Szegő [24].

3.3.1 Interpretação eletrostática dos zeros dos polinômios de Jacobi

Consideremos o campo eletrostático descrito no Capítulo 1, onde cargas p e q positivas são distribuídas uniformemente ao longo de retas perpendiculares ao eixo real e que interceptam os pontos 1 e -1, respectivamente. Tomemos n cargas unitárias distribuídas uniformemente em retas que também são perpendiculares ao eixo real e passam pelos pontos

$x_1, x_2, \ldots, x_n$, onde $-1 < x_1 < x_2 < \ldots < x_n < 1$.

Como já sabemos, a energia do campo é dada por

$$L_1(\boldsymbol{x}) = p\sum_{k=1}^{n} log\frac{1}{|1-x_k|} + q\sum_{k=1}^{n} log\frac{1}{|1+x_k|} + \sum_{1\leq i<k\leq n} log\frac{1}{|x_k - x_i|}.$$

Logo, a energia é mínima se, e somente se,

$$T_1(\boldsymbol{x}) := e^{-L_1(\boldsymbol{x})} = \prod_{k=1}^{n}(1-x_k)^p(1+x_k)^q \prod_{1\leq i<k\leq n}|x_k - x_i| \tag{3.15}$$

é máxima.

Como $T_1(\boldsymbol{x})$ é uma função contínua de $x_1, x_2, \ldots, x_n$ para $-1 \leq x_i \leq 1$, $i = 1, 2, \ldots, n$, é claro que existe um ponto de máximo.

Sabemos que o máximo de (3.15) é obtido nos pontos onde $\partial T_1(\boldsymbol{x})/(\partial x_k) = 0$, $k = 1, 2, \ldots, n$. Observe que este problema é um caso particular do problema tratado na Seção 3.1 para $m = 1$. Logo, fazendo $j = 0, 1$, $a_0 = -1$, $a_1 = 1$, $r_0 = q$ e $r_1 = p$, em (3.4), obtemos

$$\frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)} + \frac{p}{x_k - 1} + \frac{q}{x_k + 1} = 0$$

ou, então,

$$(1 - {x_k}^2)\omega''(x_k) + [2q - 2p - (2q + 2p)x_k]\omega'(x_k) = 0.$$

Portanto, a equação se anula nos zeros de $\omega(x)$. Então,

$$(1 - x^2)\omega''(x) + [2q - 2p - (2q + 2p)x]\omega'(x) = \kappa\omega(x)$$

onde κ é uma constante. Comparando os coeficientes de x^n em ambos os membros, obtemos $\kappa = -n(2p + 2q + n - 1)$. Logo,

$$(1 - x^2)f''(x) + [2q - 2p - (2q + 2p)x]f'(x) + n(2p + 2q + n - 1)f(x) = 0. \tag{3.16}$$

Comparando esta equação com a equação diferencial (2.24) satisfeita pelos polinômios de Jacobi, elas serão iguais se $q = \dfrac{\beta + 1}{2}$ e $p = \dfrac{\alpha + 1}{2}$.

Logo, tomando esses valores para q e p, do Teorema 2.22 temos que a única solução polinomial da equação diferencial acima é $cP_n^{(\alpha,\beta)}(x)$, onde c é constante, ou seja, $\omega(x) = cP_n^{(\alpha,\beta)}(x)$.

Portanto, as coordenadas do ponto de máximo de $T_1(\boldsymbol{x})$ são os zeros de $P_n^{(\alpha,\beta)}(x)$ e, como conseqüência disso, o sistema $-1 < x_1 < x_2 < \ldots < x_n < 1$ é único.

Demonstramos, então, o seguinte resultado.

Teorema 3.2 *Sejam $p > 0$, $q > 0$ e $\{x_i\}_{i=1}^n$, $-1 \leq x_i \leq 1$, $i = 1, 2, \ldots, n$, um sistema de valores para o qual a expressão (3.15) se torna máxima. Então, $\{x_i\}_{i=1}^n$ são os zeros do polinômio de Jacobi, $P_n^{(\alpha,\beta)}(x)$, onde $\alpha = 2p - 1$ e $\beta = 2q - 1$.*

Os zeros dos polinômios de Laguerre, $L_n^{(\alpha)}(x)$, e de Hermite, $H_n(x)$, também apresentam interpretações eletrostáticas semelhantes. É o que veremos a seguir.

3.3.2 Interpretação eletrostática dos zeros dos polinômios de Laguerre

A interpretação eletrostática dos zeros dos polinômios de Laguerre pode ser dada pelo resultado abaixo.

Teorema 3.3 *Consideremos uma carga positiva p fixa no ponto $x = 0$ e cargas unitárias nos pontos variáveis $x_1, x_2, \ldots, x_n$ do intervalo $(0, \infty)$ satisfazendo*

$$n^{-1}(x_1 + x_2 + \cdots + x_n) \leq K, \tag{3.17}$$

onde K é um número positivo pré-determinado. O máximo de

$$T_2(\boldsymbol{x}) = \prod_{k=1}^{n} (x_k)^p \prod_{\substack{i,k=1 \\ i<k}}^{n} |x_k - x_i|,$$

é atingido se, e somente se, os pontos x_i, $i = 1, 2, \ldots, n$, são os zeros do polinômio $L_n^{(\alpha)}(cx)$, onde $L_n^{(\alpha)}(x)$ é o polinômio de Laguerre de grau n, $\alpha = 2p - 1$ e $c = K^{-1}(n + \alpha)$.

Demonstração: A existência e a unicidade da posição de máximo são claras. Pelo Teorema de Weiertrass existe um ponto de máximo, já que $T_2(\boldsymbol{x})$ é uma função contínua num conjunto compacto.

Pelo Lema 2.1, na posição de máximo, vale o sinal de igualdade em (3.17). Portanto, este é um caso particular daquele tratado na seção anterior.

Logo, aplicando o método dos multiplicadores de Lagrange, temos que

$$\frac{\partial T_2(\boldsymbol{x})}{\partial x_k} = \rho \frac{\partial g_1(\boldsymbol{x})}{\partial x_k},$$

onde $g_1(x) = n^{-1}(x_1 + x_2 + \cdots + x_n) - K$ e ρ é um multiplicador apropriado.

Como em (3.12), obtemos

$$\frac{1}{2}\frac{\omega''(x_k)}{\omega'(x_k)} + \frac{p}{x_k} = \frac{\rho}{n}$$

ou, então,

$$x_k\omega''(x_k) + \left(2p - \frac{2\rho}{n}x_k\right)\omega'(x_k) = 0, \qquad k = 1, 2, \ldots, n.$$

Isto significa que

$$x\omega''(x) + \left(2p - \frac{2\rho}{n}x\right)\omega'(x) = a\omega(x),$$

com a é uma constante.

Igualando os coeficientes dos termos de maior grau em ambos os lados da equação acima, obtemos que $a = -2\rho$. Então, chegamos à seguinte equação diferencial de 2ª ordem

$$x\omega''(x) + \left(2p - \frac{2\rho}{n}x\right)\omega'(x) + 2\rho\omega(x) = 0. \tag{3.18}$$

Observe que (3.18) é uma equação de Lamé modificada com

$$\begin{aligned} A(x) &= x, \\ B(x) &= p, \\ C(x) &= 2\rho, \\ D(x) &= 1/n. \end{aligned}$$

Lembremos que os polinômios de Laguerre, $L_n^{(\alpha)}(y)$, satisfazem à equação diferencial

$$yz''(y) + (\alpha + 1 - y)z'(y) + nz(y) = 0, \tag{3.19}$$

onde $z'(y) = \dfrac{dz(y)}{dy}$. Se $y = cx$, onde c é uma constante, obtemos

$$\frac{dz(y(x))}{dy} = \frac{1}{c}\frac{dz(y(x))}{dx} \quad \text{e} \quad \frac{d^2z(y(x))}{dy^2} = \frac{1}{c^2}\frac{d^2z(y(x))}{dx^2}.$$

Substituindo em (3.19), obtemos

$$x\frac{d^2z(cx)}{dx^2} + (\alpha + 1 - cx)\frac{dz(cx)}{dx} + ncz(cx) = 0, \qquad z(cx) = L_n^{(\alpha)}(cx).$$

Portanto, se $\alpha = 2p - 1$ e $c = 2\rho n^{-1}$, comparando (3.19) com (3.18), concluímos que

$$\omega(x) = cte.L_n^{(\alpha)}(cx).$$

A constante c pode ser determinada da condição (3.17) quando vale o sinal de igualdade. De (2.35), sabemos que a soma dos zeros de $L_n^{(\alpha)}(y)$ é igual a $n(n+\alpha)$, ou seja, $c(x_1 + x_2 + \cdots + x_n) = n(n+\alpha)$. Logo, de (3.17) obtemos $c = \dfrac{n+\alpha}{K}$.

Mas, $c = \dfrac{2\rho}{n}$ e, então, $\rho = \dfrac{n(n+\alpha)}{2K}$. ■

3.3.3 Interpretação eletrostática dos zeros dos polinômios de Hermite

Consideremos, agora, o campo eletrostático formado por n cargas unitárias distribuídas uniformemente ao longo de fios infinitos que interceptam o eixo real nos pontos variáveis $x_1, x_2, \ldots, x_n$ pertencentes ao intervalo $(-\infty, \infty)$ e tais que

$$n^{-1}(x_1^2 + x_2^2 + \cdots + x_n^2) \leq M, \tag{3.20}$$

onde M é um número positivo pré-determinado.

A energia do campo, então, é dada por

$$L_3(\boldsymbol{x}) = \sum_{1 \leq i < k \leq n} \log \frac{1}{|x_k - x_i|}.$$

A energia é mínima se, e somente se,

$$T_3(\boldsymbol{x}) = \prod_{\substack{i,k=1 \\ i<k}}^{n} |x_k - x_i| \tag{3.21}$$

é máxima.

A existência e a unicidade do ponto de máximo são claras. Os correspondentes x_i, $i = 1, \ldots, n$, são distintos entre si. Como no caso anterior, pelo Lema 2.1, no ponto de máximo vale o sinal de igualdade em (3.20). Portanto, pelo método dos multiplicadores de Lagrange, se ρ é um multiplicador apropriado, obtemos

$$\frac{\partial T_3(\boldsymbol{x})}{\partial x_k} = \rho \frac{\partial g_2(\boldsymbol{x})}{\partial x_k}$$

onde $g_2(x) = n^{-1}(x_1^2 + x_2^2 + \cdots + x_n^2) - M$.

Procedendo como no caso de Laguerre, chegamos que

$$\frac{1}{2}\frac{\omega''(x)}{\omega'(x)} = \frac{2\rho}{n} x_k \quad \text{ou} \quad \omega''(x_k) - \frac{4\rho}{n} x_k \omega'(x_k) = 0, \quad k = 1, 2, \ldots, n.$$

A última equação acima é, então, igual a $\hat{a}\omega(x)$, onde $\hat{a}$ é uma constante. Comparando os coeficientes de x^n, obtemos $\hat{a} = -4\rho$ e, portanto,

$$\omega''(x) - \frac{4\rho}{n} x\omega'(x) + 4\rho\omega(x) = 0. \tag{3.22}$$

Novamente, observe que (3.22) é uma equação de Lamé modificada com

$$\begin{aligned} A(x) &= 1, \\ B(x) &= 0, \\ C(x) &= 4\rho, \\ D(x) &= 2x/n. \end{aligned}$$

Lembramos que os polinômios de Hermite, $H_n(y)$, satisfazem

$$z''(y) - 2yz'(y) + 2nz(y) = 0, \qquad \text{onde} \quad z'(y) = \frac{dz(y)}{dy}.$$

Fazendo $y = \hat{c}x$ e procedendo como no teorema anterior, chegamos à equação diferencial ordinária $z''(\hat{c}x) - 2\hat{c}^2 x z'(\hat{c}x) + 2n\hat{c}^2 z(\hat{c}x) = 0$, satisfeita por $z(\hat{c}x) = H_n(\hat{c}x)$.

Comparando esta equação com (3.22), concluímos que, se $\hat{c} = \left(\frac{2\rho}{n}\right)^{\frac{1}{2}}$, então $\omega(x) = cteH_n(\hat{c}x)$. A constante $\hat{c}$ pode ser determinada da condição (3.20) tomando-se o sinal de igualdade. Como, de (2.37),

$$\hat{c}^2(x_1^2 + x_2^2 + \cdots + x_n^2) = \frac{n(n-1)}{2},$$

de (3.20) obtemos $\hat{c}^2(nM) = \dfrac{n(n-1)}{2}$. Logo, $\hat{c} = \left(\dfrac{n-1}{2M}\right)^{\frac{1}{2}}$.

Mas, $\hat{c} = \left(\dfrac{2\rho}{n}\right)^{\frac{1}{2}}$ e, então, $\rho = \dfrac{n(n-1)}{4M}$. Demonstramos, então, o seguinte teorema.

Teorema 3.4 *Consideraremos n cargas unitárias nos pontos variáveis $x_1, x_2, \ldots, x_n$ pertencentes ao intervalo $(-\infty, \infty)$ e satisfazendo*

$$n^{-1}(x_1^2 + x_2^2 + \cdots + x_n^2) \leq M,$$

onde M é um número positivo pré-determinado. Então, o máximo de (3.21) é atingido se, e somente se, x_i, $i = 1, 2, \ldots, n$, são os zeros do polinômio $H_n(\hat{c}x)$, onde $H_n(x)$ é o polinômio de Hermite e $\hat{c} = \left(\dfrac{n-1}{2M}\right)^{\frac{1}{2}}$.

Capítulo 4

Interpretação eletrostática dos zeros dos polinômios de Laurent ortogonais

Neste capítulo, estudaremos a interpretação eletrostática dos zeros dos L-polinômios ortogonais. Para isso, uma relação entre os polinômios ortogonais simétricos e os L-polinômios ortogonais será apresentada.

4.1 Relação entre os polinômios ortogonais simétricos e os L-polinômios ortogonais

Sri Ranga em [20] demonstrou um criativo modo de se obter seqüências de L-polinômios ortogonais, através de uma mudança de variáveis em uma seqüência $\{\widehat{P}_n(x)\}$ de polinômios ortogonais mônicos com relação a uma função peso par num intervalo simétrico $(-d, d)$. É o que mostraremos a seguir.

Seja $Y(n)$ o espaço de todos os polinômios reais mônicos, $p_n(x)$, de grau n que satisfazem à propriedade $p_n(x) = (-1)^n p_n(-x)$. Consideremos, ainda, o espaço de todos polinômios reais mônicos, $q_n(t)$, de grau n que satisfazem à propriedade de simetria $q_n(t) = \dfrac{t^n q_n(\beta^2/t)}{(-\beta)^n}$ e que denotaremos por $Z(n; \beta)$.

Suponhamos que $\omega(x)$ seja uma função peso par em $(-d, d)$, $0 < d \leq \infty$, isto é,

$\omega(x) = \omega(-x)$. Logo, pelo Teorema 2.20, os polinômios ortogonais mônicos associados a $\omega(x)$, $\widehat{P}_n(x)$, pertencem a $Y(n)$, $n \geq 0$. Consideremos, agora, $\nu(t)$ uma função peso forte definida em $(\beta^2/b, b)$, $0 < \beta < b \leq \infty$, satisfazendo

$$\sqrt{t}\ \nu(t) = \sqrt{\beta^2/t}\ \nu(\beta^2/t).$$

Podemos facilmente mostrar que os polinômios similares aos ortogonais $B_n(t)$, $n \geq 0$, associados a $\nu(t)$ em $(\beta^2/b, b)$ são tais que $B_n(t) \in Z(n; \beta)$.

Tomemos, agora, $\alpha > 0$ e $\beta > 0$. O teorema abaixo fornece uma condição necessária e suficiente para que os polinômios $p_n(x) \in Y(n)$ sejam ortogonais em $(-d, d)$ com relação a $\omega(x)$.

Teorema 4.1 *Seja a seqüência de polinômios $\{p_n(x)\}_{n=0}^{\infty}$, tal que $p_n(x) \in Y(n)$. Então,*

$$p_n(x) = \widehat{P}_n(x), \qquad n \geq 0,$$

se, e somente se,

$$\int_{-d}^{d} \left\{\sqrt{\alpha x^2 + \beta} + \sqrt{\alpha} x\right\}^{-(n-1)+2s} \frac{p_n(x)}{\sqrt{\alpha x^2 + \beta}} \omega(x) dx = 0, \qquad 0 \leq s \leq n-1. \tag{4.1}$$

Demonstração: Como $\{\sqrt{\alpha x^2 + \beta} + \sqrt{\alpha} x\}^{-1} = \{\sqrt{\alpha x^2 + \beta} - \sqrt{\alpha} x\}/\beta$, (4.1) é equivalente a

$$\int_{-d}^{d} \left\{\sqrt{\alpha x^2 + \beta} \pm \sqrt{\alpha} x\right\}^{2l} \frac{p_{2m+1}(x)}{\sqrt{\alpha x^2 + \beta}} \omega(x) dx = 0$$

e

$$\int_{-d}^{d} \left\{\sqrt{\alpha x^2 + \beta} \pm \sqrt{\alpha} x\right\}^{2l+1} \frac{p_{2m+2}(x)}{\sqrt{\alpha x^2 + \beta}} \omega(x) dx = 0,$$

para $l = 0, 1, \ldots, m$ e $m \geq 0$.

Mas,

$$\left\{\sqrt{\alpha x^2 + \beta} \pm \sqrt{\alpha} x\right\}^{k} = \sum_{i=0}^{k} \binom{k}{i} \left(\sqrt{\alpha x^2 + \beta}\right)^{k-i} \left(\pm\sqrt{\alpha} x\right)^{i},$$

para $k \in \mathbb{N}$.

Logo, usando a propriedade de simetria de $p_n(x)$, as integrais anteriores podem ser escritas, respectivamente, como

$$\int_{-d}^{d}\left[\sum_{r=0}^{l-1}\binom{2l}{2r+1}\{\alpha x^2+\beta\}^{l-r-1}\{\pm\sqrt{\alpha}x\}^{2r+1}\right]p_{2m+1}(x)\omega(x)dx=0$$

e (4.2)

$$\int_{-d}^{d}\left[\sum_{r=0}^{l}\binom{2l+1}{2r}\{\alpha x^2+\beta\}^{l-r}\{\pm\sqrt{\alpha}x\}^{2r}\right]p_{2m+2}(x)\omega(x)dx=0,$$

para $l=0,1,\ldots,m$ e $m\geq 0$.

Agora, se $p_n(x)=\widehat{P}_n(x)$ para $n\geq 0$, então (4.2) deve valer, pois $\widehat{P}_n(x)$ é ortogonal a todos os polinômios de grau menor ou igual a n.

Reciprocamente, se (4.2) vale, então $p_{2m+1}(x)$ é ortogonal aos polinômios ímpares

$$\sum_{r=0}^{l-1}\binom{2l}{2r+1}\{\alpha x^2+\beta\}^{l-r-1}\{\sqrt{\alpha}x\}^{2r+1},$$

de grau precisamente $2l-1$, para $l=0,1,\ldots,m$. Como esses polinômios formam uma base para todos os polinômios ímpares de grau menor ou igual a $2m-1$, $p_{2m+1}(x)$ deve ser ortogonal a todos polinômios ímpares de grau menor ou igual a $2m-1$. Portanto, como $p_{2m+1}(x)$ é mônico e, também, ímpar, obtemos $p_{2m+1}(x)=\widehat{P}_{2m+1}(x)$. De modo análogo, mostramos que $p_{2m}(x)=\widehat{P}_{2m}(x)$ e a demonstração do teorema está completa. ■

Consideremos, agora, a transformação

$$t(x)=\left\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\right\}^2,\qquad x\in(-\infty,\infty),$$

que representa uma correspondência biunívoca entre $(-\infty,\infty)$ e $(0,\infty)$. A inversa de $t(x)$ é dada por

$$x(t)=\frac{1}{2\sqrt{\alpha}}\left(\sqrt{t}-\frac{\beta}{\sqrt{t}}\right),\qquad t\in(0,\infty).\tag{4.3}$$

Se $x=d$ corresponde a $t=b$, isto é, $\sqrt{b}=\sqrt{\alpha d^2+\beta}+\sqrt{\alpha}d$, então é fácil ver que $x=-d$ corresponde a $t=\beta^2/b$. Temos, assim, o seguinte resultado.

Teorema 4.2 *Considere a transformação $x(t)$ dada em (4.3). Sejam $p(x)$ e $q(x)$ polinômios mônicos satisfazendo $q(t) = (2\sqrt{\alpha t})^n p(x(t))$. Então,*

$$p(x) \in Y(n) \Leftrightarrow q(t) \in Z(n;\beta).$$

Demonstração: Já que $x(\beta^2/t) = -x(t)$, segue que

$$\frac{t^n q(\beta^2/t)}{(-\beta)^n} = \frac{t^n}{(-\beta)^n}\left(2\sqrt{\frac{\alpha\beta^2}{t}}\right)^n p(-x(t)) = (-1)^n(2\sqrt{\alpha t})^n p(-x(t)).$$

Portanto, $p(x) = (-1)^n p(-x)$ se, e somente se, $q(t) = t^n q(\beta^2/t)/(-\beta)^n$. Assim, precisamos mostrar que se $p(x)$ é um polinômio mônico de grau n, então $q(t)$ também é, e vice-versa.

Observe que se $q(t) = \sum_{i=0}^{n} b_i t^i$, onde $b_n = 1$, então, podemos escrever

$$q(t) = \begin{cases} \displaystyle\sum_{i=\lfloor\frac{n+1}{2}\rfloor}^{n} [b_i t^i + b_{n-i}t^{n-i}], & \text{se } n \text{ é impar} \\ \displaystyle\sum_{i=\lfloor\frac{n+1}{2}\rfloor}^{n} [b_i t^i + b_{n-i}t^{n-i}] - b_{\lfloor\frac{n+1}{2}\rfloor} t^{\lfloor\frac{n+1}{2}\rfloor}, & \text{se } n \text{ é par} \end{cases}.$$

Se $q(t) \in Z(n;\beta)$, então $b_{n-i} = (-\beta)^{2i-n} b_i$. Logo,

$$q(t) = \begin{cases} \displaystyle\sum_{i=\lfloor\frac{n+1}{2}\rfloor}^{n} b_i[t^i + (-\beta)^{2i-n}t^{n-i}], & \text{se } n \text{ é impar} \\ \displaystyle\sum_{i=\lfloor\frac{n+1}{2}\rfloor}^{n} b_i[t^i + (-\beta)^{2i-n}t^{n-i}] - b_{\lfloor\frac{n+1}{2}\rfloor} t^{\lfloor\frac{n+1}{2}\rfloor}, & \text{se } n \text{ é par} \end{cases}.$$

Mas, para $i = \lfloor(n+1)/2\rfloor, \ldots, n$

$$\begin{aligned} \left[2\sqrt{\alpha t(x)}\right]^{-n} b_i\{[t(x)]^i + (-\beta)^{2i-n}[t(x)]^{n-i}\} \\ &= \left[2\sqrt{\alpha}\right]^{-n} b_i\{[\sqrt{t(x)}]^{2i-n} + (-\beta)^{2i-n}[\sqrt{t(x)}]^{n-2i}\} \\ &= \left[2\sqrt{\alpha}\right]^{2(i-n)} b_i \left(\frac{1}{2\sqrt{\alpha}}\right)^{2i-n} \left\{[\sqrt{t(x)}]^{2i-n} + (\frac{-\beta}{\sqrt{t(x)}})^{2i-n}\right\} \\ &= (4\alpha)^{i-n} b_i x^{2i-n} + r(x), \end{aligned}$$

onde $r(x)$ é um polinômio de grau menor que $2i-n$. Como $p(x) = [2\sqrt{\alpha t(x)}]^{-n}q(t(x))$, chegamos ao resultado desejado.

Por outro lado, se $p(x) \in Y(n)$, então, para $n = 2m$, podemos escrever $p(x) = \sum_{i=0}^{m} a_{2i}x^{2i}$, onde $a_{2m} = 1$. Portanto, $(2\sqrt{\alpha t})^{2m}a_{2i}[x(t)]^{2i}$ é o polinômio

$$(4\alpha)^{m-i}a_{2i}t^{m-i}(t-\beta)^{2i},$$

para $i = 0, 1, \ldots, m$. Analogamente, para $n = 2m+1$, podemos escrever $p(x) = \sum_{i=0}^{m} a_{2i+1}x^{2i+1}$, onde $a_{2m+1} = 1$. Portanto, $(2\sqrt{\alpha t})^{2m+1}a_{2i+1}[x(t)]^{2i+1}$ é o polinômio

$$(4\alpha)^{m-i}a_{2i+1}t^{m-i}(t-\beta)^{2i+1},$$

para $i = 0, 1, \ldots, m$. ■

Veremos, agora, um resultado que relaciona as funções peso.

Teorema 4.3 *Sejam b e d tais que $\sqrt{b} = \sqrt{\alpha d^2 + \beta} + \sqrt{\alpha}d$ e*

$$V(t) = At^{-1/2}W(x(t)),$$

onde A é uma constante positiva. Então, $W(x)$ é uma função peso em $(-d, d)$ tal que $W(x) = W(-x)$ se, e somente se, $V(t)$ é uma função peso forte em $(\beta^2/b, b)$ tal que $\sqrt{t}\ V(t) = \sqrt{\beta^2/t}\ V(\beta^2/t)$.

Demonstração: Podemos verificar facilmente que $W(x)$ é positiva e satisfaz $W(x) = W(-x)$ em $(-d, d)$ se, e somente se, $V(t)$ é positiva e satisfaz $\sqrt{t}\ V(t) = \sqrt{\beta^2/t}\ V(\beta^2/t)$ em $(\beta^2/b, b)$.

Portanto, precisamos apenas mostrar que se os momentos (incluindo os negativos para $V(t)$) existem e são finitos para uma das funções peso, o mesmo é verdadeiro para a outra.

Temos que

$$\begin{aligned}
\int_{\beta^2/b}^{b} t^i V(t)dt &= \int_{-d}^{d} \Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{2i} A\Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{-1} 2\Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\} \\
&\quad \times \left\{\frac{2\alpha x}{\sqrt{\alpha x^2+\beta}}+\sqrt{\alpha}\right\} W(x)dx \\
&= 2A\int_{-d}^{d} \Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{2i} \left\{\frac{\alpha x+\sqrt{\alpha x^2+\beta}\sqrt{\alpha}}{\sqrt{\alpha x^2+\beta}} W(x)dx\right\} \\
&= 2A\sqrt{\alpha}\int_{-d}^{d} \Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{2i+1} \frac{W(x)}{\sqrt{\alpha x^2+\beta}} dx.
\end{aligned}$$

Mas, se $i \in \mathbb{Z}$,

$$\frac{\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\}^{2i+1}}{\sqrt{\alpha x^2+\beta}} = \begin{cases} \displaystyle\sum_{j=0}^{i} \binom{i}{j} (\alpha x^2+\beta)^{i-j}(\alpha x^2)^j, & \text{se} \quad i \geq 0, \\ \displaystyle\sum_{j=0}^{|i|-1} \binom{|i|-1}{j} (\alpha x^2+\beta)^{|i|-1-j}(\alpha x^2)^j, & \text{se} \quad i \leq -1. \end{cases}$$

Como $W(x) = W(-x)$, então

$$\int_{\beta^2/b}^{b} t^i V(t)dt = \int_{-d}^{d} S_i(x)W(x)dx,$$

onde $S_i(x)$ é um polinômio par de grau exatamente $2i$ para $i \geq 0$ e é um polinômio par de grau exatamente $|2i|-2$ para $i \leq -1$. Como os monômios x^{2i} podem ser expressos como combinação linear dos polinômios $S_i(x)$, e vice-versa, o teorema está demonstrado. ∎

Apresentaremos, agora, a relação entre os polinômios ortogonais simétricos e os L-polinômios ortogonais.

Teorema 4.4 *Sejam $W(x)$ e $V(t)$ um par de funções peso dadas pelo Teorema 4.3 Então, para $n \geq 0$,*

$$B_n(t) = (2\sqrt{\alpha t})^n \widehat{P}_n(x(t)).$$

Demonstração: Pelos Teorema 4.1 e 4.3, segue que

$$\begin{aligned}\int_{\beta^2/b}^{b} t^{-n+s}(2\sqrt{\alpha t})^n \widehat{P}_n(x(t))V(t)dt &= \int_{-d}^{d}\Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{-n+2s} A(2\sqrt{\alpha})^n \\ &\times \frac{1}{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x}\widehat{P}_n(x(t)) \\ &\times 2\Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^2\sqrt{\alpha}\frac{W(x)dx}{\sqrt{\alpha x^2+\beta}} \\ &= 2A\sqrt{\alpha}(2\sqrt{\alpha})^n\int_{-d}^{d}\Big\{\sqrt{\alpha x^2+\beta}+\sqrt{\alpha}x\Big\}^{-(n-1)+2s} \\ &\times \frac{\widehat{P}_n(x(t))}{\sqrt{\alpha x^2+\beta}}W(x)dx = 0, \qquad 0 \le s \le n-1.\end{aligned}$$

Portanto, de (2.41) e do Teorema 4.2, o resultado do teorema segue imediatamente. ∎

4.2 Interpretação eletrostática dos zeros dos polinômios de Gegenbauer-Laurent

Considere os polinômios de Gegenbauer (ultrasféricos) clássicos, $C_n^{(\lambda)}(x)$, ortogonais em $[-1,1]$ com relação à função peso $(1-t^2)^{\lambda-1/2}$. Como vimos na seção anterior, os polinômios

$$B_n^{(\lambda)}(t) = C_n^{(\lambda)}(x(t))(2\sqrt{\alpha t})^n, \tag{4.4}$$

onde $x(t)$ é dado em (4.3), são L-polinômios ortogonais com relação à função peso forte $\nu(t) = t^{-\lambda}(b-t)^{\lambda-1/2}(t-a)^{\lambda-1/2}$ no intervalo (a,b), $0<a<b$, onde $a=\beta^2/b$ e $2\sqrt{\alpha} = \sqrt{b}-\sqrt{a}$. A função peso $\nu(t)$ é facilmente obtida do Teorema 4.3.

Temos, ainda, que

$$\begin{aligned}(1-x^2) &= \frac{(b-t)(t-a)}{(\sqrt{b}-\sqrt{a})^2 t}, \\ \frac{dx(t)}{dt} &= \frac{1}{\sqrt{b}-\sqrt{a}}\left(\frac{t+\sqrt{ab}}{2t\sqrt{t}}\right), \\ \frac{d^2x(t)}{dt^2} &= \frac{-1}{\sqrt{b}-\sqrt{a}}\left(\frac{3\sqrt{ab}+t}{4t^2\sqrt{t}}\right).\end{aligned}$$

Considerando $z(t) = B_n^{(\lambda)}(t)$ e $y(x) = C_n^{(\lambda)}(x)$, de (4.4) obtemos

$$\begin{aligned} z'(t) &= (\sqrt{b}-\sqrt{a})^n \frac{d}{dt}[y(x)t^{n/2}] = (\sqrt{b}-\sqrt{a})^n \left[y'(x)\frac{dx(t)}{dt}t^{n/2} + y(x)\frac{nt^{n/2}}{2t} \right] \\ &= (\sqrt{b}-\sqrt{a})^n t^{n/2}\frac{dx(t)}{dt}y'(x) + \frac{n}{2t}z(t). \end{aligned}$$

Daí,

$$y'(x) = (\sqrt{b}-\sqrt{a})^{-n}t^{-n/2}(\sqrt{b}-\sqrt{a})\left(\frac{2t\sqrt{t}}{t+\sqrt{ab}}\right)\left[z'(t) - \frac{n}{2t}z(t)\right]. \tag{4.5}$$

Como

$$z''(t) = (\sqrt{b}-\sqrt{a})^n t^{n/2}\left[y''(x)\left(\frac{dx}{dt}\right)^2 + \left(\frac{n}{t}\frac{dx}{dt} + \frac{d^2x}{dt^2}\right)y'(x) + \left(\frac{n^2}{4t^2} - \frac{n}{2t^2}\right)y(x)\right],$$

de (4.5) concluímos que

$$\begin{aligned} y''(x) &= (\sqrt{b}-\sqrt{a})^{-n}t^{-n/2}(\sqrt{b}-\sqrt{a})^2\left(\frac{2t\sqrt{t}}{t+\sqrt{ab}}\right)^2\left\{ z''(t) - \left[\frac{2n(t+\sqrt{ab}) - (3\sqrt{ab}+t)}{2t(t+\sqrt{ab})}\right]\right. \\ &\quad \left.\times\left(z'(t) - \frac{n}{2t}z(t)\right) - \left(\frac{n^2}{4t^2} - \frac{n}{2t^2}\right)z(t)\right\}. \end{aligned} \tag{4.6}$$

Substituindo (4.3), (4.5) e (4.6) em (2.33), isto é, na equação diferencial satisfeita pelos polinômios de Gegenbauer e, em seguida, multiplicando por $[(t+\sqrt{ab})^3(\sqrt{b}-\sqrt{a})^{-n}t^{-n/2}]/(2t)$, obtemos a equação diferencial

$$A(t)z''(t) + 2B(t)z'(t) + C(t)z(t) = 0,$$

com

$$\begin{aligned} A(t) &= 2t(b-t)(t-a)(t+\sqrt{ab}), \\ B(t) &= (n-\lambda-1)t^3 - \big[(n-1/2)(a+b) - (n-\lambda-2)\sqrt{ab}\big]t^2 \\ &\quad -\big[(n-3/2)(a+b)\sqrt{ab} - (n+\lambda)ab\big]t + (n+\lambda-1)(ab)^{3/2}, \\ C(t) &= 2\lambda nt^2 + n\big[(n+4\lambda+1)\sqrt{ab} + (n+1)(a+b)/2\big]t \\ &\quad +n(n+2\lambda-1)ab + n(n-1)(a+b)\sqrt{ab}/2. \end{aligned}$$

A decomposição em frações parciais de $B(t)/A(t)$ é dada por

$$\frac{B(t)}{A(t)} = \frac{2\lambda+1}{4}\frac{1}{t-a} + \frac{2\lambda+1}{4}\frac{1}{t-b} - \frac{n+\lambda-1}{2}\frac{1}{t} - \frac{1}{2}\frac{1}{t+\sqrt{ab}}.$$

Do Teorema 1.1 *a)* obtemos, imediatamente, o resultado a seguir.

Teorema 4.5 *Considere o campo eletrostático que obedece à lei do potencial logarítmico e é gerado por duas cargas positivas de mesmo valor* $\frac{2\lambda+1}{4}$ *nos pontos* a *e* b, $0 < a < b$, *e cargas negativas* $-1/2$ *no ponto* $-\sqrt{ab}$ *e* $-(n+\lambda-1)/2$ *na origem. Então, os zeros do polinômio de Gegenbauer-Laurent,* $B_n^{(\lambda)}(t)$, *são os únicos pontos de equilíbrio de* n *cargas unitárias móveis em* (a,b).

As duas cargas negativas, uma em $-\sqrt{ab}$ e a outra na origem, estão, ambas, à esquerda do intervalo $[a,b]$ e atraem os zeros de $B_n^{(\lambda)}(t)$. Isto significa que os zeros estão mais concentrados perto de a do que de b e sua distribuição assintótica não é uma distribuição arcsen em $[a,b]$. De fato, Sri Ranga [20] demonstrou que os zeros de $B_n^{(\lambda)}(t)$ são simétricos com relação a $\sqrt{ab}$ no seguinte sentido: se t_k é um zero de $B_n^{(\lambda)}(t)$, então ab/t_k é também um zero.

4.3 Interpretação eletrostática dos zeros dos polinômios de Hermite-Laurent

Lembremos que os polinômios de Hermite, $H_n(x)$, são ortogonais em $(-\infty,\infty)$ com relação à função peso $\omega(x) = e^{-x^2}$. Então, os polinômios

$$B_n^H(t) = H_n(x(t))(2\sqrt{\alpha t})^n \tag{4.7}$$

são L-polinômios ortogonais em $(0,\infty)$, com relação à função peso forte $\nu(t) = t^{-1/2}exp\left(-\frac{(t+\beta^2/t)}{4\alpha}\right)$. A função peso $\nu(t)$ também pode facilmente ser obtida do Teorema 4.3.

Vamos considerar, aqui, idéia similar à usada por Dimitov e Van Assche em [6] para o estudo da interpretação eletrostática dos zeros dos polinômios de Gegenbauer-Laurent, que apresentamos na seção anterior.

Consideremos $z(t) = B_n^H(t)$ e $y(x) = H_n(x)$. De (4.7), concluimos que

$$z'(t) \quad = \quad (2\sqrt{\alpha t})^n \frac{dx(t)}{dt} y'(x) + \frac{n}{2t} z(t).$$

Daí, utilizando a transformação (4.3), temos que

$$y'(x) = (2\sqrt{\alpha} t)^{-n} 2(\sqrt{\alpha}) \left(\frac{2t\sqrt{t}}{t+\beta} \right) \left[z'(t) - \frac{n}{2t} z(t) \right]. \tag{4.8}$$

Como

$$z''(t) \quad = \quad (2\sqrt{\alpha t})^n \left[y''(x) \left(\frac{dx}{dt} \right)^2 + \left(\frac{n}{t} \frac{dx}{dt} + \frac{d^2 x}{dt^2} \right) y'(x) + \left(\frac{n^2}{4t^2} - \frac{n}{2t^2} \right) y(x) \right],$$

de (4.8) e (4.3), obtemos

$$\begin{aligned} y''(x) \quad &= \quad (2\sqrt{\alpha t})^{-n} (2\sqrt{\alpha})^2 \left(\frac{2t\sqrt{t}}{t+\beta} \right)^2 \left\{ z''(t) - \left[\frac{2n(t+\beta) - (3\beta + t)}{2t(t+\beta)} \right] \right. \\ & \left. \times \left(z'(t) - \frac{n}{2t} z(t) \right) - \left(\frac{n^2}{4t^2} - \frac{n}{2t^2} \right) z(t) \right\}. \end{aligned} \tag{4.9}$$

Substituindo (4.3), (4.8) e (4.9) na equação diferencial de 2ª ordem (2.38) satisfeita pelos polinômios de Hermite e, em seguida, multiplicando por $(2\sqrt{\alpha t})^n (t+\beta)^3/4t$, concluímos que $z(t) = B_n^H(t)$ é solução da equação diferencial de Lamé modificada

$$A(t) z''(t) + 2 \left\{ B(t) - \frac{n}{2} D(t) \right\} z'(t) + C(t) z(t) = 0,$$

onde

$$\begin{aligned} A(t) \quad &= \quad 4\alpha t^2 (t+\beta), \\ B(t) \quad &= \quad -(2n-1)\alpha t^2 - (2n-3)\alpha\beta t, \\ C(t) \quad &= \quad n t^2 + n(n\alpha + \alpha + 2\beta) t + n\beta \big[\alpha(n-1) + \beta \big], \\ D(t) \quad &= \quad n^{-1} (t^3 + \beta t^2 - \beta^2 t - \beta^3). \end{aligned}$$

A decomposição de $B(t)/A(t)$ em frações parciais é dada por

$$\frac{B(t)}{A(t)} = -\left(\frac{2n-3}{4} \right) \frac{1}{t} - \frac{1}{2} \frac{1}{t+\beta}.$$

Aplicando, agora, a transformação (4.3) à restrição (3.20), pelo Teorema 1.2 obtemos o seguinte resultado.

Teorema 4.6 *Considere o campo eletrostático que obedece à lei do potencial logarítmico e é gerado por duas cargas negativas,* $-\left(\dfrac{2n-3}{4}\right)$ *em* 0 *e* $-\dfrac{1}{2}$ *em* $-\beta$, $\beta > 0$, *e* n *cargas unitárias livres nos pontos* $0 < t_1 < \ldots < t_n < \infty$, *sujeitos à restrição*

$$\frac{1}{4\alpha n}\sum_{k=1}^{n}\left(t_k + \frac{\beta^2}{t_k}\right) = N, \qquad \alpha > 0.$$

Então, o ponto $\boldsymbol{t} = (t_1, \ldots, t_n)$, *onde* t_k, $k = 1, \ldots, n$, *são os zeros do polinômio de Hermite-Laurent,* $B_n^H(t)$, *é o ponto de mínimo local para a energia do campo acima descrito.*

Como no caso dos polinômios de Gegenbauer-Laurent, aqui, também, as duas cargas negativas, uma em $-\beta$ e outra na origem, atraem os zeros de $B_n^H(t)$, que são simétricos com relação a β, isto é, se t_k é zero de $B_n^H(t)$, então β^2/t_k é também um zero. Isto significa que os zeros estão mais concentrados perto da origem.

Referências Bibliográficas

[1] ALAM, M., Zeros of Stieltjes and Van Vleck polinomials. *Trans. Amer. Math. Soc.*, **252**, 197–204, 1979.

[2] ANDRADE, E. X. L., *Sobre Polinômios Similares aos Ortogonais Associados a uma Classe Especial de Distribuições,* Tese de Doutorado, UNICAMP, Campinas, SP, 1995.

[3] BERTESEKAS, D. P., *Nonlinear Programming*. Belmont, MA, Athena Scientific, 1995.

[4] BÔCHER, M., The roots of polynomials that satisfy certain differential equations of the second order, *Bull. Amer. Math. Soc.*, **4**, 256–258, 1897.

[5] CHIHARA, T. S., *An Indroduction to Orthogonal Polynomials*, Gordon and Breach, New York, 1978.

[6] DIMITROV, D. K. and VAN ASSCHE, W., Lamé differential equations and electrostatics, *Proc. Amer. Math. Soc.*, **128**, 3621–3628, 2000.

[7] GRÜNBAUM, F. A., Variations on a theme Heine and Stieltjes: an electrostatic interpretation of the zeros of certain polynomials, *J. Comput. Appl. Math.,* **99**, 189–194, 1998.

[8] GRÜNBAUM, F. A., Electrostatic interpretation for the zeros of certain polynomials and the Darboux process, *J. Comput. Appl. Math.,* **133**, 397–412, 2001.

[9] HEINE, E., *Handbuch der Kugelfunctionen*, vol I, 2ª ed., Berlim, 1878.

[10] ISMAIL, M. E. H., An electrostatics model for zeros of general orthogonal polynomials, *Pacific J. Math.*, **193**, 355–369, 2000.

[11] JONES, W. B., THRON, W. J. and WAADELAND, H., A strong Stieltjes moment problem, *Trans. Amer. Math. Soc.*, **261**, 503–528, 1980.

[12] JONES, W. B., NJASTAD, O. and THRON, W. J., Two-point Padé expansions for a family of analytic functions, *J. Comput. Appl. Math.*, **9**, 105–126, 1983.

[13] KLEIN, F., Über die Nullstellen von den Polynomen und den Potenzreihen, Göttingen, 1894, pp. 211-218.

[14] LIMA, E. L., *Curso de Análise*, Volume 2, Segunda edição, IMPA, Rio de Janeiro, 1981.

[15] MARDEN, M., On Stieltjes polynomials, *Trans. Amer. Math. Soc.*, **33**, 934–944, 1931.

[16] MARDEN, M., *Geometry of Polynomials*, Math. Surveys 3, Amer. Math. Soc., Providence, R.I. 1966.

[17] PÓLYA, G., Sur un théorème de Stiltjes., *C. R. Acad. Sci. Paris*, **155**, 767–769, 1912.

[18] SRI RANGA, A., *Continued fractions with correspond to two series expansions, and the strong Hamburger moment problem*. Tese de doutorado, Univ. of St. Andrews, St. Andrews, Escócia, 1984.

[19] SRI RANGA, A., Another quadrature rule of highest algebraic degree of precision, *Numer. Math.*, **68**, 283–294, 1994.

[20] SRI RANGA, A., Symmetric orthogonal polynomials and the associated orthogonal L-polynomials, *Proc. Amer. Math. Soc.*, **123**, 3135–3141, 1995.

[21] STIELTJES, T. J., Sur les quelques théorémes d'algébre, *C.R. Acad. Sci. Paris*, **100**, 439–440, 1885.

[22] STIELTJES, T. J., Sur les polynômes de Jacobi, *C.R. Acad. Sci. Paris*, **100**, 620–622, 1885.

[23] STIELTJES, T. J., Sur les racines de l'équation $X_n = 0$, *Acta Math.*, **9**, 385–400, 1886.

[24] SZEGŐ, G., *Orthogonal polynomials*, 4th ed., Amer. Math. Soc. Coll. Publ., Vol. 23, Providence, RI, 1975.

[25] VALENT, G. and VAN ASSCHE, W., The impact of Stieltjes' work on continued fraction and orthogonal polynomials: additional material, *J. Comput. Appl. Math.*, **65**, 419–447, 1995.

[26] VAN ASSCHE, W., The impact of Stieltjes' work on continued fraction and orthogonal polynomials, em "T. J. Stieltjes: Collected papers, Vol. I" (G. van Dijk, ed.), pp. 5-37, Springer-Verlag, Berlin, 1993.

[27] VAN VLECK, E. B., On the polynomials of Stieltjes, *Bull. Amer. Math. Soc.*, **4**, 426–438, 1898.