

unesp



UNIVERSIDADE ESTADUAL PAULISTA

”JÚLIO DE MESQUITA FILHO”

Campus de São José do Rio Preto

Fabio Nosse Niime

Polinômios Ortogonais em Várias Variáveis

São José do Rio Preto

2011

Fabio Nosse Niime

Polinômios Ortogonais em Várias Variáveis

Dissertação apresentada para obtenção do título de Mestre em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

Orientador: Prof^ª. Dr^ª. Cleonice Fátima Bracciali

São José do Rio Preto

2011

Fabio Nosse Niime

Polinômios Ortogonais em Várias Variáveis

Dissertação apresentada para obtenção do título de Mestre em Matemática do Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Campus de São José do Rio Preto.

BANCA EXAMINADORA

Profa. Dra. Cleonice Fátima Bracciali
Professor Adjunto
UNESP - São José do Rio Preto
Orientadora

Prof. Dr. Fernando Rodrigo Rafaeli
Professor Assistente Doutor
UNESP - Presidente Prudente

Profa. Dra. Eliana Xavier Linhares de Andrade
Professor Adjunto
UNESP- São José do Rio Preto

São José do Rio Preto

24/fevereiro/2011

Dedico este trabalho

Aos meus pais, Calixto e Olga, e ao meu amor, Jaqueline, pelo apoio e pelo amor.

Agradecimentos

Primeiramente, e principalmente, a Deus, pois sempre me ajudou e colocou em minha vida pessoas muito importantes.

À minha orientadora Profa. Cleonice, pela atenção, sabedoria e compreensão dedicadas ao meu auxílio.

Aos professores Ranga e Eliana, pela ajuda, idéias e sugestões.

Ao grupo de estudo de Funções Especiais e Polinômios Ortogonais pelo apoio e aprendizado que me deram.

Aos meus pais, Olga e Calixto, pela vida e pela criação que me deram, pelo apoio e por sempre terem me guiado.

A minha noiva Jaqueline pela ajuda, pela força e principalmente pelo companheirismo. Agradeço por sempre estar ao meu lado e me compreender em tudo.

À CAPES, pelo auxílio financeiro.

A todos que fizeram parte desse trabalho, deixo meu muito obrigado.

“Na maior parte das ciências,
uma geração põe abaixo o que a outra construiu,
e o que a outra estabeleceu a outra desfaz.
Somente na Matemática é que cada geração constrói
um novo andar sobre a antiga estrutura.”

Hermann Hankel

Resumo

O objetivo deste trabalho é estudar os polinômios ortogonais em várias variáveis com relação a um funcional linear, \mathcal{L} e suas propriedades análogas às dos polinômios ortogonais em uma variável, tais como: a relação de três termos, a relação de recorrência de três termos, o teorema de Favard, os zeros comuns e a cubatura gaussiana. Além disso, apresentamos um método para gerar polinômios ortonormais em duas variáveis e alguns exemplos.

Palavras-chave: Polinômios ortogonais em várias variáveis, zeros comuns de polinômios ortogonais em várias variáveis, fórmula de cubatura gaussiana.

Abstract

The aim here is to study the orthogonal polynomials in several variables with respect to a linear functional, \mathcal{L} . Also, to study its properties analogous to orthogonal polynomials in one variable, such as the three term relation, the three term recurrence relation, Favard's theorem, the common zeros and Gaussian cubature. A method to generating orthonormal polynomials in two variables and some examples are presented.

Keywords: Orthogonal polynomials in several variables, common zeros of orthogonal polynomials in several variables, Gaussian cubature formula.

Sumário

1	Introdução	1
2	Preliminares	3
2.1	Álgebra linear	3
2.2	Polinômios ortogonais em uma variável	5
2.2.1	Polinômios ortogonais clássicos	10
3	Polinômios em Várias Variáveis	16
3.1	Notações	16
3.2	Polinômios ortogonais	19
3.3	Polinômios ortogonais definidos a partir de matrizes de momento	29
3.4	Relação de três termos	32
3.5	Teorema de Favard	36
4	Propriedades dos Polinômios Ortogonais em Várias Variáveis	40
4.1	Integrais centralmente simétricas	40
4.2	Matrizes de Jacobi em blocos	44
4.3	Propriedades da relação de três termos	46
4.4	Soluções gerais da relação de três termos	55
4.5	Núcleos reprodutores e série ortogonal de Fourier	57
4.5.1	Núcleos reprodutores	58
4.5.2	Série ortogonal de Fourier	63
5	Zeros Comuns dos Polinômios Ortogonais em Várias Variáveis e Cubatura Gaussiana	66
5.1	Zeros comuns dos polinômios ortogonais em várias variáveis	66
5.2	Fórmulas de cubatura gaussianas	72
5.2.1	Caracterização das fórmulas de cubatura gaussianas	72
6	Exemplos	80
6.1	Produto tensor	80

6.2	Método de Koornwinder	82
6.3	Exemplos de matrizes coeficientes da relação de três termos	84
6.3.1	Polinômios ortonormais clássicos no quadrado	84
6.3.2	Polinômios ortonormais clássicos no triângulo	88
	Referências Bibliográficas	93
	Glossário	95

Capítulo 1

Introdução

A teoria de polinômios ortogonais na reta real tem vasta aplicação em todos os tipos de problemas da Matemática Pura e Ciências Aplicadas. Esses polinômios satisfazem diversas propriedades interessantes, como a relação de recorrência de três termos, o polinômio de grau n possui todos os seus zeros reais, distintos e dentro do intervalo de ortogonalidade, que caracteriza um resultado muito importante para as fórmulas de quadratura Gaussianas, além do entrelaçamento dos zeros de polinômios ortogonais de graus consecutivos.

Esse estudo é de verificar quais propriedades e resultados são extensões dos polinômios ortogonais em uma variável. Porém, durante este trabalho percebemos que muitas dessas estruturas se tornam complexas como, por exemplo, os zeros, que no caso de duas variáveis podem ser curvas algébricas, e essas estruturas são mais complicadas no caso de mais variáveis.

A seguir apresentamos, brevemente, os conteúdos dos capítulos desta dissertação, que foi baseada nos textos de Dunkl e Xu [5] e Krall e Sheffer [11].

No Capítulo 2 apresentamos algumas propriedades e definições sobre matrizes, descrevemos também algumas definições e resultados dos polinômios ortogonais, como a fórmula de recorrência de três termos, Teorema de Favard, zeros dos polinômios e fórmulas de quadratura Gaussiana. Por fim, apresentamos alguns polinômios clássicos.

No Capítulo 3 apresentamos as definições de polinômios ortogonais em várias variáveis e notações necessárias para essas definições. Além disso, apresentamos algumas propriedades sobre estes polinômios e os análogos à relação de três termos e ao Teorema de Favard.

No Capítulo 4 estudamos a extensão das propriedades das funções pesos pares dos polinômios ortogonais em uma variável, denotadas por integrais centralmente simétricas, e estendemos também as matrizes de Jacobi e os núcleos reprodutores. Apresentamos um

teorema sobre a condição necessária e suficiente para um sistema de polinômios em várias variáveis que satisfaz a relação de recorrência ser um sistema de polinômios ortogonais.

No Capítulo 5 descrevemos a definição de zero comum para polinômios ortogonais em várias variáveis e suas propriedades, por exemplo, que os zeros comuns dos polinômios ortogonais são pontos do \mathbb{R}^d , distintos e dois polinômios ortogonais não tem zeros comuns em comum. Apresentamos a fórmula de cubatura Gaussiana e seus teoremas relacionados com os zeros comuns dos polinômios ortogonais em várias variáveis.

Finalmente, no Capítulo 6, apresentamos os exemplos conhecidos como produto tensor, um método para gerar polinômios ortogonais em duas variáveis dado por Koornwinder [10] e alguns polinômios gerados por ele, além de exemplos das matrizes coeficientes da relação de três termos.

Capítulo 2

Preliminares

Neste capítulo veremos alguns assuntos preliminares importantes para a compreensão do trabalho. Esses assuntos são referentes a definições de corpo, resultados sobre matrizes, pois os coeficientes que acompanham o vetor polinomial coluna ou o vetor monomial coluna são matrizes, e sobre os polinômios ortogonais na reta real.

2.1 Álgebra linear

Nesta seção veremos alguns pré-requisitos importantes para o decorrer deste trabalho. As demonstrações serão omitidas, pois são bem conhecidas na literatura como em Horn e Johnson [8], Callioli [2], Coelho e Lourenço [4] e Lima [12].

Definição 2.1.1. *Um conjunto não vazio \mathbb{K} é um corpo se em \mathbb{K} pudermos definir duas operações, denotadas por $+$ (adição) e \cdot (multiplicação), satisfazendo as seguintes propriedades:*

(A1) $a + b = b + a, \quad \forall a, b \in \mathbb{K}$ (propriedade comutativa).

(A2) $a + (b + c) = (a + b) + c, \quad \forall a, b, c \in \mathbb{K}$ (propriedade associativa).

(A3) *Existe um elemento em \mathbb{K} , denotado por 0 e chamado de elemento neutro da adição, que satisfaz $0 + a = a + 0 = a, \quad \forall a \in \mathbb{K}$.*

(A4) *Para cada $a \in \mathbb{K}$, existe um elemento em \mathbb{K} , denotado por $-a$ e chamado de oposto de a (ou inverso aditivo de a), tal que $a + (-a) = (-a) + a = 0$.*

(M1) $a \cdot b = b \cdot a$, $a, b \in \mathbb{K}$ (propriedade comutativa).

(M2) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$, $a, b, c \in \mathbb{K}$ (propriedade associativa).

(M3) Existe um elemento em \mathbb{K} , denotado por 1 e chamado de elemento neutro da multiplicação, tal que $1 \cdot a = a \cdot 1 = a$, $a \in \mathbb{K}$.

(M4) Para cada elemento não nulo, $a \in \mathbb{K}$, existe um elemento em \mathbb{K} , denotado por a^{-1} e chamado de inverso multiplicativo de a , tal que $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

(D) $(a + b) \cdot c = a \cdot c + a \cdot b$, $\forall a, b, c \in \mathbb{K}$ (propriedade distributiva).

Apresentamos a seguir alguns resultados sobre matrizes.

Definição 2.1.2. Uma matriz quadrada $N \in \mathbb{R}^{n \times n}$ é definida positiva se, para todo $\mathbf{x} \in \mathbb{R}^n$ não-nulo,

$$\mathbf{x}^T N \mathbf{x} > 0.$$

Proposição 2.1.3. Seja N uma matriz quadrada inversível. Então,

a) se N é simétrica, sua inversa também é simétrica,

b) sua transposta também é inversível.

Proposição 2.1.4. Se N é uma matriz quadrada com $\det N \neq 0$, então

$$N^{-1} = \frac{1}{\det N} \text{adj}(N),$$

onde $\text{adj}(N)$ é a matriz adjunta de N .

Proposição 2.1.5. Se N é uma matriz simétrica, então seus autovalores são reais.

Proposição 2.1.6. N é uma matriz definida positiva se, e somente se, sua inversa é definida positiva.

Definição 2.1.7. Duas matrizes quadradas M e N são simultaneamente diagonalizáveis se existe uma única matriz quadrada S , tal que $S^{-1}MS$ e $S^{-1}NS$ são ambas matrizes diagonais.

Proposição 2.1.8. Sejam M e N matrizes quadradas de mesma ordem e diagonalizáveis. Então, M e N comutam se, e somente se, elas são simultaneamente diagonalizáveis.

Definição 2.1.9. *Seja $N \in \mathbb{R}^{m \times n}$. O posto de N é o número máximo de linhas, ou colunas, linearmente independentes dessa matriz.*

Definição 2.1.10. *A matriz $N \in \mathbb{R}^{m \times n}$ tem posto completo quando seu posto é igual ao $\min\{m, n\}$.*

Proposição 2.1.11. *Suponha que N é uma matriz de dimensão $m \times n$, então*

(1.) $\text{posto}(N) \leq \min(m, n)$;

(2.) $\text{posto}(N^T N) = \text{posto}(N N^T) = \text{posto}(N) = \text{posto}(N^T)$;

(3.) *se N é uma matriz quadrada ($m = n$), então N é inversível se, e somente se, N tem posto n , isto é, N tem posto completo;*

(4.) *se M é uma matriz qualquer de dimensão $n \times k$, então*

$$\text{posto}(NM) \leq \min\{\text{posto } N, \text{posto } M\};$$

(5.) *se M é uma matriz de dimensão $n \times k$ com posto n , então $\text{posto}(NM) = \text{posto}(N)$;*

(6.) *se M é uma matriz de dimensão $l \times m$ com posto m , então $\text{posto}(MN) = \text{posto}(N)$.*

A seguir apresentaremos uma desigualdade muito importante sobre posto, a desigualdade de posto de Sylvester.

Proposição 2.1.12. *Se N é uma matriz de dimensão $m \times n$ e M é de dimensão $n \times k$, então $\text{posto}(N) + \text{posto}(M) - n \leq \text{posto}(NM)$.*

Teorema 2.1.13. *$M \in \mathbb{R}^{n \times n}$ é simétrica se, e somente se, existe uma matriz ortogonal $O \in \mathbb{R}^{n \times n}$ e uma matriz diagonal $D \in \mathbb{R}^{n \times n}$ tal que $N = O^T D O$.*

2.2 Polinômios ortogonais em uma variável

Nesta seção veremos uma breve apresentação dos polinômios ortogonais na reta real. As demonstrações serão omitidas mas são facilmente encontradas em textos clássicos como os de Chihara [3] e Szegő [15], ou, ainda, em Ismail [9].

Definição 2.2.1. *Seja $\{\mu_n\}_{n=0}^{\infty}$ uma sequência de números reais e seja \mathcal{L} um funcional linear definido no espaço vetorial dos polinômios por*

$$\begin{aligned}\mathcal{L}(x^n) &= \mu_n, \quad n = 0, 1, 2, \dots \\ \mathcal{L}(\alpha\pi_1(x) + \beta\pi_2(x)) &= \alpha\mathcal{L}(\pi_1(x)) + \beta\mathcal{L}(\pi_2(x))\end{aligned}$$

para todo $\alpha, \beta \in \mathbb{R}$ e $\pi_1(x), \pi_2(x)$ polinômios.

Os números μ_n são chamados de momentos de ordem n e \mathcal{L} é dito funcional de momento determinado pela sequência de momentos $\{\mu_n\}_{n=0}^{\infty}$.

Definição 2.2.2. *Uma sequência $\{p_n(x)\}_{n=0}^{\infty}$ é chamada de sequência de polinômios ortogonais com respeito a um funcional de momento \mathcal{L} se, para todo n e m inteiros não-negativos,*

a) $p_n(x)$ é um polinômio de grau exatamente n ,

b) $\mathcal{L}(p_m(x)p_n(x)) = 0$ para $m \neq n$,

c) $\mathcal{L}(p_n^2(x)) \neq 0$.

Como $\{x^k\}_{k=0}^{\infty}$ é uma base para o espaço dos polinômios, então temos que os polinômios ortogonais podem ser escritos como

$$p_n(x) = k_n x^n + k_{n-1} x^{n-1} + k_{n-2} x^{n-2} + \dots + k_1 x + k_0,$$

onde dizemos que k_n é o coeficiente do termo de maior grau do polinômio ortogonal $p_n(x)$.

Teorema 2.2.3. *Seja $\{p_n(x)\}_{n=0}^{\infty}$ uma sequência de polinômios ortogonais com respeito a \mathcal{L} . Então, todo polinômio $\pi(x)$ de grau n pode ser escrito como combinação linear dos polinômios ortogonais até grau n , isto é,*

$$\pi(x) = \sum_{k=0}^n c_k p_k(x),$$

onde $c_k = \frac{\mathcal{L}(\pi(x)p_k(x))}{\mathcal{L}(p_n^2(x))}$, $k = 0, 1, \dots, n$.

Teorema 2.2.4. *Sejam \mathcal{L} um funcional linear e $\{p_n(x)\}_{n=0}^{\infty}$ uma sequência de polinômios. Então as seguintes afirmações são equivalentes:*

(a) $\{p_n(x)\}_{n=0}^{\infty}$ é uma sequência de polinômios ortogonais.

(b) $\mathcal{L}(x^m p_n(x)) = 0$ se $m < n$ e $\mathcal{L}(x^n p_n(x)) = \kappa_n$, onde $\kappa_n \neq 0$, $n = 0, 1, 2, \dots$

(c) $\mathcal{L}(\pi(x)p_n(x)) = 0$ se $\text{grau}(\pi(x)) < n$.

Se $\{p_n(x)\}_{n=0}^{\infty}$ é uma sequência de polinômios ortogonais e satisfazem a condição adicional $\mathcal{L}(p_n^2(x)) = 1$, então $\{p_n(x)\}_{n=0}^{\infty}$ é chamada de uma sequência de polinômios ortonormais e denotamos por $\{\tilde{p}_n(x)\}_{n=0}^{\infty}$. Uma sequência de polinômios ortogonais, $\{p_n(x)\}$, tal que os $p_n(x)$ são mônicos, é chamada sequência de polinômios ortogonais mônicos e será denotada por $\{\bar{p}_n(x)\}$.

Teorema 2.2.5. *Seja \mathcal{L} um funcional de momento e seja $\{p_n(x)\}_{n=0}^{\infty}$ uma sequência de polinômios ortogonais. Então*

(i) $\tilde{p}_n(x) = \frac{p_n(x)}{(\mathcal{L}(p_n^2(x)))^{\frac{1}{2}}}$ formam uma sequência de polinômios ortonormais,

(ii) $\bar{p}_n(x) = k_n^{-1}p_n(x)$, onde k_n é o coeficiente do termo de maior grau de $p_n(x)$, formam uma sequência de polinômios ortogonais mônicos.

Definição 2.2.6. *Seja $\{\mu_n\}_{n=0}^{\infty}$ uma sequência de momentos. Definimos o determinante de Hankel por*

$$\Delta_n = \det \begin{pmatrix} \mu_0 & \mu_1 & \cdots & \mu_n \\ \mu_1 & \mu_2 & \cdots & \mu_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_n & \mu_{n+1} & \cdots & \mu_{2n} \end{pmatrix}$$

Teorema 2.2.7. *Seja \mathcal{L} um funcional de momento determinado pela sequência de momentos $\{\mu_n\}_{n=0}^{\infty}$. Então, existe uma sequência de polinômios ortogonais se, e somente se, $\Delta_n \neq 0$.*

Definição 2.2.8. *Um funcional de momento \mathcal{L} é dito positivo definido se $\mathcal{L}(\pi(x)) > 0$ para todo polinômio $\pi(x)$ não identicamente nulo e não-negativo para todo x .*

Definição 2.2.9. \mathcal{L} é quase-definido se, e somente se, $\Delta_n \neq 0$ para $n \geq 0$.

Teorema 2.2.10. \mathcal{L} é um funcional de momento definido positivo se, e somente se, $\Delta_n > 0$, para todo $n \geq 0$.

Uma das propriedades mais importantes dos polinômios ortogonais é o fato de três polinômios de graus consecutivos estarem conectados por uma relação simples.

Teorema 2.2.11. *Seja \mathcal{L} um funcional de momento e seja $\{p_n(x)\}_{n=0}^{\infty}$ uma sequência de polinômios ortogonais. Então, existem constantes α_n , β_n e γ_n tais que*

$$xp_n(x) = \alpha_n p_{n+1}(x) + \beta_n p_n(x) + \gamma_n p_{n-1}(x), \quad n = 0, 1, 2, \dots, \quad (2.2.1)$$

onde definimos $p_{-1}(x) = 0$ e $p_0(x) = \gamma_0$. Além disso,

$$\begin{aligned} \alpha_n &= \frac{\mathcal{L}(xp_n(x)p_{n+1}(x))}{\mathcal{L}(p_{n+1}^2(x))}, \\ \beta_n &= \frac{\mathcal{L}(xp_n^2(x))}{\mathcal{L}(p_n^2(x))}, \\ \gamma_n &= \alpha_{n-1} \frac{\mathcal{L}(p_n^2(x))}{\mathcal{L}(p_{n-1}^2(x))}. \end{aligned}$$

Desse teorema obtemos a relação equivalente

$$p_{n+1}(x) = \left(\frac{1}{\alpha_n} x - \frac{\beta_n}{\alpha_n} \right) p_n(x) - \frac{\gamma_n}{\alpha_n} p_{n-1}(x).$$

Caso os polinômios sejam ortonormais, temos a seguinte relação de recorrência de três termos.

Teorema 2.2.12. *Seja \mathcal{L} um funcional de momento e seja $\{\tilde{p}_n(x)\}_{n=0}^{\infty}$ a sequência de polinômios ortonormais. Então, existem constantes $\tilde{\alpha}_n$ e $\tilde{\beta}_n$ tais que*

$$x\tilde{p}_n(x) = \tilde{\alpha}_n \tilde{p}_{n+1}(x) + \tilde{\beta}_n \tilde{p}_n(x) + \tilde{\alpha}_{n-1} \tilde{p}_{n-1}(x), \quad n = 0, 1, 2, \dots, \quad (2.2.2)$$

onde definimos $\tilde{p}_{-1}(x) = 0$ e $\tilde{p}_0(x) = \tilde{\alpha}_{-1}$.

O próximo teorema afirma que qualquer sequência de polinômios que satisfaz a relação (2.2.1) é uma sequência de polinômios ortogonais. Este teorema recebe o nome de Teorema de Favard em homenagem a J. Favard que o enunciou em 1935.

Teorema 2.2.13. *Sejam $\{\alpha_n\}_{n=0}^{\infty}$, $\{\beta_n\}_{n=0}^{\infty}$ e $\{\gamma_n\}_{n=0}^{\infty}$ sequências arbitrárias de números reais e seja $\{p_n(x)\}_{n=0}^{\infty}$ definida pela relação de recorrência de três termos*

$$\begin{aligned} xp_n(x) &= \alpha_n p_{n+1}(x) + \beta_n p_n(x) + \gamma_n p_{n-1}(x), \quad n = 0, 1, 2, \dots, \\ p_{-1}(x) &= 0, \quad p_0(x) = \gamma_0. \end{aligned} \quad (2.2.3)$$

Então, existe um único funcional de momento \mathcal{L} tal que

$$\mathcal{L}(1) = 1, \quad \mathcal{L}(p_m(x)p_n(x)) = 0, \quad \text{para } m \neq n, \quad n, m = 0, 1, 2, \dots$$

Definição 2.2.14. Um funcional de momento é simétrico se todos os momentos ímpares são iguais a zero.

Teorema 2.2.15. Seja $\{\bar{p}_n(x)\}_{n=0}^{\infty}$ uma sequência de polinômios ortogonais mônicos com respeito ao funcional quase-definido \mathcal{L} . Então, seguem as equivalências:

(i) \mathcal{L} é simétrico.

(ii) $\bar{p}_n(-x) = (-1)^n \bar{p}_n(x)$, $n \geq 0$.

(iii) Na relação de recorrência de três termos, $\beta_n = 0$ para todo $n \geq 1$.

A seguir, apresentamos a identidade de Christoffel-Darboux para polinômios ortogonais mônicos.

Teorema 2.2.16. Se $\{\bar{p}_n(x)\}_{n=0}^{\infty}$ satisfaz

$$\begin{aligned}\bar{p}_n(x) &= (x - \bar{\beta}_n)\bar{p}_{n-1}(x) - \bar{\gamma}_n\bar{p}_{n-2}(x), \quad n = 1, 2, 3, \dots \\ \bar{p}_{-1}(x) &= 0, \quad \bar{p}_0(x) = 1.\end{aligned}$$

com $\lambda_n \neq 0$ para $n \geq 1$. Então,

$$\sum_{k=0}^n \frac{\bar{p}_k(x)\bar{p}_k(u)}{\bar{\gamma}_1\bar{\gamma}_2 \cdots \bar{\gamma}_{k+1}} = \bar{\gamma}_1\bar{\gamma}_2 \cdots \bar{\gamma}_{n+1} \frac{\bar{p}_{n+1}(x)\bar{p}_n(u) - \bar{p}_n(x)\bar{p}_{n+1}(u)}{x - u}. \quad (2.2.4)$$

No Teorema 2.2.16, caso os polinômios sejam ortonormais, então a identidade de Christoffel-Darboux é da forma

$$\sum_{k=0}^n \tilde{p}_k(x)\tilde{p}_k(u) = \frac{\tilde{k}_n}{\tilde{k}_{n+1}} \frac{\tilde{p}_{n+1}(x)\tilde{p}_n(u) - \tilde{p}_n(x)\tilde{p}_{n+1}(u)}{x - u}. \quad (2.2.5)$$

Teorema 2.2.17. A seguinte fórmula confluyente de (2.2.4) é válida

$$\sum_{k=0}^n \frac{\bar{p}_k^2(x)}{\bar{\gamma}_1\bar{\gamma}_2 \cdots \bar{\gamma}_{k+1}} = \bar{\gamma}_1\bar{\gamma}_2 \cdots \bar{\gamma}_{n+1} \frac{\bar{p}'_{n+1}(x)\bar{p}_n(x) - \bar{p}'_n(x)\bar{p}_{n+1}(x)}{x - u}. \quad (2.2.6)$$

Definição 2.2.18. Os polinômios núcleos são definidos, a partir dos polinômios ortonormais, da seguinte forma

$$\mathbf{K}_n(x, u) = \sum_{k=0}^n \tilde{p}_k(x)\tilde{p}_k(u).$$

Teorema 2.2.19. Seja \mathcal{L} um funcional definido positivo e seja z_0 um número real arbitrário. Então,

$$[\mathbf{K}_n(z_0, z_0)]^{-1} = \min \mathcal{L}[|\pi^2(x)|],$$

onde o mínimo é sobre todos os polinômios de grau no máximo n e $\pi(z_0) = 1$.

Um resultado muito interessante sobre os polinômios ortogonais diz respeito a seus zeros que satisfazem muitas propriedades. Apresentamos a seguir alguns deles pertencentes ao suporte de \mathcal{L} .

Teorema 2.2.20. *Os zeros dos polinômios ortogonais $p_n(x)$, $n \geq 1$, são reais e distintos.*

Ordenando os zeros de $p_n(x)$ de forma crescente e os denotando por $x_{n,i}$, para $i = 1, 2, \dots, n$, então podemos enunciar o seguinte teorema sobre o entrelaçamento dos zeros dos polinômios ortogonais.

Teorema 2.2.21. *Os zeros de $p_n(x)$ e $p_{n+1}(x)$ se entrelaçam, isto é,*

$$x_{n+1,i} < x_{n,i} < x_{n+1,i+1}, \quad i = 1, 2, \dots, n. \quad (2.2.7)$$

Um resultado importante sobre os zeros dos polinômios ortogonais é a fórmula de quadratura Gaussiana.

Teorema 2.2.22. *Seja \mathcal{L} um funcional definido positivo. Então, existem números $a_{n,1}, a_{n,2}, \dots, a_{n,n}$ tais que, para todo polinômio $\pi(x)$ de grau no máximo $2n - 1$,*

$$\mathcal{L}(\pi(x)) = \sum_{k=1}^n a_{n,k} \pi(x_{n,k}).$$

Os números $a_{n,k}$, $k = 1, \dots, n$, são todos positivos e satisfazem a condição $\sum_{k=1}^n a_{n,k} = \mu_0$.

2.2.1 Polinômios ortogonais clássicos

Veremos, agora, os polinômios ortogonais clássicos que serão úteis no último capítulo dessa dissertação. Para isso, necessitamos de algumas definições preliminares.

Definição 2.2.23. *O símbolo de Pochhammer é definido para todo $x \in \mathbb{K}$, onde \mathbb{K} é um corpo, por*

$$(x)_0 = 1, \quad (x)_n = \prod_{i=1}^n (x + i - 1), \quad \text{para } n = 1, 2, 3, \dots$$

Uma forma recursiva para esta definição é $(x)_0 = 1$ e $(x)_{n+1} = (x)_n(x + n)$ para $n = 1, 2, 3, \dots$

Definição 2.2.24. *A função Gama é definida a partir da integral de Euler de segunda espécie, para $x \in \mathbb{C}$ e $\text{Re}(x) > 0$, como*

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

Dessa definição, obtemos propriedade de recorrência a seguir.

Propriedade 2.2.25. Para $x \in \mathbb{C}$ e $\operatorname{Re}(x) > 0$, temos que a função Gama satisfaz

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

Definição 2.2.26. A função Beta é definida, para $x, y \in \mathbb{C}$, $\operatorname{Re}(x) > 0$ e $\operatorname{Re}(y) > 0$, pela integral de Euler de primeira espécie

$$B(x, y) = \int_0^1 t^{x-1}(1-t)^{y-1} dt.$$

A função Beta pode ser escrita em termos da função Gama.

Teorema 2.2.27. Para $x, y \in \mathbb{C}$, $\operatorname{Re}(x) > 0$ e $\operatorname{Re}(y) > 0$, temos

$$B(x, y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Polinômios de Jacobi

Os polinômios de Jacobi são ortogonais no intervalo $[-1, 1]$ com relação à função peso $(1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1$, $\alpha, \beta \in \mathbb{R}$. Esses polinômios serão denotados por $P_n^{(\alpha, \beta)}$ e satisfazem

$$\int_{-1}^1 P_n^{(\alpha, \beta)}(x) P_m^{(\alpha, \beta)}(x) (1-x)^\alpha (1+x)^\beta dx = \begin{cases} 0, & \text{se } m \neq n, \\ \frac{2^{\alpha+\beta+1} \Gamma(\alpha+n+1) \Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)n! \Gamma(\alpha+\beta+n+1)} \neq 0, & \text{se } m = n. \end{cases} \quad (2.2.8)$$

Os polinômios ortonormais de Jacobi são dados por

$$\tilde{P}_n^{(\alpha, \beta)}(x) = \left(\frac{(\alpha + \beta + 2n + 1)n! \Gamma(\alpha + \beta + n + 1)}{2^{\alpha+\beta+1} \Gamma(\alpha + n + 1) \Gamma(\beta + n + 1)} \right)^{1/2} P_n^{(\alpha, \beta)}(x).$$

A relação de recorrência de três termos dos polinômios ortonormais de Jacobi é

$$x\tilde{P}_n^{(\alpha, \beta)}(x) = \alpha_n \tilde{P}_{n+1}^{(\alpha, \beta)}(x) + \beta_n \tilde{P}_n^{(\alpha, \beta)}(x) + \alpha_{n-1} \tilde{P}_{n-1}^{(\alpha, \beta)}(x),$$

onde $\tilde{P}_{-1}^{(\alpha, \beta)}(x) = 0$, $\tilde{P}_0^{(\alpha, \beta)}(x) = (2^{\alpha+\beta+1} B(\alpha+1, \beta+1))^{-1/2}$ e

$$\begin{cases} \alpha_n = \left(\frac{(\alpha + n + 1)(\beta + n + 1)(n + 1)(\alpha + \beta + n + 1)}{(\alpha + \beta + 2n + 1)(\alpha + \beta + 2n + 3)} \right)^{1/2} \frac{2}{(\alpha + \beta + 2n + 2)}, \\ \beta_n = \frac{(\beta^2 - \alpha^2)}{(\alpha + \beta + 2n)(\alpha + \beta + 2n + 2)}. \end{cases}$$

Os polinômios ortogonais no intervalo $[0, 1]$ com relação a função peso $(1-x)^\alpha x^\beta$, podem ser escritos em função dos polinômios de Jacobi, pois se aplicarmos a mudança de variável $x \rightarrow \frac{y+1}{2}$ em (2.2.8), obtemos

$$\int_0^1 P_n^{(\alpha,\beta)}(2x-1)P_m^{(\alpha,\beta)}(2x-1)(1-x)^\alpha x^\beta dx = \begin{cases} 0, & \text{se } m \neq n, \\ \frac{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)n!\Gamma(\alpha+\beta+n+1)} \neq 0, & \text{se } m = n. \end{cases}$$

Nesse caso, os polinômios ortonormais são dados por

$$\tilde{P}_n^{(\alpha,\beta)}(2x-1) = \left(\frac{(\alpha+\beta+2n+1)n!\Gamma(\alpha+\beta+n+1)}{\Gamma(\alpha+n+1)\Gamma(\beta+n+1)} \right)^{1/2} P_n^{(\alpha,\beta)}(2x-1).$$

A relação de recorrência de três termos, para esses polinômios, é dada por

$$x\tilde{P}_n^{(\alpha,\beta)}(2x-1) = \alpha_n\tilde{P}_{n+1}^{(\alpha,\beta)}(2x-1) + \beta_n\tilde{P}_n^{(\alpha,\beta)}(2x-1) + \alpha_{n-1}\tilde{P}_{n-1}^{(\alpha,\beta)}(2x-1),$$

onde $\tilde{P}_{-1}^{(\alpha,\beta)}(2x-1) = 0$, $\tilde{P}_0^{(\alpha,\beta)}(2x-1) = (B(\alpha+1, \beta+1))^{-1/2}$ e

$$\begin{cases} \alpha_n = \left(\frac{(\alpha+n+1)(\beta+n+1)(n+1)(\alpha+\beta+n+1)}{(\alpha+\beta+2n+1)(\alpha+\beta+2n+3)} \right)^{1/2} \frac{1}{(\alpha+\beta+2n+2)}, \\ \beta_n = \left(\frac{(\beta^2 - \alpha^2)}{2(\alpha+\beta+2n+2)(\alpha+\beta+2n)} + \frac{1}{2} \right). \end{cases}$$

Polinômios de Gegenbauer

Os polinômios que são ortogonais com relação a função peso $(1-x^2)^{\lambda-1/2}$ em $[-1, 1]$, são conhecidos como polinômios de Gegenbauer. Esses polinômios são denotados por $G_n^{(\lambda)}(x)$ e satisfazem

$$\int_{-1}^1 G_n^{(\lambda)}(x)G_m^{(\lambda)}(x)(1-x^2)^{\lambda-1/2} dx = \begin{cases} 0, & \text{se } m \neq n, \\ \frac{2^{2\lambda-1}\Gamma(n+2\lambda)}{(\lambda+n)n!} \left(\frac{\Gamma(\lambda+1/2)}{\Gamma(2\lambda)} \right)^2 \neq 0, & \text{se } m = n. \end{cases}$$

Os polinômios ortonormais de Gegenbauer são dados por

$$\tilde{G}_n^{(\lambda)}(x) = \left(\frac{(\lambda+n)n!}{2^{2\lambda-1}\Gamma(n+2\lambda)} \right)^{1/2} \frac{\Gamma(2\lambda)}{\Gamma(\lambda+1/2)} G_n^{(\lambda)}(x).$$

A sua relação de recorrência de três termos por

$$x\tilde{G}_n^{(\lambda)}(x) = \frac{1}{2} \left(\frac{(n+1)(n+2\lambda)}{(n+\lambda)(n+\lambda+1)} \right)^{1/2} \tilde{G}_{n+1}^{(\lambda)}(x) + \frac{1}{2} \left(\frac{n(n+2\lambda-1)}{(n+\lambda-1)(n+\lambda)} \right)^{1/2} \tilde{G}_{n-1}^{(\lambda)}(x),$$

onde $\tilde{G}_{-1}^{(\lambda)}(x) = 0$ e $\tilde{G}_0^{(\lambda)}(x) = \left(\frac{\lambda\Gamma(2\lambda)}{2^{2\lambda-1}} \right)^{1/2} \frac{1}{\Gamma(\lambda+1/2)}$.

Polinômios de Hermite

Os polinômios de Hermite são ortogonais em \mathbb{R} com relação a função peso e^{-x^2} . Os polinômios de Hermite serão denotados por H_n e satisfazem

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

Os polinômios ortonormais de Hermite são dados por

$$\tilde{H}_n(x) = (2^n n! \sqrt{\pi})^{-1/2} H_n(x).$$

Sua relação de recorrência de três termos é

$$x\tilde{H}_n(x) = \left(\frac{n+1}{2}\right)^{1/2} \tilde{H}_{n+1}(x) + \left(\frac{n}{2}\right)^{1/2} \tilde{H}_{n-1}(x),$$

onde $\tilde{H}_{-1}(x) = 0$ e $\tilde{H}_0(x) = (\sqrt{\pi})^{-1/2}$.

Polinômios de Laguerre

Os polinômios ortogonais no intervalo $\mathbb{R}_+ = \{x : x \geq 0\}$ com relação a função peso $x^\alpha e^{-x}$, $\alpha > -1$, e denotados por $L_n^{(\alpha)}$, satisfazem

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

Os polinômios ortonormais de Laguerre são dados por

$$\tilde{L}_n^{(\alpha)}(x) = (\Gamma(\alpha + n + 1)n!)^{-1/2} L_n^{(\alpha)}(x).$$

A relação de recorrência de três termos dos polinômios ortonormais de Laguerre é

$$x\tilde{L}_n^{(\alpha)}(x) = -(n+1)(\alpha+n+1)^{1/2}\tilde{L}_{n+1}^{(\alpha)}(x) + (2n+\alpha+1)\tilde{L}_n^{(\alpha)}(x) - n(\alpha+n)^{1/2}\tilde{L}_{n-1}^{(\alpha)}(x),$$

onde $\tilde{L}_{-1}^{(\alpha)}(x) = 0$ e $\tilde{L}_0^{(\alpha)}(x) = (\Gamma(\alpha + 1))^{-1/2}$.

Polinômios de Legendre

Os polinômios de Legendre são ortogonais com relação a função peso 1 no intervalo $[-1, 1]$. Esses polinômios são denotados por $P_n^{(0)}(x)$, pois é um caso particular dos polinômios de Jacobi com $\alpha = \beta = 0$, e satisfazem

$$\int_{-1}^1 P_n^{(0)}(x)P_m^{(0)}(x)dx = \begin{cases} 0, & \text{se } m \neq n, \\ \frac{2}{2n+1} \neq 0, & \text{se } m = n. \end{cases}$$

Os polinômios ortonormais de Legendre são dados por

$$\tilde{P}_n^{(0)} = \left(\frac{2n+1}{2} \right)^{1/2} P_n^{(0)}.$$

A relação de recorrência de três termos dos polinômios ortonormais de Legendre é

$$x\tilde{P}_n^{(0)} = ((2n+3)(2n+1))^{-1/2} (n+1)\tilde{P}_{n+1}^{(0)}(x) + ((2n+1)(2n-1))^{-1/2} n\tilde{P}_{n-1}^{(0)}(x),$$

onde $\tilde{P}_{-1}^{(0)}(x) = 0$ e $\tilde{P}_0^{(0)}(x) = 2^{-1/2}$.

Polinômios de Chebyshev de 1ª Espécie

Os polinômios de Chebyshev de 1ª Espécie são ortogonais com relação a função peso $(1-x^2)^{-1/2}$ no intervalo $[-1, 1]$. Esses polinômios serão denotados por $T_n(x)$ e satisfazem

$$\int_{-1}^1 T_n(x)T_m(x) \frac{1}{\sqrt{1-x^2}} dx = \begin{cases} 0, & \text{se } m \neq n, \\ \pi, & \text{se } m = n = 0, \\ \frac{\pi}{2}, & \text{se } m = n > 0. \end{cases}$$

Os polinômios ortonormais de Chebyshev de 1ª Espécie são dados por

$$\tilde{T}_0(x)(\pi)^{-1/2}T_0(x) \quad \text{e} \quad \tilde{T}_n(x) = \left(\frac{2}{\pi} \right)^{1/2} T_n(x) \quad \text{se } n > 0.$$

A relação de recorrência de três termos dos polinômios ortonormais de Chebyshev de 1ª Espécie é

$$x\tilde{T}_n(x) = \frac{1}{2}\tilde{T}_{n+1}(x) + \frac{1}{2}\tilde{T}_{n-1}(x),$$

onde $\tilde{T}_{-1}(x) = 0$ e $\tilde{T}_0(x) = (\pi)^{-1/2}$.

Polinômios de Chebyshev de 2ª Espécie

Os polinômios de Chebyshev de 2ª Espécie são ortogonais com relação a função peso $\sqrt{1-x^2}$ no intervalo $[-1, 1]$. Esses polinômios serão denotados por $\tilde{U}_n(x)$ e satisfazem

$$\int_{-1}^1 U_n(x)U_m(x)\sqrt{1-x^2} dx = \begin{cases} 0, & \text{se } m \neq n, \\ \frac{\pi}{2}, & \text{se } m = n. \end{cases}$$

Os polinômios ortonormais de Chebyshev de 2ª Espécie são dados por

$$\tilde{U}_n(x) = \left(\frac{2}{\pi} \right)^{1/2} U_n(x).$$

A relação de recorrência de três termos dos polinômios ortonormais de Chebyshev de 2ª Espécie é da forma

$$x\tilde{U}_n(x) = \frac{1}{2}\tilde{U}_{n+1}(x) + \frac{1}{2}\tilde{U}_{n-1}(x),$$

onde $\tilde{U}_{-1}(x) = 0$ e $\tilde{U}_0(x) = \left(\frac{2}{\pi}\right)^{1/2}$.

Capítulo 3

Polinômios em Várias Variáveis

Nesse capítulo vemos algumas definições necessárias para o estudo dos polinômios ortogonais em várias variáveis e as propriedades decorrentes dessas definições. Apresentamos também os polinômios construídos por meio de matrizes de momentos, a relação de três termos para polinômios ortogonais em várias variáveis e o Teorema de Favard para várias variáveis. Esse estudo foi baseado nos textos de Dunkl e Xu [5] e Krall e Sheffer [11].

3.1 Notações

Iniciemos com algumas notações para polinômios em várias variáveis e algumas definições importantes para o bom entendimento do trabalho.

Denotamos por \mathbb{N}_0 o conjunto dos inteiros não-negativos. Um multi-índice α é dado por $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}_0^d$, onde \mathbb{N}_0^d é o conjunto das d -uplas cujas componentes pertencem a \mathbb{N}_0 . Definimos a norma do multi-índice α por

$$|\alpha| = \alpha_1 + \dots + \alpha_d.$$

Para os multi-índices $\alpha, \beta \in \mathbb{N}_0^d$, definimos $\delta_{\alpha, \beta} = \delta_{\alpha_1, \beta_1} \dots \delta_{\alpha_d, \beta_d}$, onde $\delta_{\alpha_i, \beta_i}$ é o delta de Kronecker, isto é

$$\delta_{\alpha, \beta} = \begin{cases} 1, & \text{se } \alpha_i = \beta_i \text{ para todo } 1 \leq i \leq d, \\ 0, & \text{se } \alpha_i \neq \beta_i \text{ para algum } 1 \leq i \leq d. \end{cases}$$

Sejam $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$ e $\alpha \in \mathbb{N}_0^d$. Então, um monômio \mathbf{x}^α é definido pelo produto

$$\mathbf{x}^\alpha = x_1^{\alpha_1} \dots x_d^{\alpha_d},$$

onde x_1, \dots, x_d são as variáveis e o número $|\alpha|$ é o grau total do monômio \mathbf{x}^α .

Exemplo 3.1.1. No caso $d = 2$, se $\mathbf{x} = (x, y)$, um monômio com o multi-índice $\alpha = (h, k)$ é dado por $\mathbf{x}^\alpha = x^h y^k$ e seu grau total é $|\alpha| = h + k$.

Desse modo, um polinômio P em d variáveis é uma combinação linear dos monômios,

$$P(\mathbf{x}) = \sum_{\alpha} c_{\alpha} \mathbf{x}^{\alpha},$$

onde cada coeficiente c_{α} é uma constante que depende somente do multi-índice α e pertence a um corpo \mathbb{K} , onde \mathbb{K} usualmente é o conjunto dos números racionais \mathbb{Q} , dos números reais \mathbb{R} ou dos números complexos \mathbb{C} .

O grau do polinômio é o maior grau total dos monômios. Denotamos por Π^d o conjunto dos polinômios de d variáveis e por Π_n^d o conjunto dos polinômios de d variáveis e grau no máximo n .

Exemplo 3.1.2. Para $d = 2$, um exemplo de polinômio é

$$P(x, y) = 3x^2y^2 - \frac{4}{27}x^3 + 2x^2y - xy^2 + 11x^2 - 8xy - 7y^2 + 13x - 2y + \frac{1}{6},$$

cujos graus são 4.

Definição 3.1.3. Um polinômio é homogêneo se todos os seus monômios têm o mesmo grau total e denotamos o conjunto dos polinômios homogêneos de grau n em d variáveis por \mathcal{P}_n^d , isto é,

$$\mathcal{P}_n^d = \left\{ P : P(\mathbf{x}) = \sum_{|\alpha|=n} c_{\alpha} \mathbf{x}^{\alpha} \right\}.$$

Assim, todo polinômio em Π^d pode ser escrito como combinação linear de polinômios homogêneos. Para $P \in \Pi_n^d$,

$$P(\mathbf{x}) = \sum_{k=1}^n \sum_{|\alpha|=k} c_{\alpha} \mathbf{x}^{\alpha}.$$

Exemplo 3.1.4. Um exemplo de polinômio homogêneo de grau 3 em 3 variáveis é

$$P(x, y, z) = 3x^3 + y^3 - 2z^3 - \sqrt{2}x^2z + \frac{2}{5}xy^2 + xyz.$$

Definição 3.1.5. Um polinômio mônico é aquele que possui um único termo mônico de maior grau.

Denotemos por \bar{P}_n um polinômio mônico de grau n . Então, este polinômio é da seguinte forma:

$$\bar{P}_n(\mathbf{x}) = \mathbf{x}^\alpha + Q_{n-1}(\mathbf{x}),$$

onde $|\alpha| = n$ e $Q_{n-1} \in \Pi_{n-1}^d$.

Denotemos por r_n^d a dimensão de \mathcal{P}_n^d . Observe que $\{\mathbf{x}^\alpha : |\alpha| = n\}$ é uma base de \mathcal{P}_n^d e, assim, $r_n^d = \#\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$. De $\frac{1}{1-t} = \sum_{n=0}^{\infty} t^n$, segue que

$$\begin{aligned} \frac{1}{(1-t)^d} &= \left(\sum_{n=0}^{\infty} t^n \right)^d = \left(\sum_{n_1=0}^{\infty} t^{n_1} \right) \left(\sum_{n_2=0}^{\infty} t^{n_2} \right) \dots \left(\sum_{n_d=0}^{\infty} t^{n_d} \right) \\ &= \sum_{n=0}^{\infty} \left(\sum_{|\alpha|=n} 1 \right) t^n = \sum_{n=0}^{\infty} r_n^d t^n. \end{aligned}$$

Por outro lado, pela expansão em série de Taylor, temos

$$\begin{aligned} \frac{1}{(1-t)^d} &= \sum_{n=0}^{\infty} \frac{(d)_n}{n!} t^n = \sum_{n=0}^{\infty} \frac{d(d+1)\dots(d+n-1)}{n!} t^n \\ &= \sum_{n=0}^{\infty} \binom{n+d-1}{n} t^n. \end{aligned}$$

Comparando os coeficientes de t^n das duas séries, concluímos que

$$r_n^d = \dim \mathcal{P}_n^d = \binom{n+d-1}{n}.$$

Note que $\#\{\alpha \in \mathbb{N}_0^d : |\alpha| \leq n\} = \#\{\alpha \in \mathbb{N}_0^{d+1} : |\alpha| = n\}$. De fato, podemos reescrever o segundo conjunto como $\alpha_1 + \alpha_2 + \dots + \alpha_d = n - \alpha_{d+1}$. Assim, como $0 \leq \alpha_{d+1} \leq n$, segue a igualdade. Portanto,

$$\dim \Pi_n^d = \binom{n+d}{n}.$$

Definição 3.1.6. A ordem lexicográfica adotada aqui é dada por $\mathbf{x}^\alpha < \mathbf{x}^\beta$ se $|\alpha| < |\beta|$ ou se $|\alpha| = |\beta|$ e a primeira componente não-nula de $\alpha - \beta$ é positiva, onde $\alpha - \beta = (\alpha_1 - \beta_1, \alpha_2 - \beta_2, \dots, \alpha_d - \beta_d)$.

Exemplo 3.1.7. No caso $d = 2$, a ordem lexicográfica nas variáveis x, y é

$$1 < x < y < x^2 < xy < y^2 < \dots < x^n < x^{n-1}y < \dots < y^n < \dots$$

De acordo com a definição da ordem lexicográfica, temos que $x < y^2$, pois a norma do multi-índice do primeiro monômio, $\alpha = (1, 0)$, é menor que a norma do multi-índice do segundo monômio, $\beta = (0, 2)$. Além disso, $x^2y^2 < xy^3$, pois como $\alpha = (2, 2)$ e $\beta = (1, 3)$, temos $|\alpha| = |\beta|$ e a primeira componente não-nula de $\alpha - \beta = (2, 2) - (1, 3) = (1, -1)$ é positiva.

Há outras maneiras de se definir ordem lexicográfica, como, por exemplo a ordem lexicográfica definida por $\mathbf{x}^\alpha < \mathbf{x}^\beta$ se a primeira componente da diferença $\alpha - \beta = (\alpha_1 - \beta_1, \dots, \alpha_d - \beta_d)$ é positiva. Mas, esta ordem não preserva a ordem natural do grau total do monômio. De fato, para $d = 2$ note que $x^2y^3 < xy$ pois $\alpha - \beta = (2 - 1, 3 - 1) = (1, 2)$. Contudo, $|\alpha| > |\beta|$. Por isso, usamos a ordem lexicográfica da Definição 3.1.6.

3.2 Polinômios ortogonais

Denotamos por r_n , em vez de r_n^d , a dimensão de \mathcal{P}_n^d quando d estiver subentendido. Uma sequência $s : \mathbb{N}_0^d \mapsto \mathbb{R}$ pode ser escrita da forma $s = (s_\alpha)_{\alpha \in \mathbb{N}_0^d}$.

Definição 3.2.1. Para cada sequência $s = (s_\alpha)_{\alpha \in \mathbb{N}_0^d}$, seja \mathcal{L}_s o funcional linear definido em Π^d por

$$\mathcal{L}_s(\mathbf{x}^\alpha) = s_\alpha, \quad \alpha \in \mathbb{N}_0^d, \quad (3.2.1)$$

onde s_α é uma constante real que depende somente do multi-índice α . Chamamos \mathcal{L}_s de funcional de momento definido pela sequência s .

Sejam os elementos do conjunto $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$ arranjados da forma $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(r_n^d)}$, de acordo com a ordem lexicográfica.

Definição 3.2.2. Para cada $n \in \mathbb{N}_0$, seja \mathbb{X}^n o vetor monomial coluna

$$\mathbb{X}^n = (\mathbf{x}^\alpha)_{|\alpha|=n} = \left(\mathbf{x}^{\alpha^{(1)}}, \dots, \mathbf{x}^{\alpha^{(r_n^d)}} \right)^T, \quad (3.2.2)$$

isto é, \mathbb{X}^n é um vetor cujos elementos são os monômios \mathbf{x}^α , com $|\alpha| = n$ arranjados na ordem lexicográfica.

Exemplo 3.2.3. O vetor monomial coluna, no caso $d = 2$, é da forma

$$\mathbb{X}^n = (x^n, x^{n-1}y, x^{n-2}y^2, \dots, xy^{n-1}, y^n)^T.$$

Definição 3.2.4. Para $k, j \in \mathbb{N}_0$ definimos o vetor de momentos \mathbf{s}_k e o momento matricial $\mathbf{s}_{\{k\}+\{j\}}$ por

$$\mathbf{s}_k = \mathcal{L}_s(\mathbb{X}^k) \quad e \quad \mathbf{s}_{\{k\}+\{j\}} = \mathcal{L}_s(\mathbb{X}^k(\mathbb{X}^j)^T), \quad (3.2.3)$$

ou seja, o vetor de momento \mathbf{s}_k é definido aplicando-se o funcional linear em cada componente do vetor monomial coluna e o momento matricial $\mathbf{s}_{\{k\}+\{j\}}$ é definido aplicando-se o

funcional linear em cada componente da matriz $\mathbb{X}^k(\mathbb{X}^j)^T$. Pela definição, $\mathbf{s}_{\{k\}+\{j\}}$ é uma matriz de dimensão $r_k^d \times r_j^d$, cujos elementos são $\mathcal{L}_s(\mathbf{x}^{\alpha+\beta})$ para $|\alpha| = k$ e $|\beta| = j$.

Exemplo 3.2.5. No caso $d = 2$, o vetor de momentos é dado por

$$\begin{aligned} \mathbf{s}_k &= \mathcal{L}_s(\mathbb{X}^k) = \mathcal{L}_s((x^k, x^{k-1}y, x^{k-2}y^2, \dots, xy^{k-1}, y^k)^T) \\ &= (\mathcal{L}_s(x^k), \mathcal{L}_s(x^{k-1}y), \mathcal{L}_s(x^{k-2}y^2), \dots, \mathcal{L}_s(xy^{k-1}), \mathcal{L}_s(y^k))^T \\ &= (s_{(k,0)}, s_{(k-1,1)}, s_{(k-2,2)}, \dots, s_{(1,k-1)}, s_{(0,k)}) \end{aligned}$$

e o momento matricial $\mathbf{s}_{\{k\}+\{j\}}$ por

$$\begin{aligned} \mathbf{s}_{\{k\}+\{j\}} &= \mathcal{L}_s(\mathbb{X}^k(\mathbb{X}^j)^T) \\ &= \mathcal{L}_s((x^k, x^{k-1}y, \dots, y^k)^T(x^j, x^{j-1}y, \dots, y^j)) \\ &= \begin{pmatrix} \mathcal{L}_s(x^{k+j}) & \mathcal{L}_s(x^{k+j-1}y) & \dots & \mathcal{L}_s(x^k y^j) \\ \mathcal{L}_s(x^{k+j-1}y) & \mathcal{L}_s(x^{k+j-2}y^2) & \dots & \mathcal{L}_s(x^{k-1}y^{j+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}_s(x^j y^k) & \mathcal{L}_s(x^{j-1}y^{k+1}) & \dots & \mathcal{L}_s(y^{k+j}) \end{pmatrix} \\ &= \begin{pmatrix} s_{(k+j,0)} & s_{(k+j-1,1)} & \dots & s_{(k,j)} \\ s_{(k+j-1,1)} & s_{(k+j-2,2)} & \dots & s_{(k-1,j+1)} \\ \vdots & \vdots & \ddots & \vdots \\ s_{(j,k)} & s_{(j-1,k+1)} & \dots & s_{(0,k+j)} \end{pmatrix}. \end{aligned}$$

Definição 3.2.6. Para cada $n \in \mathbb{N}_0$, definimos a matriz de momentos, $M_{n,d}$, por

$$M_{n,d} = (\mathbf{s}_{\{k\}+\{j\}})_{k,j=0}^n \quad e \quad \Delta_{n,d} = \det M_{n,d}. \quad (3.2.4)$$

Os elementos da matriz de momentos são $\mathcal{L}_s(\mathbf{x}^{\alpha+\beta})$ para $|\alpha| \leq n$ e $|\beta| \leq n$.

Exemplo 3.2.7. Para $d = 2$ e $n = 2$ temos que a matriz de momentos é dada por:

$$\begin{aligned}
M_{2,2} &= (\mathbf{s}_{\{k\}+\{j\}})_{k,j=0}^2 = \begin{pmatrix} \mathbf{s}_{\{0\}+\{0\}} & \mathbf{s}_{\{0\}+\{1\}} & \mathbf{s}_{\{0\}+\{2\}} \\ \mathbf{s}_{\{1\}+\{0\}} & \mathbf{s}_{\{1\}+\{1\}} & \mathbf{s}_{\{1\}+\{2\}} \\ \mathbf{s}_{\{2\}+\{0\}} & \mathbf{s}_{\{2\}+\{1\}} & \mathbf{s}_{\{2\}+\{2\}} \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{L}_s(1) & \mathcal{L}_s(1(\mathbb{X}^1)^T) & \mathcal{L}_s(1(\mathbb{X}^2)^T) \\ \mathcal{L}_s((\mathbb{X}^1)1) & \mathcal{L}_s((\mathbb{X}^1)(\mathbb{X}^1)^T) & \mathcal{L}_s((\mathbb{X}^1)(\mathbb{X}^2)^T) \\ \mathcal{L}_s((\mathbb{X}^2)1) & \mathcal{L}_s((\mathbb{X}^2)(\mathbb{X}^1)^T) & \mathcal{L}_s((\mathbb{X}^2)(\mathbb{X}^2)^T) \end{pmatrix} \\
&= \begin{pmatrix} \mathcal{L}_s(1) & \mathcal{L}_s(x) & \mathcal{L}_s(y) & \mathcal{L}_s(x^2) & \mathcal{L}_s(xy) & \mathcal{L}_s(y^2) \\ \mathcal{L}_s(x) & \mathcal{L}_s(x^2) & \mathcal{L}_s(xy) & \mathcal{L}_s(x^3) & \mathcal{L}_s(x^2y) & \mathcal{L}_s(xy^2) \\ \mathcal{L}_s(y) & \mathcal{L}_s(xy) & \mathcal{L}_s(y^2) & \mathcal{L}_s(x^2y) & \mathcal{L}_s(xy^2) & \mathcal{L}_s(y^3) \\ \mathcal{L}_s(x^2) & \mathcal{L}_s(x^3) & \mathcal{L}_s(x^2y) & \mathcal{L}_s(x^4) & \mathcal{L}_s(x^3y) & \mathcal{L}_s(x^2y^2) \\ \mathcal{L}_s(xy) & \mathcal{L}_s(x^2y) & \mathcal{L}_s(xy^2) & \mathcal{L}_s(x^3y) & \mathcal{L}_s(x^2y^2) & \mathcal{L}_s(xy^3) \\ \mathcal{L}_s(y^2) & \mathcal{L}_s(xy^2) & \mathcal{L}_s(y^3) & \mathcal{L}_s(x^2y^2) & \mathcal{L}_s(xy^3) & \mathcal{L}_s(y^4) \end{pmatrix} \\
&= \begin{pmatrix} s_{(0,0)} & s_{(1,0)} & s_{(0,1)} & s_{(2,0)} & s_{(1,1)} & s_{(0,2)} \\ s_{(1,0)} & s_{(2,0)} & s_{(1,1)} & s_{(3,0)} & s_{(2,1)} & s_{(1,2)} \\ s_{(0,1)} & s_{(1,1)} & s_{(0,2)} & s_{(2,1)} & s_{(1,2)} & s_{(0,3)} \\ s_{(2,0)} & s_{(3,0)} & s_{(2,1)} & s_{(4,0)} & s_{(3,1)} & s_{(2,2)} \\ s_{(1,1)} & s_{(2,1)} & s_{(1,2)} & s_{(3,1)} & s_{(2,2)} & s_{(1,3)} \\ s_{(0,2)} & s_{(1,2)} & s_{(0,3)} & s_{(2,2)} & s_{(1,3)} & s_{(0,4)} \end{pmatrix}
\end{aligned}$$

Note que a matriz $M_{n,d}$ é simétrica. De fato,

$$\begin{aligned}
M_{n,d}^T &= \left((\mathbf{s}_{\{k\}+\{j\}})_{k,j=0}^n \right)^T = (\mathbf{s}_{\{j\}+\{k\}}^T)_{j,k=0}^n = \left(\left(\mathcal{L}_s(\mathbb{X}^j(\mathbb{X}^k)^T)^T \right) \right)_{j,k=0}^n \\
&= \left(\mathcal{L}_s(\mathbb{X}^k(\mathbb{X}^j)^T) \right)_{k,j=0}^n = (\mathbf{s}_{\{k\}+\{j\}})_{k,j=0}^n = M_{n,d}.
\end{aligned}$$

Por simplicidade denotaremos, a partir de agora, \mathcal{L}_s por \mathcal{L} .

Definição 3.2.8. Dizemos que P é um polinômio ortogonal de grau n com respeito ao funcional de momento \mathcal{L} , se P é um polinômio e $\mathcal{L}(PQ) = 0$ para todo polinômio Q de grau menor ou igual a $n - 1$.

Denotamos por \mathcal{V}_n^d o espaço de polinômios ortogonais de grau exatamente n , isto é,

$$\mathcal{V}_n^d = \{P \in \Pi_n^d : \mathcal{L}(P, Q) = 0, \forall Q \in \Pi_{n-1}^d\}. \quad (3.2.5)$$

A dimensão de \mathcal{V}_n^d é a mesma de \mathcal{P}_n^d . Daí, é natural usar um multi-índice para indexar o elemento de uma base ortogonal de \mathcal{V}_n^d .

Podemos ainda considerar os monômios $\{\mathbf{x}^\alpha : \alpha \in \mathbb{N}_0^d\}$ como base de Π^d , ordená-los de acordo com a ordem lexicográfica e, pelo processo de ortogonalização de Gram-Schmidt, gerar uma base que é mutuamente ortogonal com respeito a \mathcal{L} . Porém, como não é natural ordenar o conjunto $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$, iremos considerar a ortogonalidade somente em relação aos polinômios de grau menor, isto é, polinômios de mesmo grau são ortogonais a polinômios de grau menor, porém podem não ser ortogonais entre si.

Definição 3.2.9. *Um funcional linear de momento \mathcal{L} é definido positivo se*

$$\mathcal{L}(P^2) > 0, \quad \forall P \in \Pi^d, \quad P \neq 0.$$

Também dizemos que a sequência $\{s_\alpha\}$ é definida positiva quando \mathcal{L} o é.

Se $P = \sum_{\alpha} a_{\alpha} \mathbf{x}^{\alpha}$ é um polinômio em Π^d , então $\mathcal{L}(P) = \sum_{\alpha} a_{\alpha} s_{\alpha}$. Em termos da sequência s , para \mathcal{L} ser definido positivo é necessário que

$$\sum_{\alpha, \beta} a_{\alpha} a_{\beta} s_{\alpha + \beta} > 0,$$

onde $\alpha + \beta = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \dots, \alpha_d + \beta_d)$, para cada sequência $a = (a_{\alpha})_{\alpha \in \mathbb{N}_0^d}$ em que $a_{\alpha} = 0$ para um número finitos de α .

Definição 3.2.10. *Um funcional linear \mathcal{L} definido em Π^d é quase-definido se existe uma base $B \in \Pi^d$ tal que, para qualquer $P, Q \in B$,*

$$\mathcal{L}(PQ) = 0, \quad \text{se} \quad P \neq Q \quad \text{e} \quad \mathcal{L}(P^2) \neq 0.$$

Note que um funcional linear \mathcal{L} definido positivo é quase-definido, porém quase-definido não implica em definido positivo.

Definição 3.2.11. *Denotamos por \mathbb{P}_n o vetor polinomial coluna definido por*

$$\mathbb{P}_n = (P_{\alpha}^n)_{|\alpha|=n} = (P_{\alpha^{(1)}}^n, P_{\alpha^{(2)}}^n, \dots, P_{\alpha^{(r_n)}}^n)^T, \quad (3.2.6)$$

onde $\{P_{\alpha}^n\}_{|\alpha|=n}$ é uma sequência de polinômios em Π_n^d e $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(r_n)}$ são os elementos de $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$ arranjados na ordem lexicográfica.

Exemplo 3.2.12. Um vetor polinomial coluna em duas variáveis, $d = 2$, e de grau 3 é dado por

$$\mathbb{P}_3 = \begin{pmatrix} P_{(3,0)}^3 \\ P_{(2,1)}^3 \\ P_{(1,2)}^3 \\ P_{(0,3)}^3 \end{pmatrix} = \begin{pmatrix} 2x^3 + 4x^2y - x^2 + 3y + \frac{2}{5} \\ 5x^3 + 2x^2y + y^3 - y^2 + x - 7y + 6 \\ 4x^2y + 8xy^2 + 4x - \frac{9}{8}y + 6 \\ x^2y - y^3 + 4xy - x + 8y \end{pmatrix}.$$

Definição 3.2.13. *Seja \mathcal{L} um funcional de momento. Uma sequência de polinômios $\{P_\alpha^n : |\alpha| = n, n \in \mathbb{N}_0\}$, $P_\alpha^n \in \Pi_n^d$, é ortogonal com respeito a \mathcal{L} se*

$$\mathcal{L}(\mathbb{X}^m \mathbb{P}_n^T) = 0, \quad m < n \quad e \quad \mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) = S_n, \quad (3.2.7)$$

onde S_n é uma matriz inversível de ordem r_n^d .

Por simplicidade dizemos que o vetor polinomial coluna \mathbb{P}_n , que satisfaz (3.2.7), é um polinômio ortogonal com respeito a \mathcal{L} . O conjunto formado pelos polinômios ortogonais, $\{\mathbb{P}_n\}_{n=0}^\infty$ é chamado sistema de polinômios ortogonais. Note que, pela definição de polinômios ortogonais,

$$\mathcal{L}(\mathbb{X}^m \mathbb{P}_n^T) = 0 \iff \mathcal{L}(\mathbf{x}^\beta P_\alpha^n) = 0, \quad |\alpha| = n, \quad |\beta| = m.$$

Definição 3.2.14. *Um vetor polinomial coluna é mônico se todos os seus elementos P_α são polinômios mônicos.*

Para simplificar, chamamos de polinômio mônico o vetor da definição anterior e o denotamos por $\bar{\mathbb{P}}_n$.

Veremos a seguir algumas consequências imediatas da definição de polinômios ortogonais.

Proposição 3.2.15. *Sejam \mathcal{L} um funcional de momento e $\{\mathbb{P}_n\}$ um sistema de polinômios ortogonais com relação a \mathcal{L} . Então, $\{\mathbb{P}_0, \mathbb{P}_1, \dots, \mathbb{P}_n\}$ forma uma base para Π_n^d .*

Dem.: Sejam $\mathbf{a}_0^T \mathbb{P}_0 + \mathbf{a}_1^T \mathbb{P}_1 + \dots + \mathbf{a}_n^T \mathbb{P}_n = 0$, onde $\mathbf{a}_i \in \mathbb{R}^{r_i}$ e a aplicação

$$\mathcal{L}(\mathbb{X}^k (\mathbf{a}_0^T \mathbb{P}_0 + \mathbf{a}_1^T \mathbb{P}_1 + \dots + \mathbf{a}_n^T \mathbb{P}_n)^T).$$

Assim, para $k = 0$, temos

$$0 = \mathcal{L}(\mathbb{X}^0 (\mathbf{a}_0^T \mathbb{P}_0 + \mathbf{a}_1^T \mathbb{P}_1 + \dots + \mathbf{a}_n^T \mathbb{P}_n)^T) = \mathcal{L}(\mathbb{X}^0 \mathbb{P}_0^T) \mathbf{a}_0 = S_0 \mathbf{a}_0.$$

Pela definição de polinômios ortogonais, como S_0 é inversível, segue que $\mathbf{a}_0 = 0$. Para $k = 1$, como $\mathbf{a}_0 = 0$, segue

$$0 = \mathcal{L}(\mathbb{X}^1 (\mathbf{a}_0^T \mathbb{P}_0 + \mathbf{a}_1^T \mathbb{P}_1 + \dots + \mathbf{a}_n^T \mathbb{P}_n)^T) = \mathcal{L}(\mathbb{X}^1 \mathbb{P}_0^T) \mathbf{a}_0 + \mathcal{L}(\mathbb{X}^1 \mathbb{P}_1^T) \mathbf{a}_1 = S_1 \mathbf{a}_1.$$

Como S_1 é inversível, temos que $\mathbf{a}_1 = 0$. Analogamente, para $k = 2, 3, \dots, n$, obtemos $\mathbf{a}_k = 0$. Portanto, $\{\mathbb{P}_0, \mathbb{P}_1, \dots, \mathbb{P}_n\}$ é linearmente independente e forma uma base para Π_n^d . \square

O polinômio ortogonal \mathbb{P}_n pode ser escrito em função dos vetores \mathbb{X}^k como

$$\mathbb{P}_n = G_{n,n}\mathbb{X}^n + G_{n,n-1}\mathbb{X}^{n-1} + \dots + G_{n,0}\mathbb{X}^0, \quad (3.2.8)$$

onde $G_{n,i}$ são matrizes coeficientes de dimensão $r_n^d \times r_i^d$ e $G_n = G_{n,n}$ é chamada de matriz coeficiente do termo de maior grau de \mathbb{P}_n .

Proposição 3.2.16. *Seja $\{\mathbb{P}_n\}$ um sistema de polinômios ortogonais com respeito a \mathcal{L} . Então, a matriz coeficiente do termo de maior grau G_n é inversível.*

Dem.: Pela Proposição 3.2.15, existe uma matriz G'_n tal que

$$\mathbb{X}^n = G'_n\mathbb{P}_n + \mathbb{Q}_{n-1},$$

onde \mathbb{Q}_{n-1} é um vetor cujas componentes pertencem a Π_{n-1}^d . Mas, de (3.2.8), temos que

$$\mathbb{X}^n = G'_n(G_n\mathbb{X}^n + G_{n,n-1}\mathbb{X}^{n-1} + \dots + G_{n,0}\mathbb{X}^0) + \mathbb{Q}_{n-1} = G'_n G_n \mathbb{X}^n + \mathbb{Q}'_{n-1}$$

e comparando as componentes \mathbb{X}^n , segue que $G'_n G_n = I$. Portanto, G_n é inversível. \square

Proposição 3.2.17. *Seja \mathbb{P}_n o polinômio ortogonal com relação a \mathcal{L} . Então, a matriz $H_n = \mathcal{L}(\mathbb{P}_n\mathbb{P}_n^T)$ é inversível e simétrica.*

Dem.: Como $H_n = \mathcal{L}(\mathbb{P}_n\mathbb{P}_n^T) = G_n\mathcal{L}(\mathbb{X}^n\mathbb{P}_n^T) = G_n S_n$, que é inversível pela definição de S_n e pela Proposição 3.2.16, obtemos

$$H_n^T = (\mathcal{L}(\mathbb{P}_n\mathbb{P}_n^T))^T = \mathcal{L}\left((\mathbb{P}_n\mathbb{P}_n^T)^T\right) = (\mathcal{L}(\mathbb{P}_n\mathbb{P}_n^T)) = H_n.$$

Portanto, H_n é simétrica. \square

Teorema 3.2.18. *Sejam \mathcal{L} um funcional de momento quase-definido e $\{\mathbb{P}_n\}$ um sistema de polinômios. Então, as seguintes afirmações são equivalentes:*

1. *Os polinômios são ortogonais de acordo com a Definição 3.2.13.*
2. *Os polinômios satisfazem*

$$\mathcal{L}(\mathbb{P}_m\mathbb{P}_n^T) = \begin{cases} 0, & \text{se } n \neq m \\ H_n, & \text{se } n = m \end{cases},$$

cuja matriz H_n é inversível.

Dem.: Mostremos que (1) implica (2). De (3.2.8) segue que $\mathcal{L}(\mathbb{P}_m \mathbb{P}_n^T) = 0$ caso $m < n$ e $\mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) = H_n$. Portanto, basta mostrar que $\mathcal{L}(\mathbb{P}_m \mathbb{P}_n^T) = 0$ caso $m > n$, pois, da Proposição 3.2.17, segue que H_n é inversível.

Da Definição 3.2.13 temos $\mathcal{L}(\mathbb{P}_m(\mathbb{X}^n)^T) = 0$, $m > n$, então

$$\begin{aligned} \mathcal{L}(\mathbb{P}_m \mathbb{P}_n^T) &= \mathcal{L}(\mathbb{P}_m(G_n \mathbb{X}^n + \dots + G_{n,0} \mathbb{X}^0)^T) \\ &= \mathcal{L}(\mathbb{P}_m(\mathbb{X}^n)^T)G_n^T + \dots + \mathcal{L}(\mathbb{P}_m(\mathbb{X}^0)^T)G_{n,0}^T = 0. \end{aligned}$$

Por fim, provemos que (2) implica (1). Pela Proposição 3.2.15, segue que

$$\begin{aligned} \mathcal{L}(\mathbb{X}^m \mathbb{P}_n^T) &= \mathcal{L}((M_m \mathbb{P}_m + M_{m-1} \mathbb{P}_{m-1} + \dots + M_0 \mathbb{P}_0) \mathbb{P}_n^T) \\ &= \sum_{i=0}^m M_i \mathcal{L}(\mathbb{P}_i \mathbb{P}_n^T). \end{aligned} \quad (3.2.9)$$

Caso $m < n$, segue $\mathcal{L}(\mathbb{X}^m \mathbb{P}_n^T) = 0$. Caso $n = m$, temos $\mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) = M_n H_n$, onde M_n é inversível. De fato, fazendo uma demonstração semelhante à da Proposição 3.2.16, de (3.2.8) temos $\mathbb{P}_n = G_n \mathbb{X}^n + \mathbb{Q}_{n-1}$, onde \mathbb{Q}_{n-1} é um vetor cujas componentes pertencem a Π_{n-1}^d . Mas, de (3.2.9), obtemos $\mathbb{P}_n = G_n M_n \mathbb{P}_n + \mathbb{Q}'_{n-1}$ e, portanto, M_n é inversível. Logo, $\mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) = S_n$ é inversível. \square

Do Teorema 3.2.18 temos outra definição para polinômios ortogonais.

Definição 3.2.19. *Seja \mathcal{L} um funcional de momento. Um sistema de polinômios, $\{\mathbb{P}_n\}$, é ortogonal se*

$$\mathcal{L}(\mathbb{P}_m \mathbb{P}_n^T) = 0, \quad n \neq m, \quad e \quad \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) = H_n,$$

onde H_n é inversível de ordem r_n^d .

Definição 3.2.20. *Seja \mathcal{L} um funcional de momento. Um sistema de polinômios ortogonais mônicos, $\{\bar{\mathbb{P}}_n\}$, é definido por*

$$\mathcal{L}(\bar{\mathbb{P}}_n \bar{\mathbb{P}}_m^T) = \begin{cases} 0, & n \neq m \\ \bar{H}_n, & n = m \end{cases}, \quad (3.2.10)$$

onde \bar{H}_n é inversível e $\bar{\mathbb{P}}_n$ é um polinômio mônico.

Proposição 3.2.21. *Se $\{\mathbb{P}_n\}$ é um sistema de polinômios ortogonais, então $\bar{\mathbb{P}}_n = G_n^{-1} \mathbb{P}_n$ forma um sistema de polinômios ortogonais mônicos.*

Dem.: Provemos que $\{\bar{\mathbb{P}}_n\}$ é um sistema de polinômios ortogonais mônicos.

$$\mathcal{L}(\bar{\mathbb{P}}_n \bar{\mathbb{P}}_m^T) = \mathcal{L}(G_n^{-1} \mathbb{P}_n (G_m^{-1} \mathbb{P}_m)^T) = G_n^{-1} (\mathcal{L}(\mathbb{P}_n \mathbb{P}_m^T)) (G_m^{-1})^T.$$

Logo,

$$\mathcal{L}(\bar{\mathbb{P}}_n \bar{\mathbb{P}}_m^T) = \begin{cases} 0, & n \neq m \\ G_n^{-1} H_n (G_m^{-1})^T, & n = m \end{cases},$$

onde $G_n^{-1} H_n (G_m^{-1})^T$ é inversível pelas Proposições 2.1.3 e 3.2.16. Por (3.2.8), temos

$$\begin{aligned} \bar{\mathbb{P}}_n &= G_n^{-1} \mathbb{P}_n = G_n^{-1} (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + \dots + G_{n0} \mathbb{X}^0) \\ &= \mathbb{X}^n + G_n^{-1} G_{n,n-1} \mathbb{X}^{n-1} + \dots + G_n^{-1} G_{n0} \mathbb{X}^0. \end{aligned}$$

Portanto, $\{\bar{\mathbb{P}}_n\}$ é um sistema de polinômios ortogonais mônicos. \square

Lema 3.2.22. *Sejam \mathcal{L} um funcional de momento e $\{\mathbb{P}_n\}$ um sistema de polinômios ortogonais com respeito a \mathcal{L} . Então, \mathbb{P}_n é unicamente determinado pela matriz S_n .*

Dem.: Suponha que existam \mathbb{P}_n e \mathbb{P}_n^* que satisfazem a condição de ortogonalidade de (3.2.7) com respeito ao mesmo S_n . Sejam G_n e G_n^* as matrizes coeficientes dos termos de maior grau de \mathbb{P}_n e \mathbb{P}_n^* , respectivamente. Pela Proposição 3.2.15, $\{\mathbb{P}_0, \dots, \mathbb{P}_n\}$ forma uma base de Π_n^d e, portanto, existem $C_i : r_n^d \times r_i^d$ tais que

$$\mathbb{P}_n^* = C_n \mathbb{P}_n + C_{n-1} \mathbb{P}_{n-1} + \dots + C_0 \mathbb{P}_0.$$

Assim, multiplicando ambos os lados da equação acima, à direita, por \mathbb{P}_k^T e aplicando o funcional de momento \mathcal{L} , segue, da definição de polinômios ortogonais, que

$$0 = \mathcal{L}(\mathbb{P}_n^* \mathbb{P}_k^T) = C_k \mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T) = C_k H_k, \quad 0 \leq k \leq n-1.$$

Da Proposição 3.2.17, temos $C_k = 0$ para $0 \leq k \leq n-1$. Portanto, $\mathbb{P}_n^* = C_n \mathbb{P}_n$. Agora, comparando os coeficientes de \mathbb{X}^n segue que $G_n^* = C_n G_n$, ou seja, $C_n = G_n^* G_n^{-1}$. Então, $\mathbb{P}_n^* = G_n^* G_n^{-1} \mathbb{P}_n$ e, assim,

$$S_n = \mathcal{L}(\mathbb{X}^n \mathbb{P}_n^{*T}) = \mathcal{L}(\mathbb{X}^n (G_n^* G_n^{-1} \mathbb{P}_n)^T) = \mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) (G_n^* G_n^{-1})^T = S_n (G_n^* G_n^{-1})^T.$$

Portanto, $G_n^* G_n^{-1} = I$ e $\mathbb{P}_n = \mathbb{P}_n^*$. \square

Teorema 3.2.23. *Seja \mathcal{L} um funcional de momento. Um sistema de polinômios ortogonais em várias variáveis existe se, e somente se,*

$$\Delta_{n,d} \neq 0, \quad n \in \mathbb{N}_0,$$

onde $\Delta_{n,d}$ é definido em (3.2.4).

Dem.: De (3.2.8) e da notação $\mathbf{s}_{\{k\}+\{j\}}$ em (3.2.3), temos

$$\begin{aligned}\mathcal{L}(\mathbb{X}^k \mathbb{P}_n^T) &= \mathcal{L}(\mathbb{X}^k (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + \dots + G_{n,0} \mathbb{X}^0)^T) \\ &= \mathbf{s}_{\{k\}+\{n\}} G_n^T + \mathbf{s}_{\{k\}+\{n-1\}} G_{n,n-1}^T + \dots + \mathbf{s}_{\{k\}+\{0\}} G_{n,0}^T.\end{aligned}$$

Para $k = 0, 1, \dots, n$, obtemos

$$\begin{aligned}\mathbf{s}_{\{0\}+\{0\}} G_{n,0}^T + \mathbf{s}_{\{0\}+\{1\}} G_{n,1}^T + \dots + \mathbf{s}_{\{0\}+\{n\}} G_n^T &= \mathcal{L}(\mathbb{X}^0 \mathbb{P}_n^T) = 0, \\ \mathbf{s}_{\{1\}+\{0\}} G_{n,0}^T + \mathbf{s}_{\{1\}+\{1\}} G_{n,1}^T + \dots + \mathbf{s}_{\{1\}+\{n\}} G_n^T &= \mathcal{L}(\mathbb{X}^1 \mathbb{P}_n^T) = 0, \\ &\vdots \\ \mathbf{s}_{\{n\}+\{0\}} G_{n,0}^T + \mathbf{s}_{\{n\}+\{1\}} G_{n,1}^T + \dots + \mathbf{s}_{\{n\}+\{n\}} G_n^T &= \mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) = S_n.\end{aligned}$$

Assim, da definição de $M_{n,d}$ em (3.2.4), segue o sistema de equações lineares

$$M_{n,d} \begin{bmatrix} G_{n,0}^T \\ \vdots \\ G_{n,n-1}^T \\ G_n^T \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ S_n \end{bmatrix}. \quad (3.2.11)$$

Mostremos a condição necessária. Se um sistema de polinômios ortogonais existe, então, pelo Lema 3.2.22, para todo S_n existe exatamente um \mathbb{P}_n . Portanto, o sistema de equações (3.2.11) tem uma única solução, o que implica que $M_{n,d}$ é inversível e, então, $\Delta_{n,d} \neq 0$.

Por outro lado, se $\Delta_{n,d} \neq 0$, então para cada matriz S_n em (3.2.11), há uma única solução $\{G_{n,0}, G_{n,1}, \dots, G_n\}$. Seja $\mathbb{P}_n = \sum_{k=0}^n G_{n,k} \mathbb{X}^k$. Então (3.2.11) é equivalente a $\mathcal{L}(\mathbb{X}^k \mathbb{P}_n^T) = 0$, $k < n$, e $\mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T) = S_n$. \square

Lema 3.2.24. *Se \mathcal{L} é definido positivo, então o determinante $\Delta_{n,d} > 0$.*

Dem.: Suponha que \mathcal{L} seja definido positivo. Seja \mathbf{a} um auto-vetor da matriz $M_{n,d}$ correspondente ao auto-valor λ . Então, $\mathbf{a}^T M_{n,d} \mathbf{a} = \lambda \|\mathbf{a}\|^2$. Por outro lado, seja $\mathbf{a} = (\mathbf{a}_0, \mathbf{a}_1, \dots, \mathbf{a}_n)$ tal que $\mathbf{a}_j \in \mathbb{R}^{r_j}$ e, assim, $p(x) = \sum_{j=0}^n \mathbf{a}_j^T \mathbb{X}^j$. Portanto, como \mathcal{L}_s é definido positivo,

$$\begin{aligned}0 &< \mathcal{L}_s(p^2) = \mathcal{L}_s \left(\left(\sum_{j=0}^n \mathbf{a}_j^T \mathbb{X}^j \right) \left(\sum_{i=0}^n \mathbf{a}_i^T \mathbb{X}^i \right) \right) \\ &= \mathcal{L}_s \left(\left(\sum_{j=0}^n \mathbf{a}_j^T \mathbb{X}^j \right) \left(\sum_{i=0}^n \mathbf{a}_i^T \mathbb{X}^i \right)^T \right) \\ &= \sum_{j=0}^n \sum_{i=0}^n \mathbf{a}_j^T \mathcal{L}_s \left(\mathbb{X}^j (\mathbb{X}^i)^T \right) \mathbf{a}_i = \mathbf{a}^T M_{n,d} \mathbf{a}.\end{aligned}$$

Logo, $\lambda > 0$. Como todos os auto-valores são positivos e $M_{n,d}$ é simétrica, então, pelo Teorema 2.1.13, $\Delta_{n,d} = \det(M_{n,d}) = \det(O^T D_n O) = \det(D_n) > 0$, onde O é uma matriz ortogonal e D_n é a matriz diagonal cujos elementos são os auto-valores de $M_{n,d}$. \square

O corolário a seguir é consequência do Teorema 3.2.23 e do Lema 3.2.24.

Corolário 3.2.25. *Se \mathcal{L} é um funcional de momento definido positivo, então existe um sistema de polinômios ortogonais com respeito a \mathcal{L} .*

Definição 3.2.26. *Seja \mathcal{L} um funcional de momento. Uma sequência de polinômios $\{P_\alpha^n : |\alpha| = n, n \in \mathbb{N}_0\}$, $P_\alpha^n \in \Pi_n^d$, é ortonormal com respeito a \mathcal{L} se*

$$\mathcal{L}(P_\alpha^n P_\beta^m) = \delta_{\alpha,\beta}.$$

Em notação vetorial, a equação acima torna-se

$$\mathcal{L}(\mathbb{P}_m \mathbb{P}_n^T) = 0, \quad m \neq n, \quad e \quad \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) = I_{r_n}, \quad (3.2.12)$$

onde I_{r_n} é a matriz identidade de ordem r_n .

Teorema 3.2.27. *Se \mathcal{L} é um funcional de momento definido positivo, então existe uma base ortonormal com respeito a \mathcal{L} .*

Dem.: Pelo Corolário 3.2.25, existe uma base de polinômios ortogonais \mathbb{P}_n com respeito a \mathcal{L} . Seja $H_n = \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T)$. Para qualquer vetor não-nulo \mathbf{a} , $P = \mathbf{a}^T \mathbb{P}_n$ é um polinômio não-nulo pela Proposição 3.2.15. Por isso, $\mathbf{a} H_n \mathbf{a}^T = \mathcal{L}((\mathbf{a}^T \mathbb{P}_n)(\mathbb{P}_n^T \mathbf{a})) = \mathcal{L}(P^2) > 0$, o que mostra que H_n é uma matriz definida positiva. Pela Proposição 3.2.17 e pelo Teorema 2.1.13, segue que $H_n = O^T D_n O$, onde O é uma matriz ortogonal e D_n é a matriz diagonal formada pelos auto-valores. Denotamos por $H_n^{\frac{1}{2}} = O^T D_n^{\frac{1}{2}} O$ a matriz raiz quadrada positiva de H_n , onde $D_n^{\frac{1}{2}}$ é a matriz diagonal formada pelas raízes quadradas dos auto-valores. Assim, $H_n = H_n^{\frac{1}{2}} H_n^{\frac{1}{2}}$ e $\det(H_n^{\frac{1}{2}}) \neq 0$. Seja então $\mathbb{Q}_n = (H_n^{\frac{1}{2}})^{-1} \mathbb{P}_n$. Então,

$$\mathcal{L}(\mathbb{Q}_n \mathbb{Q}_n^T) = (H_n^{\frac{1}{2}})^{-1} \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) (H_n^{\frac{1}{2}})^{-1} = (H_n^{\frac{1}{2}})^{-1} H_n (H_n^{\frac{1}{2}})^{-1} = I,$$

o que mostra que os elementos de \mathbb{Q}_n formam uma base ortonormal com respeito a \mathcal{L} . \square

3.3 Polinômios ortogonais definidos a partir de matrizes de momento

Seja \mathcal{L} um funcional de momento definido positivo. Para cada $n \in \mathbb{N}_0$, seja $M_{n,d}$ a matriz de momentos de \mathcal{L} como definido em (3.2.4). Então, pelo Lema 3.2.24, $M_{n,d}$ tem determinante positivo. Pela Definição 3.2.6, cada linha de $M_{n,d}$ é determinada por um multi-índice e a linha com o índice α pode ser escrita com relação a \mathbb{X}^k da forma

$$\left(\mathcal{L}(\mathbf{x}^\alpha), \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^T), \dots, \mathcal{L}(\mathbf{x}^\alpha (\mathbb{X}^n)^T) \right) = \mathcal{L}\left(\mathbf{x}^\alpha \left(1, \mathbb{X}^T, \dots, (\mathbb{X}^n)^T\right)\right). \quad (3.3.1)$$

Primeiramente, definimos o polinômio ortogonal monomial em termos da matriz de momentos. Para cada $\alpha \in \mathbb{N}_0, |\alpha| = n$, denotemos por $M_\alpha(\mathbf{x})$ a matriz obtida de $M_{n-1,d}$ adicionando-se uma linha abaixo e uma coluna a direita da matriz $M_{n-1,d}$. A linha é dada por $\left(1, \mathbb{X}^T, \dots, (\mathbb{X}^{n-1})^T\right)$ e a coluna por $\left(\mathcal{L}(\mathbf{x}^\alpha), \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^T), \dots, \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^{n-1})^T\right)^T$, e adicionamos o elemento \mathbf{x}^α no canto inferior direito para tornar a matriz quadrada, ou seja,

$$M_\alpha(\mathbf{x}) = \begin{pmatrix} 1 & \dots & \mathcal{L}(\mathbb{X}^0(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha) \\ \mathcal{L}(\mathbb{X}^1(\mathbb{X}^0)^T) & \dots & \mathcal{L}(\mathbb{X}^1(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^1) \\ \vdots & \ddots & \vdots & \vdots \\ \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^0)^T) & \dots & \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^{n-1}) \\ 1 & \dots & (\mathbb{X}^{n-1})^T & \mathbf{x}^\alpha \end{pmatrix}.$$

Então, definimos o polinômio P_α^n por

$$P_\alpha^n(\mathbf{x}) = \frac{\det M_\alpha(\mathbf{x})}{\Delta_{n-1,d}}. \quad (3.3.2)$$

Teorema 3.3.1. *Os polinômios P_α^n definidos em (3.3.2) são polinômios ortogonais mônicos com relação a \mathcal{L} .*

Dem.: Expandindo-se o determinante de acordo com a última linha, vemos que os P_α^n assim definidos são polinômios mônicos, $P_\alpha^n(\mathbf{x}) = \mathbf{x}^\alpha + \mathbb{Q}_{n-1}(\mathbf{x})$, onde $\mathbb{Q}_{n-1} \in \Pi_{n-1}^d$. Além disso, multiplicando-se $P_\alpha(\mathbf{x})$ por \mathbf{x}^β com $|\beta| \leq n-1$,

$$\mathcal{L}(\mathbf{x}^\beta P_\alpha^n(\mathbf{x})) = \mathcal{L}\left(\mathbf{x}^\beta \left(\frac{\det M_\alpha(\mathbf{x})}{\Delta_{n-1,d}}\right)\right) = \frac{\mathcal{L}(\mathbf{x}^\beta \det M_\alpha(\mathbf{x}))}{\Delta_{n-1,d}}.$$

Agora,

$$\mathbf{x}^\beta \det M_\alpha = \det \begin{pmatrix} 1 & \dots & \mathcal{L}(\mathbb{X}^0(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha) \\ \mathcal{L}(\mathbb{X}^1(\mathbb{X}^0)^T) & \dots & \mathcal{L}(\mathbb{X}^1(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^1) \\ \vdots & \ddots & \vdots & \vdots \\ \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^0)^T) & \dots & \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^{n-1})^T) & \mathcal{L}(\mathbf{x}^\alpha \mathbb{X}^{n-1}) \\ \mathbf{x}^\beta & \dots & (\mathbf{x}^\beta \mathbb{X}^{n-1})^T & \mathbf{x}^{\alpha+\beta} \end{pmatrix}.$$

Portanto, aplicando \mathcal{L} à última linha, esta linha coincide com alguma linha acima e, conseqüentemente, $\mathcal{L}(\mathbf{x}^\beta P_\alpha^n(\mathbf{x})) = 0$. \square

Definiremos a seguir uma seqüência de base ortonormais em termos das matrizes de momentos. Para isso, seja \mathcal{L} um funcional de momento definido positivo e, para $\alpha \in \mathbb{N}_0^d$ com $|\alpha| = n$, definimos $\tilde{M}_\alpha(\mathbf{x})$ como a matriz obtida de $M_{n,d}$, em (3.2.4), substituindo-se a linha (3.3.1) do índice α por $(1, \mathbb{X}^T, \dots, (\mathbb{X}^n)^T)$, ou seja, com $\alpha^{(1)}, \alpha^{(2)}, \dots, \alpha^{(r_n)}$ sendo os elementos de $\{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$ arranjados na ordem lexicográfica, temos que se $\alpha = \alpha^{(j)}$ para algum $1 \leq j \leq r_n$, então

$$\tilde{M}_\alpha(\mathbf{x}) = \begin{pmatrix} 1 & \mathcal{L}(\mathbb{X}^0(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbb{X}^0(\mathbb{X}^n)^T) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^0)^T) & \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbb{X}^{n-1}(\mathbb{X}^n)^T) \\ \mathcal{L}(\mathbf{x}^{\alpha^{(1)}}(\mathbb{X}^0)^T) & \mathcal{L}(\mathbf{x}^{\alpha^{(1)}}(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbf{x}^{\alpha^{(1)}}(\mathbb{X}^n)^T) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(\mathbf{x}^{\alpha^{(j-1)}}(\mathbb{X}^0)^T) & \mathcal{L}(\mathbf{x}^{\alpha^{(j-1)}}(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbf{x}^{\alpha^{(j-1)}}(\mathbb{X}^n)^T) \\ 1 & \mathbb{X}^T & \cdots & (\mathbb{X}^n)^T \\ \mathcal{L}(\mathbf{x}^{\alpha^{(j+1)}}(\mathbb{X}^0)^T) & \mathcal{L}(\mathbf{x}^{\alpha^{(j+1)}}(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbf{x}^{\alpha^{(j+1)}}(\mathbb{X}^n)^T) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(\mathbf{x}^{\alpha^{(r_n)}}(\mathbb{X}^0)^T) & \mathcal{L}(\mathbf{x}^{\alpha^{(r_n)}}(\mathbb{X}^1)^T) & \cdots & \mathcal{L}(\mathbf{x}^{\alpha^{(r_n)}}(\mathbb{X}^n)^T) \end{pmatrix}. \quad (3.3.3)$$

Assim, definimos o polinômio $\tilde{P}_\alpha^n \in \Pi_n^d$ por

$$\tilde{P}_\alpha^n(\mathbf{x}) = \frac{\det \tilde{M}_\alpha(\mathbf{x})}{\Delta_{n,d}}, \quad \alpha \in \mathbb{N}_0^n \quad \text{e} \quad |\alpha| = n. \quad (3.3.4)$$

Com uma análise análoga à feita no teorema anterior, obtemos

$$\mathcal{L}(\mathbf{x}^\beta \tilde{P}_\alpha^n) = \begin{cases} 0, & \text{se } |\beta| \leq n-1, \\ \delta_{\alpha,\beta}, & \text{se } |\beta| = n. \end{cases} \quad (3.3.5)$$

Então, $\left\{ \tilde{P}_\alpha^n \right\}_{|\alpha|=n}^{\infty}$ é uma seqüência de polinômios ortogonais com respeito a \mathcal{L} .

Seja $M_{n,d}^{-1}(|\alpha| = n)$ a submatriz de $M_{n,d}^{-1} = (m_{\alpha,\beta})_{|\alpha|,|\beta| \leq n}$, que consiste dos elementos $m_{\alpha,\beta}$ cujos índices satisfazem a condição $|\alpha| = |\beta| = n$. De (2.1.4), os elementos de $M_{n,d}^{-1}(|\alpha| = n)$ são os cofatores dos elementos de $\mathcal{L}(\mathbb{X}^n(\mathbb{X}^n)^T)$ de $M_{n,d}$. Como $M_{n,d}$ é positiva definida, então, pela Proposição 2.1.6, $M_{n,d}^{-1}$ também é definida positiva. Segue, então, que $M_{n,d}^{-1}(|\alpha| = n)$ é positiva definida.

Teorema 3.3.2. *Seja \mathcal{L} um funcional de momento definido positivo. Então, os polinômios*

$$\mathbb{P}_n = (M_{n,d}^{-1}(|\alpha| = n))^{-\frac{1}{2}} \tilde{\mathbb{P}}_n = G_n \mathbb{X}^n + \dots \quad (3.3.6)$$

são polinômios ortonormais com respeito a \mathcal{L} e G_n é definida positiva.

Dem.: Pela definição de $\tilde{M}_\alpha(x)$ em (3.3.3), expandindo o determinante sobre a linha $(1, \mathbb{X}^T, \dots, (\mathbb{X}^n)^T)$, obtemos

$$\det \tilde{M}_\alpha(\mathbf{x}) = \sum_{|\beta|=n} C_{\alpha,\beta} \mathbf{x}^\beta + Q_\alpha^{n-1}, \quad Q_\alpha^{n-1} \in \Pi_{n-1}^d,$$

onde $C_{\alpha,\beta}$ é o cofator do elemento $\mathcal{L}(\mathbf{x}^\alpha \mathbf{x}^\beta)$ em $M_{n,d}$, isto é, $\text{adj}(M_{n,d}) = (C_{\alpha,\beta})_{|\alpha|,|\beta| \leq n}$. Portanto, como $\text{adj}(M_{n,d}) = \Delta_{n,d} M_{n,d}^{-1}$, podemos escrever

$$\tilde{\mathbb{P}}_n = \frac{1}{\Delta_{n,d}} \Delta_{n,d} M_{n,d}^{-1} (|\alpha| = n) \mathbb{X}^n + \mathbb{Q}_{n-1} = M_{n,d}^{-1} (|\alpha| = n) \mathbb{X}^n + \mathbb{Q}_{n-1}.$$

Segue, de (3.3.5), que $\mathcal{L}(\mathbb{X}^n \tilde{\mathbb{P}}_n^T) = I_{r_n}$ e $\mathcal{L}(\mathbb{Q} \tilde{\mathbb{P}}_n^T) = 0$ e, por isso,

$$\mathcal{L}(\tilde{\mathbb{P}}_n \tilde{\mathbb{P}}_n^T) = M_{n,d}^{-1} (|\alpha| = n).$$

Logo, $\mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) = I$ e então \mathbb{P}_n é ortonormal. Além disso,

$$G_n = (M_{n,d}^{-1} (|\alpha| = n))^{\frac{1}{2}}. \quad (3.3.7)$$

Portanto, G_n é definida positiva. □

Proposição 3.3.3. *Se $\tilde{\mathbb{P}}_n$ é ortonormal com relação a \mathcal{L} , então para qualquer matriz ortogonal O_n de ordem r_n^d , os polinômios $\mathbb{P}_n^* = O_n \tilde{\mathbb{P}}_n$ também são ortonormais.*

Dem.: De fato,

$$\mathcal{L}(\mathbb{P}_n^* (\mathbb{P}_m^*)^T) = \mathcal{L}(O_n \tilde{\mathbb{P}}_n (O_m \tilde{\mathbb{P}}_m)^T) = O_n \mathcal{L}(\tilde{\mathbb{P}}_n \tilde{\mathbb{P}}_m^T) O_m^T = \begin{cases} 0, & n \neq m, \\ I_{r_n^d}, & n = m. \end{cases}$$

□

Teorema 3.3.4. *Seja \mathcal{L} definido positivo e seja Q_α^n uma sequência de polinômios ortogonais com respeito a \mathcal{L} . Então, existe uma matriz ortogonal O_n tal que $\mathbb{Q}_n = O_n \mathbb{P}_n$, onde \mathbb{P}_n são os polinômios ortonormais definidos em (3.3.6). Além disso, a matriz coeficiente do termo de maior grau G_n^* de \mathbb{Q}_n satisfaz $G_n^* = O_n G_n$, onde G_n é a matriz definida positiva dada em (3.3.7). Mais ainda,*

$$(\det G_n)^2 = \frac{\Delta_{n-1,d}}{\Delta_{n,d}} \quad (3.3.8)$$

Dem.: Pela Proposição 3.2.15, temos que $\mathbb{Q}_n = O_n \mathbb{P}_n + \dots + O_0 \mathbb{P}_0$. Multiplicando-se por \mathbb{P}_k^T à direita e aplicando \mathcal{L} , segue que $O_k = 0, 0 \leq k \leq n-1$. Logo, $\mathbb{Q}_n = O_n \mathbb{P}_n$. Como \mathbb{P}_n e \mathbb{Q}_n são ortonormais, temos

$$I = \mathcal{L}(\mathbb{Q}_n \mathbb{Q}_n^T) = \mathcal{L}(O_n \mathbb{P}_n (O_n \mathbb{P}_n)^T) = O_n O_n^T.$$

Portanto, O_n é uma matriz ortogonal. Comparando as matrizes coeficientes do termo de maior grau de $\mathbb{Q}_n = O_n \mathbb{P}_n$ e \mathbb{P}_n segue que $G_n^* = O_n G_n$. A igualdade (3.3.8) segue de (3.3.7) e do resultado sobre os menores da matriz inversa em [8], isto é,

$$\begin{aligned} (\det G_n)^2 &= \det G_n^2 = \det(M_{n,d}^{-1}(|\alpha| = n)) = \frac{\det M_{n,d}^{-1}(1 \leq |\alpha| \leq n-1)}{\Delta_{n,d}} \\ &= \frac{\Delta_{n-1,d}}{\Delta_{n,d}} \end{aligned}$$

□

Neste teorema, $\det G_n$ pode ser visto como uma extensão do coeficiente do termo de maior grau k_n dos polinômios ortogonais em uma variável.

3.4 Relação de três termos

A relação de recorrência de três termos é uma importante propriedade dos polinômios ortogonais de uma variável e há o análogo à essa relação para os polinômios de várias variáveis, que toma a forma de vetor-matriz. Seja \mathcal{L} um funcional de momento, os polinômios ortogonais com respeito ao funcional \mathcal{L} serão denotados por \mathbb{P}_n , os polinômios ortonormais serão denotados por $\tilde{\mathbb{P}}_n$ e $H_n = \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T)$.

Teorema 3.4.1. *Para $n \geq 0$, existem únicas matrizes $A_{n,i} : r_n^d \times r_{n+1}^d$, $B_{n,i} : r_n^d \times r_n^d$, e $C_{n,i} : r_n^d \times r_{n-1}^d$, tais que*

$$x_i \mathbb{P}_n = A_{n,i} \mathbb{P}_{n+1} + B_{n,i} \mathbb{P}_n + C_{n,i} \mathbb{P}_{n-1}, \quad 1 \leq i \leq d, \quad (3.4.1)$$

onde $\mathbb{P}_{-1} = 0$ e $C_{-1,i} = 0$. Além disso,

$$\begin{cases} A_{n,i} H_{n+1} = \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_{n+1}^T), \\ B_{n,i} H_n = \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_n^T), \\ A_{n,i} H_{n+1} = H_n C_{n+1,i}^T. \end{cases} \quad (3.4.2)$$

Dem.: Como $x_i \mathbb{P}_n$ é um vetor cujas componentes são polinômios de grau $n+1$, pela Proposição 3.2.15, ele pode ser escrito como combinação linear de polinômios ortogonais. Assim, existem matrizes $M_{k,i}$ tais que

$$x_i \mathbb{P}_n = M_{n+1,i} \mathbb{P}_n + M_{n,i} \mathbb{P}_n + \dots + M_{0,i} \mathbb{P}_0.$$

Multiplicando ambos os lados à direita por \mathbb{P}_k^T e aplicando \mathcal{L} , segue que $M_{k,i}H_k = \mathcal{L}(x_i\mathbb{P}_n\mathbb{P}_k^T)$. Pela ortogonalidade e pela Proposição 3.2.17, obtemos $M_{k,i} = 0$, $k = 0, 1, \dots, n-2$. Portanto,

$$x_i\mathbb{P}_n = A_{n,i}\mathbb{P}_{n+1} + B_{n,i}\mathbb{P}_n + C_{n,i}\mathbb{P}_{n-1}.$$

De fato, $A_{n,i}H_{n+1} = \mathcal{L}(x_i\mathbb{P}_n\mathbb{P}_{n+1}^T)$, $B_{n,i}H_n = \mathcal{L}(x_i\mathbb{P}_n\mathbb{P}_n^T)$ e $C_{n,i}H_{n-1} = \mathcal{L}(x_i\mathbb{P}_n\mathbb{P}_{n-1}^T)$. Logo, basta mostrar que $A_{n,i}H_{n+1} = H_n C_{n+1,i}^T$,

$$\begin{aligned} A_{n,i}H_{n+1} &= \mathcal{L}(x_i\mathbb{P}_n\mathbb{P}_{n+1}^T) = \left(\mathcal{L} \left((x_i\mathbb{P}_n\mathbb{P}_{n+1}^T)^T \right) \right)^T \\ &= \left(\mathcal{L} (x_i\mathbb{P}_{n+1}\mathbb{P}_n^T) \right)^T = (C_{n+1,i}H_n)^T \\ &= H_n^T C_{n+1,i}^T, \end{aligned}$$

pois da Proposição 3.2.17, H_n é simétrica. \square

Para os polinômios ortonormais $\tilde{\mathbb{P}}_n$, $H_n = I$, segue a seguinte relação de três termos.

Teorema 3.4.2. *Para $n \geq 0$, existem únicas matrizes $A_{n,i} : r_n^d \times r_{n+1}^d$ e $B_{n,i} : r_n^d \times r_n^d$, tais que*

$$x_i\tilde{\mathbb{P}}_n = A_{n,i}\tilde{\mathbb{P}}_{n+1} + B_{n,i}\tilde{\mathbb{P}}_n + A_{n-1,i}^T\tilde{\mathbb{P}}_{n-1}, \quad 1 \leq i \leq d, \quad (3.4.3)$$

onde $\tilde{\mathbb{P}}_{-1} = 0$ e $A_{-1,i} = 0$. Além disso, as matrizes $B_{n,i}$, $1 \leq i \leq d$, são simétricas.

Dem.: Segue, do teorema anterior, que (3.4.3) é válido. Basta mostrar que $B_{n,i}$ é simétrica. De fato,

$$B_{n,i} = \mathcal{L}(x_i\tilde{\mathbb{P}}_n\tilde{\mathbb{P}}_n^T) = \left(\mathcal{L}((x_i\tilde{\mathbb{P}}_n\tilde{\mathbb{P}}_n^T)^T) \right)^T = \left(\mathcal{L}(x_i\tilde{\mathbb{P}}_n\tilde{\mathbb{P}}_n^T) \right)^T = B_{n,i}^T.$$

\square

As matrizes coeficientes da relação de três termos satisfazem várias propriedades. Vejamos algumas, seja G_n a matriz coeficiente do termo de maior grau de \mathbb{P}_n (ou $\tilde{\mathbb{P}}_n$), como anteriormente definida. Comparando o maior coeficiente de ambos os lados de (3.4.1), obtemos

$$A_{n,i}G_{n+1} = G_n L_{n,i}, \quad (3.4.4)$$

onde $L_{n,i}$ são matrizes de dimensão $r_n^d \times r_{n+1}^d$ definidas por

$$L_{n,i}\mathbb{X}^{n+1} = x_i\mathbb{X}^n. \quad (3.4.5)$$

Exemplo 3.4.3. No caso de duas variáveis, $d = 2$, temos que os vetores monomiais coluna são $\mathbb{X}^{n+1} = (x^{n+1}, x^n y, x^{n-1} y^2, \dots, x^2 y^{n-1}, x y^n, y^{n+1})$ e $\mathbb{X}^n = (x^n, x^{n-1} y, \dots, x y^{n-1}, y^n)$.

Assim, as matrizes $L_{n,i}$ são dadas por

$$L_{n,1} = \begin{pmatrix} 1 & 0 & 0 \\ & \ddots & \vdots \\ 0 & 1 & 0 \end{pmatrix} \quad e \quad L_{n,2} = \begin{pmatrix} 0 & 1 & 0 \\ \vdots & \ddots & \\ 0 & 0 & 1 \end{pmatrix}.$$

Definição 3.4.4. Dadas as matrizes N_1, \dots, N_d de mesma dimensão $p \times q$, definimos a matriz junção N por

$$N = (N_1^T, \dots, N_d^T)^T, \quad N : dp \times q. \quad (3.4.6)$$

Com essa definição, consideramos as matrizes $L_n = (L_{n,1}^T, \dots, L_{n,d}^T)^T$ e $A_n = (A_{n,1}^T, \dots, A_{n,d}^T)^T$ de dimensão $dr_n^d \times r_{n+1}^d$. A seguir, demonstraremos algumas propriedades das matrizes $L_n, L_{n,i}, A_n$ e $A_{n,i}$.

Proposição 3.4.5. Para cada $i, 1 \leq i \leq d$, as matrizes $L_{n,i}$ satisfazem a equação $L_{n,i} L_{n,i}^T = I$. Além disso,

$$\text{posto } L_{n,i} = r_n^d, \quad 1 \leq i \leq d, \quad e \quad \text{posto } L_n = r_{n+1}^d.$$

Dem.: Por definição, cada linha de $L_{n,i}$ contém exatamente um elemento igual a um e os demais são iguais a zero. Por esta razão, $L_{n,i} L_{n,i}^T = I$ e, pela Proposição 2.1.11, segue que $\text{posto } L_{n,i} = r_n^d$. Agora, sejam

$$\mathcal{N}_n = \{\alpha \in \mathbb{N}_0^d : |\alpha| = n\} \quad e \quad \mathcal{N}_{n,i} = \{\alpha \in \mathbb{N}_0^d : |\alpha| = n, \alpha_i \neq 0\},$$

e consideremos $\mathbf{a} = (a_\alpha)_{\alpha \in \mathcal{N}_n}$ um vetor de dimensão r_{n+1}^d . Então, $L_{n,i}$ pode ser vista como uma aplicação que projeta \mathbf{a} na restrição em $\mathcal{N}_{n+1,i}$, isto é, $L_{n,i} \mathbf{a} = \mathbf{a}|_{\mathcal{N}_{n+1,i}}$. Para provar que L_n tem posto completo, devemos mostrar que se $L_n \mathbf{a} = 0$, então $\mathbf{a} = 0$. De fato, pela definição de L_n , temos que, se $L_n \mathbf{a} = 0$, então $L_{n,i} \mathbf{a} = 0, 1 \leq i \leq d$. Do fato de $L_{n,i}$ ser uma projeção e $\bigcup_i \mathcal{N}_{n,i} = \mathcal{N}_n$, $L_{n,i} \mathbf{a} = 0$ implica que $\mathbf{a} = 0$. \square

Teorema 3.4.6. Para $n \geq 0$ e $1 \leq i \leq d$,

$$\text{posto } A_{n,i} = \text{posto } C_{n+1,i} = r_n^d. \quad (3.4.7)$$

Além disso, para as matrizes junção A_n e C_n^T ,

$$\text{posto } A_n = r_{n+1}^d = \text{posto } C_{n+1} = r_{n+1}^d. \quad (3.4.8)$$

Dem.: Da relação (3.4.4) e do fato de G_n ser inversível, segue que

$$A_{n,i} = G_n L_{n,i} G_{n+1}^{-1}.$$

Assim, das Proposições 2.1.11 e 3.4.5, temos

$$\text{posto } A_{n,i} = \text{posto } L_{n,i} = r_n^d.$$

Como H_n é inversível, pela equação (3.4.2), $A_{n,i} H_{n+1} = H_n C_{n+1,i}^T$, e, da Proposição 2.1.11, segue que $\text{posto } C_{n+1,i} = r_n^d$.

Vamos provar, agora, (3.4.8). De $A_{n,i} G_{n+1} = G_n L_{n,i}$ implica, pela notação de matriz junção, que $A_n G_{n+1} = \text{diag}\{G_n, \dots, G_n\} L_n$, onde $\text{diag}\{G_n, \dots, G_n\}$ é de ordem dr_n^d . Como G_n é inversível, segue que $\text{diag}\{G_n, \dots, G_n\}$ é inversível. Das Proposições 3.4.5 e 2.1.11 segue que $\text{posto } A_n = r_{n+1}^d$. Além disso, da terceira equação de (3.4.2) obtemos $A_n H_{n+1} = \text{diag}\{H_n, \dots, H_n\} C_{n+1}^T$, onde $\text{diag}\{H_n, \dots, H_n\}$ é uma matriz de ordem dr_n^d , e, como H_n é inversível, temos $\text{posto } C_{n+1} = \text{posto } A_n = r_{n+1}^d$. \square

Portanto, pelo teorema anterior, segue que a matriz A_n tem posto completo, e, assim, A_n tem uma inversa generalizada D_n^T , escrita como

$$D_n^T = (D_{n,1}^T, \dots, D_{n,d}^T), \quad r_{n+1}^d \times dr_n^d,$$

onde $D_{n,i}^T : r_{n+1}^d \times r_n^d$, ou seja,

$$D_n^T A_n = \sum_{n=1}^d D_{n,i}^T A_{n,i} = I. \quad (3.4.9)$$

Note que a matriz D que satisfaz (3.4.9) não é única. De fato, seja A uma matriz de dimensão $s \times t$, $s > t$. Considere $\text{posto } A = t$, e seja a decomposição em valores singulares de A dada por

$$A = W^T \begin{bmatrix} D \\ 0 \end{bmatrix} U,$$

onde $D : t \times t$ é uma matriz diagonal inversível, $W : s \times s$ e $U : t \times t$ são matrizes unitárias. Então, a matriz D^T definida por

$$D^T = U^T (D^{-1}, D_1) W,$$

onde $D_1 : s \times t$ pode ser qualquer matriz, satisfaz $D^T A = I$. A matriz D^T é chamada de inversa generalizada de A . Caso $D_1 = 0$, então a matriz D^T é a única matriz inversa generalizada, conhecida como inversa generalizada de Moore-Penrose, que é denotado por A^+ (veja Ben-Israel e Greville [1]).

A seguir obtemos uma generalização da relação de recorrência de três termos com a matriz generalizada D_n^T .

Teorema 3.4.7. *Seja D_n^T uma inversa generalizada de A_n . Então, existem matrizes $E_n : r_{n+1}^d \times r_n^d$ e $F_n : r_{n+1}^d \times r_{n-1}^d$ tais que*

$$\mathbb{P}_{n+1} = \sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n + E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1}, \quad (3.4.10)$$

onde

$$\sum_{i=1}^d D_{n,i}^T B_{n,i} = -E_n \quad e \quad \sum_{i=1}^d D_{n,i}^T C_{n,i}^T = -F_n.$$

Dem.: Multiplicando a relação de três termos (3.4.3) por $D_{n,i}^T$ e somando para $i = 1, \dots, d$, obtemos

$$\sum_{i=1}^d D_{n,i}^T A_{n,i} \mathbb{P}_{n+1} = \sum_{i=1}^d D_{n,i}^T (x_i \mathbb{P}_n - B_{n,i} \mathbb{P}_n - C_{n,i} \mathbb{P}_{n-1})$$

e, portanto, como D_n^T é uma inversa generalizada de A_n ,

$$\mathbb{P}_{n+1} = \sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n + \sum_{i=1}^d (-D_{n,i}^T B_{n,i}) \mathbb{P}_n + \sum_{i=1}^d (-D_{n,i}^T C_{n,i}^T) \mathbb{P}_{n-1}.$$

Portanto, (3.4.10) vale com $\sum_{i=1}^d D_{n,i}^T B_{n,i} = -E_n$ e $\sum_{i=1}^d D_{n,i}^T C_{n,i}^T = -F_n$. \square

A equação (3.4.10) é análoga à relação de recorrência de três termos para polinômios ortogonais em uma variável. Porém, vale ressaltar que, no caso de várias variáveis, a relação de três termos implica na relação de recorrência de três termos, porém a relação de recorrência de três termos não implica na relação de três termos, como no caso de polinômios ortogonais em uma variável, ou seja, se escolhermos matrizes $A_{n,i}$, $B_{n,i}$ e $C_{n,i}$ e usarmos (3.4.10) para gerar uma sequência de polinômios, esses polinômios em geral não satisfazem a relação de três termos.

3.5 Teorema de Favard

Os próximos teoremas são para várias variáveis um análogo ao Teorema de Favard em uma variável. O Teorema de Favard para várias variáveis fornece uma condição necessária e suficiente para que um sistema de polinômios sejam ortogonais ou ortonormais.

Teorema 3.5.1. *Seja $\{\mathbb{P}_n\}_{n=0}^\infty = \{P_\alpha^n : |\alpha| = n, n \in \mathbb{N}_0\}$, $\mathbb{P}_0 = 1$, um sistema arbitrário de polinômios de Π^d . Então, as seguintes afirmações são equivalentes:*

- (1.) *Existe um funcional linear \mathcal{L} que é quase-definido em Π^d e que torna $\{\mathbb{P}_n\}_{n=0}^\infty$ um sistema de polinômios ortogonais em Π^d .*

(2.) Para $n \geq 0$, $1 \leq i \leq d$, existem matrizes $A_{n,i}$, $B_{n,i}$ e $C_{n,i}$ tais que

(a) os polinômios \mathbb{P}_n satisfazem a relação de três termos (3.4.1),

(b) as matrizes da relação de três termos (3.4.1) satisfazem as condições de posto (3.4.7) e (3.4.8).

Dem.: Dos Teoremas 3.4.1 e 3.4.6 obtemos que a afirmação (1) implica (2). Para provar que (2) implica (1), primeiro provamos que o sistema $\{\mathbb{P}_n\}$ forma uma base de Π^d . Para isso é suficiente mostrar que a matriz coeficiente do termo de maior grau, G_n , de \mathbb{P}_n é inversível.

Provemos, por indução sobre n , que G_n é inversível. Para $n = 0$, temos, por hipótese, que $\mathbb{P}_0 = 1$. Então, $G_0 = 1$ e é inversível. Agora, suponha que G_n é inversível e, então, uma vez que \mathbb{P}_n satisfaz a relação de três termos, comparando as matrizes coeficientes de \mathbb{X}^{n+1} de ambos os lados da igualdade, segue que $A_{n,i}G_{n+1} = G_n L_{n,i}$, $1 \leq i \leq d$. Assim, $A_n G_{n+1} = \text{diag}\{G_n, \dots, G_n\} L_n$, onde a matriz $\text{diag}\{G_n, \dots, G_n\}$ é de ordem dr_n^d , que tem a matriz G_n na diagonal principal. Portanto, $\text{diag}\{G_n, \dots, G_n\}$ é inversível. Das Proposições 2.1.11 e 3.4.5 temos que $\text{posto } A_n G_{n+1} = \text{posto } \text{diag}\{G_n, \dots, G_n\} L_n = r_{n+1}^d$. Das Proposições 2.1.11 e 2.1.12, e de (3.4.8), segue que

$$\text{posto } G_{n+1} \geq \text{posto } A_n G_{n+1} \geq \text{posto } A_n + \text{posto } G_{n+1} - r_{n+1}^d = \text{posto } G_{n+1}.$$

Então, $\text{posto } G_{n+1} = \text{posto } A_n G_{n+1} = r_{n+1}^d$. Portanto, da Proposição 2.1.11 temos que G_{n+1} é inversível.

Assim, $\{\mathbb{P}_n\}$ é uma base de Π^d . Então, podemos definir o funcional linear \mathcal{L} em Π^d por

$$\mathcal{L}(1) = 1, \quad \mathcal{L}(\mathbb{P}_n) = 0, \quad n \geq 1.$$

Agora, provemos, por indução, que

$$\mathcal{L}(\mathbb{P}_k \mathbb{P}_j^T) = 0, \quad k \neq j. \quad (3.5.1)$$

Seja $n \geq 0$ um inteiro. Suponha que (3.5.1) vale para quaisquer k e j tais que $0 \leq k \leq n$ e $j > k$. Como na demonstração de (3.4.10), usa-se somente a relação de três termos e o fato do posto $A_n = r_{n+1}^d$. Segue, de (a) e (b), que (3.4.10) vale para \mathbb{P}_n . Portanto, para $l > n + 1$, segue, da hipótese de indução, que

$$\begin{aligned} \mathcal{L}(\mathbb{P}_{n+1} \mathbb{P}_l^T) &= \mathcal{L} \left(\left(\sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n + E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1} \right) \mathbb{P}_l^T \right) \\ &= \mathcal{L} \left(\sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n \mathbb{P}_l^T \right) = \mathcal{L} \left(\sum_{i=1}^d D_{n,i}^T \mathbb{P}_n x_i \mathbb{P}_l^T \right) \\ &= \mathcal{L} \left(\sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n (A_{l,i} \mathbb{P}_{l+1} + B_{l,i} \mathbb{P}_l + C_{l,i} \mathbb{P}_{l-1})^T \right) = 0, \end{aligned}$$

o que completa a indução.

Por fim, provemos que $H_n = \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T)$ é inversível, para assim provar que $\{\mathbb{P}_n\}$ é uma base ortogonal com relação ao funcional linear \mathcal{L} . Para $n = 0$ segue da definição que $H_0 = \mathcal{L}(\mathbb{P}_0 \mathbb{P}_0^T) = \mathcal{L}(1) = 1$. Logo H_0 é inversível. Suponha que H_n seja inversível. De (3.5.1) e da demonstração da relação de três termos, segue que $A_{n,i} H_{n+1} = H_n C_{n+1,i}^T$ e, assim,

$$A_n H_{n+1} = \text{diag}\{H_n, \dots, H_n\} C_{n+1}^T. \quad (3.5.2)$$

Então, $\text{diag}\{H_n, \dots, H_n\}$ é inversível e, por (b) e da Proposição 2.1.11,

$$\text{posto } A_n H_{n+1} = \text{posto}(\text{diag}\{H_n, \dots, H_n\} C_{n+1}^T) = r_{n+1}^d.$$

Mas, da hipótese (b) e das Proposições 2.1.11 e 2.1.12, temos

$$\text{posto } H_{n+1} \geq \text{posto}(A_n H_{n+1}) \geq \text{posto } A_n + \text{posto } H_{n+1} - r_{n+1}^d = \text{posto } H_{n+1}.$$

Portanto, segue que $\text{posto } H_{n+1} = \text{posto}(A_n H_{n+1}) = r_{n+1}^d$ e, pela Proposição 2.1.11, temos que H_{n+1} é inversível. Isto completa a indução e, assim, $\{\mathbb{P}_n\}$ é uma base ortogonal com relação a \mathcal{L} .

Assim, juntamente com a Proposição 3.2.17 obtemos que H_n é simétrica e inversível. Logo, existem matrizes M_n e D_n tais que $H_n = M_n D M_n^T$ e D_n é diagonal. Para $\mathbb{Q}_n = M_n^{-1} \mathbb{P}_n$, obtemos

$$\mathcal{L}(\mathbb{Q}_n \mathbb{Q}_n^T) = M_n^{-1} \mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) (M_n^{-1})^T = D_n.$$

Isto prova que \mathcal{L} define um funcional linear quase-definido em Π^d . \square

Se os polinômios do teorema acima satisfazem (3.4.3) em vez de (3.4.1), então eles são polinômios ortonormais com relação a um funcional linear definido positivo.

Teorema 3.5.2. *Seja $\{\mathbb{P}_n\}_{n=0}^\infty = \{P_\alpha^n : |\alpha| = n, n \in \mathbb{N}_0\}$, $\mathbb{P}_0 = 1$, um sistema arbitrário de polinômios de Π^d . Então, as seguintes afirmações são equivalentes:*

- (1.) *Existe um funcional linear \mathcal{L} que é definido positivo em Π^d e que torna $\{\mathbb{P}_n\}_{n=0}^\infty$ uma base ortonormal em Π^d .*
- (2.) *Para $n \geq 0$, $1 \leq i \leq d$, existem matrizes $A_{n,i}$ e $B_{n,i}$ tais que*
 - (a) *os polinômios \mathbb{P}_n satisfazem a relação de três termos (3.4.3),*
 - (b) *as matrizes da relação de três termos (3.4.3) satisfazem as condições de posto (3.4.7) e (3.4.8).*

Dem.: Dos Teoremas 3.4.2 e 3.4.6 segue que (1) implica (2).

Provemos que (2) implica (1). Do Teorema 3.5.2, temos que existe um funcional linear \mathcal{L} quase-definido com relação ao qual $\{\mathbb{P}_n\}_{n=0}^{\infty}$ é uma base ortogonal em Π^d . Assim, basta provar que este funcional linear é definido positivo e torna $\{\mathbb{P}_n\}_{n=0}^{\infty}$ uma base ortonormal em Π^d , ou seja, H_n é a matriz identidade. Provemos, inicialmente, que $\{\mathbb{P}_n\}$ é ortonormal. Por indução provemos que $H_n = I_{r_n^d}$, onde $I_{r_n^d}$ é a matriz identidade de ordem r_n^d . Temos que $H_0 = \mathcal{L}(\mathbb{P}_0\mathbb{P}_0^T) = \mathcal{L}(1) = 1$. Suponha, agora, que $H_n = I_{r_n^d}$. Então, $\text{diag}\{H_n, \dots, H_n\} = I_{dr_n^d}$ e, como $C_{n+1}^T = A_n$, a equação (3.5.2) pode ser escrita como $A_n H_{n+1} = A_n$. Da condição de posto (3.4.8), existe D_n^T , a inversa generalizada de A_n . Isto implica que H_{n+1} é a matriz identidade, o que completa a prova de indução e mostra que $\{\mathbb{P}_n\}_{n=0}^{\infty}$ é uma base ortonormal em Π^d com relação ao funcional \mathcal{L} .

Provemos, agora, que \mathcal{L} define um funcional linear definido positivo. Assim, seja P um polinômio não nulo de grau n . Então, P pode ser escrito como $P = \mathbf{a}_n^T \mathbb{P}_n + \dots + \mathbf{a}_0^T \mathbb{P}_0$ e, desse modo,

$$\begin{aligned} \mathcal{L}(P^2) &= \mathcal{L}\left((\mathbf{a}_n^T \mathbb{P}_n + \dots + \mathbf{a}_0^T \mathbb{P}_0)(\mathbf{a}_n^T \mathbb{P}_n + \dots + \mathbf{a}_0^T \mathbb{P}_0)^T\right) \\ &= \mathcal{L}\left((\mathbf{a}_n^T \mathbb{P}_n + \dots + \mathbf{a}_0^T \mathbb{P}_0)(\mathbb{P}_n^T \mathbf{a}_n + \dots + \mathbb{P}_0^T \mathbf{a}_0)\right) \\ &= \mathbf{a}_n^T L(\mathbb{P}_n \mathbb{P}_n^T) \mathbf{a}_n + \dots + \mathbf{a}_0^T L(\mathbb{P}_0 \mathbb{P}_0^T) \mathbf{a}_0 \\ &= \|\mathbf{a}_n\|^2 + \dots + \|\mathbf{a}_0\|^2 > 0. \end{aligned}$$

Isto conclui a prova do teorema. □

Vale ressaltar, do Teorema de Favard em várias variáveis (Teorema 3.5.1), que a relação de três termos e a condição de posto caracterizam a ortogonalidade. Porém, extrair informações essenciais dos polinômios ortogonais a partir da relação de três termos é mais complicado, pois os coeficientes de (3.4.1) e (3.4.3) são matrizes e eles são únicos a menos de uma matriz similar.

Capítulo 4

Propriedades dos Polinômios

Ortogonais em Várias Variáveis

Neste capítulo estudamos algumas propriedades dos polinômios ortogonais em várias variáveis que são extensões das propriedades de polinômios ortogonais em uma variável, mas com uma complexidade maior, como, por exemplo, as dos zeros de polinômios ortogonais em várias variáveis. Para mais detalhes veja Dunkl e Xu [5].

4.1 Integrais centralmente simétricas

Seja \mathcal{L} um funcional linear definido positivo. Nesta seção estamos interessados em exemplos de \mathcal{L} que sejam representados como integrais de uma função peso não-negativa com momentos finitos, isto é, $\mathcal{L}(f) = \int f(\mathbf{x})W(\mathbf{x})d\mathbf{x}$. Nesses casos referimo-nos a função W como função peso. Esta seção é similar àquela dos polinômios ortogonais de uma variável com relação a uma função peso par.

Definição 4.1.1. *Seja $\Omega \subset \mathbb{R}^d$ o suporte compacto da função peso W . Então, W é centralmente simétrica se*

$$\mathbf{x} \in \Omega \Rightarrow -\mathbf{x} \in \Omega \quad e \quad W(\mathbf{x}) = W(-\mathbf{x}).$$

Um funcional linear \mathcal{L} é centralmente simétrico se

$$\mathcal{L}(\mathbf{x}^\alpha) = 0, \quad \alpha \in \mathbb{N}^d, \quad |\alpha| \text{ é um número ímpar.}$$

Note que se \mathcal{L} é expresso como a integral com relação a função peso W , então as duas definições são equivalentes.

Lema 4.1.2. *Se as matrizes coeficientes $B_{n,i}$ da relação de três termos são nulas para $n \geq 0$ e $1 \leq i \leq d$, então os polinômios de grau ímpar só têm componentes monomiais de grau ímpar e os polinômios de grau par só têm componentes monomiais de grau par.*

Dem.: Provemos por indução a afirmação. Para $n = 0$, temos que $\mathbb{P}_0 = 1 = x_i^0$ que só possui monômios de grau par. Para $n = 1$, temos, da relação de recorrência (3.4.10), que $\mathbb{P}_1 = \sum_{i=1}^d x_i D_{0,i}^T$, o que implica que não há termos constantes nas componentes de \mathbb{P}_1 e só há monômios de grau ímpar. Agora, como $B_{n,i} = 0$ para $n \geq 0, 1 \leq i \leq d$, então, da relação de recorrência (3.4.10), $E_n = 0$ para $n \geq 0$. Suponha, por indução, que \mathbb{P}_{2n-1} só tem monômios de grau ímpar e \mathbb{P}_{2n} só tem monômios de grau par. Da relação de recorrência (3.4.10), obtemos

$$\mathbb{P}_{2n+1} = \sum_{i=1}^d x_i D_{2n}^T \mathbb{P}_{2n} - F_n \mathbb{P}_{2n-1}, \text{ que só tem monômios de grau ímpar e}$$

$$\mathbb{P}_{2n+2} = \sum_{i=1}^d x_i D_{2n+1}^T \mathbb{P}_{2n+1} - F_n \mathbb{P}_{2n}, \text{ que só tem monômios de grau par,}$$

o que completa a demonstração. \square

Um funcional centralmente simétrico tem uma estrutura relativamente simples que nos permite obter mais informações sobre a estrutura de polinômios ortogonais. Iniciaremos com um teorema que conecta os funcionais centralmente simétricos com as matrizes coeficientes da relação de três termos.

Teorema 4.1.3. *Seja \mathcal{L} um funcional linear definido positivo e sejam $B_{n,i}$, $1 \leq i \leq d$, $n \geq 0$, as matrizes coeficientes em (3.4.3) dos polinômios ortogonais com respeito a \mathcal{L} . Então, \mathcal{L} é centralmente simétrico se, e somente se, $B_{n,i} = 0$, para todo $n \in \mathbb{N}_0$ e $1 \leq i \leq d$.*

Dem.: Provemos, inicialmente, a condição necessária. Suponha que \mathcal{L} seja centralmente simétrico. Provemos, por indução, que $B_{n,i} = 0$. De (3.4.2),

$$B_{0,i} = \mathcal{L}(x_i \mathbb{P}_0 \mathbb{P}_0^T) = \mathcal{L}(x_i) = 0.$$

Suponha que $B_{k,i} = 0$ para $0 \leq k \leq n-1$ e $1 \leq i \leq d$ e provemos que $B_{n,i} = 0$. Pela

definição de $B_{n,i}$ e da relação de recorrência (3.4.10),

$$\begin{aligned}
B_{n,i} &= \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_n^T) = \mathcal{L}\left(x_i \left(\sum_{j=1}^d D_{n-1,j}^T (x_j \mathbb{P}_{n-1} - A_{n-2,j}^T \mathbb{P}_{n-2})\right)\right) \\
&\quad \times \left(\sum_{k=1}^d D_{n-1,k}^T (x_k \mathbb{P}_{n-1} - A_{n-2,k}^T \mathbb{P}_{n-2})\right)^T \\
&= \sum_{j=1}^d \sum_{k=1}^d \mathcal{L}(x_i D_{n-1,j}^T (x_j \mathbb{P}_{n-1} - A_{n-2,j}^T \mathbb{P}_{n-2}) \\
&\quad \times (x_k \mathbb{P}_{n-1} - A_{n-2,k}^T \mathbb{P}_{n-2})^T D_{n-1,k}).
\end{aligned}$$

Note que, pelo Lema 4.1.2, a componente $(x_j \mathbb{P}_{n-1} - A_{n-2,j}^T \mathbb{P}_{n-2})(x_k \mathbb{P}_{n-1} - A_{n-2,k}^T \mathbb{P}_{n-2})^T$ tem somente somas de monômios de grau par. Portanto, o que está no funcional linear \mathcal{L} são somas de monômios de grau ímpar. Desse modo, como \mathcal{L} é centralmente simétrico segue que $B_{n,i} = 0$, o que completa a indução.

Por outro lado, suponha que $B_{n,i} = 0$, para $n \geq 0$ e provemos, por indução, que $\mathcal{L}(\mathbf{x}^\alpha) = 0$, se $|\alpha| = \text{ímpar}$. Assim, mostremos que $\mathcal{L}(\mathbf{x}^\alpha) = 0$, se $|\alpha| = 1$. De fato, temos $0 = B_{0,i} = \mathcal{L}(x_i \mathbb{P}_0 \mathbb{P}_0^T) = \mathcal{L}(x_i)$. Suponha que $\mathcal{L}(x_1^{\alpha_1} \dots x_d^{\alpha_d}) = 0$ para $|\alpha| = \alpha_1 + \dots + \alpha_d = 2k - 1, k = 1, \dots, n$ e, provemos que $\mathcal{L}(x^\alpha) = 0$, para $|\alpha| = 2n + 1$. De (3.4.2) e (3.4.3), para $1 \leq j, k \leq d$,

$$\begin{aligned}
0 &= A_{n-1,j} B_{n,i} A_{n-1,k}^T = A_{n-1,j} \mathcal{L}(x_i \mathbb{P}_n \mathbb{P}_n^T) A_{n-1,k}^T \\
&= \mathcal{L}(x_i A_{n-1,j} \mathbb{P}_n (A_{n-1,k} \mathbb{P}_n)^T) \\
&= \mathcal{L}(x_i (x_j \mathbb{P}_{n-1} - A_{n-2,j}^T \mathbb{P}_{n-2})(x_k \mathbb{P}_{n-1} - A_{n-2,k}^T \mathbb{P}_{n-2})^T) \\
&= \mathcal{L}(x_i x_j x_k \mathbb{P}_{n-1} \mathbb{P}_{n-1}^T) - \mathcal{L}(x_i x_j \mathbb{P}_{n-1} \mathbb{P}_{n-2}^T) A_{n-2,k} \\
&\quad - A_{n-2,j}^T \mathcal{L}(x_i x_k \mathbb{P}_{n-2} \mathbb{P}_{n-1}^T) + A_{n-2,j}^T \mathcal{L}(x_i \mathbb{P}_{n-2} \mathbb{P}_{n-2}^T) A_{n-2,k}.
\end{aligned}$$

Pelo Lema 4.1.2 e pela hipótese de indução, segue que $\mathcal{L}(x_i x_j \mathbb{P}_{n-1} \mathbb{P}_{n-2}^T) = 0$ e $\mathcal{L}(x_i x_k \mathbb{P}_{n-2} \mathbb{P}_{n-1}^T) = 0$, pois são polinômios que só têm monômios de grau ímpar até $2n - 1$. Além disso, $\mathcal{L}(x_i \mathbb{P}_{n-2} \mathbb{P}_{n-2}^T) = B_{n-2,i} = 0$. Portanto, $\mathcal{L}(x_i x_j x_k \mathbb{P}_{n-1} \mathbb{P}_{n-1}^T) = 0$. Multiplicando $\mathcal{L}(x_i x_j x_k \mathbb{P}_{n-1} \mathbb{P}_{n-1}^T)$ à esquerda e à direita por $A_{n-2,l}$ e $A_{n-2,m}^T$, respectivamente, temos para $1 \leq l, m \leq d$, obtemos

$$\begin{aligned}
0 &= \mathcal{L}(x_i x_j x_k \mathbb{P}_{n-1} \mathbb{P}_{n-1}^T) = A_{n-2,l} \mathcal{L}(x_i x_j x_k \mathbb{P}_{n-1} \mathbb{P}_{n-1}^T) A_{n-2,m}^T \\
&= \mathcal{L}(x_i x_j x_k A_{n-2,l} \mathbb{P}_{n-1} (A_{n-2,m} \mathbb{P}_{n-1})^T) \\
&= \mathcal{L}(x_i x_j x_k (x_l \mathbb{P}_{n-2} - A_{n-3,l}^T \mathbb{P}_{n-3})(x_m \mathbb{P}_{n-2} - A_{n-3,m}^T \mathbb{P}_{n-3})^T) \\
&= \mathcal{L}(x_i x_j x_k x_l x_m \mathbb{P}_{n-2} \mathbb{P}_{n-2}^T) - \mathcal{L}(x_i x_j x_k x_l \mathbb{P}_{n-2} \mathbb{P}_{n-3}^T) A_{n-3,m} \\
&\quad - A_{n-3,l}^T \mathcal{L}(x_i x_j x_k x_m \mathbb{P}_{n-3} \mathbb{P}_{n-2}^T) + A_{n-3,l}^T \mathcal{L}(x_i x_j x_k \mathbb{P}_{n-3} \mathbb{P}_{n-3}^T) A_{n-3,m}.
\end{aligned}$$

Analogamente, segue do Lema 4.1.2 e da hipótese de indução, que $0 = \mathcal{L}(x_i x_j x_k x_l x_m \mathbb{P}_{n-2} \mathbb{P}_{n-2}^T)$. Multiplicando à esquerda e à direita por $A_{p,q}$ e $A_{p,r}^T$, respectivamente, para $1 \leq q, r \leq d$ e $p = n - 3, \dots, 1$ e usando o mesmo argumento segue que $\mathcal{L}(x_1^{\alpha_1} \dots x_d^{\alpha_d}) = 0$, com $\alpha_1 + \dots + \alpha_d = 2n - 1$, o que completa a indução. \square

Do Lema 4.1.2 e do teorema anterior podemos enunciar o seguinte resultado.

Teorema 4.1.4. *Seja \mathcal{L} um funcional linear centralmente simétrico. Então, um polinômio ortogonal de grau n com respeito a \mathcal{L} é uma soma de monômios de grau par, se n for par, e uma soma de monômios de grau ímpar, se n for ímpar.*

Se um funcional linear \mathcal{L} ou a função peso W não forem centralmente simétricos, mas tornam-se centralmente simétrico através de uma transformação linear não-singular, então os polinômios ortogonais têm estrutura semelhante ao caso de centralmente simétricos. Porém, os coeficientes $B_{n,i}$ da relação de três termos não serão nulos. Eles devem satisfazer a seguinte condição de comutatividade, que sugere a seguinte definição.

Definição 4.1.5. *Se as matrizes coeficientes $B_{n,i}$ da relação de três termos satisfazem*

$$B_{n,i} B_{n,j} = B_{n,j} B_{n,i}, \quad 1 \leq i, j \leq d, \quad n \in \mathbb{N}_0, \quad (4.1.1)$$

então o funcional linear \mathcal{L} ou o função peso W é quase-centralmente simétrico.

Uma vez que a condição $B_{n,i} = 0$ implica (4.1.1), então centralmente simétrico implica em quase-centralmente simétrico.

Teorema 4.1.6. *Seja W uma função peso definida em $\Omega \subset \mathbb{R}^d$. Suponha que W seja centralmente simétrica sob uma transformação linear não-singular*

$$\mathbf{u} \mapsto \mathbf{x}, \quad \mathbf{x} = T\mathbf{u} + \mathbf{a}, \quad \det T > 0, \quad \mathbf{x}, \mathbf{u}, \mathbf{a} \in \mathbb{R}^d.$$

Então, W é quase-centralmente simétrica.

Dem.: Sejam $W^*(\mathbf{u}) = W(T\mathbf{u} + \mathbf{a})$ e $\Omega^* = \{\mathbf{u} : \mathbf{x} = T\mathbf{u} + \mathbf{a}, \mathbf{x} \in \Omega\}$. Denotemos por $\tilde{\mathbb{P}}_n$ os polinômios ortonormais associados a W . Então, usando a mudança de variáveis $\mathbf{x} \rightarrow T\mathbf{u} + \mathbf{a}$, temos

$$\int_{\Omega} \tilde{\mathbb{P}}_n(\mathbf{x}) \tilde{\mathbb{P}}_m(\mathbf{x}) W(\mathbf{x}) d\mathbf{x} = \int_{\Omega^*} \tilde{\mathbb{P}}_n(T\mathbf{u} + \mathbf{a}) \tilde{\mathbb{P}}_m^T(T\mathbf{u} + \mathbf{a}) W(T\mathbf{u} + \mathbf{a}) \det T d\mathbf{u}.$$

Assim, os polinômios ortonormais com respeito a W^* são dados por $\mathbb{P}^*(\mathbf{u}) = \sqrt{\det T} \tilde{\mathbb{P}}_n(T\mathbf{u} + \mathbf{a})$. Mostremos que as matrizes coeficientes $B_{n,i}(W)$ associadas a W são quase-centralmente

simétricas. De (3.4.2), fazendo a mudança de variáveis $\mathbf{x} \rightarrow T\mathbf{u} + \mathbf{a}$, obtemos

$$\begin{aligned}
B_{n,i}(W) &= \int_{\Omega} x_i \tilde{\mathbb{P}}_n(\mathbf{x}) \tilde{\mathbb{P}}_n^T(\mathbf{x}) W(\mathbf{x}) d\mathbf{x} \\
&= \int_{\Omega^*} \left(\sum_{j=1}^d t_{ij} u_j + a_j \right) \mathbb{P}_n^*(\mathbf{u}) \mathbb{P}_n^{*T}(\mathbf{u}) W^*(\mathbf{u}) d\mathbf{u} \\
&= \sum_{i=1}^d t_{ij} \int_{\Omega^*} u_j \mathbb{P}_n^*(\mathbf{u}) \mathbb{P}_n^{*T}(\mathbf{u}) W^*(\mathbf{u}) d\mathbf{u} \\
&\quad + a_i \int_{\Omega^*} \mathbb{P}_n^*(u) \mathbb{P}_n^{*T}(u) W^*(u) du \\
&= \sum_{i=1}^d t_{ij} B_{n,j}(W^*) + a_i I.
\end{aligned}$$

Por hipótese, W^* é centralmente simétrica em Ω^* , o que, pelo Teorema 4.1.3, implica que $B_{n,i}(W^*) = 0$. Portanto, $B_{n,i}(W) = a_i I$ implica que $B_{n,i}(W)$ satisfaz (4.1.1). Logo, W é quase-centralmente simétrica. \square

4.2 Matrizes de Jacobi em blocos

A seguir apresentamos um teorema que nos fornece uma condição de comutatividade para os polinômios ortogonais em várias variáveis.

Teorema 4.2.1. *Seja $\{\tilde{\mathbb{P}}_n\}$ uma sequência de polinômios ortonormais que satisfazem a relação de três termos (3.4.3). Então, as matrizes coeficientes satisfazem as relações de comutatividade*

$$\begin{aligned}
A_{k,i} A_{k+1,j} &= A_{k,j} A_{k+1,i}, \\
A_{k,i} B_{k+1,j} + B_{k,i} A_{k,j} &= A_{k,j} B_{k+1,i} + B_{k,j} A_{k,i}, \\
A_{k-1,i}^T A_{k-1,j} + B_{k,i} B_{k,j} + A_{k,i} A_{k,j}^T &= A_{k-1,j}^T A_{k-1,i} + B_{k,j} B_{k,i} + A_{k,j} A_{k,i}^T,
\end{aligned} \tag{4.2.1}$$

para $i \neq j$, $1 \leq i, j \leq d$, e $k \geq 0$, onde $A_{-1,i} = 0$.

Dem.: Pelo Teorema 3.5.2, existe um funcional linear \mathcal{L} com relação ao qual $\{\tilde{\mathbb{P}}_k\}$ é ortonormal. Assim, da relação de três termos (3.4.3), temos

$$\begin{aligned}
\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_{k+2}^T) &= \mathcal{L}(x_i \tilde{\mathbb{P}}_k x_j \tilde{\mathbb{P}}_{k+2}^T) = \mathcal{L}[(A_{k,i} \tilde{\mathbb{P}}_{k+1} + B_{k,i} \tilde{\mathbb{P}}_k + A_{k-1,i}^T \tilde{\mathbb{P}}_{k-1}) \\
&\quad \times (A_{k+2,j} \tilde{\mathbb{P}}_{k+3} + B_{k+2,j} \tilde{\mathbb{P}}_{k+2} + A_{k+1,j}^T \tilde{\mathbb{P}}_{k+1})^T] \\
&= A_{k,i} A_{k+1,j}.
\end{aligned}$$

Por outro lado,

$$\begin{aligned}\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_{k+2}^T) &= \mathcal{L}(x_j \tilde{\mathbb{P}}_k x_i \tilde{\mathbb{P}}_{k+2}^T) = \mathcal{L}[(A_{k,j} \tilde{\mathbb{P}}_{k+1} + B_{k,j} \tilde{\mathbb{P}}_k + A_{k-1,j}^T \tilde{\mathbb{P}}_{k-1}) \\ &\quad \times (A_{k+2,i} \tilde{\mathbb{P}}_{k+3} + B_{k+2,i} \tilde{\mathbb{P}}_{k+2} + A_{k+1,i}^T \tilde{\mathbb{P}}_{k+1})^T] \\ &= A_{k,j} A_{k+1,i}.\end{aligned}$$

Portanto, temos a primeira igualdade.

A segunda igualdade segue de $\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_{k+1}^T)$. De fato,

$$\begin{aligned}\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_{k+1}^T) &= \mathcal{L}(x_i \tilde{\mathbb{P}}_k x_j \tilde{\mathbb{P}}_{k+1}^T) = \mathcal{L}[(A_{k,i} \tilde{\mathbb{P}}_{k+1} + B_{k,i} \tilde{\mathbb{P}}_k + A_{k-1,i}^T \tilde{\mathbb{P}}_{k-1}) \\ &\quad \times (A_{k+1,j} \tilde{\mathbb{P}}_{k+2} + B_{k+1,j} \tilde{\mathbb{P}}_{k+1} + A_{k,j}^T \tilde{\mathbb{P}}_k)^T] \\ &= A_{k,i} B_{k+1,j} + B_{k,i} A_{k,j}.\end{aligned}$$

Além disso,

$$\begin{aligned}\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_{k+1}^T) &= \mathcal{L}(x_j \tilde{\mathbb{P}}_k x_i \tilde{\mathbb{P}}_{k+1}^T) = \mathcal{L}[(A_{k,j} \tilde{\mathbb{P}}_{k+1} + B_{k,j} \tilde{\mathbb{P}}_k + A_{k-1,j}^T \tilde{\mathbb{P}}_{k-1}) \\ &\quad \times (A_{k+1,i} \tilde{\mathbb{P}}_{k+2} + B_{k+1,i} \tilde{\mathbb{P}}_{k+1} + A_{k,i}^T \tilde{\mathbb{P}}_k)^T] \\ &= A_{k,j} B_{k+1,i} + B_{k,j} A_{k,i}.\end{aligned}$$

Por fim, a terceira igualdade segue de

$$\begin{aligned}\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_k^T) &= \mathcal{L}(x_i \tilde{\mathbb{P}}_k x_j \tilde{\mathbb{P}}_k^T) = \mathcal{L}[(A_{k,i} \tilde{\mathbb{P}}_{k+1} + B_{k,i} \tilde{\mathbb{P}}_k + A_{k-1,i}^T \tilde{\mathbb{P}}_{k-1}) \\ &\quad \times (A_{k,j} \tilde{\mathbb{P}}_{k+1} + B_{k,j} \tilde{\mathbb{P}}_k + A_{k-1,j}^T \tilde{\mathbb{P}}_{k-1})^T] \\ &= A_{k,i} A_{k,j}^T + B_{k,i} B_{k,j} + A_{k-1,i}^T A_{k-1,j}^T\end{aligned}$$

e

$$\begin{aligned}\mathcal{L}(x_i x_j \tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_k^T) &= \mathcal{L}(x_j \tilde{\mathbb{P}}_k x_i \tilde{\mathbb{P}}_k^T) = \mathcal{L}[(A_{k,j} \tilde{\mathbb{P}}_{k+1} + B_{k,j} \tilde{\mathbb{P}}_k + A_{k-1,j}^T \tilde{\mathbb{P}}_{k-1}) \\ &\quad \times (A_{k,i} \tilde{\mathbb{P}}_{k+1} + B_{k,i} \tilde{\mathbb{P}}_k + A_{k-1,i}^T \tilde{\mathbb{P}}_{k-1})^T] \\ &= A_{k,j} A_{k,i}^T + B_{k,j} B_{k,i} + A_{k-1,j}^T A_{k-1,i}^T.\end{aligned}$$

Portanto, todas as igualdades estão demonstradas. \square

As matrizes de Jacobi em uma variável, J , são definidas através dos coeficientes da relação de recorrência de três termos dos polinômios ortogonais. Para polinômios em várias variáveis, também temos o análogo à matriz de Jacobi.

Definição 4.2.2. *Suponha que as matrizes coeficientes da relação de três termos (3.4.1) satisfazem as condições de posto (3.4.7) e (3.4.8), e as condições de comutatividade (4.2.1).*

Definimos a matriz de Jacobi em blocos como a família de matrizes tri-diagonais em bloco J_i , $1 \leq i \leq d$, dadas por

$$J_i = \begin{bmatrix} B_{0,i} & A_{0,i} & & 0 \\ A_{0,i}^T & B_{1,i} & A_{1,i} & \\ & A_{1,i}^T & B_{2,i} & \ddots \\ 0 & & \ddots & \ddots \end{bmatrix}, \quad 1 \leq i \leq d. \quad (4.2.2)$$

Definição 4.2.3. Para cada $n \in \mathbb{N}_0$ fixo, definimos a i -ésima matriz de Jacobi em bloco truncada, através das matrizes coeficientes da relação de três termos (3.4.3), como

$$J_{n,i} = \begin{bmatrix} B_{0,i} & A_{0,i} & 0 & \cdots & 0 \\ A_{0,i}^T & B_{1,i} & A_{1,i} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & A_{n-3,i}^T & B_{n-2,i} & A_{n-2,i} \\ 0 & \cdots & 0 & A_{n-2,i}^T & B_{n-1,i} \end{bmatrix}, \quad 1 \leq i \leq d.$$

Note que $J_{n,i}$ é uma matriz quadrada de ordem $N = \dim \Pi_{n-1}^d$.

4.3 Propriedades da relação de três termos

A relação de recorrência em várias variáveis não é tão forte quanto em uma variável. Isso decorre da relação de recorrência não implicar a relação de três termos, pois o problema é que a matriz D_n^T não é a única inversa à esquerda de A_n . Porém $A_n D_n^T$ geralmente não é igual à identidade. Assim, devemos compreender melhor o espaço núcleo de $I - A_n D_n^T$. A relação de recorrência não implica no Teorema de Favard sem acréscimo de condições.

Discutiremos essas condições através de polinômios ortonormais, pois caso seja de um sistema de polinômios ortogonais, $\{\mathbb{P}_n\}$, pela demonstração do Teorema 3.2.27, temos que $\tilde{\mathbb{P}}_n = \left(H_n^{\frac{1}{2}}\right)^{-1} \mathbb{P}_n$ é uma sequência de polinômios ortonormais.

Serão necessárias algumas condições e definições. Iniciemos com uma definição.

Definição 4.3.1. Para qualquer sequência de matrizes N_1, \dots, N_d , onde cada N_i é de dimensão $s \times t$, definimos a matriz junção Ξ_N como segue. Seja $\Xi_{i,j}$, $1 \leq i, j \leq d$, $i \neq j$, a matriz em blocos definida por

$$\Xi_{i,j} = (\cdots 0 N_j^T \cdots - N_i^T 0 \cdots), \quad \Xi_{i,j} : t \times ds,$$

isto é, os únicos elementos em blocos não-nulos são as matrizes N_j^T no i -ésimo bloco e $-N_i^T$ no j -ésimo bloco. Assim, a matriz Ξ_N é definida a partir das matrizes $\Xi_{i,j}$ como blocos em ordem lexicográfica de $\{(i, j) : 1 \leq i < j \leq d\}$,

$$\Xi_N = [\Xi_{1,2}^T | \Xi_{1,3}^T | \dots | \Xi_{d-1,d}^T].$$

A matriz Ξ_N é de dimensão $ds \times \binom{d}{2}t$.

O próximo lema segue imediatamente da definição de Ξ_N .

Lema 4.3.2. *Seja $Y = (Y_1^T, \dots, Y_d^T)^T$, $Y_i \in \mathbb{R}^s$. Então, $\Xi_N^T Y = 0$ se, e somente se, $N_i^T Y_j = N_j^T Y_i$, $1 \leq i < j \leq d$.*

Por definição, Ξ_{A_n} e $\Xi_{A_n^T}$ são matrizes de dimensões diferentes. De fato, como $A_{n,i} : r_n^d \times r_{n+1}^d$ e $A_{n,i}^T : r_{n+1}^d \times r_n^d$ então,

$$\Xi_{A_n} : dr_n^d \times \binom{d}{2}r_{n+1}^d \quad \text{e} \quad \Xi_{A_n^T} : dr_{n+1}^d \times \binom{d}{2}r_n^d.$$

Necessitamos encontrar os postos dessas duas matrizes. Para isso, como anteriormente, da relação (3.4.4) precisamos compreender melhor as matrizes Ξ_{L_n} e $\Xi_{L_n^T}$, onde as matrizes $L_{n,i}$ são definidas por (3.4.5).

Lema 4.3.3. *Para $n \in \mathbb{N}_0$ e $1 \leq i < j \leq d$, $L_{n,i}L_{n+1,j} = L_{n,j}L_{n+1,i}$.*

Dem.: De (3.4.5) segue que, para qualquer $\mathbf{x} \in \mathbb{R}^d$,

$$\begin{aligned} L_{n,i}L_{n+1,j}\mathbb{X}^{n+2} &= L_{n,i}x_j\mathbb{X}^{n+1} = x_jL_{n,i}\mathbb{X}^{n+1} = x_jx_i\mathbb{X}^n = x_ix_j\mathbb{X}^n \\ &= x_iL_{n,j}\mathbb{X}^{n+1} = L_{n,j}x_i\mathbb{X}^{n+1} = L_{n,j}L_{n+1,i}\mathbb{X}^{n+2}. \end{aligned}$$

Tomando as derivadas parciais com respeito às componentes de \mathbf{x} na identidade acima, segue que $(L_{n,i}L_{n+1,j} - L_{n,j}L_{n+1,i})\mathbf{e}_k = 0$ para todo \mathbf{e}_k da base de $\mathbb{R}^{r_{n+2}}$. Portanto, segue a identidade desejada. \square

Lema 4.3.4. *Para $d \geq 2$ e $n \geq 1$, posto $\Xi_{L_n^T} = dr_{n+1}^d - r_{n+2}^d$.*

Dem.: Vamos provar que a dimensão do espaço nulo de $\Xi_{L_n^T}$ é igual a r_{n+2}^d . Seja $Y = (Y_1^T, \dots, Y_d^T)^T \in \mathbb{R}^{dr_{n+1}}$, onde $Y_i \in \mathbb{R}^{r_{n+1}}$. Considere a equação homogênea em dr_{n+1} variáveis $\Xi_{L_n^T} Y = 0$, que, pelo Lema 4.3.2, é equivalente a

$$L_{n,i}Y_j = L_{n,j}Y_i, \quad 1 \leq i < j \leq d. \quad (4.3.1)$$

Há exatamente um elemento igual a 1 em cada linha de $L_{n,i}$ e o posto de $L_{n,i}$ é r_n^d . Lembremos $\mathcal{N}_n = \{\alpha \in \mathbb{N}_0^d : |\alpha| = n\}$ e, para cada $1 \leq i \leq d$, $\mathcal{N}_{n,i} = \{\alpha \in \mathbb{N}_0^d : |\alpha| = n \text{ e } \alpha_i \neq 0\}$.

Seja $\mu(\mathcal{N})$ o número de elementos do conjunto \mathcal{N} . Contando o número de soluções inteiras de $|\alpha| = n$, isto mostra que $\mu(\mathcal{N}_n) = r_n^d$ e $\mu(\mathcal{N}_{n,i}) = r_{n-1}^d$. Lembrando que $L_{n,i}$ é uma projeção de \mathcal{N}_n em $\mathcal{N}_{n,i}$. Fixamos a correspondência um a um entre os elementos \mathcal{N}_n e os elementos de um vetor em \mathbb{R}^{r_n} , e escrevemos $Y_i|_{\mathcal{N}_{n+1,j}} = L_{n,j}Y_i$. O sistema das equações lineares (4.3.1) pode ser escrito como

$$Y_i|_{\mathcal{N}_{n+1,j}} = Y_j|_{\mathcal{N}_{n+1,i}}, \quad 1 \leq i < j \leq d. \quad (4.3.2)$$

Estes produzem $\binom{d}{2}r_n^d$ equações em dr_{n+1}^d variáveis em Y , mas nem todos eles são independentes. Para quaisquer inteiros distintos i, j e k ,

$$\begin{aligned} Y_i|_{\mathcal{N}_{n+1,j} \cap \mathcal{N}_{n+1,k}} &= Y_j|_{\mathcal{N}_{n+1,i} \cap \mathcal{N}_{n+1,k}}, \\ Y_i|_{\mathcal{N}_{n+1,k} \cap \mathcal{N}_{n+1,j}} &= Y_k|_{\mathcal{N}_{n+1,i} \cap \mathcal{N}_{n+1,j}}, \\ Y_j|_{\mathcal{N}_{n+1,k} \cap \mathcal{N}_{n+1,i}} &= Y_k|_{\mathcal{N}_{n+1,j} \cap \mathcal{N}_{n+1,i}}. \end{aligned}$$

Assim, há exatamente $\mu(\mathcal{N}_{n+1,j} \cap \mathcal{N}_{n+1,k}) = r_{n-1}^d$ equações repetidas entre estes três sistemas de equações. Contando todas as combinações dos três sistemas gerados em (4.3.2), temos $\binom{d}{3}r_{n-1}^d$ equações. Contudo, entre elas algumas são contadas mais de uma vez. Repetindo o argumento anterior para quatro sistemas distintos gerados pelas equações em (4.3.1), temos que há $\mu(\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j} \cap \mathcal{N}_{n,k}) = r_{n-2}^d$ equações repetidas. Há $\binom{d}{4}$ combinações de quatro sistemas diferentes de equações. Então precisamos considerar cinco sistemas de equações e assim por diante. Desta maneira, segue que há

$$\binom{d}{3}r_{n-1}^d - \binom{d}{4}r_{n-2}^d + \cdots + (-1)^{d+1}r_{n-d+2}^d = \sum_{k=3}^d (-1)^{k+1} \binom{d}{k} r_{n-k+2}^d$$

equações repetidas em (4.3.2). Então, entre as $\binom{d}{2}r_n^d$ equações de (4.3.1), o número de equações independentes é

$$\binom{d}{2}r_n^d - \sum_{k=3}^d (-1)^{k+1} \binom{d}{k} r_{n-k+2}^d = \sum_{k=2}^d (-1)^k \binom{d}{k} r_{n-k+2}^d.$$

Como a dimensão do espaço nulo é igual ao número de variáveis menos o número de equações independentes, e como há dr_{n+1}^d variáveis, temos que

$$\dim\{\text{núcleo}\Xi_{L_n^T}^T\} = dr_{n+1}^d - \sum_{k=2}^d (-1)^k \binom{d}{k} r_{n-k+2}^d = \sum_{k=1}^d (-1)^{k+1} \binom{d}{k} r_{n-k+2}^d = r_{n+2}^d,$$

onde a última igualdade segue da soma de Chu-Vandermonde. Portanto, posto $\Xi_{L_n^T} = dr_{n+1}^d - \dim\{\text{núcleo}\Xi_{L_n^T}^T\} = dr_{n+1}^d - r_{n+2}^d$. \square

Lema 4.3.5. Para $d \geq 2$ e $n \geq 1$, posto $\Xi_{L_n} = dr_n^d - r_{n-1}^d$

Dem.: Como no lema anterior, provemos que a dimensão do espaço nulo de Ξ_{L_n} é r_{n-1}^d . Para $Y = (Y_1^T, \dots, Y_d^T)^T \in \mathbb{R}^{dr_n^d}$ onde $Y_i \in \mathbb{R}^{r_n}$. Considere a equação homogênea $\Xi_{L_n}^T Y = 0$. Pelo Lema 4.3.2, esta equação é equivalente ao sistema de equações lineares

$$L_{n,i}^T Y_j = L_{n,j}^T Y_i, \quad 1 \leq i < j \leq d. \quad (4.3.3)$$

Usando \mathcal{N}_n e $\mathcal{N}_{n,i}$ da demonstração do lema anterior. Note que $L_{n,i}^T Y_j$ é um vetor em $\mathbb{R}^{r_{n+1}}$, com as coordenadas que correspondem à $\mathcal{N}_{n,i}$ iguais às de Y_j e as demais coordenadas são zeros. Assim, a equação (4.3.3) implica que os elementos de Y_j são não nulos somente onde eles correspondem à $\mathcal{N}_{n,j} \cap \mathcal{N}_{n,i}$ em $L_{n,i}^T Y_j$, isto é, $(L_{n,i}^T Y_j)|_{\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j}}$, e os elementos não nulos de Y_i são $(L_{n,j}^T Y_i)|_{\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j}}$. Além disso, estes dois vetores são iguais. Como para qualquer $X \in \mathbb{R}^{r_{n+1}}$, $L_{n-1,j} L_{n,i} X = X|_{\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j}}$, que segue de $L_{n-1,i} L_{n,j} \mathbb{X}^{n+1} = x_i x_j \mathbb{X}^{n-1}$, do fato de $L_{n,i} L_{n,i}^T = I$,

$$(L_{n,i}^T Y_j)|_{\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j}} = L_{n-1,j} L_{n,i} (L_{n,i}^T Y_j) = L_{n-1,j} Y_j = Y_j|_{\mathcal{N}_{n-1,j}}.$$

Entretanto, os elementos não nulos de Y_i e Y_j satisfazem $Y_i|_{\mathcal{N}_{n-1,i}} = Y_j|_{\mathcal{N}_{n-1,j}}$, $1 \leq i < j \leq d$. Assim, há exatamente $\mu(\mathcal{N}_{n-1,i}) = r_{n-1}^d$ variáveis independentes na solução de (4.3.3). Isto é, a dimensão do espaço nulo de $\Xi_{L_n}^T$ é r_{n-1}^d . \square

Desses dois lemas segue o posto de Ξ_{A_n} e $\Xi_{A_n^T}$.

Proposição 4.3.6. Sejam $A_{n,i}$ as matrizes coeficientes da relação de três termos. Então,

$$\text{posto } \Xi_{A_n} = dr_n^d - r_{n-1}^d \text{ e } \text{posto } \Xi_{A_n^T} = dr_{n+1}^d - r_{n+2}^d.$$

Dem.: De (3.4.4) e da definição de Ξ_N , segue que

$$\Xi_{A_n} \text{diag}\{G_{n+1}, \dots, G_{n+1}\} = \text{diag}\{G_n, \dots, G_n\} \Xi_{L_n},$$

onde as matrizes diagonais em blocos da esquerda são de ordem $\binom{d}{2} r_{n+1}^d$, e as da direita de ordem $dr_n^d \times dr_n^d$. Como G_n e G_{n+1} são ambas inversíveis, então estas matrizes em bloco são inversíveis. Portanto, $\text{posto } \Xi_{A_n} = \text{posto } \Xi_{L_n}$. Assim, segue a primeira igualdade. Analogamente, temos que $\Xi_{A_n^T} \text{diag}\{G_{n+1}, \dots, G_{n+1}\} = \text{diag}\{G_n, \dots, G_n\} \Xi_{L_n^T}$ e, assim, $\text{posto } \Xi_{A_n^T} = \text{posto } \Xi_{L_n^T}$. \square

O próximo lema envolve o espaço núcleo de $I - A_n D_n^T$.

Lema 4.3.7. Sejam $\{A_{n,i}\}$ as matrizes coeficientes da relação de três termos (3.4.1) e D_n^T a inversa à direita de A_n definida em (3.4.9). Então, $Y = (Y_1^T, \dots, Y_d^T)^T$, com $Y_i \in \mathbb{R}^{r_n}$, pertence ao espaço nulo de $I - A_n D_n^T$ se, e somente se, $A_{n-1,i} Y_j = A_{n-1,j} Y_i$, $1 \leq i, j \leq d$, ou, equivalentemente, $\Xi_{A_{n-1}^T} Y = 0$.

Dem.: Seja a decomposição em valores singulares de A_n e D_n^T como em (3.4.9). A_n e D_n são ambos de mesma dimensão $dr_n^d \times r_{n+1}^d$. Segue que

$$I - A_n D_n^T = I - W^T \begin{pmatrix} \Lambda \\ 0 \end{pmatrix} U U^T (\Lambda^{-1}, \Lambda_1) W = W^T \begin{pmatrix} 0 & -\Lambda \Lambda^{-1} \\ 0 & I \end{pmatrix} W,$$

onde I é a matriz identidade de ordem $dr_n^d - r_{n+1}^d$, daí segue que o posto $(I - A_n D_n^T) = dr_n^d - r_{n+1}^d$. Por isso,

$$\dim\{\text{núcleo}(I - A_n D_n^T)\} = dr_n^d - \text{posto}(I - A_n D_n^T) = r_{n+1}^d.$$

Como $D_n^T A_n = I$, segue que $(I - A_n D_n^T) A_n = A_n - A_n = 0$, que implica que as colunas de A_n formam uma base para o espaço nulo de $I - A_n D_n^T$. Portanto, se $(I - A_n D_n^T) Y = 0$, então existe um vetor \mathbf{c} tal que $Y = A_n \mathbf{c}$. Mas pela primeira condição de comutatividade em (4.2.1) é equivalente a $\Xi_{A_{n-1}^T}^T A_n = 0$, isto segue que $\Xi_{A_{n-1}^T}^T Y = 0$. Por outro lado, suponha $\Xi_{A_{n-1}^T}^T Y = 0$. Pela Proposição 4.3.6,

$$\dim\{\text{núcleo}\Xi_{A_{n-1}^T}^T\} = dr_n^d - \text{posto}\Xi_{A_{n-1}^T}^T = r_{n+1}^d.$$

De $\Xi_{A_{n-1}^T}^T Y = 0$ segue que as colunas de A_n formam uma base do núcleo de $\Xi_{A_{n-1}^T}^T$. Portanto, existe um vetor \mathbf{d} tal que $Y = A_n \mathbf{d}$, que por, $D_n^T A_n = I$, implica que $(I - A_n D_n^T) Y = 0$. \square

A seguir enunciamos o teorema sobre a relação de recorrência ser equivalente a um sistema de polinômios ortonormais com algumas condições adicionais.

Teorema 4.3.8. *Seja $\{\mathbb{P}_k\}_{k=0}^\infty$ definido por (3.4.10). Então, existe um funcional linear definido positivo, \mathcal{L} , com relação ao qual $\{\mathbb{P}_k\}_{k=0}^\infty$ uma base ortonormal em Π^d se, e somente se, $B_{n,i}$ são simétricos, $A_{n,i}$ satisfazem a condição de posto (3.4.7), e essas matrizes coeficientes satisfazem as condições de comutatividade (4.2.1).*

Dem.: A condição suficiente já foi provada anteriormente. Mostremos, então, a condição necessária. Suponha que $\{\mathbb{P}_k\}_{k=0}^\infty$ seja definido recursivamente por (3.4.10) e que as matrizes $A_{n,i}$ e $B_{n,i}$ satisfaçam a condição de posto (3.4.7) e as condições de comutatividade (4.2.1). Provemos que esses polinômios satisfazem a relação de três termos. Desse modo, pelo Teorema 3.5.2, segue a implicação. Provemos por indução. Para $n = 0$, como $\mathbb{P}_0 = 1$ e $\mathbb{P}_{-1} = 0$, então

$$\mathbb{P}_1 = \sum_{i=1}^d D_{0,i}^T x_i - \sum_{i=1}^d D_{0,i}^T B_{0,1}. \quad (4.3.4)$$

Como A_0 e D_0 são ambas matrizes $d \times d$ e $D_0^T A_0 = I$, segue que $A_0 D_0^T = I$, o que implica que $A_{0,i} D_{0,j}^T = \delta_{i,j}$. Portanto, multiplicando por $A_{0,j}$ à esquerda em ambos os membros de (4.3.4), obtemos

$$A_{0,j} \mathbb{P}_1 = x_j - B_{0,j} = x_j \mathbb{P}_0 - B_{0,j} \mathbb{P}_0.$$

Suponha que

$$A_{k,j} \mathbb{P}_{k+1} = x_j \mathbb{P}_k - B_{k,j} \mathbb{P}_k - A_{k-1,j}^T \mathbb{P}_{k-1}, 0 \leq k \leq n-1. \quad (4.3.5)$$

Agora, mostremos que esta equação vale para $k = n$. Para isso, provemos que

$$A_{n,j} \mathbb{P}_{n+1} = x_j \mathbb{P}_n - B_{n,j} \mathbb{P}_n - A_{n-1,j}^T \mathbb{P}_{n-1} + \mathbb{Q}_{n-2,j}, \quad (4.3.6)$$

onde as componentes do $\mathbb{Q}_{n-2,j}$ são elementos de Π_{n-2}^d e $\mathbb{Q}_{n-2,j} \in \mathbb{R}^{r_j}$. Como cada componente de \mathbb{P}_n é um polinômio de grau no máximo n , podemos escrever

$$\mathbb{P}_n = G_{n,n} \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots,$$

para algumas matrizes $G_{n,k}$, e novamente denotaremos $G_{n,n}$ por G_n . Logo, como \mathbb{P}_{n+1} é definido pela relação de recorrência (3.4.10) e expandindo o lado esquerdo de (4.3.6) em componentes de \mathbb{X}^k , lembrando que $L_{n,i} \mathbb{X}^{n+1} = x_i \mathbb{X}^n$, temos

$$\begin{aligned} A_{n,j} \mathbb{P}_{n+1} &= A_{n,j} \left(\sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n + E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1} \right) \\ &= A_{n,j} \sum_{i=1}^d x_i D_{n,i}^T (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots) \\ &\quad - A_{n,j} \sum_{i=1}^d D_{n,i}^T B_{n,i} (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots) \\ &\quad - A_{n,j} \sum_{i=1}^d D_{n,i}^T A_{n-1,i}^T (G_{n-1} \mathbb{X}^{n-1} + G_{n-1,n-2} \mathbb{X}^{n-2} + \dots) \\ &= \left(A_{n,j} \sum_{i=1}^d D_{n,i}^T G_n L_{n,i} \right) \mathbb{X}^{n+1} \\ &\quad + \left[A_{n,j} \sum_{i=1}^d D_{n,i}^T (G_{n,n-1} L_{n-1,i} - B_{n,i} G_n) \right] \mathbb{X}^n \\ &\quad + \left[A_{n,j} \sum_{i=1}^d D_{n,i}^T (G_{n,n-2} L_{n-2,i} - B_{n,i} G_{n,n-1} - A_{n-1,i}^T G_{n-1}) \right] \mathbb{X}^{n-1} + \dots \end{aligned} \quad (4.3.7)$$

Agora, expandindo o lado direito de (4.3.6) em componentes de \mathbb{X}^k , segue que

$$\begin{aligned}
x_j \mathbb{P}_n - B_{n,j} \mathbb{P}_n &= A_{n-1,j}^T \mathbb{P}_{n-1} + \mathbb{Q}_{n-2,j} \\
&= x_j (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots) \\
&\quad - B_{n,j} (G_n \mathbb{X}^n + G_{n,n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots) \\
&\quad - A_{n-1,j}^T (G_{n-1} \mathbb{X}^{n-1} + G_{n,n-2} \mathbb{X}^{n-2} + \dots) + \dots \\
&= (G_n L_{n,j}) \mathbb{X}^{n+1} + (G_{n,n-1} L_{n-1,j} - B_{n,j} G_n) \mathbb{X}^n \\
&\quad + (G_{n,n-2} L_{n-2,j} - B_{n,j} G_{n,n-1} - A_{n-1,j}^T G_{n-1}) \mathbb{X}^{n-1} + \dots
\end{aligned} \tag{4.3.8}$$

Comparando (4.3.7) e (4.3.8) é suficiente mostrar que os três maiores coeficientes na expansão são iguais. Assim, precisamos estabelecer $A_{n,j} \sum_{n,i}^d D_{n,i}^T X_{n,i} = X_{n,j}$, $1 \leq i \leq d$, onde $X_{n,j}$ é uma das seguintes matrizes:

$$X_{n,j} = \begin{cases} Q_{n,j} = G_n L_{n,j}, \\ R_{n,j} = G_{n,n-1} L_{n-1,j} - B_{n,j} G_n, \\ S_{n,j} = G_{n,n-2} L_{n-2,j} - B_{n,j} G_{n,n-1} - A_{n-1,j}^T G_{n-1}. \end{cases} \tag{4.3.9}$$

Retomando a notação da matriz junção de (3.4.6), podemos escrever as igualdades acima de forma mais compacta, $(I - A_n D_n^T) X_n$. Desse modo, as equações que precisamos mostrar são

$$A_n D_n^T Q_n = Q_n, \quad A_n D_n^T R_n = R_n \quad \text{e} \quad A_n D_n^T S_n = S_n.$$

Continuamos provando que as colunas de Q_n , R_n e S_n pertencem ao espaço nulo de $I - A_n D_n^T$. Mas, pelos Lemas 4.3.7 e 4.3.2, se reduzem a mostrar que

$$\begin{cases} A_{n-1,i} Q_{n,j} = A_{n-1,j} Q_{n,i}, \\ A_{n-1,i} R_{n,j} = A_{n-1,j} R_{n,i}, \\ A_{n-1,i} S_{n,j} = A_{n-1,j} S_{n,i}. \end{cases} \tag{4.3.10}$$

Para demonstrar esta afirmação, note que, se fizermos $k = n - 1$ em (4.3.5) e compararmos os coeficiente de \mathbb{X}^n , \mathbb{X}^{n-1} e \mathbb{X}^{n-2} dos dois lados da igualdade, fornecem, respectivamente,

$$\begin{cases} A_{n-1,j} G_n = G_{n-1} L_{n-1,j}, \\ A_{n-1,j} G_{n,n-1} = G_{n-1,n-2} L_{n-2,j} - B_{n-1,j} G_{n-1}, \\ A_{n-1,j} G_{n,n-2} = G_{n-1,n-3} L_{n-3,j} - B_{n-1,j} G_{n-1,n-2} - A_{n-2,j}^T G_{n-2}. \end{cases} \tag{4.3.11}$$

Demonstremos, a seguir, cada igualdade de (4.3.10). A primeira igualdade segue da primeira igualdade de (4.3.11) e do Lema 4.3.3,

$$\begin{aligned}
A_{n-1,i} Q_{n,j} &= A_{n-1,i} G_n L_{n,j} = G_{n-1} L_{n-1,i} L_{n,j} = G_{n-1} L_{n-1,j} L_{n,i} \\
&= A_{n-1,j} G_n L_{n,i} = A_{n-1,j} Q_{n,i}.
\end{aligned}$$

A segunda igualdade segue das duas primeiras igualdade de (4.3.11), do Lema 4.3.3 e da segunda condição de comutatividade (4.2.1) com $k = n - 1$,

$$\begin{aligned}
A_{n-1,i}R_{n,j} &= A_{n-1,i}(G_{n,n-1}L_{n-1,j} - B_{n,j}G_n) \\
&= A_{n-1,i}G_{n,n-1}L_{n-1,j} - A_{n-1,i}B_{n,j}G_n \\
&= (G_{n-1,n-2}L_{n-2,i} - B_{n-1,i}G_{n-1})L_{n-1,j} - A_{n-1,i}B_{n,j}G_n \\
&= G_{n-1,n-2}L_{n-2,i}L_{n-1,j} - B_{n-1,i}G_{n-1}L_{n-1,j} + A_{n-1,i}B_{n,j}G_n \\
&= G_{n-1,n-2}L_{n-2,j}L_{n-1,i} - B_{n-1,i}A_{n-1,j}G_n - A_{n-1,i}B_{n,j}G_n \\
&= G_{n-1,n-2}L_{n-2,j}L_{n-1,i} - (B_{n-1,i}A_{n-1,j} + A_{n-1,i}B_{n,j})G_n \\
&= G_{n-1,n-2}L_{n-2,j}L_{n-1,i} - (A_{n-1,j}B_{n,i} + B_{n-1,j}A_{n-1,i})G_n \\
&= A_{n-1,j}R_{n,i}.
\end{aligned}$$

Por fim, a última igualdade de (4.3.10) segue de (4.3.11), do Lema 4.3.3 e da segunda e terceira condições de comutatividade (4.2.1) com $k = n - 1$,

$$\begin{aligned}
A_{n-1,i}S_{n,j} &= A_{n-1,i}(G_{n,n-2}L_{n-2,j} - B_{n,j}G_{n,n-1} - A_{n-1,j}^T G_{n-1}) \\
&= A_{n-1,i}G_{n,n-2}L_{n-2,j} - A_{n-1,i}B_{n,j}G_{n,n-1} - A_{n-1,i}A_{n-1,j}^T G_{n-1} \\
&= (G_{n-1,n-3}L_{n-3,i} - B_{n-1,i}G_{n-1,n-2} - A_{n-2,i}^T G_{n-2})L_{n-2,j} \\
&\quad - A_{n-1,i}B_{n,j}G_{n,n-1} - A_{n-1,i}A_{n-1,j}^T G_{n-1} \\
&= G_{n-1,n-3}L_{n-3,i}L_{n-2,j} - B_{n-1,i}G_{n-1,n-2}L_{n-2,j} \\
&\quad - A_{n-2,i}^T G_{n-2}L_{n-2,j} - A_{n-1,i}B_{n,j}G_{n,n-1} - A_{n-1,i}A_{n-1,j}^T G_{n-1} \\
&= G_{n-1,n-3}L_{n-3,j}L_{n-2,i} - B_{n-1,i}(A_{n-1,j}G_{n,n-1} + B_{n-1,j}G_{n-1}) \\
&\quad - A_{n-2,i}^T (A_{n-2,j}G_{n-1}) - A_{n-1,i}B_{n,j}G_{n,n-1} - A_{n-1,i}A_{n-1,j}^T G_{n-1} \\
&= G_{n-1,n-3}L_{n-3,j}L_{n-2,i} - (B_{n-1,i}A_{n-1,j} + A_{n-1,i}B_{n,j})G_{n,n-1} \\
&\quad - (B_{n-1,i}B_{n-1,j} + A_{n-2,i}^T A_{n-2,j} + A_{n-1,i}A_{n-1,j}^T)G_{n-1} \\
&= A_{n-1,j}S_{n,i}.
\end{aligned}$$

Portanto, temos (4.3.6). Em particular, de (4.3.5) e (4.3.6), obtemos que

$$A_{k,i}G_{k+1} = G_k L_{k,i}, \quad 0 \leq k \leq n, \quad 1 \leq i \leq d, \quad (4.3.12)$$

e, como $A_{n,i}$ satisfaz a condição de posto, implica que G_{n+1} é inversível, como na prova do Teorema 3.5.1. Para completar a demonstração, devemos mostrar que $\mathbb{Q}_{n-2,i} = 0$. Temos, de (4.3.6), que

$$\mathbb{Q}_{n-2,j} = A_{n,j}\mathbb{P}_{n+1} - (x_j\mathbb{P}_n - B_{n,j}\mathbb{P}_n - A_{n-1,j}^T\mathbb{P}_{n-1}). \quad (4.3.13)$$

Agora, multiplicando por $D_{n,j}^T$ e somando para $j = 1, \dots, d$, segue, de $D_n^T A_n = I$ e da relação de recorrência de três termos (3.4.10), que

$$\sum_{j=1}^d D_{n,j}^T \mathbb{Q}_{n-2,j} = \sum_{j=1}^d D_{n,j}^T A_{n,j} \mathbb{P}_{n+1} - \mathbb{P}_{n+1} = 0. \quad (4.3.14)$$

Por outro lado, multiplicando (4.3.13) por $A_{n-1,i}$ à esquerda, segue, de (4.3.5), que

$$\begin{aligned} A_{n-1,i} \mathbb{Q}_{n-2,j} &= A_{n-1,i} A_{n,j} \mathbb{P}_{n+1} - x_j A_{n-1,i} \mathbb{P}_n + A_{n-1,i} B_{n,j} \mathbb{P}_n \\ &\quad + A_{n-1,i} A_{n-1,j}^T \mathbb{P}_{n-1} \\ &= A_{n-1,i} A_{n,j} \mathbb{P}_{n+1} - x_j (x_i \mathbb{P}_{n-1} - B_{n-1,i} \mathbb{P}_{n-1} - A_{n-2,i}^T \mathbb{P}_{n-2}) \\ &\quad + A_{n-1,i} B_{n,j} \mathbb{P}_n + A_{n-1,i} A_{n-1,j}^T \mathbb{P}_{n-1} \\ &= A_{n-1,i} A_{n,j} \mathbb{P}_{n+1} - x_i x_j \mathbb{P}_{n-1} \\ &\quad + B_{n-1,i} (A_{n-1,j} \mathbb{P}_n + B_{n-1,j} \mathbb{P}_{n-1} + A_{n-2,j}^T \mathbb{P}_{n-2}) \\ &\quad + A_{n-2,i}^T (A_{n-2,j} \mathbb{P}_{n-1} + B_{n-2,j} \mathbb{P}_{n-2} + A_{n-3,j}^T \mathbb{P}_{n-3}) \\ &\quad + A_{n-1,i} B_{n,j} \mathbb{P}_n + A_{n-1,i} A_{n-1,j}^T \mathbb{P}_{n-1} \\ &= -x_i x_j \mathbb{P}_{n-1} + A_{n-1,i} A_{n-1,j} \mathbb{P}_{n+1} \\ &\quad + (A_{n-1,i} B_{n,j} + B_{n-1,i} A_{n-1,j}) \mathbb{P}_n \\ &\quad + (A_{n-1,i} A_{n-1,j}^T + B_{n-1,i} B_{n-1,j} + A_{n-2,i}^T A_{n-2,j}) \mathbb{P}_{n-1} \\ &\quad + (A_{n-2,i}^T B_{n-2,j} + B_{n-1,i} A_{n-2,j}^T) \mathbb{P}_{n-2} + A_{n-2,i}^T A_{n-3,j}^T \mathbb{P}_{n-3}, \end{aligned}$$

o que, pela condição de comutatividade (4.2.1) e lembrando que $B_{n,i}$ é simétrica, implica que

$$A_{n-1,i} \mathbb{Q}_{n-2,j} = A_{n-1,j} \mathbb{Q}_{n-2,i}. \quad (4.3.15)$$

De (4.3.12), temos $A_{n-1,i} = G_{n-1} L_{n-1,i} G_n^{-1}$. Portanto, (4.3.15) é equivalente a $L_{n-1,i} X_j = L_{n-1,j} X_i$, com $X_i = G_n^{-1} \mathbb{Q}_{n-2,i}$. Definimos X como um vetor que satisfaz $L_{n,i} X = X_i$. Em outras palavras, recorrendo a notação usado na prova do Lema 4.3.4, X é definido como um vetor que satisfaz $X|_{\mathcal{N}_{n,i}} = X_i$. Como $\mathcal{N} = \cup \mathcal{N}_{n,i}$ e

$$X|_{\mathcal{N}_{n,i} \cap \mathcal{N}_{n,j}} = L_{n-1,i} L_{n,j} X = L_{n-1,i} X_j = L_{n-1,j} X_i = X|_{\mathcal{N}_{n,j} \cap \mathcal{N}_{n,i}},$$

vemos que X está bem definido. Desse modo, temos que \mathbb{Q}_n , dado por $\mathbb{Q}_n = G_{n+1} X$, é bem definido. Assim, por (4.3.12), segue que

$$\begin{aligned} L_{n,i} X = X_i &\Rightarrow L_{n,i} G_{n+1}^{-1} \mathbb{Q}_n = G_n^{-1} \mathbb{Q}_{n-2,i} \\ &\Rightarrow G_n^{-1} A_{n,i} \mathbb{Q}_n = G_n^{-1} \mathbb{Q}_{n-2,i} \\ &\Rightarrow A_{n,i} \mathbb{Q}_n = \mathbb{Q}_{n-2,i}. \end{aligned}$$

Logo, usando (4.3.14) e o fato que $D_n^T A_n = I$, obtemos $\mathbb{Q}_n = 0$. \square

4.4 Soluções gerais da relação de três termos

Nesta seção discutiremos mais uma diferença entre polinômios ortogonais de uma e de várias variáveis. No caso de uma variável, pelo Teorema de Favard e a relação de três termos, os polinômios ortogonais podem ser considerados como a solução da equação de diferenças,

$$a_k y_{k+1} + b_k y_k + a_{k-1} y_{k-1} = x y_k, \quad k \geq 1,$$

onde $a_k > 0$ e as condições iniciais são dadas por $y_0 = 1$, $y_1 = a_0^{-1}(x - b_0)$. Sabemos que a equação de diferenças tem outras soluções que satisfazem a condição inicial $y_0 = 0$ e $y_1 = a_0^{-1}$. As componentes da solução, denotada por q_k , são costumeiramente chamados de polinômios associados ou polinômios de segunda espécie. Juntas, essas duas soluções compartilham de muitas propriedades interessantes e $\{q_n\}$ desempenha um papel importante em áreas como o problema de momentos, teoria espectral das matrizes de Jacobi e frações contínuas.

No caso de várias variáveis, pela relação de três termos (3.4.3), podemos considerar os polinômios ortogonais em várias variáveis como solução das equações de diferenças de finitos multi-parâmetros

$$x_i Y_k = A_{k,i} Y_{k+1} + B_{k,i} Y_k + A_{k-1,i}^T Y_{k-1}, \quad 1 \leq i \leq d, k \geq 1, \quad (4.4.1)$$

onde $B_{k,i}$ são matrizes simétricas e $A_{k,i}$ satisfazem as condições de posto (3.4.8) e (3.4.7), com valores iniciais

$$Y_0 = a, Y_1 = b, \quad a \in \mathbb{R}, b \in \mathbb{R}^d. \quad (4.4.2)$$

Mas, ao procurar outras soluções linearmente independentes de (4.4.1) que podem ser definidas como os polinômios associados em várias variáveis, encontramos que não há tal solução além do sistema de polinômios ortogonais.

Teorema 4.4.1. *Se as equações de diferenças de multi-parâmetros (4.4.1) têm solução $\mathbb{P} = \{\mathbb{P}_k\}_{k=0}^\infty$ para os valores iniciais*

$$Y_0^* = 1, \quad Y_1^* = A_0^{-1}(x - B_0), \quad (4.4.3)$$

então todas as outras soluções de (4.4.1) e (4.4.2) são múltiplas de \mathbb{P}_n com a possível excessão da primeira componente. Mais precisamente, se $Y = \{Y_k\}_{k=0}^\infty$ é solução de (4.4.1) e (4.4.2), então $Y_k = h \mathbb{P}_k$ para todo $k \geq 1$, onde h é uma função independente de k .

Dem.: Por hipótese e pelo Teorema de Favard (Teorema 3.5.2), temos que $\mathbb{P}_n = \{\mathbb{P}_k\}_{k=0}^\infty$ forma uma sequência de polinômios ortonormais. Portanto, as matrizes coeficientes de (4.4.1) satisfazem as condições de comutatividade (4.2.1). Suponha que uma sequência de vetores $\{Y_k\}$ satisfaz (4.4.1) e os valores iniciais (4.4.2). Para a equação (4.4.1),

$$A_{k,i}Y_{k+1} = x_iY_k - B_{k,i}Y_k - A_{k-1,i}^T Y_{k-1}, \quad 1 \leq i \leq d. \quad (4.4.4)$$

Multiplicando $A_{k,j}Y_{k+1}$ à esquerda por $A_{k-1,i}$ e $A_{k,i}Y_{k+1}$ à esquerda por $A_{k-1,j}$, segue, da primeira condição de comutatividade (4.2.1) e de (4.4.4) que

$$\begin{aligned} A_{k-1,i}(x_jY_k - B_{k,j}Y_k - A_{k-1,j}^T Y_{k-1}) \\ = A_{k-1,j}(x_iY_k - B_{k,i}Y_k - A_{k-1,i}^T Y_{k-1}), \end{aligned} \quad (4.4.5)$$

para $1 \leq i, j \leq d$, e $k \geq 1$. Em particular, no caso $k = 1$ temos uma relação entre Y_0 e Y_1 , que, pela segunda equação de (4.2.1) com $k = 0$ e o fato de que $A_{0,i}A_{0,j}^T$ e $B_{0,i}$ são números, pode ser reescrita como

$$\begin{aligned} A_{0,i}(x_jY_1 - B_{1,j}Y_1 - A_{0,j}^T Y_0) &= A_{0,j}(x_iY_1 - B_{1,i}Y_1 - A_{0,i}^T Y_0) \\ x_jA_{0,i}Y_1 - A_{0,i}B_{1,j}Y_1 - A_{0,i}A_{0,j}^T Y_0 &= x_iA_{0,j}Y_1 - A_{0,j}B_{1,i}Y_1 - A_{0,j}A_{0,i}^T Y_0 \\ x_jA_{0,i}Y_1 - A_{0,i}B_{1,j}Y_1 &= x_iA_{0,j}Y_1 - A_{0,j}B_{1,i}Y_1 \\ (x_i - B_{0,i})A_{0,j}Y_1 &= (x_j - B_{0,j})A_{0,i}Y_1. \end{aligned} \quad (4.4.6)$$

Contudo, como x_i e x_j são variáveis independentes, Y_1 só pode ser função de \mathbf{x} . Além disso, tem que satisfazer $A_{0,i}Y_1 = (x_i - b_i)h(\mathbf{x})$, onde h é uma função de \mathbf{x} . Substituindo esta última equação na equação em (4.4.6), obtemos que $b_i = B_{0,i}$. No caso $h(\mathbf{x}) = 1$ corresponde a solução do polinômio ortogonal \mathbb{P} . Como $D_0^T = A_0^{-1}$, de (4.4.3) segue que

$$Y_1 = \sum_{i=1}^d D_{0,i}^T (x_i - B_{0,i})h(\mathbf{x}) = A_0^{-1}(x - B_0)h(\mathbf{x}) = h(\mathbf{x})\mathbb{P}_1.$$

Portanto, usando a inversa à esquerda, D_k^T , de A_k em (3.4.9), vemos que toda solução de (4.4.1) satisfaz

$$Y_{k+1} = \sum_{i=1}^d D_{k,i}^T x_i Y_k - E_k Y_k - F_k Y_{k-1}, \quad (4.4.7)$$

onde E_k e F_k são definidos como em (3.4.10). Consequentemente, para $k = 1$,

$$\begin{aligned} Y_2 &= \sum_{i=1}^d D_{1,i}^T (x_i I - B_{1,i})h(\mathbf{x})\mathbb{P}_1 - F_1 Y_0 \\ &= \left(\sum_{i=1}^d D_{1,i}^T (x_i I - B_{1,i})h(\mathbf{x})\mathbb{P}_1 - F_1 h(\mathbf{x})\mathbb{P}_0 \right) + F_1 h(\mathbf{x})\mathbb{P}_0 - F_1 Y_0 \\ &= h(\mathbf{x})\mathbb{P}_2 + F_1 (h(\mathbf{x}) - Y_0). \end{aligned}$$

Como \mathbb{P}_n é solução da equação (4.4.1) com as condições iniciais (4.4.3), segue, de (4.4.5) com $k = 2$, que

$$A_{1,i}(x_j\mathbb{P}_2 - B_{2,j}\mathbb{P}_2 - A_{1,j}^T\mathbb{P}_1) = A_{1,j}(x_i\mathbb{P}_2 - B_{2,i}\mathbb{P}_2 - A_{1,i}^T\mathbb{P}_1).$$

Multiplicando esta equação por $h(\mathbf{x})$ e subtraindo de (4.4.5) com $k = 2$, concluímos, das fórmulas para Y_1 e Y_2 , que

$$\begin{aligned} & A_{1,i}(x_jY_2 - B_{2,j}Y_2 - A_{1,j}^TY_1) - A_{1,i}h(\mathbf{x})(x_j\mathbb{P}_2 - B_{2,j}\mathbb{P}_2 - A_{1,j}^T\mathbb{P}_1) \\ &= A_{1,i}((x_jI - B_{2,j})Y_2 - (x_jI - B_{2,j})\mathbb{P}_2h(\mathbf{x})) - A_{1,j}^TY_1 + A_{1,j}^T\mathbb{P}_1h(\mathbf{x}) \\ &= A_{1,i}((x_jI - B_{2,j})(h(\mathbf{x})\mathbb{P}_2 + F_1(h(\mathbf{x}) - Y_0) - h(\mathbf{x})\mathbb{P}_2)) \\ &= A_{1,i}(x_jI - B_{2,j})F_1(h(\mathbf{x}) - Y_0). \end{aligned}$$

Analogamente, obtemos

$$\begin{aligned} & A_{1,j}(x_iY_2 - B_{2,i}Y_2 - A_{1,i}^TY_1) - A_{1,j}h(\mathbf{x})(x_i\mathbb{P}_2 - B_{2,i}\mathbb{P}_2 - A_{1,i}^T\mathbb{P}_1) \\ &= A_{1,j}(x_iI - B_{2,i})F_1(h(\mathbf{x}) - Y_0). \end{aligned}$$

Portanto, segue que

$$A_{1,i}(x_jI - B_{2,j})F_1(h(\mathbf{x}) - Y_0) = A_{1,j}(x_iI - B_{2,i})F_1(h(\mathbf{x}) - Y_0).$$

Se $Y_0 = h(\mathbf{x})$, então segue, de (4.4.7) e $Y_1 = h(\mathbf{x})\mathbb{P}_1$, que $Y_k = h(\mathbf{x})\mathbb{P}_k$ para todo $k \geq 0$. Agora, suponha que $h(\mathbf{x}) \neq Y_0$. Logo, $h(\mathbf{x}) - Y_0$ é um número não nulo e, então, da fórmula anterior,

$$A_{1,i}(x_jI - B_{2,j})F_1 = A_{1,j}(x_iI - B_{2,i})F_1.$$

Contudo, como x_i e x_j são variáveis independente, podemos concluir que para manter essa igualdade é necessário que $A_{1,i}F_1 = 0$, o que implica, por $D_n^T A_n = 1$, $F_1 = 0$. Assim, obtemos da fórmula para Y_2 que $Y_2 = h(\mathbf{x})\mathbb{P}_2$. Portanto, por (4.4.7), $Y_k = h(\mathbf{x})\mathbb{P}_k$ para todo $k \geq 1$, o que conclui a demonstração. \square

4.5 Núcleos reprodutores e série ortogonal de Fourier

Seja $\mathcal{L}(f) = \int f d\mu$ um funcional linear definido positivo, onde $d\mu$ é uma medida de Borel não-negativa com momentos finitos. Seja $\{\tilde{P}_\alpha^n\}$ uma sequência de polinômios ortonormais com respeito a \mathcal{L} . Para qualquer função f em $L^2(d\mu)$, podemos considerar a série ortogonal de Fourier com respeito a $\{\tilde{P}_\alpha^n\}$,

$$f \sim \sum_{n=0}^{\infty} \sum_{|\alpha|=n} a_\alpha^n(f) \tilde{P}_\alpha^n \quad \text{com} \quad a_\alpha^n(f) = \int f(\mathbf{x}) \tilde{P}_\alpha^n(\mathbf{x}) d\mu.$$

Contudo, as bases de polinômios ortonormais não são únicas e a série de Fourier, de fato, é independente de uma base particular. De fato, usando a notação vetorial $\tilde{\mathbb{P}}_n$, a definição acima de série de Fourier pode ser escrita como

$$f \sim \sum_{n=0}^{\infty} \mathbf{a}_n^T(f) \tilde{\mathbb{P}}_n, \quad \text{onde} \quad \mathbf{a}_n(f) = \int f(\mathbf{x}) \tilde{\mathbb{P}}_n(\mathbf{x}) d\mu. \quad (4.5.1)$$

Além disso, lembrando que \mathcal{V}_n^d denota o espaço dos polinômios ortogonais de grau exatamente n como em (3.2.5), segue, da ortonormalidade, que a expansão pode ser considerada como

$$f \sim \sum_{n=0}^{\infty} \text{proj}_{\mathcal{V}_n^d} f, \quad \text{onde} \quad \text{proj}_{\mathcal{V}_n^d} f = \mathbf{a}_n^T(f) \tilde{\mathbb{P}}_n.$$

Denotaremos o operador projeção $\text{proj}_{\mathcal{V}_n^d} f$ por $\mathbf{P}_n(f)$. Em termos de uma base ortonormal, podemos escrever como

$$\mathbf{P}_n(f; \mathbf{x}) = \int f(\mathbf{y}) \mathbf{P}_n(\mathbf{x}, \mathbf{y}) d\mu(\mathbf{y}), \quad \mathbf{P}_n(\mathbf{x}, \mathbf{y}) = \tilde{\mathbb{P}}_n^T(\mathbf{x}) \tilde{\mathbb{P}}_n(\mathbf{y}). \quad (4.5.2)$$

O operador projeção é independente de uma base particular. De fato, pelo Teorema 3.3.4, como duas bases ortonormais diferem por uma matriz ortogonal, é evidente de (4.5.2) que $\mathbf{P}_n(f)$ depende de \mathcal{V}_n^d e não de uma base particular de \mathcal{V}_n^d . Definimos a n -ésima soma parcial da expansão ortogonal de Fourier (4.5.1) por

$$S_n(f) = \sum_{k=0}^n \mathbf{a}_k^T(f) \mathbb{P}_k = \int \mathbf{K}_n(\cdot, \mathbf{y}) f(\mathbf{y}) d\mu, \quad (4.5.3)$$

onde o núcleo $\mathbf{K}_n(\cdot, \cdot)$ é definido por

$$\mathbf{K}_n(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^n \sum_{|\alpha|=k} P_\alpha^k(\mathbf{x}) P_\alpha^k(\mathbf{y}) = \sum_{k=0}^n \mathbf{P}_k(\mathbf{x}, \mathbf{y}), \quad (4.5.4)$$

que muitas vezes é chamado de n -ésimo núcleo reprodutor pela razões dadas na sub-seção abaixo.

4.5.1 Núcleos reprodutores

Para a definição do núcleo reprodutor \mathbf{K}_n não é necessário usar uma base ortonormal, qualquer base ortogonal servirá. Além disso, a definição faz sentido para qualquer funcional de momento.

Definição 4.5.1. *Sejam \mathcal{L} um funcional de momento e seja \mathbb{P}_n um sistema de polinômios ortogonais com respeito a \mathcal{L} . Definimos o núcleo reprodutor \mathbf{K}_n , em termos de \mathbb{P}_n , como*

$$\mathbf{K}_n(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^n \mathbb{P}_k^T(\mathbf{x}) H_k^{-1} \mathbb{P}_k(\mathbf{y}), \quad H_k = \mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T).$$

A seguir mostraremos que \mathbf{K}_n não depende de uma base particular de \mathcal{V}_k^d .

Teorema 4.5.2. *Seja \mathcal{V}_k^d , $k \geq 0$, definido por meio do funcional de momento \mathcal{L} . Então, $\mathbf{K}_n(\cdot, \cdot)$ depende somente de \mathcal{V}_k^d e não de uma base particular de \mathcal{V}_k^d .*

Dem.: Seja \mathbb{P}_k uma base de \mathcal{V}_k^d . Se \mathbb{Q}_k é outra base, então existe uma matriz inversível M_k , independente de \mathbf{x} , tal que $\mathbb{P}_k(\mathbf{x}) = M_k \mathbb{Q}_k(\mathbf{x})$. Sejam $H_k(\mathbb{P}) = \mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T)$ e $H_k(\mathbb{Q}) = \mathcal{L}(\mathbb{Q}_k \mathbb{Q}_k^T)$. Então,

$$\begin{aligned} H_k^{-1}(\mathbb{P}) &= (\mathcal{L}(\mathbb{P}_k \mathbb{P}_k^T))^{-1} = \left(\mathcal{L}((M_k \mathbb{Q}_k) (M_k \mathbb{Q}_k)^T) \right)^{-1} \\ &= (M_k \mathcal{L}(\mathbb{Q}_k \mathbb{Q}_k^T) M_k^T)^{-1} = (M_k H_k(\mathbb{Q}) M_k^T)^{-1} \\ &= (M_k^T)^{-1} H_k^{-1}(\mathbb{Q}) M_k^{-1}. \end{aligned}$$

Portanto,

$$\begin{aligned} \mathbb{P}_k^T H_k^{-1}(\mathbb{P}) \mathbb{P}_k &= \mathbb{P}_k^T \left[(M_k^T)^{-1} H_k^{-1}(\mathbb{Q}) M_k^{-1} \right] \mathbb{P}_k \\ &= (M_k^{-1} \mathbb{P}_k)^T H_k^{-1}(\mathbb{Q}) (M_k^{-1} \mathbb{P}_k) \\ &= \mathbb{Q}_k^T H_k^{-1}(\mathbb{Q}) \mathbb{Q}_k, \end{aligned}$$

o que conclui a demonstração. \square

Assim, como a definição de \mathbf{K}_n não depende de uma base em particular, portanto é conveniente utilizar a base de polinômios ortonormais com relação ao funcional definido positivo \mathcal{L} . O próximo teorema justifica o nome de núcleo reprodutor.

Teorema 4.5.3. *Seja \mathcal{L} um funcional linear definido positivo em Π^d . Então, para todo $P \in \Pi_n^d$,*

$$P(\mathbf{x}) = \mathcal{L}(\mathbf{K}_n(\mathbf{x}, \cdot) P(\cdot)).$$

Dem.: Seja \mathbb{P}_n um sistema de polinômios ortogonais com relação a \mathcal{L} e seja $\mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T) = H_n$. Para qualquer $P \in \Pi_n^d$, podemos expandí-lo em termos da base $\{\mathbb{P}_0, \dots, \mathbb{P}_n\}$ como

$$P(\mathbf{x}) = \sum_{k=0}^n [\mathbf{a}_k(P)]^T H_k^{-1} \mathbb{P}_k(\mathbf{x}), \text{ com } \mathbf{a}_k(P) = \mathcal{L}(P \mathbb{P}_k).$$

De fato, seja $\mathbf{b}_k \in \mathbb{R}^{r_k}$. Então, pela Proposição 3.2.15, temos

$$P(\mathbf{x}) = \sum_{k=0}^n \mathbf{b}_k^T \mathbb{P}_k(\mathbf{x}).$$

Multiplicando por $\mathbb{P}_l^T(\mathbf{x})$ à direita e aplicando \mathcal{L} , segue que

$$\begin{aligned} \mathcal{L}(P(\mathbf{x}) \mathbb{P}_l^T(\mathbf{x})) &= \mathcal{L} \left(\left(\sum_{k=0}^n \mathbf{b}_k^T \mathbb{P}_k(\mathbf{x}) \right) \mathbb{P}_l^T(\mathbf{x}) \right) \\ &= \mathbf{b}_l^T H_l. \end{aligned}$$

Portanto, pela Proposição 2.1.3 e tomando $\mathcal{L}(P(\mathbf{x})\mathbb{P}_l(\mathbf{x})) = \mathbf{a}_l(P)$, segue o fato. Além disso, como \mathbf{K}_n é uma função escalar, temos

$$\begin{aligned} \mathcal{L}(\mathbf{K}_n(\mathbf{x}, \cdot)P(\cdot)) &= \mathcal{L}\left(\sum_{k=0}^n \mathbb{P}_k^T(\mathbf{x})H_k^{-1}\mathbb{P}_k(\cdot)P(\cdot)\right) \\ &= \mathcal{L}\left(\sum_{k=0}^n P(\cdot)\mathbb{P}_k^T(\cdot)H_k^{-1}\mathbb{P}_k(\mathbf{x})\right) \\ &= \sum_{k=0}^n \mathcal{L}(P\mathbb{P}_k)^T H_k^{-1}\mathbb{P}_k(\mathbf{x}) = \sum_{k=0}^n [\mathbf{a}_k(P)]^T H_k^{-1}\mathbb{P}_k(\mathbf{x}) \\ &= P(\mathbf{x}), \end{aligned}$$

o que conclui a demonstração do teorema. \square

Os núcleos reprodutores para polinômios ortogonais em uma variável têm uma fórmula compacta, chamada de fórmula de Christoffel-Darboux. O teorema a seguir é uma extensão dessa fórmula para várias variáveis.

Teorema 4.5.4. *Sejam \mathcal{L} um funcional linear definido positivo e $\{\mathbb{P}_k\}_{k=0}^\infty$ um sistema de polinômios ortogonais com respeito a \mathcal{L} . Então, para qualquer inteiro $n \geq 0$, \mathbf{x} e $\mathbf{y} \in \mathbb{R}^d$,*

$$\begin{aligned} \mathbf{K}_n(\mathbf{x}, \mathbf{y}) &= \sum_{k=0}^n \mathbb{P}_k^T(\mathbf{x})H_k^{-1}\mathbb{P}_k(\mathbf{y}) \\ &= \frac{[A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1}\mathbb{P}_n(\mathbf{y}) - \mathbb{P}_n^T(\mathbf{x})H_n^{-1}[A_{n,i}\mathbb{P}_{n+1}(\mathbf{y})]}{x_i - y_i}, \end{aligned} \quad (4.5.5)$$

para $1 \leq i \leq d$, onde $\mathbf{x} = (x_1, \dots, x_d)$ e $\mathbf{y} = (y_1, \dots, y_d)$.

Dem.: Pelo Teorema 3.4.1, o polinômio ortogonal \mathbb{P}_n satisfaz a relação de três termos (3.4.1). Vamos denotar

$$\sum_k = [A_{k,i}\mathbb{P}_{k+1}(\mathbf{x})]^T H_k^{-1}\mathbb{P}_k(\mathbf{y}) - \mathbb{P}_k^T(\mathbf{x})H_k^{-1}[A_{k,i}\mathbb{P}_{k+1}(\mathbf{y})].$$

Da relação de três termos (3.4.1),

$$\begin{aligned} \sum_k &= [x_i\mathbb{P}_k(\mathbf{x}) - B_{k,i}\mathbb{P}_k(\mathbf{x}) - C_{k,i}\mathbb{P}_{k-1}(\mathbf{x})]^T H_k^{-1}\mathbb{P}_k(\mathbf{y}) \\ &\quad - \mathbb{P}_k^T(\mathbf{x})H_k^{-1}[y_i\mathbb{P}_k(\mathbf{y}) - B_{k,i}\mathbb{P}_k(\mathbf{y}) - C_{k,i}\mathbb{P}_{k-1}(\mathbf{y})] \\ &= (x_i - y_i)\mathbb{P}_k^T(\mathbf{x})H_k^{-1}\mathbb{P}_k(\mathbf{y}) - \mathbb{P}_k^T(\mathbf{x})[B_{k,i}^T H_k^{-1} - H_k^{-1}B_{k,i}]\mathbb{P}_k(\mathbf{y}) \\ &\quad - [\mathbb{P}_{k-1}^T(\mathbf{x})C_{k,i}^T H_k^{-1}\mathbb{P}_k(\mathbf{y}) - \mathbb{P}_k^T(\mathbf{x})H_k^{-1}C_{k,i}\mathbb{P}_{k-1}(\mathbf{y})]. \end{aligned}$$

Pela Proposição 3.2.17 e (3.4.2), ambas, H_k e $B_{k,i}H_k$, são matrizes simétricas. Daí, $B_{k,i}H_k = H_k B_{k,i}^T$, o que implica que $B_{k,i}^T H_k^{-1} = H_k^{-1}B_{k,i}$. Por isso, o segundo termo da igualdade à direita é zero. Da terceira equação de (3.4.2),

$$A_{k-1,i}H_k = H_{k-1}C_{k,i}^T \quad \Leftrightarrow \quad H_{k-1}^{-1}A_{k-1,i} = C_{k,i}^T H_k^{-1}.$$

Por esta razão, o terceiro termo da igualdade à direita é dada por

$$-\mathbb{P}_{k-1}^T(\mathbf{x})H_{k-1}^{-1}A_{k-1,i}\mathbb{P}_k(\mathbf{y}) + \mathbb{P}_k^T(\mathbf{x})A_{k-1,i}^T H_{k-1}^{-1}\mathbb{P}_{k-1}(\mathbf{y}) = \sum_{k-1}.$$

Assim, a expressão pode ser escrita como

$$(x_i - y_i)\mathbb{P}_k^T(\mathbf{x})H_{k-1}^{-1}\mathbb{P}_k(\mathbf{y}) = \sum_k - \sum_{k-1}.$$

Note que $\sum_{-1} = 0$. De fato,

$$\sum_{-1} = [A_{-1,i}\mathbb{P}_0(\mathbf{x})]^T H_{-1}^{-1}\mathbb{P}_{-1}(\mathbf{y}) - \mathbb{P}_{-1}^T(\mathbf{x})H_{-1}^{-1}[A_{-1,i}\mathbb{P}_0(\mathbf{y})] = 0,$$

pois $\mathbb{P}_{-1}(\cdot) = 0$. Desse modo, somando esta identidade de 0 a n , obtemos a igualdade (4.5.5). \square

Corolário 4.5.5. *Seja $\{\mathbb{P}_k\}_{k=0}^\infty$ como no Teorema 4.5.4. Então,*

$$\begin{aligned} \mathbf{K}_n(\mathbf{x}, \mathbf{x}) &= \sum_{k=0}^n \mathbb{P}_k^T(\mathbf{x})H_k^{-1}\mathbb{P}_k(\mathbf{x}) \\ &= \mathbb{P}_n^T(\mathbf{x})H_n^{-1}[A_{n,i}\partial_i\mathbb{P}_{n+1}(\mathbf{x})] - [A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1}\partial_i\mathbb{P}_n(\mathbf{x}), \end{aligned}$$

para $i = 1, \dots, d$, onde $\partial_i = \partial/\partial x_i$ denotada a derivada parcial com relação a x_i .

Dem.: Como $\mathbb{P}_n^T(\mathbf{x})H_n^{-1}[A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})]$ é uma função escalar, é igual à sua própria transposta. Assim,

$$\mathbb{P}_n^T(\mathbf{x})H_n^{-1}[A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})] = [A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1}\mathbb{P}_n(\mathbf{x}).$$

Portanto, o numerador do lado direito da fórmula de Christoffel-Darboux pode ser escrito como

$$[A_{n,i}\mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1}[\mathbb{P}_n(\mathbf{y}) - \mathbb{P}_n(\mathbf{x})] - \mathbb{P}_n^T(\mathbf{x})H_n^{-1}A_{n,i}[\mathbb{P}_{n+1}(\mathbf{y}) - \mathbb{P}_{n+1}(\mathbf{x})].$$

O resultado segue, então, da fórmula do Teorema 4.5.4 fazendo $y_i \rightarrow x_i$. \square

Se $\tilde{\mathbb{P}}_n$ são polinômios ortonormais, então a fórmula do Teorema 4.5.4 e o Corolário 4.5.5 valem com $H_k = I$. Seguem estes casos no teorema a seguir, que são facilmente verificados.

Teorema 4.5.6. *Seja $\tilde{\mathbb{P}}_n$ um sistema de polinômios ortonormais. Então,*

$$\mathbf{K}_n(\mathbf{x}, \mathbf{y}) = \frac{[A_{n,i}\tilde{\mathbb{P}}_{n+1}(\mathbf{x})]^T \tilde{\mathbb{P}}_n(\mathbf{y}) - \tilde{\mathbb{P}}_n^T(\mathbf{x})[A_{n,i}\tilde{\mathbb{P}}_{n+1}(\mathbf{y})]}{x_i - y_i}, \quad (4.5.6)$$

e

$$\mathbf{K}_n(\mathbf{x}, \mathbf{x}) = \tilde{\mathbb{P}}_n^T(\mathbf{x})[A_{n,i}\partial_i\tilde{\mathbb{P}}_{n+1}(\mathbf{x})] - [A_{n,i}\tilde{\mathbb{P}}_{n+1}(\mathbf{x})]^T \partial_i\tilde{\mathbb{P}}_n(\mathbf{x}). \quad (4.5.7)$$

É interessante notar que, embora o lado direito da fórmula parece depender de i , o lado esquerdo mostra que não depende.

Para uma variável, a função $1/\mathbf{K}_n(\mathbf{x}, \mathbf{x})$ é chamada, frequentemente, função de Christoffel. Mantemos este nome para várias variáveis e definimos

$$\Lambda_n(\mathbf{x}) = [\mathbf{K}_n(\mathbf{x}, \mathbf{x})]^{-1}. \quad (4.5.8)$$

Claramente esta função é positiva e, além disso, satisfaz a propriedade a seguir.

Teorema 4.5.7. *Seja \mathcal{L} um funcional linear definido positivo. Para um ponto arbitrário $\mathbf{x} \in \mathbb{R}^d$,*

$$\Lambda_n(\mathbf{x}) = \min\{\mathcal{L}[P^2] : P(\mathbf{x}) = 1, P \in \Pi_n^d\},$$

isto é, o mínimo é tomado entre todos os polinômios $P \in \Pi_n^d$ sujeito à condição $P(\mathbf{x}) = 1$.

Dem.: Como \mathcal{L} é definido positivo, existe um sistema $\{\tilde{\mathbb{P}}_k\}$ de polinômios ortonormais. Se P é um polinômio de grau no máximo n , então P pode ser escrito em termos da base ortonormal como

$$P(\mathbf{x}) = \sum_{k=0}^n \mathbf{a}_k^T(P) \tilde{\mathbb{P}}_k(\mathbf{x}), \quad \text{onde } \mathbf{a}_k(P) = \mathcal{L}(P \tilde{\mathbb{P}}_k).$$

Das propriedades de ortonormalidade de $\tilde{\mathbb{P}}_k$, segue que

$$\begin{aligned} \mathcal{L}[P^2] &= \mathcal{L} \left[\left(\sum_{k=0}^n \mathbf{a}_k^T(P) \tilde{\mathbb{P}}_k \right) \left(\sum_{l=0}^n \mathbf{a}_l^T(P) \tilde{\mathbb{P}}_l \right) \right] \\ &= \mathcal{L} \left[\left(\sum_{k=0}^n \mathbf{a}_k^T(P) \tilde{\mathbb{P}}_k \right) \left(\sum_{l=0}^n \tilde{\mathbb{P}}_l^T \mathbf{a}_l(P) \right) \right] \\ &= \sum_{k=0}^d \sum_{l=0}^d \mathbf{a}_k^T(P) \mathcal{L}(\tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_l^T) \mathbf{a}_l(P) = \sum_{k=0}^n \mathbf{a}_k^T(P) \mathbf{a}_k(P) = \sum_{k=0}^n \|\mathbf{a}_k(P)\|^2 \end{aligned}$$

Se $P(\mathbf{x}) = 1$, então, pela desigualdade de Cauchy-Schwarz,

$$\begin{aligned} 1 &= P(\mathbf{x}) = \sum_{k=0}^n \mathbf{a}_k^T(P) \tilde{\mathbb{P}}_k(\mathbf{x}) \leq \sum_{k=0}^n \|\mathbf{a}_k(P)\| \|\tilde{\mathbb{P}}_k(\mathbf{x})\| \\ &\leq \sum_{k=0}^n \|\mathbf{a}_k(P)\|^2 \sum_{k=0}^n \|\tilde{\mathbb{P}}_k(\mathbf{x})\|^2 = \mathcal{L}[P^2] \mathbf{K}_n(\mathbf{x}, \mathbf{x}), \end{aligned}$$

onde a igualdade vale se, e somente se, $\mathbf{a}_k(P) = [\mathbf{K}_n(\mathbf{x}, \mathbf{x})]^{-1} \tilde{\mathbb{P}}_k(\mathbf{x})$, o que conclui a demonstração. \square

O teorema afirma que a função Λ_n é solução de um problema extremo. Podemos também tomar a fórmula do Teorema 4.5.7 como a definição da função de Christoffel.

4.5.2 Série ortogonal de Fourier

Seja \mathcal{L} um funcional linear definido positivo e $\mathcal{H}_{\mathcal{L}}$ o espaço de Hilbert de funções de valores reais definido em \mathbb{R}^d com produto interno $\langle f, g \rangle = \mathcal{L}(fg)$. Então, o teorema padrão do espaço de Hilbert aplica-se à na série de Fourier definida em (4.5.1).

Teorema 4.5.8. *Sejam \mathcal{L} um funcional linear definido positivo e $f \in \mathcal{H}_{\mathcal{L}}$. Então, sobre todos os polinômios P em Π_n^d , o valor de*

$$\mathcal{L} (|f(\mathbf{x}) - P(\mathbf{x})|^2)$$

será mínimo se, e somente se, $P = S_n(f)$.

Dem.: Seja $\{\tilde{\mathbb{P}}_k\}$ uma base ortonormal de \mathcal{V}_k . Para qualquer $P \in \Pi_n^d$, existem \mathbf{b}_k , $k = 1, \dots, n$, tais que

$$P(\mathbf{x}) = \sum_{k=0}^n \mathbf{b}_k^T \tilde{\mathbb{P}}_k(\mathbf{x})$$

e $S_n(f)$ satisfaz a equação (4.5.3). Segue, do argumento padrão, que

$$\begin{aligned} 0 &\leq \mathcal{L} (|f - P|^2) = \mathcal{L} (f^2) - 2 \sum_{k=0}^n \mathbf{b}_k^T \mathcal{L}(f \tilde{\mathbb{P}}_k) + \sum_{k=0}^n \sum_{j=0}^n \mathbf{b}_k^T \mathcal{L}(\tilde{\mathbb{P}}_k \tilde{\mathbb{P}}_j^T) \mathbf{b}_j \\ &= \mathcal{L} (f^2) - 2 \sum_{k=0}^n \mathbf{b}_k^T \mathbf{a}_k(f) + \sum_{k=0}^n \mathbf{b}_k^T \mathbf{b}_k \\ &= \mathcal{L} (f^2) - \sum_{k=0}^n \mathbf{a}_k^T(f) \mathbf{a}_k(f) + \sum_{k=0}^n [\mathbf{a}_k^T(f) \mathbf{a}_k(f) + \mathbf{b}_k^T \mathbf{b}_k - 2\mathbf{b}_k^T \mathbf{a}_k(f)] \\ &= \mathcal{L} (f^2) - \sum_{k=0}^n \mathbf{a}_k^T(f) \mathbf{a}_k(f) + \sum_{k=0}^n [(\mathbf{a}_k(f) - \mathbf{b}_k)^T (\mathbf{a}_k(f) - \mathbf{b}_k)]. \end{aligned}$$

Pela desigualdade de Cauchy-Schwarz o terceiro termo do lado direito da igualdade é não-negativo. Além disso, o valor de $\mathcal{L} (|f - P|^2)$ é mínimo se, e somente se, $\mathbf{b}_k = \mathbf{a}_k$, ou seja, $P = S_n(f)$. \square

No caso do mínimo ser atingido, temos a desigualdade de Bessel,

$$\sum_{k=0}^{\infty} \mathbf{a}_k^T(f) \mathbf{a}_k(f) \leq \mathcal{L} (|f|^2).$$

O próximo resultado segue do teorema padrão do espaço de Hilbert.

Teorema 4.5.9. *Se Π_n^d é denso em $\mathcal{H}_{\mathcal{L}}$, então $S_n(f)$ converge para f em $\mathcal{H}_{\mathcal{L}}$ e obtemos a igualdade de Parseval*

$$\sum_{k=0}^{\infty} \mathbf{a}_k^T(f) \mathbf{a}_k(f) = \mathcal{L} (|f|^2).$$

A definição da série ortogonal de Fourier não depende de uma base particular de \mathcal{V}_n^d . Se \mathbb{P}_n é uma sequência de polinômios ortogonais, ao invés de ortonormais, então a série de Fourier toma a forma

$$f \sim \sum_{n=0}^{\infty} \mathbf{a}_n^T(f) H_n^{-1} \mathbb{P}_n, \quad \mathbf{a}_n(f) = \mathcal{L}(f \mathbb{P}_n), \quad H_n = \mathcal{L}(\mathbb{P}_n^T \mathbb{P}_n). \quad (4.5.9)$$

A matriz inversa H_n^{-1} é o que torna difícil encontrar a série ortogonal de Fourier usando uma base ortogonal, ao invés de ortonormal.

Um método alternativo é usar um polinômio biortogonal para encontrar a série ortogonal de Fourier. Sejam $\mathbb{P}_n = \{P_\alpha^n\}$ e $\mathbb{Q}_n = \{Q_\alpha^n\}$ dois sistemas de polinômios ortogonais com respeito a \mathcal{L} . Eles são biortogonais se $\mathcal{L}(P_\alpha^n Q_\beta^m) = d_\alpha \delta_{\alpha,\beta}$, onde d_α é diferente de zero. Se todos os d_α são iguais a um, os dois sistemas são chamados de ortonormais. Se $\tilde{\mathbb{P}}_n$ e $\tilde{\mathbb{Q}}_n$ são biortonormais, então o coeficiente da série ortogonal de Fourier em termos de $\tilde{\mathbb{P}}_n$ pode ser encontrados como segue

$$f \sim \sum_{n=0}^{\infty} \mathbf{a}_n^T(f) \tilde{\mathbb{P}}_n, \quad \mathbf{a}_n(f) = \mathcal{L}(f \tilde{\mathbb{Q}}_n),$$

isto é, os coeficientes de Fourier associados a \tilde{P}_α^n são dados por $\mathcal{L}(f \tilde{Q}_\alpha^n)$. A idéia de usar polinômios biortogonais para estudar a expansão ortogonal de Fourier foi primeiro aplicado no caso dos polinômios ortogonais clássicos na bola unitária e podem ser vistos em [6].

Por fim, veremos um problema extremo para polinômios em várias variáveis. Seja G_n a matriz coeficiente do termo de maior grau do vetor polinomial ortonormal $\tilde{\mathbb{P}}_n$ como anteriormente. Denotemos por $\|G_n\| = \max_{x \neq 0} \frac{\|G_n x\|_2}{\|x\|_2}$ a norma espectral de G_n . Assim, $\|G_n\|^2$ é igual ao maior autovalor de $G_n G_n^T$.

Teorema 4.5.10. *Seja \mathcal{L} um funcional linear definido positivo. Então,*

$$\min \{ \mathcal{L}(P^2) : P = \mathbf{a}^T \mathbb{X}^n + \dots \in \Pi_n^d, \quad \|\mathbf{a}\|_2 = 1 \} = \|G_n\|^{-2}.$$

Dem.: Como $\{\tilde{\mathbb{P}}_0, \dots, \tilde{\mathbb{P}}_n\}$ forma uma base de Π_n^d , podemos reescrever P como

$$P(\mathbf{x}) = \mathbf{a}^T G_n^{-1} \tilde{\mathbb{P}}_n + \mathbf{a}_{n-1}^T \tilde{\mathbb{P}}_{n-1} + \dots + \mathbf{a}_0^T \tilde{\mathbb{P}}_0.$$

Para $\|\mathbf{a}\|_2 = 1$, segue, da desigualdade de Bessel e de $P(\mathbf{x})$ ser uma função escalar, que

$$\begin{aligned}
\mathcal{L}(P^2) &\geq \sum_{k=0}^n \mathbf{a}_k^T(P) \mathbf{a}_k(P) = \sum_{k=0}^n \left[\mathcal{L}(P\tilde{\mathbb{P}}_k) \right]^T \left[\mathcal{L}(P\tilde{\mathbb{P}}_k) \right] \\
&= \sum_{k=0}^n \left[\mathcal{L} \left(\left(\mathbf{a}^T G_n^{-1} \tilde{\mathbb{P}}_n + \mathbf{a}_{n-1}^T \tilde{\mathbb{P}}_{n-1} + \cdots + \mathbf{a}_0^T \tilde{\mathbb{P}}_0 \right) \tilde{\mathbb{P}}_k \right) \right] \\
&\quad \times \left[\mathcal{L} \left(\left(\mathbf{a}^T G_n^{-1} \tilde{\mathbb{P}}_n + \mathbf{a}_{n-1}^T \tilde{\mathbb{P}}_{n-1} + \cdots + \mathbf{a}_0^T \tilde{\mathbb{P}}_0 \right) \tilde{\mathbb{P}}_k \right) \right]^T \\
&= \mathbf{a}^T G_n^{-1} (G_n^{-1})^T \mathbf{a} + \|\mathbf{a}_{n-1}\|^2 + \cdots + \|\mathbf{a}_0\|^2 \\
&\geq \mathbf{a}^T (G_n^T G_n)^{-1} \mathbf{a} = \|\mathbf{a}^T G_n^{-1}\|^2 \geq \lambda_{\min},
\end{aligned}$$

onde λ_{\min} é o menor autovalor de $(G_n^T G_n)^{-1}$, que é igual ao inverso do maior autovalor de $G_n G_n^T$. Então, provamos que

$$\mathcal{L}(P^2) \geq \|G_n\|^{-2}, \quad P(x) = \mathbf{a}^T \mathbb{X}^n + \cdots, \quad \|\mathbf{a}\| = 1.$$

Escolhendo \mathbf{a} tal que $\|\mathbf{a}\| = 1$ e $\mathbf{a}^T G_n^{-1} (G_n^{-1})^T \mathbf{a} = \lambda_{\min}$, vemos que o polinômio $P = \mathbf{a}^T G_n^{-1} \tilde{\mathbb{P}}_n$ atinge o limite inferior. \square

O teorema anterior mostra que $\|G_n\|$ é também uma extensão natural do coeficiente do termo de maior grau k_n dos polinômios ortogonais em uma variável.

Capítulo 5

Zeros Comuns dos Polinômios Ortogonais em Várias Variáveis e Cubatura Gaussiana

5.1 Zeros comuns dos polinômios ortogonais em várias variáveis

O n -ésimo polinômio ortogonal em uma variável tem exatamente n zeros reais e distintos, e estes zeros são os autovalores da matriz truncada de Jacobi, J_n . Este fato tem importantes aplicações em fórmulas de quadratura para cálculo de integrais definidas.

Em geral, um zero de polinômio em duas variáveis pode ser um ponto ou uma curva algébrica no plano. A estrutura dos zeros em várias variáveis são muito mais complexas.

Definição 5.1.1. *Seja \mathcal{L} um funcional linear definido positivo e $\tilde{\mathbb{P}}_n = \{\tilde{P}_\alpha^n\}$ um vetor polinomial coluna ortonormal associado a \mathcal{L} . Um zero comum de $\tilde{\mathbb{P}}_n$ é um zero para todo \tilde{P}_α^n , ou seja, um zero comum é um zero de todas as componentes polinomiais do vetor polinomial coluna.*

Definição 5.1.2. *Dizemos que \mathbf{z} é um zero simples de $\tilde{\mathbb{P}}_n$, se \mathbf{z} é um zero de $\tilde{\mathbb{P}}_n$ e pelo menos uma derivada parcial de $\tilde{\mathbb{P}}_n$ em \mathbf{z} não se anula em \mathbf{z} .*

Agora, veremos duas propriedades simples dos zeros comuns.

Teorema 5.1.3. *Todos os zeros comuns de $\tilde{\mathbb{P}}_n$ são distintos. Além disso, os zeros comuns de $\tilde{\mathbb{P}}_n$ são diferentes dos zeros comuns de $\tilde{\mathbb{P}}_{n-1}$.*

Dem.: Da fórmula de Christoffel-Darboux do Teorema 4.5.4, temos

$$\sum_{k=0}^{n-1} \tilde{\mathbb{P}}_n^T(\mathbf{x}) \tilde{\mathbb{P}}_k(\mathbf{x}) = \tilde{\mathbb{P}}_{n-1}^T(\mathbf{x}) A_{n-1,i} \partial_i \tilde{\mathbb{P}}_n(\mathbf{x}) - \tilde{\mathbb{P}}_n^T(\mathbf{x}) A_{n-1,i}^T \partial_i \tilde{\mathbb{P}}_{n-1}(\mathbf{x}),$$

onde $\partial_i = \partial/\partial x_i$ denota a derivada parcial com respeito a x_i . Se \mathbf{z} é um zero de $\tilde{\mathbb{P}}_n$, então

$$0 < \sum_{k=0}^{n-1} \left\| \tilde{\mathbb{P}}_k(\mathbf{z}) \right\|^2 = \sum_{k=0}^{n-1} \tilde{\mathbb{P}}_k^T(\mathbf{z}) \tilde{\mathbb{P}}_k(\mathbf{z}) = \tilde{\mathbb{P}}_{n-1}^T(\mathbf{z}) A_{n-1,i} \partial_i \tilde{\mathbb{P}}_n(\mathbf{z}).$$

Portanto, nem $\tilde{\mathbb{P}}_{n-1}(\mathbf{z})$ nem $\partial_i \tilde{\mathbb{P}}_n(\mathbf{z})$ podem ser nulos. \square

Definição 5.1.4. *Dizemos que $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$ é um autovalor junção de $J_{n,1}, \dots, J_{n,d}$, onde $J_{n,i}$ é a matriz da Definição 4.2.3, se existe $\epsilon \neq 0$, $\epsilon \in \mathbb{R}^N$, tal que $J_{n,i}\epsilon = \lambda_i\epsilon$ para $i = 1, \dots, d$. O vetor ϵ é chamado de autovetor junção associados a Λ .*

Os zeros comuns de $\tilde{\mathbb{P}}_n$ são caracterizados pelo próximo teorema.

Teorema 5.1.5. *Um ponto $\Lambda = (\lambda_1, \dots, \lambda_d)^T \in \mathbb{R}^d$ é um zero comum de $\tilde{\mathbb{P}}_n$ se, e somente se, é um autovalor junção de $J_{n,1}, \dots, J_{n,d}$. Além disso, $\left(\tilde{\mathbb{P}}_0^T(\Lambda), \dots, \tilde{\mathbb{P}}_{n-1}^T(\Lambda) \right)^T$ é um autovetor junção de Λ .*

Dem.: Se $\tilde{\mathbb{P}}_n(\Lambda) = 0$, então segue da relação de três termos que

$$\begin{aligned} B_{0,i} \tilde{\mathbb{P}}_0(\Lambda) + A_{0,i} \tilde{\mathbb{P}}_1(\Lambda) &= \lambda_i \tilde{\mathbb{P}}_0(\Lambda), \\ A_{k-1,i}^T \tilde{\mathbb{P}}_{k-1}(\Lambda) + B_{k,i} \tilde{\mathbb{P}}_k(\Lambda) + A_{k,i} \tilde{\mathbb{P}}_{k+1}(\Lambda) &= \lambda_i \tilde{\mathbb{P}}_k(\Lambda), \quad 1 \leq k \leq n-2, \\ A_{n-2,i}^T \tilde{\mathbb{P}}_{n-2}(\Lambda) + B_{n-1,i} \tilde{\mathbb{P}}_{n-1}(\Lambda) &= \lambda_i \tilde{\mathbb{P}}_{n-1}(\Lambda), \end{aligned}$$

para $1 \leq i \leq d$. Da definição de $J_{n,i}$, segue que

$$J_{n,i}\epsilon = \lambda_i\epsilon, \quad \epsilon = \left(\tilde{\mathbb{P}}_0^T(\Lambda), \dots, \tilde{\mathbb{P}}_{n-1}^T(\Lambda) \right)^T.$$

Então, Λ é o autovalor de $J_{n,i}$ com o autovetor junção ϵ .

Por outro lado, suponha que $\Lambda = (\lambda_1, \dots, \lambda_d)^T$ é um autovalor de $J_{n,i}$ e $J_{n,i}$ tem um autovetor junção para Λ . Queremos encontrar esse autovetor. Escrevendo $\epsilon = (\mathbf{x}_0^T, \dots, \mathbf{x}_{n-1}^T)$, $\mathbf{x}_j \in \mathbb{R}^{r_j}$, como $J_{n,i}\epsilon = \lambda_i\epsilon$, segue que $\{\mathbf{x}_j\}$ satisfaz a relação de três termos.

$$\begin{aligned} B_{0,i} \mathbf{x}_0 + A_{0,i} \mathbf{x}_1 &= \lambda_i \mathbf{x}_0, \\ A_{k-1,i}^T \mathbf{x}_{k-1} + B_{k,i} \mathbf{x}_k + A_{k,i} \mathbf{x}_{k+1} &= \lambda_i \mathbf{x}_k, \quad 1 \leq k \leq n-2, \\ A_{n-2,i}^T \mathbf{x}_{n-2} + B_{n-1,i} \mathbf{x}_{n-1} &= \lambda_i \mathbf{x}_{n-1}, \end{aligned} \tag{5.1.1}$$

para $1 \leq i \leq d$. Mostremos que $\epsilon \neq 0$. De fato, suponha que $\mathbf{x}_0 = 0$. Então, segue, da primeira equação da relação de três termos em (5.1.2), que $A_{0,i}\mathbf{x}_1 = 0$, o que implica que $A_0\mathbf{x}_1 = 0$. Como A_0 é uma matriz $d \times d$ e tem posto completo, temos que $\mathbf{x}_1 = 0$. Logo, como $\mathbf{x}_0 = 0$ e $\mathbf{x}_1 = 0$, temos da próxima relação de três termos de (5.1.2), que $A_{1,i}\mathbf{x}_2 = 0$. Portanto, $A_1\mathbf{x}_2 = 0$ e, como A_1 tem posto completo, obtemos que $\mathbf{x}_2 = 0$. Continuando este argumento, provamos que $\mathbf{x}_i = 0$ para $0 \leq i \leq n-1$. Portanto, temos que $\epsilon = 0$, o que contradiz o fato de ϵ ser um autovetor.

Agora, provaremos que $\mathbf{x}_j = \tilde{\mathbb{P}}_j(\Lambda)$ para todo $0 \leq j \leq n$. Suponha que $\mathbf{x}_0 = 1 = \tilde{\mathbb{P}}_0$ e definimos $\mathbf{x}_n \in \mathbb{R}_{r_n}$ como $\mathbf{x}_n = 0$. Portanto, a última equação da relação de três termos em (5.1.2) pode ser escrita como

$$A_{n-2,i}^T \mathbf{x}_{n-2} + B_{n-1,i} \mathbf{x}_{n-1} + A_{n-1,i} \mathbf{x}_n = \lambda_i \mathbf{x}_{n-1}.$$

Logo, $\{\mathbf{x}_k\}_{k=0}^n$ e $\{\tilde{\mathbb{P}}_k(\Lambda)\}_{k=0}^n$ satisfazem a mesma relação de três termos. Então, construímos $\{\mathbf{y}_k\} = \{\tilde{\mathbb{P}}_k(\Lambda) - \mathbf{x}_k\}$. Mas $\mathbf{y}_0 = 0$ e, usando o mesmo argumento anterior, obtemos $\mathbf{y}_k = 0$, para $0 \leq k \leq n$ e, em particular, $\mathbf{y}_n = \tilde{\mathbb{P}}_n(\Lambda) = 0$, o que prova o teorema. \square

Deste teorema seguem vários resultados interessantes.

Corolário 5.1.6. *Todos os zeros comuns de $\tilde{\mathbb{P}}_n$ são pontos pertencentes a \mathbb{R}^d .*

Dem.: Pela Proposição 2.1.5, como $J_{n,i}$ são matrizes simétricas e os autovalores de matrizes simétricas são reais, portanto os zeros comuns são de \mathbb{R}^d . \square

Corolário 5.1.7. *Os polinômios ortogonais \mathbb{P}_n têm no máximo $N = \dim \Pi_{n-1}^d$ zeros comuns.*

Dem.: Como $J_{n,i}$ é uma matriz quadrada de ordem $N = \dim \Pi_{n-1}^d$, temos no máximo esses números de autovetores. \square

Vemos que estes corolários generalizam o caso de polinômios ortogonais em uma variável, pois o n -ésimo polinômio ortogonal p_n tem $n = \dim \Pi_{n-1}^1$ zeros distintos. Veremos alguns resultados que caracterizam quando \mathbb{P}_n tem o número máximo de zeros comuns.

Lema 5.1.8. *Sejam $J_{n,i}$ as matrizes de Jacobi em bloco truncadas. Então, $J_{n,1}, \dots, J_{n,d}$ têm $N \dim \Pi_{n-1}^d$ autovetores linearmente independentes distintos se, e somente se, $J_{n,1}, \dots, J_{n,d}$ podem ser simultaneamente diagonalizáveis por uma matriz inversível.*

Dem.: Mostremos a condição necessária. Se $J_{n,1}, \dots, J_{n,d}$ têm N autovetores linearmente independentes, $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(N)}$, então a matriz S , cujas colunas são esses autovetores, é

não-singular. Então,

$$\begin{aligned} S^{-1}J_{n,i}S &= S^{-1} \begin{bmatrix} J_{n,i}\mathbf{x}^{(1)} & J_{n,i}\mathbf{x}^{(2)} & \cdots & J_{n,i}\mathbf{x}^{(N)} \end{bmatrix} \\ &= S^{-1} \begin{bmatrix} \lambda_1^{(i)}\mathbf{x}^{(1)} & \lambda_2^{(i)}\mathbf{x}^{(2)} & \cdots & \lambda_N^{(i)}\mathbf{x}^{(N)} \end{bmatrix} \\ &= S^{-1} \begin{bmatrix} \mathbf{x}^{(1)} & \mathbf{x}^{(2)} & \cdots & \mathbf{x}^{(N)} \end{bmatrix} \Lambda_i = S^{-1}S\Lambda_i = \Lambda_i, \end{aligned}$$

onde $\Lambda_i = \begin{bmatrix} \lambda_1^{(i)} & & 0 \\ & \ddots & \\ 0 & & \lambda_N^{(i)} \end{bmatrix}$ e $\lambda_j^{(i)}$ são os autovalores de $J_{n,i}$.

Agora, demonstremos a condição suficiente. Suponha que exista uma matriz S tal que $S^{-1}J_{n,i}S = \Lambda_i$ para $1 \leq i \leq d$ e Λ_i é diagonal. Então, $J_{n,i}S = S\Lambda_i$. Daí, temos que a j -ésima coluna de $J_{n,i}S$, isto é, $J_{n,i}$ vezes a j -ésima coluna de S , é igual à j -ésima coluna de $S\Lambda_i$, ou seja, a j -ésima diagonal de Λ_i vezes a j -ésima coluna de S . Logo, a j -ésima coluna de S é um autovetor de $J_{n,i}$ associado com a j -ésima diagonal de Λ_i . Como S é não-singular, há N autovetores linearmente independentes. \square

A partir desses lemas temos o teorema que caracteriza o número máximo de zeros comuns de \mathbb{P}_n .

Teorema 5.1.9. *O polinômio ortogonal \mathbb{P}_n tem $N = \dim \Pi_{n-1}^d$ zeros comuns reais e distintos se, e somente se,*

$$A_{n-1,i}A_{n-1,j}^T = A_{n-1,j}A_{n-1,i}^T, \quad 1 \leq i, j \leq d. \quad (5.1.2)$$

Dem.: Segue, do Teorema 5.1.5, que \mathbb{P}_n tem $N = \dim \Pi_{n-1}^d$ zeros distintos se, e somente se, $J_{n,1}, \dots, J_{n,d}$ tem $N = \dim \Pi_{n-1}^d$ autovalores junção distintos, o que é equivalente a dizer que $J_{n,1}, \dots, J_{n,d}$ têm N autovetores linearmente independentes, uma vez que autovetores associados a autovalores distintos são ortogonais. Assim, pelo Lema 5.1.8, temos que $J_{n,1}, \dots, J_{n,d}$ são simultaneamente diagonalizadas por uma matriz inversível. Porém, pelo Lema 2.1.8, esta afirmação é equivalente a equação,

$$J_{n,i}J_{n,j} = J_{n,j}J_{n,i},$$

para $1 \leq i, j \leq d$. Logo, pela definição de $J_{n,i}$ e pelas condições de comutatividade (4.2.1), a equação acima é equivalente à condição

$$A_{n-2,i}^T A_{n-2,j} + B_{n-1,i} B_{n-1,j} = A_{n-2,j}^T A_{n-2,i} + B_{n-1,j} B_{n-1,i}.$$

Da terceira equação de (4.2.1), chegamos ao resultado desejado. \square

A condição (5.1.2) é trivial para polinômios ortogonais em uma variável, pois as matrizes coeficientes, $A_i = (a_i)$, da relação de três termos para uma variável são escalares. Mas, para várias variáveis, esta condição não é usualmente satisfeita, pois a multiplicação matricial é não comutativa. Portanto, \mathbb{P}_n usualmente não tem $N = \dim \Pi_{n-1}^d$ zeros comuns. O próximo teorema, dá uma condição necessária para \mathbb{P}_n ter o número máximo de zeros comuns. Para isso, usaremos a notação Ξ_N da Definição 4.3.1.

Teorema 5.1.10. *Se os polinômios ortogonais \mathbb{P}_n têm $N = \dim \Pi_{n-1}^d$ zeros comuns reais e distintos, então*

$$\text{posto} [\Xi_{B_n}^T B_n] = \text{posto} [\Xi_{A_{n-1}}^T A_{n-1}] = \binom{n+d-2}{n} + \binom{n+d-3}{n-1}.$$

Em particular, para $d = 2$ a condição torna-se

$$\text{posto} [B_{n,1}B_{n,2} - B_{n,2}B_{n,1}] = \text{posto} [A_{n-1,1}^T A_{n-1,2} - A_{n-1,2}^T A_{n-1,1}] = 2.$$

Dem.: Se \mathbb{P}_n tem $N = \dim \Pi_{n-1}^d$ zeros comuns reais e distintos, então, pelo Teorema 5.1.9, as matrizes $A_{n-1,i}$ satisfazem a equação (5.1.2). Segue que a condição de comutatividade (4.2.1) se reduz a

$$B_{n,i}B_{n,j} - B_{n,j}B_{n,i} = A_{n-1,j}^T A_{n-1,i} - A_{n-1,i}^T A_{n-1,j},$$

que, usando o fato de $B_{n,i}$ ser matriz simétrica, podemos escrever como $\Xi_{B_n}^T B_n = \Xi_{A_{n-1}}^T A_{n-1}$. Usando a condição de posto (3.4.8), a Proposição 4.3.6 e a desigualdade de posto da Proposição 2.1.12,

$$\begin{aligned} \text{posto} \Xi_{B_n}^T B_n &= \text{posto} \Xi_{A_{n-1}}^T A_{n-1} \geq \text{posto} A_{n-1} + \text{posto} \Xi_{A_{n-1}} - dr_{n-1}^d \\ &= \binom{r_n^d}{n} + (dr_{n-1}^d - r_{n-2}^d) - dr_{n-1}^d = r_n^d - r_{n-2}^d. \end{aligned}$$

Para demonstrar o resultado desejado, mostremos a desigualdade contrária. Usando a notação Ξ_N , a primeira equação em (4.2.1) pode ser escrito como $\Xi_{A_n}^T A_n = 0$. Pela Proposição 4.3.6 e a condição de posto (3.4.8), segue que as colunas de A_n formam uma base do espaço nulo de $\Xi_{A_n}^T$. Por outro lado, a condição (5.1.2) pode ser escrito como $\Xi_{A_{n-1}}^T (A_{n-1,1} | \cdots | A_{n-1,d})^T = 0$, o que mostra que as colunas da matriz $(A_{n-1,1} | \cdots | A_{n-1,d})^T$ pertencem ao espaço nulo de $\Xi_{A_{n-1}}^T$. Portanto, existe uma matriz $S_n : r_{n+1}^d \times r_{n-1}^d$ tal que $(A_{n-1,1} | \cdots | A_{n-1,d})^T = A_n S_n$. Daí, com a condição de posto (3.4.7) e a Proposição 2.1.11, segue que

$$\begin{aligned} r_{n-1}^d &= \text{posto} (A_{n-1,1} | \cdots | A_{n-1,d})^T = \text{posto} A_n S_n \\ &= \min\{\text{posto} A_n, \text{posto} S_n\} \leq \text{posto} S_n \leq r_{n-1}^d. \end{aligned}$$

Logo, posto $S_n = r_{n-1}^d$. Usando a condição (5.1.2), mas substituindo $n-1$ por n , obtemos

$$\Xi_{B_{n+1}}^T B_{n+1} S_n = -\Xi_{A_n}^T A_n S_n = -\Xi_{A_n}^T (A_{n-1,1} | \cdots | A_{n-1,d})^T = 0,$$

onde a última igualdade segue da primeira equação de (4.2.1). Portanto,

$$\begin{aligned} \text{posto} [\Xi_{B_{n+1}}^T B_{n+1}] &= r_{n+1}^d - \dim(\text{núcleo} \Xi_{B_{n+1}}^T B_{n+1}) \\ &\leq r_{n+1}^d - \text{posto} S_n = r_{n+1}^d - r_{n-1}^d. \end{aligned}$$

Logo, posto $[\Xi_{B_n}^T B_n] = r_n^d - r_{n-2}^d$. Provemos, por fim, que

$$r_n^d - r_{n-2}^d = \binom{n+d-2}{n} + \binom{n+d-3}{n-1}.$$

De fato,

$$\begin{aligned} r_n^d - r_{n-2}^d &= \binom{n+d-1}{n} - \binom{n+d-3}{n-2} \\ &= \frac{(n+d-1)!}{n!(d-1)!} - \frac{(n+d-3)!}{(n-2)!(d-1)!} \\ &= \frac{(n+d-3)!}{n!(d-1)!} ((n+d-1)(n+d-2) - n(n-1)) \\ &= \frac{(n+d-3)!}{n!(d-1)!} (n^2 + 2nd - 2n + d^2 - 2d - n - d + 2 - n^2 + n) \\ &= \frac{(n+d-3)!}{n!(d-1)!} (2nd - 2n - 2d + d^2 - d + 2) \\ &= \frac{(n+d-3)!}{n!(d-1)!} (nd + d^2 - 2d - n - d + 2 + dn - n) \\ &= \frac{(n+d-3)!}{n!(d-1)!} ((n+d-2)(d-1) + (d-1)n) \\ &= \frac{(n+d-2)!}{n!(d-2)!} - \frac{(n+d-3)!}{(n-1)!(d-2)!} \\ &= \binom{n+d-2}{n} + \binom{n+d-3}{n-1} \end{aligned}$$

□

Se \mathcal{L} é um funcional linear quase-centralmente simétrico, então temos que $\text{posto}[\Xi_{B_n}^T B_n] = 0$. De fato, segue do Lema 4.3.2 que

$$\Xi_{B_n}^T B_n = 0 \Leftrightarrow B_{n,i} B_{n,j} = B_{n,j} B_{n,i},$$

onde usamos o fato de $B_{n,i}$ ser simétrica. Portanto, um corolário imediato do resultado acima é o resultado a seguir.

Corolário 5.1.11. *Se \mathcal{L} é um funcional linear quase-centralmente simétrico, então \mathbb{P}_n não possui $N = \dim \Pi_{n-1}^d$ zeros comuns distintos.*

As implicações desses resultados serão discutidas a seguir, quando trataremos de uma fórmula análoga à quadratura gaussiana para polinômios ortogonais em uma variável.

5.2 Fórmulas de cubatura gaussianas

Os polinômios ortogonais de uma variável são muito importantes para a teoria das fórmulas da quadratura gaussianas.

Seja ω uma função peso definida em \mathbb{R} . Uma fórmula de quadratura com respeito a ω é uma soma ponderada de um número finito de valores da função que fornecem uma aproximação para a integral definida $\int_a^b f(x)\omega(x)dx$ e o valor exato da integral para polinômios de um certo grau, que são chamados de grau de precisão da fórmula. As fórmulas de quadratura com grau máximo de precisão são chamadas fórmulas de quadratura gaussianas e têm grau $2n - 1$ de precisão para fórmulas que usam n nós. Além disso, fórmulas de quadratura gaussianas existem se, e somente se, os nós são zeros de um polinômio ortogonal de grau n com respeito a ω . Como tal polinômio sempre tem n zeros distintos, a fórmula de quadratura Gaussiana existe para cada função peso adequada.

Desse modo, estudaremos um análogo à fórmula de quadratura gaussiana em várias variáveis, denominadas de fórmulas de cubatura gaussianas. A existência de tais fórmulas dependem dos zeros comuns de polinômios ortogonais em várias variáveis.

5.2.1 Caracterização das fórmulas de cubatura gaussianas

Iniciaremos com a definição de fórmulas de cubatura de grau de precisão $2n - 1$.

Definição 5.2.1. *Seja \mathcal{L} um funcional linear definido positivo. Uma fórmula de cubatura de grau de precisão $2n - 1$ com respeito a um funcional linear \mathcal{L} é*

$$I_n(f) = \sum_{k=1}^N \lambda_k f(\mathbf{x}_k), \quad \lambda_k \in \mathbb{R}, \quad \mathbf{x}_k \in \mathbb{R}^d, \quad (5.2.1)$$

tal que $\mathcal{L}(f) = I_n(f)$ para todo $f \in \Pi_{2n-1}^d$ e $\mathcal{L}(f^) \neq I_n(f^*)$ para pelo menos um $f^* \in \mathbb{R}_{2n}^d$. Os pontos \mathbf{x}_k são chamados de nós e os números λ_k são chamados de pesos da fórmula de cubatura. Tal fórmula é chamada de mínima, se N , o número de nós, é mínimo. Se uma fórmula de cubatura tem todos os pesos positivos, é chamada de cubatura positiva.*

Estamos interessados na fórmula de cubatura mínima. Para identificar uma fórmula de cubatura mínima, é necessário saber o número mínimo de nós. O próximo teorema nos

fornece o menor limitante para o número de nós da fórmula de cubatura e fórmulas que atingem o menor limitante são mínimas.

Teorema 5.2.2. *O número de nós na fórmula de cubatura de grau de precisão $2n - 1$ satisfaz $M \geq \dim \Pi_{n-1}^d$.*

Dem.: Seja $N = \dim \Pi_{n-1}^d$. Se uma fórmula de cubatura I_n com grau de precisão $2n - 1$ tem menos que $N = \dim \Pi_{n-1}^d$ nós, então o sistema de equações lineares $P(\mathbf{x}_k) = 0$, onde \mathbf{x}_k são os nós da fórmula de cubatura e $P \in \Pi_{n-1}^d$, tem $N = \dim \Pi_{n-1}^d$ coeficientes de P desconhecidos, mas menos de $N = \dim \Pi_{n-1}^d$ equações. Portanto, haverá pelo menos um polinômio diferente do polinômio nulo $P \in \Pi_{n-1}^d$ que anula todos os nós. Como a fórmula tem de grau precisão $2n - 1$, segue que $\mathcal{L}(P^2) = I_n(P^2) = 0$, o que contradiz o fato de \mathcal{L} ser definido positivo. \square

No caso de $d = 1$, o limitante inferior do teorema anterior é atingido pela fórmula de quadratura gaussiana. Analogamente, definimos.

Definição 5.2.3. *Uma fórmula de cubatura de grau de precisão $2n - 1$ com $\dim \Pi_{n-1}^d$ nós é chamada de cubatura gaussiana.*

O resultado principal desta seção é a caracterização da fórmula de cubatura gaussiana. Iniciaremos com um lema básico de Mysovskikh [14] que detém a chave para a prova do teorema principal desta seção e que é importante por si só. Precisaremos de uma definição para enunciarmos o lema.

Definição 5.2.4. *Uma hipersuperfície algébrica de grau m em \mathbb{R}^d é o conjunto de zeros de um polinômio de grau m de d variáveis.*

Lema 5.2.5. *Os pares de pontos distintos \mathbf{x}_k , $1 \leq k \leq N = \dim \Pi_{n-1}^d$, em \mathbb{R}^d não pertencem a uma hipersuperfície de grau $n - 1$ se, e somente se, existem r_n^d polinômios linearmente independentes de grau exatamente n , para os quais esses pontos são os zeros comuns.*

Dem.: O significado dos \mathbf{x}_k não pertencerem a uma hipersuperfície algébrica de grau $n - 1$ é que a matriz, $N \times N$, (\mathbf{x}_k^β) é não singular, onde as colunas são arranjadas de acordo com a ordem lexicográfica em $\{\beta : 0 \leq |\beta| \leq n - 1\}$ e as linhas correspondem a $1 \leq k \leq N$. Portanto, podemos construir r_n^d polinômios

$$Q_\alpha^n = \mathbf{x}^\alpha - R_\alpha, \quad |\alpha| = n, \quad R_\alpha(\mathbf{x}) = \sum_{|\beta| \leq n-1} C_{\alpha,\beta} \mathbf{x}^\beta \in \Pi_{n-1}^d, \quad (5.2.2)$$

com R_α satisfazendo

$$R_\alpha(\mathbf{x}_k) = \mathbf{x}_k^\alpha, \quad 1 \leq k \leq N,$$

para cada α fixado. Claramente, o conjunto dos polinômios Q_α^n é linearmente independente e a condição $Q_\alpha^n(\mathbf{x}_k) = 0$ é satisfeita.

Por outro lado, suponha que há r_n^d polinômios linearmente independentes Q_α^n de grau n que têm \mathbf{x}_k como zeros comuns. devido aos polinômios serem linearmente independentes, podemos assumir que são da forma (5.2.2). Se os pontos \mathbf{x}_k pertencem a uma hipersuperfície algébrica de grau $n - 1$, então a matriz $X_n = (\mathbf{x}_n^\beta)$, com $1 \leq k \leq N$ e $|\beta| \leq n - 1$, é singular. Suponha que posto $X_n = M - 1$, com $M \leq N$. Assuma que as primeiras $M - 1$ linhas desta matriz são linearmente independente. Para cada k definimos um vetor $U_k = (\mathbf{x}_n^\beta)_{|\beta| \leq n-1} \in \mathbb{R}^N$. Então, U_1, \dots, U_{M-1} são linearmente independentes e existem escalares c_1, \dots, c_{M-1} , não todos nulos, tais que $\sum_{k=1}^M c_k U_k = 0$. Além disso, considerando a primeira componente desta equação vetorial, que é formada apenas por elementos 1, segue que pelo menos dois coeficientes c_k são não-nulos. Assuma, então, que $c_1 \neq 0$ e $c_M \neq 0$. Por (5.2.2), as equações $Q_\alpha^n(\mathbf{x}_k) = 0$ podem ser escritas como

$$\mathbf{x}_k^\alpha - \sum_{|\beta| \leq n-1} C_{\alpha,\beta} \mathbf{x}_k^\beta = 0, \quad |\alpha| = n, \quad 1 \leq k \leq N. \quad (5.2.3)$$

Como o termo da soma pode ser escrito como $C_\alpha^T U_k$, com C_α um vetor de \mathbb{R}^N , (5.2.3) pode ser reescrita como

$$\mathbf{x}_k^\alpha - C_\alpha^T U_k = 0.$$

Multiplicando por c_k e somando de 1 a M , obtemos

$$\sum_{k=1}^M (c_k \mathbf{x}_k^\alpha - c_k C_\alpha^T U_k) = 0 \Rightarrow \sum_{k=1}^M c_k \mathbf{x}_k^\alpha = C_\alpha^T \sum_{k=1}^M c_k U_k.$$

Segue, de $\sum_{k=1}^M c_k U_k = 0$, que $\sum_{k=1}^M c_k \mathbf{x}_k^\alpha = 0$, $|\alpha| = n$. Portanto, $\sum_{k=1}^M c_k \mathbf{x}_{k,i} U_k = 0$, $1 \leq i \leq d$, onde usamos a notação $\mathbf{x}_k = (x_{k,1}, \dots, x_{k,d})$. Agora, multiplicando a equação $\sum_{k=1}^M c_k U_k = 0$ por $x_{M,i}$ e, em seguida, subtraindo da equação $\sum_{k=1}^M c_k x_{k,i} U_k = 0$, obtemos

$$\sum_{k=1}^{M-1} c_k (x_{k,i} - x_{M,i}) U_k = 0, \quad 1 \leq i \leq d.$$

Como U_1, \dots, U_{M-1} são linearmente independentes, segue que $c_k (x_{k,i} - x_{M,i}) = 0$, $1 \leq k \leq M - 1$, $1 \leq i \leq d$. Como $c_1 \neq 0$, $x_{1,i} = x_{M,i}$ para $1 \leq i \leq d$. Portanto, $\mathbf{x}_1 = \mathbf{x}_m$, uma contradição à hipótese de que os pontos são distintos. \square

A existência da cubatura gaussiana de grau de precisão $2n - 1$ é caracterizada pelos zeros comuns dos polinômios ortogonais. O teorema seguinte foi primeiramente provado por Mysovskikh [13], mas aqui apresentamos a demonstração dada em Dunkly e Xu [5].

Teorema 5.2.6. *Seja \mathcal{L} um funcional linear definido positivo e seja \mathbb{P}_n os polinômios ortogonais correspondentes ao funcional \mathcal{L} . Então, \mathcal{L} admite uma fórmula de cubatura gaussiana de grau de precisão $2n - 1$ se, e somente se, \mathbb{P}_n possui $N = \dim \Pi_{n-1}^d$ zeros comuns.*

Dem.: Suponha que \mathbb{P}_n possui N zeros comuns. Pelo Teorema 5.1.3 e Corolário 5.1.6, todos os zeros comuns são reais e distintos. Denotando esses zeros comuns por $\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N$, da fórmula de Christoffel-Darboux (4.5.5) e da fórmula para $\mathbf{K}_n(\mathbf{x}, \mathbf{x})$, segue que o polinômio $l_k(\mathbf{x}) = \mathbf{K}_n(\mathbf{x}, \mathbf{z}_k) / \mathbf{K}_n(\mathbf{z}_k, \mathbf{z}_k)$ satisfaz a condição $l_k(\mathbf{z}_j) = \delta_{k,j}$, $1 \leq k, j \leq N$. De fato, denotando $\partial / \partial x_i$ por ∂_i ,

$$\begin{aligned} l_k(\mathbf{x}) &= \frac{\mathbf{K}_n(\mathbf{x}, \mathbf{z}_k)}{\mathbf{K}_n(\mathbf{z}_k, \mathbf{z}_k)} \\ &= \frac{[A_{n,i} \mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1} \mathbb{P}_n(\mathbf{z}_k) - \mathbb{P}_n^T(\mathbf{x}) H_n^{-1} [A_{n,i} \mathbb{P}_{n+1}(\mathbf{z}_k)]}{(\mathbf{x} - \mathbf{z}_k) [\mathbb{P}_n^T(\mathbf{z}_k) H_n^{-1} [A_{n,i} \partial_i \mathbb{P}_{n+1}(\mathbf{z}_k)] - [A_{n,i} \mathbb{P}_{n+1}(\mathbf{z}_k)]^T H_n^{-1} \partial_i \mathbb{P}_n(\mathbf{z}_k)]} \\ &= \frac{\mathbb{P}_n^T(\mathbf{x}) H_n^{-1} [A_{n,i} \mathbb{P}_{n+1}(\mathbf{z}_k)] - [A_{n,i} \mathbb{P}_{n+1}(\mathbf{x})]^T H_n^{-1} \mathbb{P}_n(\mathbf{z}_k)}{(\mathbf{x} - \mathbf{z}_k) [[A_{n,i} \mathbb{P}_{n+1}(\mathbf{z}_k)]^T H_n^{-1} \partial_i \mathbb{P}_n(\mathbf{z}_k)]}. \end{aligned}$$

Assim, $l_k(\mathbf{z}_i) = 0$ para $i \neq k$ e fazendo $\mathbf{x} \rightarrow \mathbf{z}_k$ temos $l_k(\mathbf{z}_k) = 1$. Como os polinômios l_k 's são de grau $n - 1$, eles servem como polinômios de interpolação fundamentais. Segue, então, que

$$L_n(f; \mathbf{x}) = \sum_{k=1}^N f(\mathbf{z}_k) \frac{\mathbf{K}_n(\mathbf{x}, \mathbf{z}_k)}{\mathbf{K}_n(\mathbf{z}_k, \mathbf{z}_k)}, \quad L_n(f) \in \Pi_{n-1}^d, \quad (5.2.4)$$

é uma interpolação de Lagrange baseada em \mathbf{z}_k , isto é, $L_n(f; \mathbf{z}_k) = f(\mathbf{z}_k)$, $1 \leq k \leq N$.

Note que $L_n(f)$ é a única fórmula de interpolação de Lagrange, pois supondo que exista outra fórmula de interpolação de Lagrange, $\tilde{L}_n(f; \mathbf{x}) \in \Pi_{n-1}^d$, tal que $\tilde{L}_n(f; \mathbf{z}_k) = f(\mathbf{z}_k)$, temos que $L_n(f; \mathbf{x}) - \tilde{L}_n(f; \mathbf{x}) \in \Pi_{n-1}^d$ e $L_n(f; \mathbf{z}_k) - \tilde{L}_n(f; \mathbf{z}_k) = 0$. Disso, temos que os \mathbf{z}_k , para $1 \leq k \leq N$, pertencem a uma hipersuperfície algébrica de grau $n - 1$, o que, pelo Lema 5.2.5, contradiz o fato de que \mathbb{P}_n tem N zeros comuns. Então, pelo Teorema 4.5.3, obtemos a fórmula

$$\begin{aligned} \mathcal{L}[L_n(f)] &= \mathcal{L} \left(\sum_{k=1}^N f(\mathbf{z}_k) \frac{\mathbf{K}_n(\mathbf{x}, \mathbf{z}_k)}{\mathbf{K}_n(\mathbf{z}_k, \mathbf{z}_k)} \right) \\ &= \sum_{k=1}^N \Lambda_k f(\mathbf{z}_k) \mathcal{L}(\mathbf{K}_n(\mathbf{x}, \mathbf{z}_k)) = \sum_{k=1}^N \Lambda_k f(\mathbf{z}_k), \end{aligned} \quad (5.2.5)$$

onde $\Lambda_k = [\mathbf{K}_n(\mathbf{z}_k, \mathbf{z}_k)]$, que é uma fórmula de cubatura para \mathcal{L} , exata para todos os polinômios em Π_{n-1}^d . Contudo, como \mathbb{P}_n é ortogonal a Π_{n-1}^d e os nós são zeros de \mathbb{P}_n , segue que, para qualquer vetor $\mathbf{a} \in \mathbb{R}^{r_n}$,

$$\mathcal{L}(\mathbf{a}^T \mathbb{P}_n) = 0 = \sum_{k=1}^N \Lambda_k \mathbf{a}^T \mathbb{P}_n(\mathbf{z}_k).$$

Como qualquer polinômio $P \in \Pi_n^d$ pode ser escrito como $P = \mathbf{a}^T \mathbb{P}_n + R$ para algum \mathbf{a} com $R \in \Pi_{n-1}^d$, então

$$\mathcal{L}(P) = \mathcal{L}(R) = \sum_{k=1}^N \Lambda_k R(\mathbf{z}_k) = \sum_{k=1}^N \Lambda_k P(\mathbf{z}_k), \quad \forall P \in \Pi_n^d. \quad (5.2.6)$$

Então, a fórmula de cubatura (5.2.5) é exata para todos os polinômios em Π_n^d . Provemos que é exata para todo polinômio pertencente a Π_{n+1}^d . Sejam $\mathbf{z}_k = (z_1^{(k)}, z_2^{(k)}, \dots, z_d^{(k)})$, $1 \leq k \leq N$. Pela ortogonalidade e o fato de os \mathbf{z}_k serem os zeros de \mathbb{P}_n , obtemos

$$\mathcal{L}(x_i \mathbf{a}^T \mathbb{P}_n) = 0 = \sum_{k=1}^N \Lambda_k z_i^{(k)} \mathbf{a}^T \mathbb{P}_n(\mathbf{z}_k), \quad 1 \leq i \leq d.$$

Pela relação de recorrência de três termos (3.4.10) e como que a fórmula de cubatura vale para Π_n^d , segue que

$$\begin{aligned} \mathcal{L}(\mathbf{a}^T \mathbb{P}_{n+1}) &= \mathcal{L}\left(\mathbf{a}^T \left(\sum_{i=1}^d x_i D_{n,i}^T \mathbb{P}_n + E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1}\right)\right) \\ &= \sum_{i=1}^d \mathcal{L}(x_i \mathbf{a}^T D_{n,i}^T \mathbb{P}_n) + \mathcal{L}(\mathbf{a}^T E_n \mathbb{P}_n) + \mathcal{L}(\mathbf{a}^T F_n \mathbb{P}_{n-1}) \\ &= \sum_{i=1}^d \sum_{k=1}^N \Lambda_n z_i^{(k)} \mathbf{a}^T D_{n,i}^T \mathbb{P}_n(\mathbf{z}_k) + \sum_{k=1}^N \Lambda_n \mathbf{a}^T E_n \mathbb{P}_n(\mathbf{z}_k) \\ &\quad + \sum_{k=1}^N \Lambda_n \mathbf{a}^T F_n \mathbb{P}_{n-1}(\mathbf{z}_k) \\ &= \sum_{k=1}^N \Lambda_k \mathbf{a}^T \left(\sum_{i=1}^d z_i^{(k)} D_{n,i}^T \mathbb{P}_n(\mathbf{z}_k) + E_n \mathbb{P}_n(\mathbf{z}_k) + F_n \mathbb{P}_{n-1}(\mathbf{z}_k)\right) \\ &= \sum_{k=1}^N \Lambda_k \mathbf{a}^T \mathbb{P}_{n+1}(\mathbf{z}_k). \end{aligned}$$

Portanto, pelo mesmo argumento usado para mostrar (5.2.6), a fórmula de cubatura é exata para polinômios em Π_{n+1}^d . Claramente, podemos repetir o processo. Do fato, como $\mathbb{P}_n(\mathbf{z}_k) = 0$, da propriedade de ortogonalidade obtemos

$$\mathcal{L}(\mathbf{x}^\alpha \mathbf{a}^T \mathbb{P}_n) = 0 = \sum_{k=1}^N \Lambda_k \mathbf{z}_k^\alpha \mathbf{a}^T \mathbb{P}_n(\mathbf{z}_k),$$

onde $|\alpha| = 2$. Logo, como a fórmula de cubatura vale para Π_{n+1}^d , pela relação de recorrência (3.4.10) segue que

$$\begin{aligned}
\mathcal{L}(\mathbf{a}^T \mathbb{P}_{n+2}) &= \mathcal{L} \left(\mathbf{a}^T \left(\sum_{i=1}^d x_i D_{n+1,i}^T \mathbb{P}_{n+1} + E_{n+1} \mathbb{P}_{n+1} + F_{n+1} \mathbb{P}_n \right) \right) \\
&= \sum_{i=1}^d \mathcal{L} \left(\mathbf{a}^T D_{n+1,i}^T x_i \left(\sum_{j=1}^d x_j D_{n,j}^T \mathbb{P}_n + E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1} \right) \right) \\
&\quad + \mathcal{L}(\mathbf{a}^T (E_{n+1} \mathbb{P}_{n+1} + F_{n+1} \mathbb{P}_n)) \\
&= \sum_{i=1}^d \sum_{j=1}^d \mathcal{L}(x_i x_j \mathbf{a}^T D_{n+1,i}^T D_{n,j}^T \mathbb{P}_n) \\
&\quad + \sum_{i=1}^d \mathcal{L}(x_i \mathbf{a}^T D_{n+1,i}^T (E_n \mathbb{P}_n + F_n \mathbb{P}_{n-1})) \\
&\quad + \mathcal{L}(\mathbf{a}^T (E_{n+1} \mathbb{P}_{n+1} + F_{n+1} \mathbb{P}_n)) \\
&= \sum_{k=1}^N \Lambda_k \mathbf{a}^T \left(\sum_{i=1}^d \sum_{j=1}^d z_i^{(k)} z_j^{(k)} D_{n+1,i}^T D_{n,j}^T \mathbb{P}_n(\mathbf{z}_k) \right) \\
&\quad + \sum_{k=1}^N \Lambda_k \mathbf{a}^T \left(\sum_{i=1}^d z_i^{(k)} D_{n+1,i}^T (E_n \mathbb{P}_n(\mathbf{z}_k) + F_n \mathbb{P}_{n+1}(\mathbf{z}_k)) \right) \\
&\quad + \sum_{k=1}^N \Lambda_k \mathbf{a}^T (E_{n+1} \mathbb{P}_{n+1}(\mathbf{z}_k) + F_{n+1} \mathbb{P}_n(\mathbf{z}_k)) \\
&= \sum_{k=1}^N \Lambda_k \mathbf{a}^T \sum_{i=1}^d z_i^{(k)} D_{n+1,i}^T \mathbb{P}_{n+1}(\mathbf{z}_k) + E_{n+1} \mathbb{P}_{n+1}(\mathbf{z}_k) \\
&\quad + \sum_{k=1}^N \Lambda_k \mathbf{a}^T F_n \mathbb{P}_n(\mathbf{z}_k) \\
&= \sum_{k=1}^N \Lambda_k \mathbf{a}^T \mathbb{P}_{n+2}(\mathbf{z}_k).
\end{aligned}$$

Consequentemente, a fórmula de cubatura é válida para todo polinômio pertencente a Π_{n+2}^d . Pela ortogonalidade de \mathbb{P}_n podemos repetir o processo para \mathbf{x}^α com $|\alpha| \leq n-1$ e concluir, assim, que a fórmula de cubatura (5.2.5) vale para todo polinômio em Π_{2n-1}^d .

Por outro lado, suponha que uma fórmula de cubatura gaussiana existe e tem \mathbf{x}_k , $1 \leq k \leq N = \dim \Pi_{n-1}^d$, como nós. Como \mathcal{L} é definido positivo, pela demonstração do Lema 5.2.5, estes nós não podem todos pertencer a uma hipersuperfície algébrica de grau $n-1$. Porém, haverá um $Q \in \Pi_{n-1}^d$, não identicamente nulo, que se anula em todos os nós, o que implicará que $\mathcal{L}(Q^2) = 0$. Pelo Lema 5.2.5, existem polinômios linearmente independentes P_j^n , $1 \leq j \leq r_n^d$, que têm \mathbf{x}_k como zeros comuns. Pela fórmula de cubatura gaussiana, para qualquer $R \in \Pi_{n-1}^d$,

$$\mathcal{L}(R P_j^n) = \sum_{k=1}^N \Lambda_k (R P_j^n)(\mathbf{x}_k) = 0.$$

Então, os polinômios P_j^n são ortogonais a Π_{n-1}^d e podemos escrevê-los na forma (5.2.2). \square

Vale mencionar que a prova acima é similar à prova para a fórmula de quadratura gaussiana.

O uso do polinômio interpolador conduz ao seguinte corolário.

Corolário 5.2.7. *Uma fórmula de cubatura gaussiana que toma a forma (5.2.5), em particular, é uma cubatura positiva.*

Dem.: Seja $f(\mathbf{x}) = l_k^2(\mathbf{x}) = \mathbf{K}_n^2(\mathbf{x}, \mathbf{z}_k) / \mathbf{K}_n^2(\mathbf{z}_k, \mathbf{z}_k)$, então $f \in \Pi_{2n-2}^d$. Mas, \mathcal{L} é definido positivo, segue que

$$0 < \mathcal{L}(l_k^2) = \sum_{k=1}^N \Lambda_k l_k^2(\mathbf{z}_k) = \Lambda_k,$$

para $1 \leq k \leq N$, concluindo, assim, a prova. \square

Da prova do Teorema 5.2.6, a fórmula de cubatura gaussiana (5.2.5) é obtida através da integração do polinômio de interpolação de Lagrange $L_n f$ em (5.2.5). Este processo é similar ao caso de uma variável. Para polinômios de interpolação em uma variável, sabemos que o polinômio de interpolação de Lagrange baseado em nós distintos sempre existe e é único. O mesmo, contudo, não é verdade para polinômios de interpolação em várias variáveis. Um exemplo típico é a interpolação em seis pontos pertencentes ao círculo unitário $x^2 + y^2 = 1$, onde o polinômio $p(x, y) = x^2 + y^2 - 1$ se anula nos nós. Se P é um polinômio de interpolação de grau 2 nesses nós, então $P + \mathbf{a}p$, para qualquer número \mathbf{a} , também o é. Isto mostra que o polinômio interpolador de Lagrange, se existir, não é único. Se os nós são zeros de uma fórmula de cubatura gaussiana, então eles admitem um único polinômio de interpolação de Lagrange.

Teorema 5.2.8. *Se \mathbf{x}_k , $1 \leq k \leq \dim \Pi_{n-1}^d$, são os nós da fórmula de cubatura gaussiana, então, para quaisquer dados $\{\mathbf{y}_i\}$, há um único polinômio $P \in \mathbb{P}_{n-1}^d$ tal que $P(\mathbf{x}_i) = \mathbf{y}_i$, $1 \leq i \leq \dim \Pi_{n-1}^d$.*

Dem.: O polinômio de interpolação toma a forma de (5.2.4) com $\mathbf{y}_i = f(\mathbf{x}_i)$. Se $f(\mathbf{x}_i) = 0$, $1 \leq i \leq \dim \Pi_{n-1}^d$, então $P(\mathbf{x}_i) = 0$ e a fórmula de cubatura gaussiana mostra que $\mathcal{L}(P^2) = I_n(P^2) = 0$, o que implica que $P = 0$. \square

Teorema 5.2.9. *Seja \mathcal{L} um funcional linear definido positivo e seja $\tilde{\mathbb{P}}_n$ os polinômios ortonormais correspondentes que satisfazem a relação de três termos (3.4.3) Então, a fórmula de cubatura gaussiana existe se, e somente se, (5.1.2) vale.*

Dem.: Pelo Teorema 5.2.6, a fórmula de cubatura gaussiana vale se, e somente se, $\tilde{\mathbb{P}}_n$ tem N zeros comuns, o que, pelo Teorema 5.1.9, é equivalente a (5.1.2). \square

O próximo corolário segue imediatamente do Corolário 5.1.11, pois este afirma que (5.1.2) não vale para um funcional linear quase-centralmente simétrico.

Corolário 5.2.10. *Um funcional linear quase-centralmente simétrico não admite uma fórmula de cubatura gaussiana.*

Em particular, não há fórmulas de cubatura gaussianas para função peso clássicas simétricas e para as funções peso clássica na bola unitária, pois os funcionais lineares definidos a partir das funções peso clássicas simétricas e na bola unitária são funcionais quase-centralmente simétricos.

Capítulo 6

Exemplos

Iniciamos este capítulo com alguns exemplos bem simples de polinômios ortonormais em várias variáveis, comumente conhecidos na literatura como produto tensor. Vemos a seguir um método para gerar polinômios ortonormais em duas variáveis e exemplos gerados por esse método, além de exemplos das matrizes coeficientes da relação de três termos em duas variáveis.

O funcional linear, neste capítulo, é definido através de uma função peso $W(\mathbf{x})$, ou seja, o funcional linear é dado por $\mathcal{L}(F(\mathbf{x})G(\mathbf{x})) = \int F(\mathbf{x})G(\mathbf{x})W(\mathbf{x})d\mathbf{x}$.

6.1 Produto tensor

Seja $\alpha = (\alpha_1, \dots, \alpha_d)$ o multi-índice definido anteriormente. Definimos, a seguir, o polinômio ortonormal de grau n em d variáveis, denotado por $P_\alpha(\mathbf{x})$, através dos polinômios ortonormais em uma variável, com $|\alpha| = n$.

Sejam $\omega_{\alpha_k}(x_k)$, $1 \leq k \leq d$, funções pesos em $(a_{\alpha_k}, b_{\alpha_k}) \subset \mathbb{R}$, respectivamente, onde $\infty \leq a_{\alpha_k} < b_{\alpha_k} \leq \infty$. Sejam p_{α_k} os polinômios ortonormais em uma variável com relação a função peso $\omega_{\alpha_k}(x_k)$. Definimos o polinômio ortonormal de grau n em d variáveis como

$$P_\alpha(\mathbf{x}) = p_{\alpha_1}(x_1)p_{\alpha_2}(x_2) \dots p_{\alpha_d}(x_d),$$

que é ortonormal com relação à função peso

$$W(\mathbf{x}) = \omega_{\alpha_1}(x_1)\omega_{\alpha_2}(x_2) \dots \omega_{\alpha_d}(x_d),$$

no domínio $R = (a_{\alpha_1}, b_{\alpha_1}) \times (a_{\alpha_2}, b_{\alpha_2}) \times \dots \times (a_{\alpha_d}, b_{\alpha_d})$.

A partir desse produto tensor podemos construir diversos polinômios ortonormais em várias variáveis através dos polinômios ortonormais clássicos de Jacobi, Hermite, Laguerre,

Gegenbauer, Legendre e Chebyshev. A seguir, vemos exemplos de polinômios ortonormais em duas variáveis gerados por alguns dos polinômios clássicos. Denotamos por $P_{n-k,k}(x, y)$ cada componente do vetor polinomial coluna em duas variáveis de grau n .

Exemplo 6.1.1. Sejam o multi-índice $\alpha = (\alpha_1, \alpha_2)$, com $\alpha_1 + \alpha_2 = n$ e como $d = 2$ temos que $\alpha_1 = n - k$ e $\alpha_2 = k$. A função peso de Jacobi é dada por $\omega_{\alpha_1}(x) = (1 - x)^\alpha(1 + x)^\beta$ e a de Hermite por $\omega_{\alpha_2}(y) = e^{-y^2}$, com $\alpha, \beta > -1$. Portanto, a função peso do polinômio ortonormal em duas variáveis é

$$W(x, y) = (1 - x)^\alpha(1 + x)^\beta e^{-y^2},$$

no domínio $R = (-1, 1) \times (-\infty, \infty)$ e os polinômios ortonormais são dados em termos dos polinômios ortonormais clássicos de Jacobi e Hermite, da forma

$$P_{n-k,k}(x, y) = \tilde{P}_{n-k}^{(\alpha,\beta)}(x) \tilde{H}_k(y).$$

Exemplo 6.1.2. Um outro exemplo é o caso dos polinômios ortonormais em duas variáveis formados pelos polinômios ortonormais clássicos de Chebyshev de 1ª espécie e Laguerre,

$$P_{n-k,k}(x, y) = \tilde{T}_{n-k}(x) \tilde{L}_k^{(\alpha)}(y),$$

cuja função peso é dada por

$$W(x, y) = \frac{y^\alpha e^{-y}}{\sqrt{1 - x^2}}$$

no domínio $R = (-1, 1) \times (0, \infty)$, com $\alpha > -1$.

Exemplo 6.1.3. Os polinômios ortonormais em duas variáveis deste exemplo são gerados através dos polinômios ortonormais clássicos de Gegenbauer e Legendre. Assim sua função peso é da forma,

$$W(x, y) = (1 - x^2)^{\lambda-1/2},$$

no domínio $R = (-1, 1) \times (-1, 1)$ e os polinômios ortonormais em duas variáveis são da forma

$$P_{n-k,k}(x, y) = \tilde{G}_{n-k}^{(\lambda)}(x) \tilde{P}_k^{(0)}(y).$$

6.2 Método de Koornwinder

Este método foi dado por Koornwinder [10] e gera polinômios ortonormais de duas variáveis através de polinômios ortonormais de uma variável. Novamente denotamos por $P_{n-k,k}(x, y)$ cada componente do vetor polinomial coluna em duas variáveis de grau n .

Sejam $\omega_1(x)$ uma função peso em (a, b) e $\omega_2(y)$ uma função peso em (c, d) . Seja $\rho(x)$ uma função positiva em (a, b) tal que ρ é um polinômio de grau 1 ou ρ é raiz quadrada de um polinômio não-negativo de grau no máximo 2, e, nesse caso, assumamos que $c = -d > 0$ e $\omega_2(y)$ é uma função par em $(-c, c)$.

Para cada $k \in \mathbb{N}_0$, seja $\{\tilde{p}_n^{(k)}\}_{n=0}^{\infty}$ a sequência de polinômios ortonormais com respeito à função peso $\rho^{2k+1}(x)\omega_1(x)$ e seja $\{\tilde{q}_n\}$ a sequência de polinômios ortonormais com respeito à função peso ω_2 . Então, definimos o polinômio $P_{n-k,k}(x, y)$ em duas variáveis por

$$P_{n-k,k}(x, y) = \tilde{p}_{n-k}^{(k)}(x) (\rho(x))^k \tilde{q}_k\left(\frac{y}{\rho(x)}\right), \quad 0 \leq k \leq n. \quad (6.2.1)$$

Os polinômios $P_{n-k,k}(x, y)$ são de grau n . De fato, para o caso de $\rho(x)$ ser um polinômio de grau 1, claramente o polinômio $P_{n-k,k}(x, y)$ é de grau n e, para o caso de $\rho(x)$ ser raiz quadrada de um polinômio não-negativo de grau no máximo 2, basta notar que $\tilde{q}_k(y)$ tem a mesma paridade de k , uma vez que $\omega_2(y)$ é par.

Os polinômios $P_{n-k,k}(x, y)$ são ortonormais com respeito à função peso

$$W(x, y) = \omega_1(x)\omega_2(\rho^{-1}(x)y), \quad (x, y) \in R, \quad (6.2.2)$$

onde o domínio R é definido por

$$R = \{(x, y) : a < x < b, \quad c\rho(x) < y < d\rho(x)\}. \quad (6.2.3)$$

De fato, fazendo a mudança de variáveis $z = \rho^{-1}(x)y$ na integral, temos

$$\begin{aligned} & \int \int_R P_{n-k,k}(x, y) P_{m-j,j}(x, y) W(x, y) dx dy \\ &= \int_a^b \int_{c\rho(x)}^{d\rho(x)} \tilde{p}_{n-k}^{(k)}(x) \tilde{p}_{m-j}^{(j)}(x) \tilde{q}_k\left(\frac{y}{\rho(x)}\right) \tilde{q}_j\left(\frac{y}{\rho(x)}\right) (\rho(x))^{k+j} \omega_1(x) \omega_2\left(\frac{y}{\rho(x)}\right) dx dy \\ &= \int_a^b \tilde{p}_{n-k}^{(k)}(x) \tilde{p}_{m-j}^{(j)}(x) (\rho(x))^{k+j+1} \omega_1(x) dx \int_c^d \tilde{q}_k(z) \tilde{q}_j(z) \omega_2(z) dz = \delta_{n,m} \delta_{k,j}. \end{aligned}$$

Vejam alguns exemplos de polinômios ortonormais em duas variáveis gerados por esse método.

Exemplo 6.2.1. (*Polinômios de Jacobi no quadrado*) Sejam $\omega_1(x) = (1-x)^{\alpha_1}(1+x)^{\beta_1}$ e $\omega_2(y) = (1-y)^{\alpha_2}(1+y)^{\beta_2}$ em $(-1, 1)$, e $\rho(x) = 1$. Portanto a função peso (6.2.2) é dada por

$$W(x, y) = (1-x)^{\alpha_1}(1+x)^{\beta_1}(1-y)^{\alpha_2}(1+y)^{\beta_2} \quad (6.2.4)$$

no domínio $R = (-1, 1) \times (-1, 1)$. Os polinômios $\tilde{p}_{n-k}^{(k)}(x) = \tilde{P}_{n-k}^{(\alpha_1, \beta_1)}(x)$ e os polinômios $\tilde{q}_k(y) = \tilde{P}_k^{(\alpha_2, \beta_2)}(y)$, onde $\tilde{P}_j^{(\alpha_i, \beta_i)}$ são os polinômios ortonormais de Jacobi, então temos que

$$P_{n-k, k}(x, y) = \tilde{P}_{n-k}^{(\alpha_1, \beta_1)}(x)\tilde{P}_k^{(\alpha_2, \beta_2)}(y) \quad (6.2.5)$$

são polinômios ortonormais em duas variáveis no quadrado, vale notar que este exemplo gerado pelo método de Koornwinder é um caso particular do produto tensor dado na seção acima com o produto de dois polinômios ortonormais clássicos de Jacobi.

Exemplo 6.2.2. (*Polinômios de Jacobi no disco unitário*) Sejam $\omega_1(x) = (1-x^2)^{\mu-1/2}$ e $\omega_2(y) = (1-y^2)^{\mu-1/2}$ em $(-1, 1)$, e $\rho(x) = (1-x^2)^{1/2}$. Então, a função peso (6.2.2) é da forma

$$W(x, y) = (1-x^2)^{\mu-1/2}(1-(1-x)^{-1}y^2)^{\mu-1/2} = (1-x^2-y^2)^{\mu-1/2} \quad (6.2.6)$$

no domínio $R = B^2 = \{(x, y) : x^2 + y^2 < 1\}$, $\tilde{p}_{n-k}^{(k)}(x) = \tilde{G}_{n-k}^{(\mu+k+1/2)}(x)$ e os polinômios $\tilde{q}_k(y) = \tilde{G}_k^{(\mu)}(y)$, onde $\tilde{G}_j^{(\mu)}$ são os polinômios ortonormais de Gegenbauer. Então, temos que

$$P_{n-k, k}(x, y) = \tilde{G}_{n-k}^{(\mu+k+1/2)}(x)((1-x^2)^{1/2})^k \tilde{G}_k^{(\mu)}\left(\frac{y}{(1-x^2)^{1/2}}\right) \quad (6.2.7)$$

são polinômios ortonormais em duas variáveis no disco unitário.

Exemplo 6.2.3. (*Polinômios de Jacobi no triângulo*) Sejam $\omega_1(x) = x^{\kappa_1-1/2}(1-x)^{\kappa_2+\kappa_3-1}$ e $\omega_2(y) = y^{\kappa_2-1/2}(1-y)^{\kappa_3-1/2}$ em $(0, 1)$, e $\rho(x) = 1-x$. Logo a função peso (6.2.2) é da forma

$$\begin{aligned} W(x, y) &= x^{\kappa_1-1/2}(1-x)^{\kappa_2+\kappa_3-1}(1-x)^{\kappa_2-1/2}y^{\kappa_2-1/2}(1-(1-x)y)^{\kappa_3-1/2} \\ &= x^{\kappa_1-1/2}y^{\kappa_2-1/2}(1-x-y)^{\kappa_3-1/2} \end{aligned} \quad (6.2.8)$$

em $R = T^2 = \{(x, y) : x > 0, y > 0, 1-x-y > 0\}$, $\tilde{p}_{n-k}^{(k)}(x) = \tilde{P}_{n-k}^{(\kappa_2+\kappa_3+2k, \kappa_1-1/2)}(2x-1)$ e $\tilde{q}_k(y) = \tilde{P}_k^{(\kappa_3-1/2, \kappa_2-1/2)}(2y-1)$, onde $\tilde{P}_n^{(\alpha, \beta)}(2x-1)$ são os polinômios ortonormais de

Jacobi. Portanto,

$$P_{n-k,k}(x, y) = \tilde{P}_{n-k}^{(\kappa_2+\kappa_3+2k, \kappa_1-1/2)}(2x-1)(1-x)^k \tilde{P}_k^{(\kappa_3-1/2, \kappa_2-1/2)}\left(\frac{2y}{1-x}-1\right), \quad (6.2.9)$$

que são os polinômios ortonormais em duas variáveis no triângulo.

Exemplo 6.2.4. (*Polinômios ortonormais em domínio parabólico*) Seja $\omega_1(x) = x^a(1-x)^b$ em $(0, 1)$, $\omega_2(y) = (1-y^2)^a$ em $(-1, 1)$ e $\rho(x) = \sqrt{x}$. Assim a função peso (6.2.2) é dada por

$$W(x, y) = x^a(1-x)^b(1-x^{-1}y^2)^a = (1-x)^b(x-y^2)^a. \quad (6.2.10)$$

O domínio $R = \{(x, y) : y^2 < x < 1\}$ é limitado por uma reta e uma parábola. Os polinômios $\tilde{p}_{n-k}^{(k)}(x)$ são dados por $\tilde{P}_{n-k}^{(b, a+k+1/2)}(2x-1)$ e $\tilde{q}_k(x)$ por $\tilde{G}_k^{(a+1/2)}(x)$. Portanto, os polinômios ortonormais $P_{n-k,k}$ em (6.2.1) são dados por

$$P_{n-k,k}(x, y) = \tilde{P}_{n-k}^{(b, a+k+1/2)}(2x-1)x^{k/2}\tilde{G}_k^{(a+1/2)}(y/\sqrt{x}), \quad (6.2.11)$$

onde $\tilde{P}_j^{(i,l)}$ são os polinômios ortonormais de Jacobi e $\tilde{G}_k^{(\mu)}$ são os de Gegenbauer.

6.3 Exemplos de matrizes coeficientes da relação de três termos

Apresentamos, aqui, exemplos das matrizes coeficientes da relação de três termos para polinômios ortonormais clássicos em duas variáveis no quadrado e no triângulo. Neste capítulo, o funcional \mathcal{L} é escrito como uma integral com relação a função peso W .

6.3.1 Polinômios ortonormais clássicos no quadrado

Seja $\{\tilde{p}_n\}_{n=0}^\infty$ uma sequência de polinômios ortonormais com respeito à função peso ω em $[-1, 1]$, que satisfaz a relação de recorrência de três termos

$$x\tilde{p}_n = a_n\tilde{p}_{n+1} + b_n\tilde{p}_n + a_{n-1}\tilde{p}_{n-1}, \quad n \geq 0.$$

Seja W o produto das funções peso definido por $W(x, y) = \omega(x)\omega(y)$. Uma base de polinômios ortonormais é dada por

$$P_{n-k,k}(x, y) = \tilde{p}_{n-k}(x)\tilde{p}_k(y), \quad 0 \leq k \leq n.$$

Além disso,

$$\begin{aligned}
 A_{n,1} &= \mathcal{L}(x\mathbb{P}_n\mathbb{P}_{n+1}^T) = \mathcal{L}\left(x \begin{pmatrix} P_{n,0} \\ P_{n-1,1} \\ \vdots \\ P_{0,n} \end{pmatrix} \begin{pmatrix} P_{n+1,0} & P_{n,1} & \cdots & P_{0,n+1} \end{pmatrix}\right) \\
 &= \begin{pmatrix} \mathcal{L}(xP_{n,0}P_{n+1,0}) & \mathcal{L}(xP_{n,0}P_{n,1}) & \cdots & \mathcal{L}(xP_{n,0}P_{0,n+1}) \\ \mathcal{L}(xP_{n-1,1}P_{n+1,0}) & \mathcal{L}(xP_{n-1,1}P_{n,1}) & \cdots & \mathcal{L}(xP_{n-1,1}P_{0,n+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(xP_{0,n}P_{n+1,0}) & \mathcal{L}(xP_{0,n}P_{n,1}) & \cdots & \mathcal{L}(xP_{0,n}P_{0,n+1}) \end{pmatrix}
 \end{aligned}$$

e

$$A_{n,2} = \begin{pmatrix} \mathcal{L}(yP_{n,0}P_{n+1,0}) & \mathcal{L}(yP_{n,0}P_{n,1}) & \cdots & \mathcal{L}(yP_{n,0}P_{0,n+1}) \\ \mathcal{L}(yP_{n-1,1}P_{n+1,0}) & \mathcal{L}(yP_{n-1,1}P_{n,1}) & \cdots & \mathcal{L}(yP_{n-1,1}P_{0,n+1}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(yP_{0,n}P_{n+1,0}) & \mathcal{L}(yP_{0,n}P_{n,1}) & \cdots & \mathcal{L}(yP_{0,n}P_{0,n+1}) \end{pmatrix}.$$

Assim, devemos calcular $\mathcal{L}(xP_{n-j,j}P_{n+1-l,l})$ e $\mathcal{L}(yP_{n-j,j}P_{n+1-l,l})$ para $0 \leq j \leq n$ e $0 \leq l \leq n+1$. Determinemos, inicialmente, $\mathcal{L}(xP_{n-j,j}P_{n+1-l,l})$

$$\begin{aligned}
 \mathcal{L}(xP_{n-j,j}P_{n+1-l,l}) &= \int_{-1}^1 \int_{-1}^1 x(\tilde{p}_{n-j}(x)\tilde{p}_j(y))(\tilde{p}_{(n+1)-l}(x)\tilde{p}_l(y))\omega(x)\omega(y)dx dy \\
 &= \int_{-1}^1 x\tilde{p}_{n-j}(x)\tilde{p}_{(n+1)-l}(x)\omega(x)dx \int_{-1}^1 \tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
 &= \int_{-1}^1 (a_{n-j}\tilde{p}_{n-j+1}(x) + b_{n-j}\tilde{p}_{n-j}(x) + a_{n-k-1}\tilde{p}_{n-k-1}(x))\tilde{p}_{(n+1)-l}(x) \\
 &\quad \times \omega(x)dx \int_{-1}^1 \tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
 &= \begin{cases} 0, & j \neq l \\ a_{n-j}, & j = l \end{cases}.
 \end{aligned}$$

Portanto,

$$A_{n,1} = \begin{bmatrix} a_n & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & a_0 & 0 \end{bmatrix}.$$

Agora, efetuemos $\mathcal{L}(yP_{n-j,j}P_{n+1-l,l})$

$$\begin{aligned}
\mathcal{L}(yP_{n-j,j}P_{n+1-l,l}) &= \int_{-1}^1 \int_{-1}^1 y(\tilde{p}_{n-j}(x)\tilde{p}_j(y))(\tilde{p}_{(n+1)-l}(x)\tilde{p}_l(y))\omega(x)\omega(y)dx dy \\
&= \int_{-1}^1 \tilde{p}_{n-j}(x)\tilde{p}_{(n+1)-l}(x)\omega(x)dx \int_{-1}^1 y\tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
&= \int_{-1}^1 \tilde{p}_{n-j}(x)\tilde{p}_{(n+1)-l}(x)\omega(x)dx \int_{-1}^1 a_j\tilde{p}_{j+1}(y)\tilde{p}_l(y)\omega(y)dy \\
&= \begin{cases} 0, & j+1 \neq l \\ a_j, & j+1 = l \end{cases}.
\end{aligned}$$

Logo,

$$A_{n,2} = \begin{bmatrix} 0 & a_0 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & a_n \end{bmatrix}.$$

Analogamente, de (3.4.2), temos

$$\begin{aligned}
B_{n,1} &= \mathcal{L}(x\mathbb{P}_n\mathbb{P}_n^T) = \mathcal{L}\left(x \begin{pmatrix} P_{n,0} \\ P_{n-1,1} \\ \vdots \\ P_{0,n} \end{pmatrix} \begin{pmatrix} P_{n,0} & P_{n-1,1} & \cdots & P_{0,n} \end{pmatrix}\right) \\
&= \begin{pmatrix} \mathcal{L}(xP_{n,0}P_{n,0}) & \mathcal{L}(xP_{n,0}P_{n-1,1}) & \cdots & \mathcal{L}(xP_{n,0}P_{0,n}) \\ \mathcal{L}(xP_{n-1,1}P_{n,0}) & \mathcal{L}(xP_{n-1,1}P_{n-1,1}) & \cdots & \mathcal{L}(xP_{n-1,1}P_{0,n}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(xP_{0,n}P_{n,0}) & \mathcal{L}(xP_{0,n}P_{n-1,1}) & \cdots & \mathcal{L}(xP_{0,n}P_{0,n}) \end{pmatrix}
\end{aligned}$$

e

$$B_{n,2} = \begin{pmatrix} \mathcal{L}(yP_{n,0}P_{n,0}) & \mathcal{L}(yP_{n,0}P_{n-1,1}) & \cdots & \mathcal{L}(yP_{n,0}P_{0,n}) \\ \mathcal{L}(yP_{n-1,1}P_{n,0}) & \mathcal{L}(yP_{n-1,1}P_{n-1,1}) & \cdots & \mathcal{L}(yP_{n-1,1}P_{0,n}) \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{L}(yP_{0,n}P_{n,0}) & \mathcal{L}(yP_{0,n}P_{n-1,1}) & \cdots & \mathcal{L}(yP_{0,n}P_{0,n}) \end{pmatrix}.$$

Novamente, devemos calcular $\mathcal{L}(xP_{n-j,j}P_{n-l,l})$ e $\mathcal{L}(yP_{n-j,j}P_{n-l,l})$ para $0 \leq j \leq n$ e $0 \leq l \leq n$.

$$\begin{aligned}
\mathcal{L}(xP_{n-j,j}P_{n-l,l}) &= \int_{-1}^1 \int_{-1}^1 x(\tilde{p}_{n-j}(x)\tilde{p}_j(y))(\tilde{p}_{n-l}(x)\tilde{p}_l(y))\omega(x)\omega(y)dx dy \\
&= \int_{-1}^1 x\tilde{p}_{n-j}(x)\tilde{p}_{n-l}(x)\omega(x)dx \int_{-1}^1 \tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
&= \int_{-1}^1 (a_{n-j}\tilde{p}_{n-j+1}(x) + b_{n-j}\tilde{p}_{n-j}(x) + a_{n-j-1}\tilde{p}_{n-j-1}(x))\tilde{p}_{n-l}(x)\omega(x)dx \\
&\quad \times \int_{-1}^1 \tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
&= \begin{cases} 0, j \neq l \\ b_{n-j}, j = l \end{cases}.
\end{aligned}$$

Portanto,

$$B_{n,1} = \begin{bmatrix} b_n & & 0 \\ & \ddots & \\ 0 & & b_0 \end{bmatrix}.$$

No outro caso, temos

$$\begin{aligned}
\mathcal{L}(yP_{n-j,j}P_{n-l,l}) &= \int_{-1}^1 \int_{-1}^1 y(\tilde{p}_{n-j}(x)\tilde{p}_j(y))(\tilde{p}_{n-l}(x)\tilde{p}_l(y))\omega(x)\omega(y)dx dy \\
&= \int_{-1}^1 \tilde{p}_{n-j}(x)\tilde{p}_{n-l}(x)\omega(x)dx \int_{-1}^1 y\tilde{p}_j(y)\tilde{p}_l(y)\omega(y)dy \\
&= \int_{-1}^1 \tilde{p}_{n-j}(x)\tilde{p}_{n-l}(x)\omega(x)dx \\
&\quad \times \int_{-1}^1 (a_j\tilde{p}_{j+1}(y) + b_j\tilde{p}_j(y) + a_{j-1}\tilde{p}_{j-1}(y))\tilde{p}_l(y)\omega(y)dy \\
&= \begin{cases} 0, j \neq l \\ b_j, j = l \end{cases}.
\end{aligned}$$

Logo,

$$B_{n,2} = \begin{bmatrix} b_0 & & 0 \\ & \ddots & \\ 0 & & b_n \end{bmatrix}.$$

Se ω é centralmente simétrico, então $b_k = 0$ para $0 \leq k \leq n$. Portanto, $B_{n,i} = 0$.

6.3.2 Polinômios ortonormais clássicos no triângulo

Uma sequência de polinômios ortonormais no triângulo é da forma (6.2.9) e a função peso é dada por (6.2.8). Alteramos a notação dos índices de $\kappa = (\kappa_1, \kappa_2, \kappa_3)$ para $\kappa = (\alpha, \beta, \gamma)$. Assim, o polinômio e a função peso são das formas,

$$P_{n-k,k} = \tilde{P}_{n-k}^{(\beta+\gamma+2k, \alpha-1/2)}(2x-1)(1-x)^k \tilde{P}_k^{(\gamma-1/2, \beta-1/2)}\left(\frac{2y}{1-x}-1\right) \quad \text{e}$$

$$W(x, y) = x^{\alpha-1/2} y^{\beta-1/2} (1-x-y)^{\gamma-1/2}$$

Para calcular as matrizes coeficientes precisamos, como no exemplo anterior, determinar $\mathcal{L}(xP_{n-j,j}P_{(n+1)-l,l})$, $\mathcal{L}(yP_{n-j,j}P_{(n+1)-l,l})$, $\mathcal{L}(xP_{n-j,j}P_{n-l,l})$ e $\mathcal{L}(yP_{n-j,j}P_{n-l,l})$. Calculamos, inicialmente, $A_{n,1}$. Então,

$$\begin{aligned} \mathcal{L}(xP_{n-j,j}P_{(n+1)-l,l}) &= \int_0^1 \int_0^{1-x} x \tilde{P}_{n-j}^{(\beta+\gamma+2j, \alpha-1/2)}(2x-1)(1-x)^j \\ &\quad \times \tilde{P}_j^{(\gamma-1/2, \beta-1/2)}\left(\frac{2y}{1-x}-1\right) \tilde{P}_{(n+1)-l}^{(\beta+\gamma+2l, \alpha-1/2)}(2x-1)(1-x)^l \\ &\quad \times \tilde{P}_l^{(\gamma-1/2, \beta-1/2)}\left(\frac{2y}{1-x}-1\right) x^{\alpha-1/2} y^{\beta-1/2} (1-x-y)^{\gamma-1/2} dy dx. \end{aligned}$$

Fazendo a mudança de variáveis $y \rightarrow (1-x)z$, temos

$$\begin{aligned} \mathcal{L}(xP_{n-j,j}P_{(n+1)-l,l}) &= \int_0^1 x \tilde{P}_{n-j}^{(\beta+\gamma+2j, \alpha-1/2)}(2x-1) \tilde{P}_{(n+1)-l}^{(\beta+\gamma+2l, \alpha-1/2)}(2x-1) \\ &\quad \times x^{\alpha-1/2} (1-x)^{j+l+\beta+\gamma} dx \\ &\quad \times \int_0^1 \tilde{P}_j^{(\gamma-1/2, \beta-1/2)}(2z-1) \tilde{P}_l^{(\gamma-1/2, \beta-1/2)}(2z-1) \\ &\quad \times z^{\beta-1/2} (1-z)^{\gamma-1/2} dz. \end{aligned}$$

Quando $j \neq l$ a integral com relação a z é igual a zero. Portanto, quando $j = l$, temos, da relação de recorrência de três termos de $\tilde{P}_{n-j}^{(\beta+\gamma+2j, \alpha-1/2)}(2x-1)$, que

$$\mathcal{L}(xP_{n-j,j}P_{(n+1)-j,j}) = \begin{cases} a_j, & j = l \\ 0, & j \neq l \end{cases},$$

onde $a_j = \left(\frac{4(\beta + \gamma + n + j + 1)(\alpha + n - j + 1/2)(n - j + 1)(|\kappa| + n + j + 1/2)}{(|\kappa| + 2n + 1/2)(|\kappa| + 2n + 3/2)^2(|\kappa| + 2n + 5/2)} \right)^{1/2}$, para $1 \leq j \leq n$.

Portanto,

$$A_{n,1} = \begin{bmatrix} a_0 & & 0 & 0 \\ & \ddots & & \vdots \\ 0 & & a_n & 0 \end{bmatrix}.$$

Agora, calculemos $\mathcal{L}(yP_{n-j,j}P_{(n+1)-l,l})$ para determinar $A_{n,2}$. Como no cálculo anterior vamos fazer a mudança de variáveis $y \rightarrow (1-x)z$. Assim,

$$\begin{aligned}
\mathcal{L}(yP_{n-j,j}P_{(n+1)-l,l}) &= \int_0^1 \int_0^{1-x} y \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1)(1-x)^j \\
&\quad \times \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}\left(\frac{2y}{1-x}-1\right) \tilde{P}_{(n+1)-l}^{(\beta+\gamma+2l,\alpha-1/2)}(2x-1)(1-x)^l \\
&\quad \times \tilde{P}_l^{(\gamma-1/2,\beta-1/2)}\left(\frac{2y}{1-x}-1\right) x^{\alpha-1/2} y^{\beta-1/2} (1-x-y)^{\gamma-1/2} dx dy \\
&= \int_0^1 (1-x) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{(n+1)-l}^{(\beta+\gamma+2l,\alpha-1/2)}(2x-1) \\
&\quad \times x^{\alpha-1/2} (1-x)^{j+l+\beta+\gamma} dx \\
&\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_l^{(\gamma-1/2,\beta-1/2)}(2z-1) \\
&\quad \times z^{\beta-1/2} (1-z)^{\gamma-1/2} dz.
\end{aligned}$$

Pelo Teorema 2.2.4, a integral com relação à variável z é igual a 0 se $l+1 < j$ ou $j < l-1$. Portanto, devemos calcular os outros casos. Iniciemos com $l = j-1$. Denotando $c_{j,j-1}$ para $\mathcal{L}(yP_{n-j,j}P_{n-j+2,j-1})$, temos, para $1 \leq j \leq n$, que

$$\begin{aligned}
c_{j,j-1} &= \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j+2}^{(\beta+\gamma+2j-2,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma} dx \\
&\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_{j-1}^{(\gamma-1/2,\beta-1/2)}(2z-1) z^{\beta-1/2} (1-z)^{\gamma-1/2} dz \\
&= \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j+2}^{(\beta+\gamma+2j-2,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma} dx \\
&\quad \times \left(\frac{(\gamma+j-1/2)(\beta+j-1/2)j(\beta+\gamma+j-1)}{(\beta+\gamma+2j-2)(\beta+\gamma+2j)} \right)^{1/2} \frac{1}{\beta+\gamma+2j-1}.
\end{aligned}$$

Quando $l = j$ temos, da relação de recorrência de três termos de $\tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1)$ e de $\tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1)$, denotando $\mathcal{L}(yP_{n-j,j}P_{n-j+1,j})$ por $d_{j,j}$ para $0 \leq j \leq n$, que

$$\begin{aligned}
d_{j,j} &= \int_0^1 (1-x) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j+1}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma} dx \\
&\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) z^{\beta-1/2} (1-z)^{\gamma-1/2} dz \\
&= \left(\frac{(\beta+\gamma+n+j+1)(\alpha+n-j+1/2)(n-j+1)(|\kappa|+n+j+1/2)}{(|\kappa|+2n+1/2)(|\kappa|+2n+5/2)} \right)^{1/2} \\
&\quad \times \left(\frac{(\beta-1/2)^2 - (\gamma-1/2)^2}{(\beta+\gamma+2j+1)(\beta+\gamma+2j-1)} - 1 \right) \frac{1}{2(|\kappa|+2n+3/2)}
\end{aligned}$$

Por fim, calculemos o caso $l = j+1$. Denotando novamente $\mathcal{L}(yP_{n-j,j}P_{n-j,j+1})$ por

$e_{j,j+1}$, temos, para $0 \leq j \leq n$, que

$$e_{j,j+1} = \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j}^{(\beta+\gamma+2j+2,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma+2} dx \\ \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_{j+1}^{(\gamma-1/2,\beta-1/2)}(2z-1) z^{\beta-1/2} (1-z)^{\gamma-1/2} dz.$$

Fazendo a mudança de variáveis $x \rightarrow \frac{w+1}{2}$, do item 7 da subseção 7.391 de Gradshteyn e Ryzhik [7], obtemos

$$e_{j,j+1} = \frac{\int_{-1}^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(w) \tilde{P}_{n-j}^{(\beta+\gamma+2j+2,\alpha-1/2)}(w) (1+w)^{\alpha-1/2} (1-w)^{2j+\beta+\gamma+2} dw}{2^{|\kappa|+2j+5/2}} \\ \times \left(\frac{(\gamma+j+1/2)(\beta+j+1/2)(j+1)(\gamma+\beta+j)}{(\gamma+\beta+2j)(\gamma+\beta+2j+2)} \right)^{1/2} \frac{1}{\gamma+\beta+2j+1} \\ = \left(\frac{(\gamma+j+1/2)(\beta+j+1/2)(j+1)(\gamma+\beta+j)(\beta+\gamma+n+j+2)}{(\gamma+\beta+2j)(\gamma+\beta+2j+2)(|\kappa|+n+j+3/2)(|\kappa|+n+j+1/2)} \right)^{1/2} \\ \times \frac{(\beta+\gamma+n+j+1)^{1/2} 2^{|\kappa|+2j+3/2} \Gamma(\beta+\gamma+n+j+3) (\Gamma(\alpha+n-j+1/2))^2}{\gamma+\beta+2j+1 ((n-j)!)^2 \Gamma(|\kappa|+2n+7/2) (\Gamma(|\kappa|+n+j+1/2))^2} \\ \times \frac{\Gamma(|\kappa|+2n+1/2) \Gamma(\beta+\gamma+n+j+1)}{(|\kappa|+2n+1/2)(|\kappa|+2n+5/2)}.$$

Assim, pelos coeficientes acima dados, temos a matriz da relação de três termos

$$A_{n,2} = \begin{bmatrix} d_{0,0} & e_{0,1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ c_{1,0} & d_{1,1} & e_{1,2} & 0 & \cdots & 0 & 0 & 0 \\ 0 & c_{2,1} & d_{2,2} & e_{2,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & d_{n-1,n-1} & e_{n-1,n} & 0 \\ 0 & 0 & 0 & 0 & \cdots & c_{n,n-1} & d_{n,n} & e_{n,n+1} \end{bmatrix}.$$

Agora, calculemos as matrizes coeficiente $B_{n,1}$ e $B_{n,2}$. Para determinar $B_{n,1}$ devemos obter $\mathcal{L}(xP_{n-j,j}P_{n-l,l})$. Analogamente, fazendo a mudança de variáveis $y \rightarrow (1-x)z$, obtemos

$$\mathcal{L}(xP_{n-j,j}P_{n-l,l}) = \int_0^1 \int_0^{1-x} x \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) (1-x)^j \\ \times \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}\left(\frac{2y}{1-x}-1\right) \tilde{P}_{n-l}^{(\beta+\gamma+2l,\alpha-1/2)}(2x-1) (1-x)^l \\ \times \tilde{P}_l^{(\gamma-1/2,\beta-1/2)}\left(\frac{2y}{1-x}-1\right) x^{\alpha-1/2} y^{\beta-1/2} (1-x-y)^{\gamma-1/2} dy dx \\ = \int_0^1 x \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-l}^{(\beta+\gamma+2l,\alpha-1/2)}(2x-1) \\ \times x^{\alpha-1/2} (1-x)^{j+l+\beta+\gamma} dx \\ \times \int_0^1 \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_l^{(\gamma-1/2,\beta-1/2)}(2z-1) \\ \times z^{\beta-1/2} (1-z)^{\gamma-1/2} dz.$$

Caso $j \neq l$, a integral com relação a z é igual a 0. Desse modo, basta fazer o caso $j = l$. Logo,

$$\begin{aligned} \mathcal{L}(xP_{n-j,j}P_{n-j,j}) &= \int_0^1 x \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \\ &\quad \times x^{\alpha-1/2}(1-x)^{2j+\beta+\gamma} dx \\ &\quad \times \int_0^1 \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \\ &\quad \times z^{\beta-1/2}(1-z)^{\gamma-1/2} dz \\ &= \left(\frac{(\alpha-1/2)^2 - (\beta+\gamma+2j)^2}{2(|\kappa|+2n+3/2)(|\kappa|+2n-1/2)} + \frac{1}{2} \right). \end{aligned}$$

Portanto,

$$B_{n,1} = \begin{bmatrix} b_0 & & 0 \\ & \ddots & \\ 0 & & b_n \end{bmatrix},$$

onde $b_j = \left(\frac{(\alpha-1/2)^2 - (\beta+\gamma+2j)^2}{2(|\kappa|+2n+3/2)(|\kappa|+2n-1/2)} + \frac{1}{2} \right)$, para $0 \leq j \leq n$.

Por fim, calculemos $B_{n,2}$, ou seja, temos que obter $\mathcal{L}(yP_{n-j,j}P_{n-l,l})$. Fazendo a mesma mudança de variáveis, temos

$$\begin{aligned} \mathcal{L}(yP_{n-j,j}P_{n-l,l}) &= \int_0^1 (1-x) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-l}^{(\beta+\gamma+2l,\alpha-1/2)}(2x-1) \\ &\quad \times x^{\alpha-1/2}(1-x)^{j+l+\beta+\gamma} dx \\ &\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_l^{(\gamma-1/2,\beta-1/2)}(2z-1) \\ &\quad \times z^{\beta-1/2}(1-z)^{\gamma-1/2} dz. \end{aligned}$$

Analogamente ao caso de $A_{n,2}$, temos que se $l+1 < j$ ou $j < l-1$, a integral com relação a z é igual a 0. Calculemos os outros casos.

Caso $l = j-1$, denotando $\mathcal{L}(yP_{n-j,j}P_{n-j+1,j-1})$ por $f_{j,j-1}$, obtemos

$$\begin{aligned} f_{j,j-1} &= \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j+1}^{(\beta+\gamma+2j-2,\alpha-1/2)}(2x-1) x^{\alpha-1/2}(1-x)^{2j+\beta+\gamma} dx \\ &\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_{j-1}^{(\gamma-1/2,\beta-1/2)}(2z-1) z^{\beta-1/2}(1-z)^{\gamma-1/2} dz \\ &= \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j+1}^{(\beta+\gamma+2j-2,\alpha-1/2)}(2x-1) x^{\alpha-1/2}(1-x)^{2j+\beta+\gamma} dx \\ &\quad \times \left(\frac{(\gamma+j-1/2)(\beta+j-1/2)j(\beta+\gamma+j-1)}{(\beta+\gamma+2j-2)(\beta+\gamma+2j)} \right)^{1/2} \frac{1}{(\beta+\gamma+2j-1)}. \end{aligned}$$

No caso $j = l$, denotando $g_{j,j}$ para $\mathcal{L}(yP_{n-j,j}P_{n-j,j})$, obtemos

$$\begin{aligned} g_{j,j} &= \int_0^1 (1-x) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma} dx \\ &\quad \times \int_0^1 z \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) \tilde{P}_j^{(\gamma-1/2,\beta-1/2)}(2z-1) z^{\beta-1/2} (1-z)^{\gamma-1/2} dz \\ &= \frac{1}{4} \left(\frac{(\beta-1/2)^2 - (\gamma-1/2)^2}{(\beta+\gamma+2j+1)(\beta+\gamma+2j-1)} + 1 \right) \\ &\quad \times \left(\frac{(\beta+\gamma+2j)^2 - (\alpha-1/2)^2}{(|\kappa|+2n+3/2)(|\kappa|+2n-1/2)} + 1 \right). \end{aligned}$$

O último caso é quando $l = j+1$ e denotamos, agora, $\mathcal{L}(yP_{n-j,j}P_{n-j-1,j+1})$ por $h_{j,j+1}$.

$$\begin{aligned} h_{j,j+1} &= \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j-1}^{(\beta+\gamma+2j+2,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma+2} dx \\ &\quad \int_0^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(2x-1) \tilde{P}_{n-j-1}^{(\beta+\gamma+2j+2,\alpha-1/2)}(2x-1) x^{\alpha-1/2} (1-x)^{2j+\beta+\gamma+2} dx \\ &= \frac{\int_{-1}^1 \tilde{P}_{n-j}^{(\beta+\gamma+2j,\alpha-1/2)}(w) \tilde{P}_{n-j-1}^{(\beta+\gamma+2j+2,\alpha-1/2)}(w) (1+w)^{\alpha-1/2} (1-w)^{2j+\beta+\gamma+2} dx}{2^{|\kappa|+2j+5/2}} \\ &\quad \times \left(\frac{(\gamma+j+1/2)(\beta+j+1/2)(j+1)(\gamma+\beta+n)}{(\gamma+\beta+2j)(\gamma+\beta+2j+2)} \right)^{1/2} \frac{1}{\gamma+\beta+2j+1}. \end{aligned}$$

Portanto, dos coeficientes calculados acima, temos a matriz

$$B_{n,2} = \begin{bmatrix} g_{0,0} & h_{0,1} & 0 & 0 & \cdots & 0 & 0 & 0 \\ f_{1,0} & g_{1,1} & h_{1,2} & 0 & \cdots & 0 & 0 & 0 \\ 0 & f_{2,1} & g_{2,2} & h_{2,3} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & g_{n-1,n-1} & h_{n-1,n} & 0 \\ 0 & 0 & 0 & 0 & \cdots & f_{n,n-1} & g_{n,n} & h_{n,n+1} \end{bmatrix}.$$

Referências Bibliográficas

- [1] BEN-ISRAEL, A.; GREVELLE, T.N.E. **Generalized Inverses: Theory and Applications**. New York: Wiley, 1977.
- [2] CALLIOLI, C.A.; DOMINGUES, H.H.; COSTA, R.C.F. **Álgebra Linear e Aplicações**. 6ª ed. rev.. São Paulo: Atual, 1990.
- [3] CHIHARA, T.S. **An Introduction to Orthogonal Polynomials**. New York: Gordon and Breach, 1978.
- [4] COELHO, F.U.; LOURENÇO, M.L. **Um Curso de Álgebra Linear**. São Paulo: Editora da Universidade de São Paulo, 2001.
- [5] DUNKL, C.F.; XU, Y. **Orthogonal Polynomials of Several Variables**. Cambridge: Cambridge University Press, 2001.
- [6] ERDÉLYI, A. et al. **Higher Transcendental Functions**. New York: McGraw-Hill, 1953.
- [7] GRADSTEYN, I.S.; RYZHIK, I.M. **Table of Integrals, Series, and Products**. 7ª ed. London: Academic Press, 2007.
- [8] HORN, R.A.; JOHNSON, C.R. **Matrix Analysis**. Cambridge: Cambridge University Press, 1985.
- [9] ISMAIL, M.E.H. **Classical and Quantum Orthogonal Polynomials in One Variable**. Cambridge: Cambridge University Press, 2005.
- [10] KOORNWINDER, T.H. Two-variable analogues of the classical orthogonal polynomials. **Theory and Applications of Special Functions**. p. 435-495, ed. R.A. Askey, Academic Press, New York, 1975.
- [11] KRALL, H.L.; SHEFFER, I.M. Orthogonal polynomials in two variables. **Ann. Mat. Pura Appl.** v. 76, p. 325-376, 1967.
- [12] LIMA, E.L., **Álgebra Linear**. 7ª ed. Rio de Janeiro: IMPA, 2004.

-
- [13] MYSOVSKIKH, I.P. A multidimensional analog of quadrature formula of Gaussian type and the generalized problem of Radon, **Vopr. Vychisl. i Prikl. Mat., Tashkent.** v. 38, p. 55-69, 1970.
- [14] MYSOVSKIKH, I.P., Numerical characteristics of orthogonal polynomials in two variables, **Vestnik Leningrad Univ. Math.** v. 3, p. 323-332, 1976.
- [15] SZEGŐ, G. **Orthogonal Polynomials.** 4th ed., Amer. Math. Soc. Colloq. Publ., vol. 23, Amer. Math. Soc., Providence, RI, 1975.

Glossário

$\mathbb{R}^{m \times n}$	-	espaço vetorial das matrizes reais de dimensão $m \times n$
$(x)_n$	-	símbolo de Pochhammer
$\Gamma(x)$	-	função Gama
$B(x)$	-	função Beta
$P_n^{(\alpha, \beta)}(x)$	-	polinômios ortogonais de Jacobi
$\tilde{P}_n^{(\alpha, \beta)}(x)$	-	polinômios ortonormais de Jacobi
$H_n(x)$	-	polinômios ortogonais de Hermite
$\tilde{H}_n(x)$	-	polinômios ortonormais de Hermite
$L_n^{(\alpha)}(x)$	-	polinômios ortogonais de Laguerre
$\tilde{L}_n^{(\alpha)}(x)$	-	polinômios ortonormais de Laguerre
$G_n^{(\lambda)}(x)$	-	polinômios ortogonais de Gegenbauer
$\tilde{G}_n^{(\lambda)}(x)$	-	polinômios ortonormais de Gegenbauer
$P_n^{(0)}(x)$	-	polinômios ortogonais de Legendre
$\tilde{P}_n^{(0)}(x)$	-	polinômios ortonormais de Legendre
$T_n(x)$	-	polinômios ortogonais de Chebyshev de 1ª espécie
$\tilde{T}_n(x)$	-	polinômios ortonormais de Chebyshev de 1ª espécie
$U_n(x)$	-	polinômios ortogonais de Chebyshev de 2ª espécie
$\tilde{U}_n(x)$	-	polinômios ortonormais de Chebyshev de 2ª espécie
α	-	multi-índice
\mathbb{N}_0	-	conjunto dos inteiros não-negativos
\mathbb{N}_0^d	-	conjunto das d -uplas inteiras não-negativas
$\delta_{\alpha, \beta}$	-	delta de Kronecker
\mathbf{x}	-	d -upla real
x_i	-	i -ésima variável de \mathbf{x}
\mathbb{R}^d	-	conjunto das d -uplas reais
\mathbf{x}^α	-	monômio em d variáveis
Π^d	-	conjunto dos polinômios de d variáveis
Π_n^d	-	conjunto dos polinômios de d variáveis de grau no máximo n
\mathcal{P}_n^d	-	conjunto dos polinômios homogêneos de grau n em d variáveis

r_n^d	-	dimensão de \mathcal{P}_n^d
\mathcal{L}_s ou \mathcal{L}	-	funcional linear de momento definido pela sequência s
s_α	-	$\mathcal{L}_s(\mathbf{x}^\alpha)$
\mathbb{X}^n	-	vetor monomial coluna de grau n
\mathbf{s}_k	-	vetor de momento
$\mathbf{s}_{\{k\}+\{j\}}$	-	momento matricial
$M_{n,d}$	-	matriz de momento
$\Delta_{n,d}$	-	determinante da matriz $M_{n,d}$
\mathcal{V}_n^d	-	espaço dos polinômios ortogonais de grau exatamente n
\mathbb{P}_n	-	vetor polinomial coluna de grau n
$\bar{\mathbb{P}}_n$	-	vetor polinomial mônico de grau n
$\tilde{\mathbb{P}}_n$	-	vetor polinomial ortonormal de grau n
S_n	-	$\mathcal{L}(\mathbb{X}^n \mathbb{P}_n^T)$
H_n	-	$\mathcal{L}(\mathbb{P}_n \mathbb{P}_n^T)$
$G_{n,i}$	-	matriz coeficiente
G_n ou $G_{n,n}$	-	matriz coeficiente líder
$A_{n,i}$, $B_{n,i}$ e $C_{n,i}$	-	matrizes coeficiente da relação de três termos
$L_{n,i}$	-	A matriz que satisfaz $L_{n,i} \mathbb{X}^{n+1} = x_i \mathbb{X}^n$
A_n	-	matriz junção de $\{A_{n,i}\}$
C_n	-	matriz junção de $\{C_{n,i}\}$
L_n	-	matriz junção de $\{L_{n,i}\}$
D_n^T	-	inversa generalizada de A_n
J_i	-	matriz de Jacobi em blocos
$J_{n,i}$	-	n -ésima matriz de Jacobi em blocos truncada
$\mathcal{H}_{\mathcal{L}}$	-	espaço de Hilbert de funções de valores reais em definido \mathbb{R}^d com produto interno $\langle f, g \rangle = \mathcal{L}(fg)$
\mathbf{K}_n	-	n -ésimo núcleo reprodutor
$\Lambda_n(\mathbf{x})$	-	função de Christoffel
∂_i	-	derivada parcial com relação a variável x_i
\mathbf{z}	-	zero comum do polinômio ortogonal em d variáveis
Λ	-	autovalor junção de $J_{n,1}, \dots, J_{n,d}$