

JOÃO LUÍS GONÇALVES

Introdução à Aproximação de Funções

Trabalho de Conclusão de Curso apresentado ao
Curso de Matemática - Habilitação Licenciatura
Departamento de Matemática
Centro de Ciências Físicas e Matemáticas
Universidade Federal de Santa Catarina

Orientador: Daniel Norberto Kozakevich

Florianópolis
Novembro de 2004

Sumário

Introdução	4
1 Aproximação Polinomial dos Mínimos Quadrados	5
1.1 Introdução	5
1.2 Caso Discreto	5
1.3 Caso Contínuo	9
1.4 A abordagem da Álgebra Linear	9
1.5 Avaliação da Qualidade do Ajuste	15
2 Polinômios Ortogonais	18
2.1 Introdução	18
2.2 Geração Recursiva de Polinômios Ortogonais	20
2.3 Polinômios de Legendre	22
2.4 Polinômios de Chebyshev	26
2.4.1 Zeros e Extremos dos Polinômios de Chebyshev	29
2.4.2 Minimizando o Erro na Interpolação de Lagrange	32
2.5 Polinômios de Hermite	32
2.6 Polinômios de Laguerre	34
3 Aproximação por Funções Racionais	35
3.1 Introdução	35
3.2 Frações Contínuas	36

3.3	Aproximação de Padé	38
3.3.1	Aproximação por Funções Racionais Chebyshev	43
4	Aproximação Polinomial Trigonométrica	47
	Conclusão	55
	Referências Bibliográficas	56
A	Norma de Vetores e Matrizes	57
B	Condicionamento de Matrizes	61
C	Processo de Gram-Schmidt	63
D	Série de Taylor	68
E	Teorema da Multiplicidade	71

Introdução

A teoria da aproximação envolve, em geral, dois tipos de problemas: dada uma função explicitamente, desejamos encontrar um tipo mais simples de função, para representá-la, com propriedades que facilitem sua utilização; um outro problema diz respeito a ajustar um conjunto de dados, por meio de uma função.

Ao longo deste trabalho buscaremos dentre os métodos de aproximação que estudaremos, aquele que melhor possa aproximar uma função.

No primeiro capítulo introduzimos a aproximação polinomial do método dos mínimos quadrados, tendo em vista a grande importância deste método. Apresentaremos seus casos discreto e contínuo e na última seção analisaremos brevemente este método do ponto de vista da álgebra linear.

O segundo capítulo trata dos Polinômios Ortogonais, em uma introdução à aproximação de funções estes polinômios não poderiam ficar de fora. São apresentadas as propriedades importantes desses polinômios, e introduzimos alguns conceitos importantes tais como dependência linear, função peso e ortogonalidade.

O terceiro capítulo trata da aproximação por Funções Racionais, este método é explicado utilizando-se duas técnicas para obter a função racional, sendo a primeira por frações contínuas, interpolando alguns pontos e a segunda é a aproximação de Padé onde além de um novo método utilizam-se resultados importantes da teoria da aproximação.

No quarto e último capítulo tratamos da aproximação polinomial trigonométrica. Além do novo método surgem alguns conceitos importantes tais como Séries de Fourier.

Capítulo 1

Aproximação Polinomial dos Mínimos Quadrados

1.1 Introdução

Dado um conjunto de pares ordenados ou uma função $y(x)$, desejamos encontrar uma função simples que “melhor” aproxime ou descreva o fenômeno do qual os dados foram obtidos empiricamente. É nesse tipo de problema que se aplica o método dos quadrados mínimos.

O que está na base deste método é a minimização da norma $\|y - p\|_2$ onde y representa os dados e p o polinômio de aproximação.

1.2 Caso Discreto

Para dados discretos (x_i, y_i) , com $i = 1, 2, \dots, n$, deve minimizar-se a soma:

$$S(a_0, a_1, \dots, a_m) = \sum_{i=0}^N \left(y_i - a_0 - a_1 x_i - \dots - a_m x_i^m \right)^2, \text{ onde } m < N$$

A condição $m < N$ torna muito pouco provável que o polinômio $p(x) = a_0 + a_1 x + \dots + a_m x^m$ se coloque em todos os N pontos dados e, portanto, é pouco provável que

S possa ser zero. A idéia de Gauss, que disse ter utilizado o método pela primeira vez, é tornar S tão pequeno quanto possível.

Para um conjunto de dados cujo gráfico de dispersão mostra uma distribuição na qual os pontos aparecem alinhados, o tipo de função mais adequada para representar estes dados é uma função linear, a escolha da função pode ser definida observando-se a disposição dos pontos no gráfico, ou considerando a natureza do fenômeno estudado.

Assim devemos minimizar:

$$S(a_0, a_1) = \sum_{i=0}^m \left(y_i - (a_0 + a_1 x_i) \right)^2$$

para os valores de a_0 e a_1 .

Esta minimização pode ser feita usando as técnicas de cálculo. Assim derivamos S em relação a a_1 e igualamos a zero; depois derivamos S em relação a a_0 e igualamos a zero.

Este processo gera as seguintes equações:

$$\begin{cases} \frac{\partial}{\partial a_1} \sum_{i=1}^m \left(y_i - (a_0 + a_1 x_i) \right)^2 = 0 \\ \frac{\partial}{\partial a_0} \sum_{i=1}^m \left(y_i - (a_0 + a_1 x_i) \right)^2 = 0 \end{cases}$$

que chamaremos de equações normais.

A idéia apresentada para o caso linear pode ser estendida para polinômios do tipo $\sum_{i=0}^m a_k x^k = P_m(x)$, onde m é menor que o número de pontos dados menos 1.

Procedemos como no caso linear. Determinando as derivadas em relação a a_0, a_1, \dots, a_m e igualando-as a zero obtém-se um sistema de $m + 1$ equações

$$\frac{\partial S}{\partial a_k} = -2 \sum_{i=0}^N x_i^k (y_i - a_0 - a_1 x_i - \dots - a_m x_i^m) = 0$$

com $k = 0, 1, \dots, m$. Fazendo $S_k = \sum_{i=0}^N x_i^k$ e $t_k = \sum_{i=0}^N y_i x_i^k$, estas equações tomam a forma

$$S_0 a_0 + S_1 a_1 + \dots + S_m a_m = t_0$$

$$S_1 a_0 + S_2 a_1 + \dots + S_{m+1} a_m = t_1$$

.....

$$S_m a_0 + S_{m+1} a_1 + \dots + S_{2m} a_m = t_m$$

Encontrando os a_k teremos os coeficientes do polinômio que ajusta os dados no sentido dos mínimos quadrados.

Exemplo 1.2.1 *Determinar a reta dos mínimos quadrados relativa aos dados:*

x	6	8	10	12	14	16	18	20	22	24
y	3.8	3.7	4.0	3.9	4.3	4.2	4.2	4.4	4.5	4.5

Designando $S = \sum_{i=0}^N (y_i - a_0 - a_1 x_i)^2$, os valores que conduzem ao máximo anulam as primeiras derivadas, assim as equações normais são:

$$(N + 1)a_0 + \left(\sum_{i=0}^N x_i \right) a_1 = \sum_{i=0}^N y_i$$

e

$$\left(\sum_{i=0}^N x_i \right) a_0 + \left(\sum_{i=0}^N x_i^2 \right) a_1 = \sum_{i=0}^N x_i y_i$$

Introduzindo as notações:

$$A = N + 1, B = \sum x_i, C = \sum x_i^2, t_0 = \sum y_i \text{ e } t_1 = \sum x_i y_i$$

a solução deste sistema será

$$a_1 = \frac{At_1 - Bt_0}{AC - B^2}$$

e

$$a_0 = \frac{Ct_0 - Bt_1}{AC - B^2}$$

Para aplicar estas fórmulas aos nossos dados, devemos fazer uma transformação $z = \frac{x-6}{2}$, assim os z_i são os inteiros $0,1,\dots,9$. Tem-se então:

$$A = 9 + 1 = 10, B = 45, C = 285, t_0 = 41,5 \text{ e } t_1 = 194,1.$$

Assim

$$a_0 = \frac{10 \cdot 194,1 - 45 \cdot 41,5}{10 \cdot 285 - 45^2} \simeq 0,089$$

e

$$a_1 = \frac{285 \cdot 41,5 - 45 \cdot 194,1}{10 \cdot 285 - 45^2} \simeq 3,76$$

logo $y \simeq p(z)$ sendo $p(z) = 0,09x + 3,76$.

Como $2z + 6 = x$, temos $p(x) = 0,045x + 3,49$.

Veja esta aproximação no gráfico:

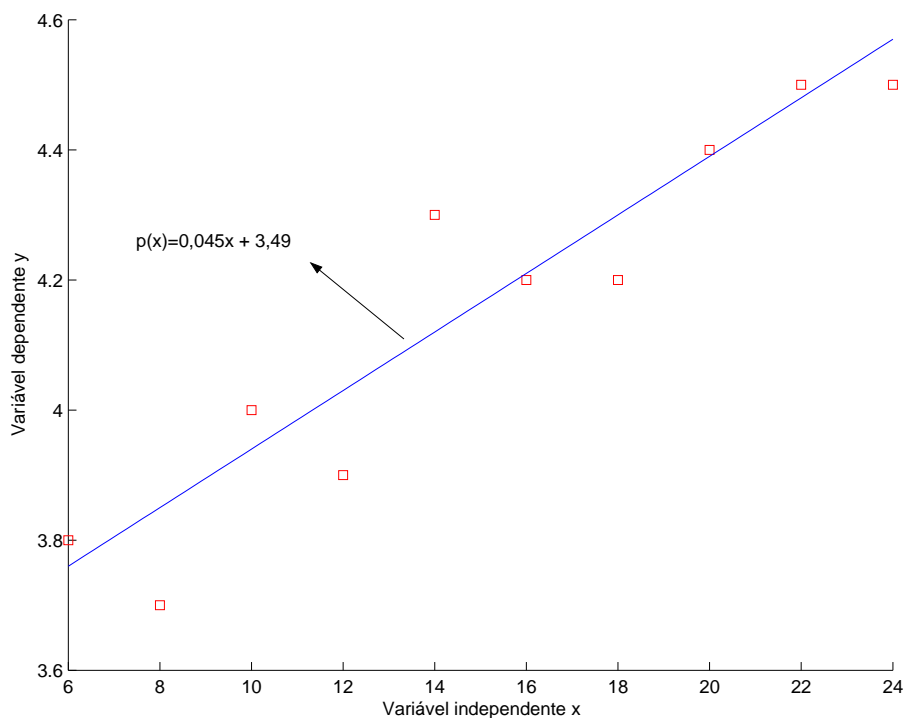


Figura 1.1: Aproximação dos dados pela reta dos mínimos quadrados

1.3 Caso Contínuo

No caso contínuo, ao invés de pares ordenados temos uma função dada explicitamente, com características inadequadas para o fim ao qual a desejamos aplicar. Nos interessa substituí-la em um intervalo desejado, por uma função que se aproxima a ela e que tenha as propriedades adequadas (por exemplo regularidade, derivabilidade e calculável em todo o domínio).

Supondo $f \in \mathcal{C}[a, b]$, o problema será determinar P_n que minimize

$$E(a_0, a_1, a_2, \dots, a_n) = \int_a^b \left(f(x) - P_n(x) \right)^2 dx$$

o que implica

$$\frac{\partial E}{\partial a_j} = 0 \text{ para cada } j=0,1,\dots,n.$$

Fazendo estas derivadas e igualando-as a zero obtemos novamente um sistema de equações, as Equações Normais, cuja solução são os a_j , que são os coeficientes do polinômio que melhor substitui a função inicial, em um determinado intervalo, no sentido dos mínimos quadrados.

Porém é muito trabalhoso e difícil fazer esta aproximação quando $P_n(x)$ não é uma combinação linear de funções ortogonais entre si.

Como exemplo do caso contínuo do método dos mínimos quadrados veja o exemplo 2.3.2, no próximo capítulo.

1.4 A abordagem da Álgebra Linear

A fim de dar um tratamento unificado aos vários métodos dos mínimos quadrados, vamos considerar o problema geral da minimização em um espaço linear. A solução é encontrada por um processo algébrico, usando a idéia de projeção ortogonal. Naturalmente o problema geral dá $p(x)$ e as equações normais.

Em primeiro lugar, lembre que em geometria euclidiana plana, dados um ponto y e uma reta r , o ponto de r mais próximo de y é o ponto p , único, tal que \overline{py} é ortogonal a r , p é a projeção ortogonal de y sobre r .

De modo análogo, no espaço euclidiano tridimensional, dados um ponto y e um plano S , o ponto de S mais próximo de y é o único p tal que \overline{py} é ortogonal a todos os vetores em S , p é a projeção ortogonal de y sobre o plano S .

Vamos estender esta idéia a um espaço linear geral.

Seja E um espaço linear com produto interno e \mathbf{y} um vetor de E , procuramos determinar um vetor \mathbf{p} de um dado subespaço S tal que $\|\mathbf{y} - \mathbf{p}\| < \|\mathbf{y} - \mathbf{q}\|$ sendo \mathbf{q} qualquer vetor de S e diferente de \mathbf{y} e sendo que esta norma em E é definida por: $\|\mathbf{v}\| = \sqrt{(\mathbf{v}, \mathbf{v})}$ onde os parênteses denotam o produto interno. Vamos mostrar em primeiro lugar que existe um único vetor \mathbf{p} tal que $\mathbf{y} - \mathbf{p}$ é ortogonal a qualquer vetor de S . Este vetor \mathbf{p} chama-se projeção ortogonal de \mathbf{y} sobre S .

Suponhamos que S é de dimensão finita e seja $\mathbf{e}_0, \mathbf{e}_1, \dots, \mathbf{e}_m$ uma base ortonormada.

Considere o vetor $\mathbf{p} = (\mathbf{y}, \mathbf{e}_0)\mathbf{e}_0 + (\mathbf{y}, \mathbf{e}_1)\mathbf{e}_1 + \dots + (\mathbf{y}, \mathbf{e}_m)\mathbf{e}_m$.

por cálculo direto obtemos $(\mathbf{p}, \mathbf{e}_k) = (\mathbf{y}, \mathbf{e}_k)$, pois $(\mathbf{e}_k, \mathbf{e}_i) = 0$ se $k \neq i$ e $(\mathbf{e}_k, \mathbf{e}_i) = 1$ se $k = i$, portanto $(\mathbf{p} - \mathbf{y}, \mathbf{e}_k) = 0$ para $k = 0, 1, \dots, m$, pelo uso das propriedades do produto interno. Vem então que $(\mathbf{p} - \mathbf{y}, \mathbf{q}) = 0$ qualquer que seja \mathbf{q} pertencente a S , pois \mathbf{q} é uma combinação linear de elementos da base ortogonal. Se existir um outro vetor \mathbf{p}' com esta mesma propriedade, isto é, se $(\mathbf{p}' - \mathbf{y}, \mathbf{q}) = 0$, então obtém-se que para qualquer \mathbf{q} de S , $(\mathbf{p} - \mathbf{p}', \mathbf{p} - \mathbf{p}') = 0$. Como $\mathbf{p} - \mathbf{p}'$ pertence a S , deverá ter-se $(\mathbf{p} - \mathbf{p}', \mathbf{p} - \mathbf{p}') = 0$, o que, em virtude das propriedades do produto interno, implica que $\mathbf{p} = \mathbf{p}'$. Para a projeção ortogonal \mathbf{p} , tem-se $\|\mathbf{y} - \mathbf{q}\|^2 = \|(\mathbf{y} - \mathbf{p}) + (\mathbf{p} - \mathbf{q})\|^2 = \|\mathbf{y} - \mathbf{p}\|^2 + \|\mathbf{p} - \mathbf{q}\|^2 + 2(\mathbf{y} - \mathbf{p}, \mathbf{p} - \mathbf{q})$ e como a última parcela do segundo membro é igual a zero, porque $\mathbf{p} - \mathbf{q}$ é elemento de S , conclui-se que $\|\mathbf{y} - \mathbf{p}\| < \|\mathbf{y} - \mathbf{q}\|$ como se pretendia.

Sendo $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_m$ uma base qualquer de W , como determinar o vetor \mathbf{p} que seja a projeção ortogonal de \mathbf{y} sobre W ?

Deveremos ter $(\mathbf{y} - \mathbf{p}, \mathbf{u}_k) = 0$, ou seja, $(\mathbf{p}, \mathbf{u}_k) = (\mathbf{y}, \mathbf{u}_k)$ para $k = 0, 1, \dots, m$. Como

\mathbf{p} tem representação única $\mathbf{p} = a_0\mathbf{u}_0 + a_1\mathbf{u}_1 + \dots + a_m\mathbf{u}_m$, por substituição obtém-se

$$(\mathbf{u}_0, \mathbf{u}_k)a_0 + (\mathbf{u}_1, \mathbf{u}_k)a_1 + \dots + (\mathbf{u}_m, \mathbf{u}_k)a_m = (\mathbf{y}, \mathbf{u}_k)$$

para $k = 0, \dots, m$. Estas são as Equações Normais para este problema, devemos resolvê-las em relação a a_0, a_1, \dots, a_m .

Como observamos anteriormente, este sistema tem uma única solução. Note que no caso especial em que $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_m$ é ortonormada, as equações normais reduzem-se a $a_i = (\mathbf{y}, \mathbf{u}_i)$.

Note ainda que se \mathbf{y} é dado numa base ortogonal de E que contenha $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_m$, digamos,

$$\mathbf{y} = a_0\mathbf{u}_0 + a_1\mathbf{u}_1 + \dots + a_m\mathbf{u}_m + a_{m+1}\mathbf{u}_{m+1} + \dots$$

então a projeção ortogonal \mathbf{p} , que é a aproximação dos mínimos quadrados, obtém-se imediatamente por truncatura da representação depois do termo $a_m\mathbf{u}_m$:

$$\mathbf{y} = a_0\mathbf{u}_0 + a_1\mathbf{u}_1 + \dots + a_m\mathbf{u}_m.$$

É interessante relacionar os conceitos entre o método para a aproximação polinomial dos mínimos quadrados e o que foi apresentado nesta abordagem até agora, vejamos:

E : espaço das funções reais de domínio x_0, x_1, \dots, x_N .

W : o subconjunto de E dos polinômios de grau menor ou igual a m .

\mathbf{y} : a função dada tendo valores y_0, y_1, \dots, y_N .

$(\mathbf{v}_1, \mathbf{v}_2)$: o produto interno $\sum_{i=0}^N \mathbf{v}_1(x_i)\mathbf{v}_2(x_i)$.

$\|\mathbf{v}\|^2$: a soma $\sum_{i=0}^N [\mathbf{v}(x_i)]^2$.

\mathbf{u}_k : a função com valores \mathbf{x}_i^k .

\mathbf{p} : o polinômio com valores $p_i = a_0 + a_1x_1 + \dots + a_mx_i^m$.

$\|\mathbf{y} - \mathbf{p}\|^2$: a soma $S = \sum_{i=0}^N (y_i - p_i)^2$.

$(\mathbf{y}, \mathbf{u}_k)$: a soma $\mathbf{t}_k = \sum_{i=0}^N y_i x_i^k$.

$$(\mathbf{u}_j, \mathbf{u}_k): \text{ a soma } \mathbf{s}_{j+k} = \sum_{i=0}^N \mathbf{x}_i^{j+k}.$$

Usando as mesmas idéias apresentadas no início desta seção só que com notação matricial temos que o problema dos mínimos quadrados é:

Sendo A uma matriz $m \times n$, \mathbf{b} um vetor pertence a \mathbb{R}^m e $A\mathbf{x} = \mathbf{b}$ um sistema que não tem solução exata, encontrar um vetor $\hat{\mathbf{x}}$ pertencente a \mathbb{R}^n tal que

$$\|\mathbf{b} - A\hat{\mathbf{x}}\| \leq \|\mathbf{b} - A\mathbf{x}\|,$$

para todo \mathbf{x} em \mathbb{R}^n .

O aspecto mais importante de um problema de mínimos quadrados é que, independente do vetor \mathbf{x} selecionado, o vetor $A\mathbf{x}$ pertence, necessariamente, ao espaço de colunas, $\text{col}A$ (veja a figura (1.2)). Procuramos, então, \mathbf{x} que torna $A\mathbf{x}$ o ponto mais próximo de \mathbf{b} em $\text{col}A$. É claro que, se \mathbf{b} está em $\text{col}A$, então \mathbf{b} é da forma $A\mathbf{x}$ para algum \mathbf{x} e tal que \mathbf{x} é a solução de mínimos quadrados.

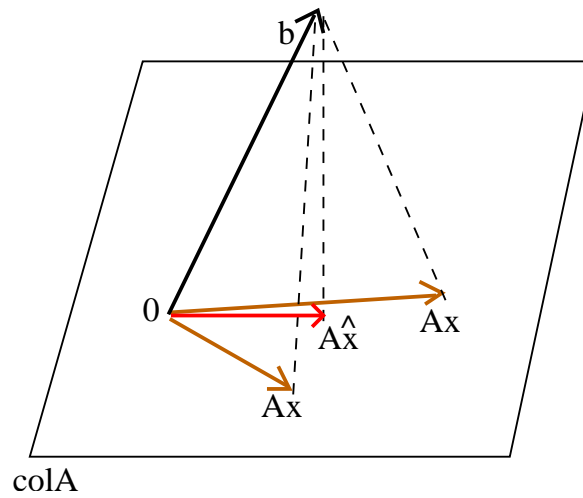


Figura 1.2: Projeção de \mathbf{b} no espaço coluna de A ($\text{col}A$)

A solução do problema dos mínimos quadrados pode ser encontrada assim, dados A e \mathbf{b} como acima aplicamos ao subespaço $\text{col}A$ o teorema da melhor aproximação que diz:

Seja W um subespaço de \mathbb{R}^n e seja $\hat{\mathbf{y}}$ a projeção ortogonal de \mathbf{y} sobre W . Então, $\hat{\mathbf{y}}$ é o ponto de W mais próximo de \mathbf{y} , isto é,

$$\|\mathbf{y} - \hat{\mathbf{y}}\| < \|\mathbf{y} - \mathbf{v}\|,$$

para todo \mathbf{v} em W diferente de $\hat{\mathbf{y}}$.

Seja

$$\hat{\mathbf{b}} = \text{proj}_{\text{col}A} \mathbf{b}$$

Como $\hat{\mathbf{b}}$ está no espaço das colunas de A , a equação $A\mathbf{x} = \hat{\mathbf{b}}$ é possível e existe $\hat{\mathbf{x}}$ em \mathbb{R}^n tal que

$$A\hat{\mathbf{x}} = \hat{\mathbf{b}}.$$

Como $\hat{\mathbf{b}}$ é o ponto mais próximo de \mathbf{b} em $\text{col}A$, um vetor $\hat{\mathbf{x}}$ é uma solução de mínimos quadrados para $A\mathbf{x} = \mathbf{b}$ se e somente se $\hat{\mathbf{x}}$ satisfaz $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Um tal $\hat{\mathbf{x}}$ em \mathbb{R}^n é uma lista de coeficientes que vai formar $\hat{\mathbf{b}}$ com as colunas de A (existem muitas soluções de $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$ se a equação tem variáveis livres).

Suponha que $\hat{\mathbf{x}}$ satisfaz $A\hat{\mathbf{x}} = \hat{\mathbf{b}}$. Sabemos da decomposição ortogonal que, a projeção $\hat{\mathbf{b}}$ tem a propriedade de que $\mathbf{b} - \hat{\mathbf{b}}$ é ortogonal a $\text{col}A$ (veja a figura (1.3)), logo $\mathbf{b} - A\hat{\mathbf{x}}$ é ortogonal a cada coluna de A . Se \mathbf{a}_j é uma coluna de A , então $\mathbf{a}_j(\mathbf{b} - A\hat{\mathbf{x}}) = 0$ (produto interno) ou $\mathbf{a}_j^T[\mathbf{b} - A\hat{\mathbf{x}}] = 0$. Como cada \mathbf{a}_j^T é uma linha de A^T , temos:

$$A^T[\mathbf{b} - A\hat{\mathbf{x}}] = 0$$

$$A^T\mathbf{b} - A^T A\hat{\mathbf{x}} = 0$$

$$A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$$

que representa os sistema de Equações Normais para $\hat{\mathbf{x}}$, e o conjunto de soluções dos mínimos quadrados de $A\mathbf{x} = \mathbf{b}$ coincide com o conjunto não-vazio de soluções das equações normais $A^T A\hat{\mathbf{x}} = A^T\mathbf{b}$.

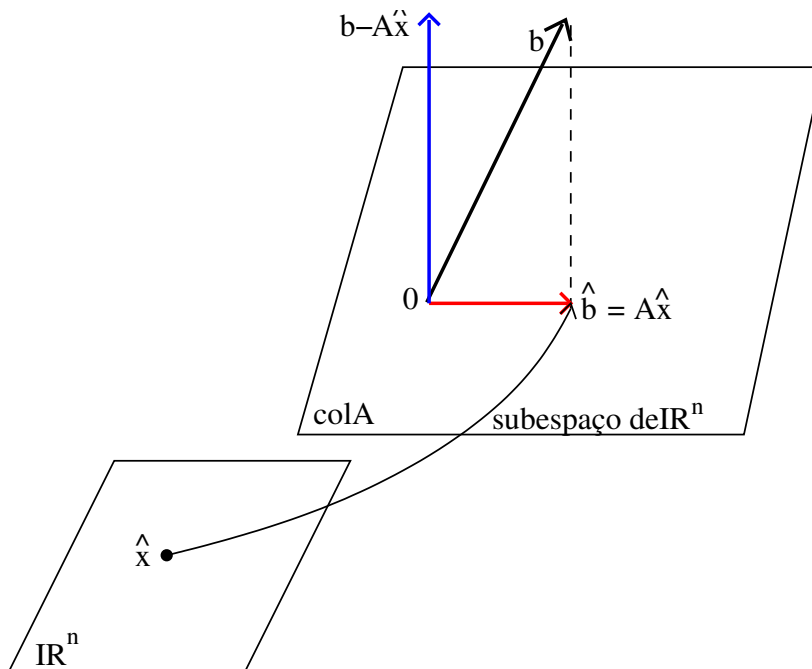


Figura 1.3: A projeção ortogonal de \mathbf{b} , vista como $A\hat{\mathbf{x}}$

Exemplo 1.4.1 Determinar a reta dos mínimos quadrados relativa aos dados:

x	6	8	10	12	14	16	18	20	22	24
y	3.8	3.7	4.0	3.9	4.3	4.2	4.2	4.4	4.5	4.5

Solução: Em forma matricial o problema denota-se assim:

$$\begin{pmatrix} 1 & 6 \\ 1 & 8 \\ 1 & 10 \\ 1 & 12 \\ 1 & 14 \\ 1 & 16 \\ 1 & 18 \\ 1 & 20 \\ 1 & 22 \\ 1 & 24 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 3.8 \\ 3.7 \\ 4.0 \\ 3.9 \\ 4.3 \\ 4.2 \\ 4.2 \\ 4.4 \\ 4.5 \\ 4.5 \end{pmatrix}$$

Então de acordo com o que foi apresentado podemos resolver este problema, multiplicando ambos os lados desta equação pela transposta da matriz 10×2 , que resulta em:

$$\begin{pmatrix} 10 & 150 \\ 150 & 2580 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 41.5 \\ 637.2 \end{pmatrix},$$

que é equivalente a

$$\begin{pmatrix} 10 & 150 \\ 0 & 330 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} 41.5 \\ 14.7 \end{pmatrix},$$

e a partir desse sistema obtemos $a_0 = 3.481818182$ e $a_1 = 0.0445454$, ou seja, a reta que representa estes dados no sentido dos mínimos quadrados é: $y(x) = 0.0445454x + 3.481818182$. Este resultado, a menos de arredondamentos, é igual ao obtido usando técnicas de cálculo.

1.5 Avaliação da Qualidade do Ajuste

Um método pelo qual podemos avaliar a qualidade de um ajuste, segundo MACHADO, é através do coeficiente de correlação de Pearson. O coeficiente de correlação de Pearson r^2 pode ser calculado na forma mais geral como:

$$r^2 = 1 - \frac{n \sum_{i=1}^n (y_i - p(x_i))^2}{n \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i \right)^2}$$

Uma outra forma do coeficiente de correlação, válido para ajuste de função do tipo $y = f(x)$, é expressa como:

$$r = \frac{\sum_{i=1}^n [(x_i - x_m)(p(x_i) - p(x_m))]}{\sqrt{\sum_{i=1}^n (x_i - x_m)^2 \sum_{i=1}^n (p(x_i) - p(x_m))^2}} = \frac{\sum_{i=1}^n x_i p(x_i) - n x_m p(x_m)}{(n-1) S_x S_y}$$

para a qual $x_m = \frac{\sum_{i=1}^n x_i}{n}$ e $p_m = \frac{\sum_{i=1}^n p(x_i)}{n}$ são os valores médios de x e $p(x)$, respectivamente.

As expressões:

$$S_x = \sqrt{\frac{\sum_{i=1}^n (x_i - x_m)^2}{n-1}} \quad \text{e} \quad S_y = \sqrt{\frac{\sum_{i=1}^n (p(x_i) - p(x_m))^2}{n-1}}$$

representam a covariância x e covariância y , respectivamente.

O coeficiente de correlação é limitado aos seguintes valores: $0 \leq r^2 \leq 1$. Quanto mais próximo de 1 for o valor de r^2 , melhor será o ajuste. Quanto mais próximo r^2 estiver de 0, para um ajuste linear, significa que mais desprezível é o coeficiente angular.

Exemplo 1.5.1 *Ajustar aos pontos tabelados uma função linear do tipo $p(x) = a_0 + a_1 x$ e determinar o coeficiente de correlação de Pearson para o ajuste.*

x	1,0	1,2	1,4	1,6	1,8	2,0
y	0,525	0,8448	1,2807	1,8634	2,6326	3,6383

x	2,2	2,4	2,6	2,8	3,0
y	4,944	6,6358	8,7768	11,5076	14,9484

Solução:

Seja $p(x) = a_0 + a_1 x$ a forma da aproximação que desejamos.

x	y	x^2	xy	$(x - x_m)^2$	$(y - y_m)^2$
1,0	0,52500	1	0,525	1	22,1863609
1,2	0,84478	1,44	1,013736	0,64	19,276139
1,4	1,28068	1,96	1,792952	0,36	15,6385448
1,6	1,86340	2,56	2,98144	0,16	11,369305
1,8	2,63260	3,24	4,73868	0,04	6,77373497
2,0	3,63856	4	7,27712	0	2,54938702
2,2	4,94400	4,84	10,8768	0,04	0,08482074
2,4	6,62580	5,76	15,90192	0,16	1,93365711
2,6	8,77679	6,76	22,819654	0,36	12,5425764
2,8	11,50759	7,84	32,221252	0,64	39,3423745
3,0	14,94844	9	44,84532	1	94,3462542
22	57,58764	48,4	144,993874	4,4	226,043155

$$a_0 = \frac{\sum x^2 \sum y - \sum x \sum xy}{n \sum x^2 - (\sum x)^2} = \frac{(48,40)(57,58764) - (22)(144,99387)}{(11)(48,40) - (22)^2} = -8,3187$$

$$a_1 = \frac{n \sum xy \sum y - \sum x \sum y^2}{n \sum x^2 - (\sum x)^2} = \frac{(11)(144,99387) - (22)(57,58764)}{(11)(48,40) - (22)^2} = 6,7770$$

$$S_x = \sqrt{\frac{\sum (x - x_m)^2}{n - 1}} = \sqrt{\frac{4,4}{10}} = 0,66332$$

$$S_y = \sqrt{\frac{\sum (y - y_m)^2}{n - 1}} = \sqrt{\frac{226,04314}{10}} = 4,7544$$

$$r = a_1 \frac{S_x}{S_y} = (6,7770) \frac{0,66332}{4,7544} = 0,95$$

$$p(x) = -8,3187 + 6,7770x \quad r^2 = 0,9$$

Capítulo 2

Polinômios Ortogonais

2.1 Introdução

Para discutirmos uma forma geral da aproximação, devemos introduzir noções de Dependência Linear, Função Peso e Ortogonalidade. Vejamos as definições:

Propriedade 2.1.1 *Se ϕ_j é polinômio de grau j para cada $j = 0, 1, 2, \dots, n$ então $\{\phi_1, \phi_2, \phi_3, \dots, \phi_n\}$ é linearmente independente em algum intervalo $[a, b]$.*

Com um conjunto desta forma qualquer polinômio de grau até n , pode ser escrito de forma única como combinação linear deste conjunto.

Definição 2.1.1 *Uma função integrável ω é chamada de função peso no intervalo I se $\omega(x) \geq 0$ para todo $x \in I$, e se ω não é identicamente nula em algum subintervalo de I .*

O propósito da função peso é atribuir importância maior ou menor para a aproximação em certas partes (porções) do intervalo. Por exemplo, a função peso:

$$\omega(x) = \frac{1}{\sqrt{1-x^2}}$$

coloca menos ênfase no centro do intervalo $(-1,1)$ e mais ênfase quando $|x|$ está perto de 1.

Suponha dado $\{\phi_1, \phi_2, \phi_3, \dots, \phi_n\}$ um conjunto linearmente independente de funções em $[a,b]$, ω uma função peso em $[a,b]$ e $f \in \mathcal{C}[a, b]$, uma combinação linear

$$P(x) = \sum_{k=0}^n a_k \phi_k(x)$$

é procurada para minimizar

$$E(a_0, a_1, \dots, a_n) = \int_a^b \omega(x) \left[f(x) - \sum_{k=0}^n a_k \phi_k(x) \right]^2 dx.$$

Usando métodos de cálculo, obtemos que isto acontece quando:

$$0 = \frac{\partial E}{\partial a_j} = 2 \int_a^b \omega(x) \left[f(x) - \sum_{k=0}^n a_k \phi_k(x) \right] \phi_j(x) dx$$

para $k = 0, 1, \dots, n$, gerando um sistema de equações, este sistema de equações pode ser escrito como:

$$\int_a^b \omega(x) f(x) \phi_j(x) dx = \sum_{k=0}^n a_k \int_a^b \omega(x) \phi_k(x) \phi_j(x) dx,$$

para cada $j = 0, 1, \dots, n$.

Se as funções $\phi_1, \phi_2, \phi_3, \dots, \phi_n$ forem escolhidas de forma que:

$$\int_a^b \omega(x) \phi_k(x) \phi_j(x) dx = \begin{cases} 0 & \text{se } j \neq k \\ \alpha_k > 0 & \text{se } j = k \end{cases} \quad (2.1)$$

para algum número α_k , então as equações normais reduzem-se a:

$$\int_a^b \omega(x) f(x) \phi_j(x) dx = a_j \int_a^b \omega(x) [\phi_j(x)]^2 dx = a_j \alpha_j$$

para cada $j = 0, 1, 2, \dots, n$.

E assim fica fácil determinar os coeficientes a_j , com $j = 0, 1, \dots, n$ da combinação linear que queríamos

$$a_j = \frac{1}{\alpha_j} \int_a^b \omega(x) f(x) \phi_j(x) dx$$

Observe que o problema de aproximação dos mínimos quadrados é grandemente simplificado quando as funções $\phi_1, \phi_2, \phi_3, \dots, \phi_n$ são escolhidas satisfazendo a equação 2.1.

Definição 2.1.2 *Dadas as funções $\phi_1, \phi_2, \phi_3, \dots, \phi_n$, o conjunto $\{ \phi_1, \phi_2, \phi_3, \dots, \phi_n \}$ é ortogonal no intervalo $[a, b]$ com respeito a função peso ω , se a equação 2.1 é satisfeita.*

Se tivermos $\alpha_k = 1$ para cada $k=0,1,2,\dots,n$, o conjunto de funções será chamado de ortonormal.

Essa definição, juntamente com o processo realizado anteriormente, implica no seguinte:

Se $\{ \phi_1, \phi_2, \phi_3, \dots, \phi_n \}$ é um conjunto ortogonal de funções em um intervalo $[a, b]$ com respeito a função peso ω , então a aproximação dos mínimos quadrados para f em $[a, b]$ com respeito à ω é:

$$P(x) = \sum_{k=0}^n a_k \phi_k(x)$$

onde

$$a_k = \frac{\int_a^b \omega(x) \phi_k(x) f(x) dx}{\int_a^b \omega(x) [\phi_k(x)]^2 dx} = \frac{1}{\alpha_k} \int_a^b \omega(x) \phi_k(x) f(x) dx$$

2.2 Geração Recursiva de Polinômios Ortogonais

Este resultado está baseado no processo de Gram-Schmidt, este descreve um processo recursivo para construção de polinômios ortogonais em $[a, b]$ com respeito à função peso ω .

O conjunto de polinômios $\{ \phi_0, \phi_1, \phi_2, \phi_3, \dots, \phi_n \}$ definido a seguir é linearmente independente e ortogonal em $[a, b]$ com respeito a função peso ω ,

$$\phi_0 \equiv 1, \phi_1 = X - B_1$$

onde

$$B_1 = \frac{\int_a^b x\omega(x)[\phi_0(x)]^2 dx}{\int_a^b \omega(x)[\phi_0(x)]^2 dx}$$

e para $k \geq 2$

$$\phi_k(x) = (X - B_k)\phi_{k-1}(x) - C_k\phi_{k-2}(x) \quad (2.2)$$

onde

$$B_k = \frac{\int_a^b x\omega(x)[\phi_{k-1}(x)]^2 dx}{\int_a^b \omega(x)[\phi_{k-1}(x)]^2 dx}$$

e

$$C_k = \frac{\int_a^b x\omega(x)\phi_{k-1}(x)\phi_{k-2}(x) dx}{\int_a^b \omega(x)[\phi_{k-2}(x)]^2 dx}$$

Assim, para algum polinômio Q_k de algum grau $k < n$

$$\int_a^b \omega(x)\phi_n(x)\phi_k(x) dx = 0$$

Exemplo 2.2.1 *O conjunto dos polinômios de Legendre (que são o assunto da nossa próxima seção), $\{P_n\}$ é ortogonal em $[-1,1]$ com respeito à função peso $\omega = 1$. A definição clássica dos polinômios de Legendre requer que $P_n(1) = 1$. Para cada n , a relação recursiva (equação 2.2) pode ser usada para gerar os polinômios quando $n \geq 2$. A normalização não será necessária na nossa discussão, e nos polinômios da aproximação por mínimos quadrados gerados em outro caso será essencialmente o mesmo.*

Usando o processo recursivo, $P_0(x) \equiv 1$, e

$$B_1 = \frac{\int_{-1}^1 x dx}{\int_{-1}^1 dx} = 0 \text{ e } P_1(x) = (x - B_1)P_0(x) = x$$

também

$$B_2 = \frac{\int_{-1}^1 x^3 dx}{\int_{-1}^1 x^2 dx} = 0 \text{ e } C_2 = \frac{\int_{-1}^1 x^2 dx}{\int_{-1}^1 1 dx} = \frac{1}{3},$$

$$\text{e } P_2(x) = (x - B_2)P_1(x) - C_2P_0(x) = (x - 0)x - \frac{1}{3} \cdot 1 = x^2 - \frac{1}{3}$$

Os outros polinômios de Legendre são obtidos da mesma maneira. Os próximos três são:

$$P_3(x) = x^3 - \frac{3}{5} \cdot x, \quad P_4(x) = x^4 - \frac{6}{7} \cdot x^2 + \frac{3}{35} \text{ e } P_5(x) = x^5 - \frac{10}{9} \cdot x^3 + \frac{5}{21}.$$

2.3 Polinômios de Legendre

Definição 2.3.1 Os Polinômios de Legendre são definidos no intervalo $[-1,1]$ pela fórmula

$$y = \begin{cases} P_0(x) = 1 \\ P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} [(x^2 - 1)^n], \quad n \geq 1 \end{cases}$$

onde o índice que aparece em P_n indica a ordem do polinômio.

A partir dessa definição, podemos encontrar algumas propriedades, que seguem:

Propriedade 2.3.1 Os polinômios de Legendre são ortogonais em relação ao produto interno usual para funções contínuas no intervalo $[-1,1]$, $(f, g) = \int_{-1}^1 f(x)g(x)dx$, tendo-se então que

$$(P_n, P_j) = y = \begin{cases} 0 & \text{se } n \neq j \\ \frac{2}{2n+1} & \text{se } n = j \end{cases}$$

cujos resultados são obtidos realizando a integração envolvida no produto interno.

Propriedade 2.3.2 Os polinômios de Legendre apresentam a seguinte simetria:

$$P_n(-x) = (-1)^n P_n(x).$$

Propriedade 2.3.3 Para $x \in [-1,1]$, tem-se que

$$|P_n(x)| \leq 1.$$

A figura abaixo nos mostra os polinômios de Legendre até a ordem 5 no intervalo $[-1,1]$.

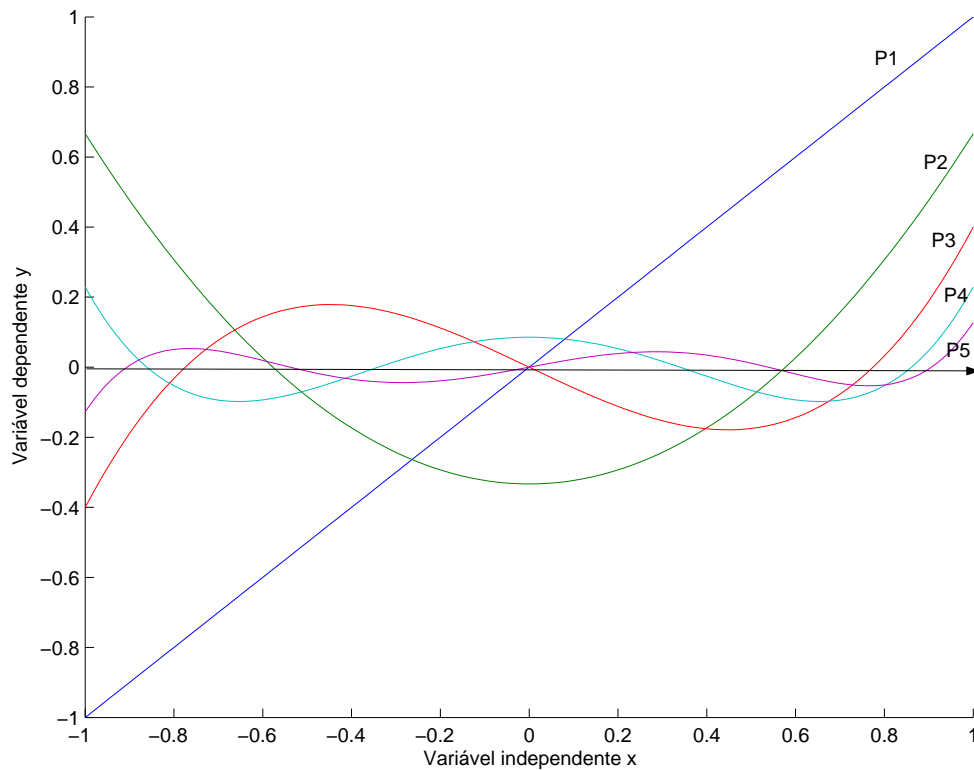


Figura 2.1: Polinômios de Legendre

Veremos alguns exemplos da utilização dos polinômios de Legendre.

Exemplo 2.3.1 (Neste exemplo podemos perceber como o uso de polinômios ortogonais simplifica a aproximação). Determinar os coeficientes a_i , para $i = 0, 1, \dots, m$ de forma que

$$E = \int_{-1}^1 [y(x) - a_0 P_0(x) - a_1 P_1(x) - \dots - a_m P_m(x)]^2 dx$$

seja o mínimo, onde P_k é o Polinômio de Legendre de grau k .

Sol.: O uso dos Polinômios de Legendre é bastante conveniente, pois vai permitir que as equações normais que permitem determinar os coeficientes a_k formem um sistema muito simples de resolver. Como qualquer polinômio pode ser expresso como uma combinação linear de Polinômios de Legendre, estamos efetivamente resolvendo o problema da aproximação polinomial dos mínimos quadrados para dados contínuos.

Igualando a zero as derivadas de E , tem-se

$$\frac{\partial E}{\partial a_k} = -2 \int_{-1}^1 [y(x) - a_0 P_0(x) - a_1 P_1(x) - \dots - a_m P_m(x)] P_k(x) dx = 0$$

para $k = 0, 1, \dots, m$. Dada a ortogonalidade destes polinômios, estas equações simplificam-se imediatamente para

$$\int_{-1}^1 [y(x) - a_k P_k(x)] P_k(x) dx = 0.$$

Assim cada equação envolve apenas um a_k e tem-se:

$$a_k = \frac{\int_{-1}^1 y(x) P_k(x) dx}{\int_{-1}^1 P_k^2(x) dx}$$

Exemplo 2.3.2 Determinar a aproximação do mínimos quadrados para $y(x) = \sin(t)$ no intervalo $(0, \pi)$ por uma parábola. Use os polinômios de Legendre. Solução: Em primeiro lugar façamos $t = \frac{\pi(x+1)}{2}$ a fim de se obter o intervalo $(-1, 1)$ no argumento x .

Teremos então $y = \sin\left[\frac{\pi(x+1)}{2}\right]$.

Assim

$$a_0 = \frac{\int_{-1}^1 \sin\left[\frac{\pi(x+1)}{2}\right] \cdot 1 dx}{\int_{-1}^1 1 dx}$$

$$a_0 = \frac{-\frac{2}{\pi}(\cos\left[\frac{\pi(x+1)}{2}\right])\Big|_{-1}^1}{x\Big|_{-1}^1}$$

$$a_0 = \frac{-\frac{2}{\pi} \cdot (-2)}{2}$$

$$a_0 = \frac{2}{\pi}$$

$$a_1 = \frac{\int_{-1}^1 \sin\left[\frac{\pi(x+1)}{2}\right] \cdot x dx}{\int_{-1}^1 x dx}$$

$$a_1 = 0$$

$$a_2 = \frac{\int_{-1}^1 \sin\left[\frac{\pi(x+1)}{2}\right] \cdot (x^2 - \frac{1}{3}) dx}{\int_{-1}^1 (x^2 - \frac{1}{3}) dx}$$

$$a_2 = \frac{10}{\pi} \left(1 - \frac{12}{\pi^2}\right)$$

sendo então a parábola dada por:

$$y = \frac{2}{\pi} + \frac{10}{\pi} \left(1 - \frac{12}{\pi^2}\right) \frac{1}{2} (3x^2 - 1)$$

$$y = \frac{2}{\pi} + \frac{10}{\pi} \left(1 - \frac{12}{\pi^2}\right) \left[\frac{6}{\pi^2} \left(t - \frac{\pi}{2}\right)^2 - \frac{1}{2}\right]$$

A parábola e a senóide estão representadas na figura a seguir, para mostrar a qualidade da aproximação:

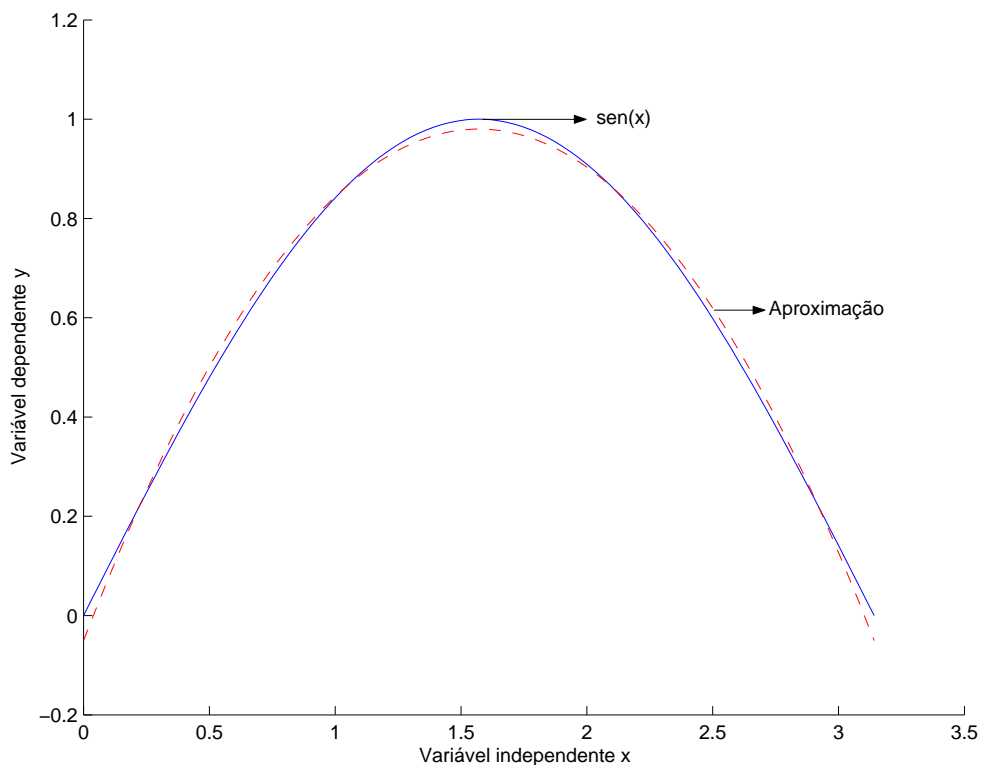


Figura 2.2: Senóide e sua aproximação dos mínimos quadrados usando polinômios de Legendre

2.4 Polinômios de Chebyshev

Os Polinômios de Chebyshev $\{T_n\}$ são ortogonais em $(-1,1)$. Com respeito a função peso $\omega(x) = \frac{1}{\sqrt{1-x^2}}$. Podemos mostrar que isso é verdade usando a definição dada de ortogonalidade.

Definição 2.4.1 Para $x \in [-1, 1]$, define-se:

$$T_n(x) = \cos(n \arccos(x)), \text{ para cada } n \geq 0$$

Primeiro note que

$$T_0(x) = \cos 0 = 1$$

e

$$T_1(x) = \cos(\arccos x) = x$$

Para $n \geq 1$ introduzimos a substituição $\theta = \arccos x$, mudando a equação para

$$T_n(\theta(x)) = T_n(\theta) = \cos(n\theta),$$

onde $\theta \in [0, \pi]$.

Uma relação de recorrência é obtida notando-se que

$$T_{n+1}(\theta) = \cos(n\theta + \theta) = \cos(n\theta) \cos \theta - \sin(n\theta) \sin(\theta)$$

e

$$T_{n-1}(\theta) = \cos(n\theta - \theta) = \cos(n\theta) \cos \theta + \sin(n\theta) \sin(\theta)$$

somando estas equações temos que

$$T_{n+1}(\theta) = 2 \cos(n\theta) \cos \theta - T_{n-1}(\theta)$$

retornando à variável x temos o seguinte:

$$T_0(x) = 1,$$

$$T_1(x) = x$$

e para $n \geq 1$

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$

A relação de recorrência implica que T_n é um polinômio de grau n com o coeficiente do termo de maior grau igual a 2^{n-1} , quando $n \geq 1$. Os próximos três polinômios de Chebyshev são:

$$T_2(x) = 2xT_1(x) - T_0(x) = 2x^2 - 1$$

$$T_3(x) = 2xT_2(x) - T_1(x) = 4x^3 - 3x, \text{ e}$$

$$T_4(x) = 2xT_3(x) - T_2(x) = 8x^4 - 8x^2 + 1$$

Os gráficos de T_1 , T_2 , T_3 e T_4 são mostrados na figura abaixo:

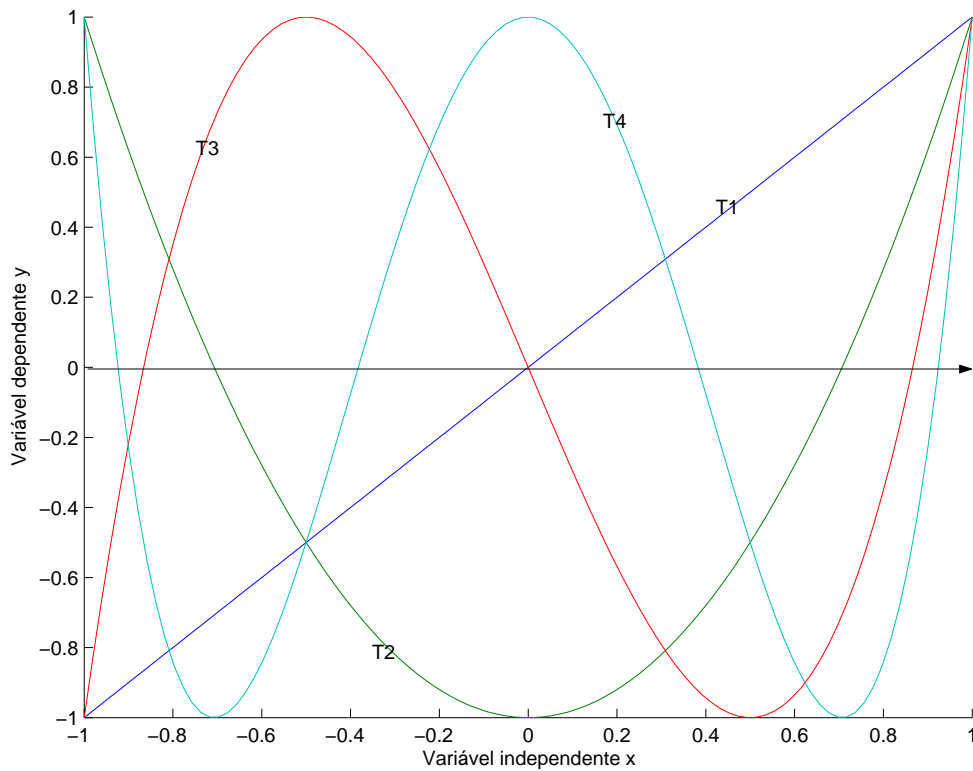


Figura 2.3: Polinômios de Chebyshev

Para mostrar a ortogonalidade dos Polinômios de Chebyshev, considere:

$$\int_{-1}^1 \frac{T_n(x) \cdot T_m(x)}{\sqrt{1-x^2}} dx = \int_{-1}^1 \frac{\cos(n \arccos x) \cos(m \arccos x)}{\sqrt{1-x^2}} dx$$

Utilizando novamente a substituição $\theta = \arccos x$, temos $d\theta = -\frac{1}{\sqrt{1-x^2}} dx$ e

$$\int_{-1}^1 \frac{T_n(x) \cdot T_m(x)}{\sqrt{1-x^2}} dx = - \int_{\pi}^0 \cos(n\theta) \cos(m\theta) d\theta = \int_0^{\pi} \cos(n\theta) \cos(m\theta) d\theta$$

Supondo $m \neq n$.

Desde então:

$$\cos(n\theta) \cos(m\theta) = \frac{1}{2} \left[\cos((n+m)\theta) + \cos((n-m)\theta) \right]$$

temos:

$$\begin{aligned} \int_{-1}^1 \frac{T_n(x) \cdot T_m(x)}{\sqrt{1-x^2}} dx &= \frac{1}{2} \int_0^\pi \cos((n+m)\theta) d\theta + \frac{1}{2} \int_0^\pi \cos((n-m)\theta) d\theta = \\ &= \left[\frac{1}{2(n+m)} \sin((n+m)\theta) + \frac{1}{2(m-n)} \sin((n-m)\theta) \right]_0^\pi = 0. \end{aligned}$$

De maneira semelhante, podemos também mostrar que:

$$\int_{-1}^1 \frac{[T_n(x)]^2}{\sqrt{1-x^2}} dx = \frac{\pi}{2}$$

para cada $n \geq 1$.

2.4.1 Zeros e Extremos dos Polinômios de Chebyshev

Um dos resultados importantes sobre os Polinômios de Chebyshev, diz respeito aos **zeros e extremos**.

Teorema 2.4.1 *O Polinômio de Chebyshev T_n de grau $n \geq 1$ tem n zeros no intervalo $[-1, 1]$ e nenhum zero fora desse intervalo.*

Prova: Como $T_n(x) = \cos n\theta$, onde $\theta = \arccos x$, ou seja $x = \cos \theta$ e $-1 \leq x \leq 1$, pode pôr-se sem perda de generalidade $0 \leq \theta \leq \pi$, o que torna a relação entre θ e x mais precisa. É claro que $T_n(x)$ é igual a zero para $\theta = \frac{(2i+1)\pi}{2n}$, ou seja

$$x_i = \cos \frac{(2i+1)\pi}{2n},$$

para $i = 0, 1, \dots, n-1$ que são n argumentos distintos entre -1 e 1 e Como $T_n(x)$ tem apenas n zeros, não pode existir qualquer zero fora do intervalo.

Além disso, T_n assume o valor máximo no intervalo $[-1, 1]$ quando $\bar{x}'_k = \cos\left(\frac{k\pi}{n}\right)$ com $T_n(\bar{x}'_k) = (-1)^k$ para cada $k = 0, 1, \dots, n$.

Esta é a principal propriedade dos Polinômios de Chebyshev e recebe vários nomes como propriedade do erro igual e propriedade de igual oscilação.

Devido a esta propriedade o polinômio de Chebyshev é muito importante na aproximação de funções, seja pelo método dos quadrados mínimos ou outros métodos.

Mas em que medida estas propriedades fazem com que a aproximação dos mínimos quadrados $y(x) \simeq a_0T_0(x) + \dots + a_mT_m(x)$, seja superior a aproximações semelhantes usando outros polinômios em vez de $T_k(x)$?

Bem, suponhamos que, para a função $y(x)$ em questão, a série obtida quando m tende ao infinito converge para $y(x)$ e converge suficientemente depressa para que $y(x) - a_0T_0(x) - \dots - a_mT_m(x) \simeq a_{m+1}T_{m+1}(x)$, isto é, o erro cometido por truncatura da série seja essencialmente o primeiro termo omitido. Como $T_{m+1}(x)$ tem a mesma propriedade da igual oscilação, o erro da nossa aproximação flutuará entre a_{m+1} e $-a_{m+1}$ ao longo do intervalo $(-1, 1)$. O erro não será essencialmente maior sobre uma parte do intervalo quando comparada com outra. Esta uniformidade do erro pode ser vista como uma recompensa por aceitar a nada agradável função peso $\omega(x) = \frac{1}{\sqrt{1-x^2}}$.

Definição 2.4.2 *O Polinômio Mônico de Chebyshev é aquele que tem o coeficiente do termo de maior grau igual a um, esse é obtido de um polinômio de Chebyshev, dividindo este por seu coeficiente do termo maior grau, que é 2^{n-1} . E este é denotado por \tilde{T}_n .*

De tal modo:

$$\tilde{T}_0(x) = 1,$$

e

$$\tilde{T}_n(x) = 2^{1-n}T_n(x),$$

para cada $n \geq 1$.

Esses polinômios satisfazem a relação de recorrência:

$$\tilde{T}_{n+1}(x) = x\tilde{T}_n(x) - \frac{1}{4}\tilde{T}_{n-1}(x),$$

para cada $n \geq 2$.

Por causa da relação linear entre \tilde{T}_n e $T_n(x)$, os zeros de \tilde{T}_n também ocorrem em:

$$\bar{x}_k = \cos\left(\frac{2k-1}{2n}\pi\right) \text{ para cada } k = 1, 2, \dots, n.$$

e os valores extremos de \tilde{T}_n ocorrem em:

$$\bar{x}'_k = \cos\left(\frac{k\pi}{n}\right) \text{ com } T_n(\bar{x}'_k) = (-1)^k \text{ para cada } k = 0, 1, \dots, n.$$

Denotaremos $\tilde{\Pi}_n$ para o conjunto de todos os Polinômios Mônicos de grau n .

Temos então uma importante propriedade de minimização que distingue os polinômios \tilde{T}_n dos outros membros de $\tilde{\Pi}_n$.

Propriedade 2.4.1 *O polinômio \tilde{T}_n , quando $n \geq 1$, tem a propriedade:*

$$\frac{1}{2^{n-1}} = \max_{x \in [-1, 1]} |\tilde{T}_n(x)| \leq \max_{x \in [-1, 1]} |P_n(x)|$$

para todo $P_n(x) \in \tilde{\Pi}_n$.

Além disso, a igualdade pode ocorrer somente se $P_n(x) = \tilde{T}_n$.

Esse resultado é usado para responder a questão de onde colocar os nós de interpolação para minimizar o erro na interpolação de Lagrange.

A forma do erro do Polinômio de Lagrange aplicado ao intervalo $[-1, 1]$ afirma que se x_0, x_1, \dots, x_n são números distintos no intervalo $[-1, 1]$ e $f \in \mathcal{C}^{n+1}[-1, 1]$, então, para cada $x \in [-1, 1]$, existe um número $\xi(x)$ em $(-1, 1)$ com

$$f(x) - P(x) = \frac{f^{n+1}(\xi(x))}{(n+1)!} (x - x_0) \dots (x - x_n),$$

onde P denota o polinômio Interpolador de Lagrange.

Como não há controle sobre $\xi(x)$, minimizar o erro por uma escolha inteligente dos nós x_0, x_1, \dots, x_n é equivalente a encontrar os x_0, x_1, \dots, x_n que minimizam a quantidade

$$|(x - x_0)(x - x_1) \dots (x - x_n)|$$

para valores de x_0, x_1, \dots, x_n no intervalo $[-1, 1]$. Desde que $(x - x_0)(x - x_1) \dots (x - x_n)$ seja um polinômio mônico de grau $n + 1$, o mínimo é obtido quando:

$$(x - x_0)(x - x_1) \dots (x - x_n) = \tilde{T}_{n+1}(x).$$

Como x_k é escolhido para ser o $(k + 1)$ -ésimo zero de \tilde{T}_{n+1} , para cada $k = 0, 1, \dots, n$, isto é,

$$x_{k+1} = \cos\left(\frac{2k + 1}{2(n + 1)}\pi\right)$$

o valor máximo de $|(x - x_0)(x - x_1)\dots(x - x_n)|$ é minimizado. Desde que $\max_{x \in [-1, 1]} |\tilde{T}_n(x)| = \frac{1}{2^n}$, isso também implica que

$$\frac{1}{2^n} = \max_{x \in [-1, 1]} |(x - \tilde{x}_1)(x - \tilde{x}_2)\dots(x - \tilde{x}_{n+1})| \leq \max_{x \in [-1, 1]} |(x - x_0)(x - x_1)\dots(x - x_n)|$$

para alguma escolha de $x_0, x_1, x_2, \dots, x_n$ no intervalo $[-1, 1]$.

2.4.2 Minimizando o Erro na Interpolação de Lagrange

Se P_n é o polinômio interpolador de grau n , com nós nas raízes de T_{n+1} , então, para alguma $f \in \mathcal{C}^{n+1}[-1, 1]$,

$$\max_{x \in [-1, 1]} |f(x) - P(x)| \leq \frac{1}{2^n(n + 1)!} \max_{x \in [-1, 1]} |f^{(n+1)}(x)|.$$

A técnica de escolher pontos que minimizem o erro pode ser facilmente estendida para um intervalo genérico fechado $[a, b]$, usando a mudança de variável

$$\tilde{x} = \frac{1}{2} \left[(b - a)x + a + b \right]$$

para transformar os números do intervalo $[-1, 1]$ nos seus correspondentes no intervalo $[a, b]$.

2.5 Polinômios de Hermite

Os polinômios de Hermite são ortogonais em $] \infty_-, \infty_+[$, em relação a função peso $\omega(x) = e^{-x^2}$.

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

Estes polinômios podem ser gerados pela seguinte propriedade recursiva.

Propriedade 2.5.1

$$H_n(x) = 2xH_{n-1}(x) - H'_{n-1}(x), n \geq 2$$

onde $H_0(x) = 1$ e $H_1(x) = 2x$.

O H_n tem grau n . Os 6 primeiros polinômios de Hermite são:

$$H_0(x) = 1, H_1(x) = 2x, H_2(x) = 4x^2 - 2,$$

$$H_3(x) = 8x^3 - 12x, H_4(x) = 16x^4 - 48x^2 + 12 \text{ e}$$

$$H_5(x) = 32x^5 - 160x^3 + 120x$$

Veja-os no gráfico abaixo:

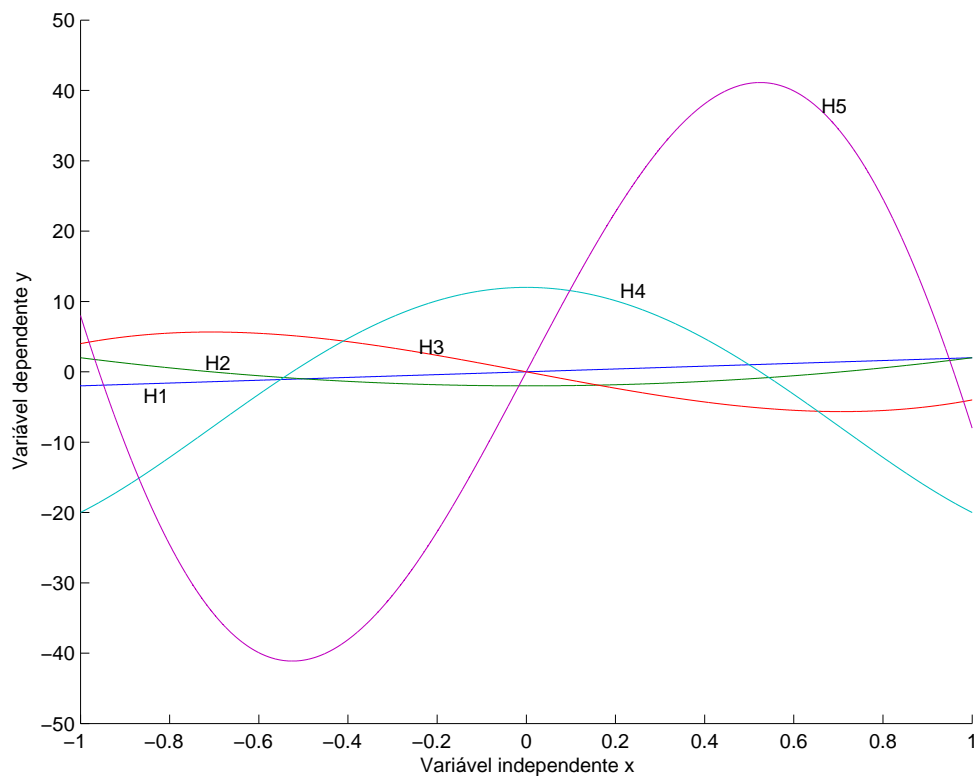


Figura 2.4: Polinômios de Hermite

2.6 Polinômios de Laguerre

Os polinômios de Laguerre são ortonormais em $[0, \infty_+]$, em relação a função peso $\omega(x) = e^{-x}$.

$$(L_n, L_m) = \int_0^{\infty_+} \omega(x) L_n(x) L_m(x) dx = \begin{cases} 0 & \text{se } n \neq m \\ 1 & \text{se } n = m \end{cases}$$

Os 5 primeiros polinômios de Laguerre são:

$$L_0 = 1, L_1 = 1 - x, L_2(x) = 1 - 2x + \frac{x^2}{2},$$

$$L_3(x) = 1 - 3x + \frac{3x^2}{2} - \frac{1}{6}x^3 \text{ e } L_4(x) = 1 - 4x + 3x^2 - \frac{2x^3}{3} + \frac{x^4}{24}$$

Veja-os abaixo:

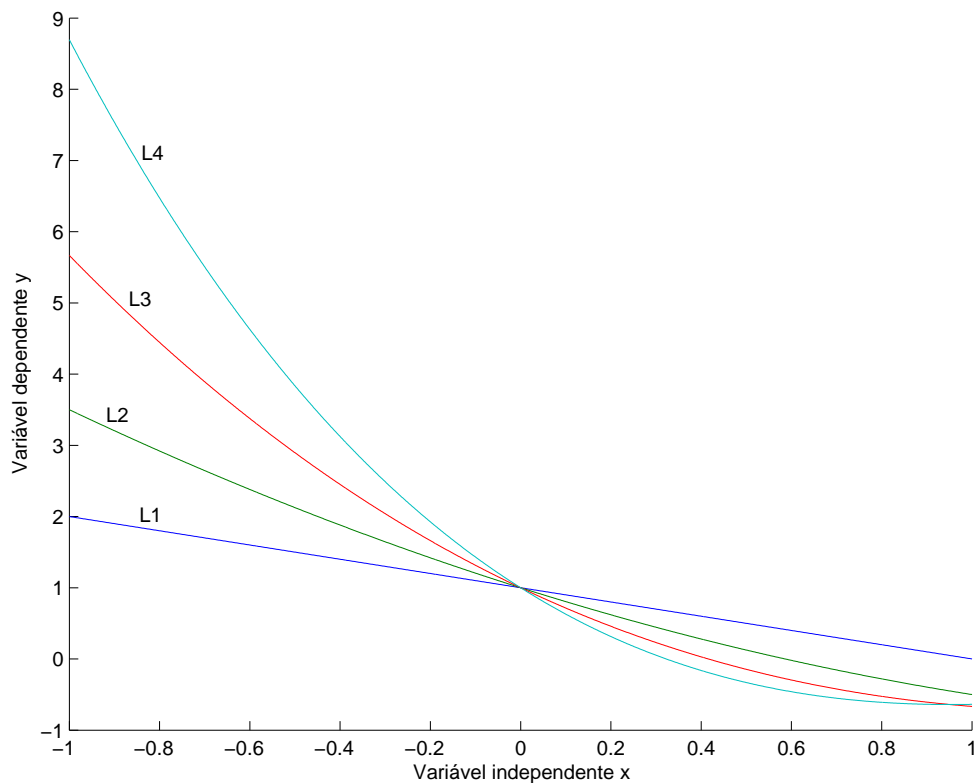


Figura 2.5: Polinômios de Laguerre

Capítulo 3

Aproximação por Funções Racionais

3.1 Introdução

As funções racionais são quocientes de polinômios e por isso formam uma classe maior que a dos polinômios, o que de certo modo aumenta as perspectivas de conduzir a melhores aproximações. Por exemplo, funções com pólos dificilmente poderão ser aproximadas por polinômios que não têm singularidades; tais funções são o principal alvo da aproximação por funções racionais. Mas mesmo com funções sem singularidades, pode-se em certas situações ser preferível a aproximação por funções racionais.

Uma função racional r de grau N tem a forma

$$r(x) = \frac{p(x)}{q(x)}$$

onde p e q são polinômios cuja soma dos graus é N .

Segundo SCHEID,

as funções racionais cujo numerador e o denominador tem o mesmo ou aproximadamente o mesmo grau, geralmente produzem aproximações com resultados superiores aos dos métodos polinomiais, para a mesma quantidade de esforço computacional.

As funções racionais têm a vantagem adicional de permitir a eficiente aproximação

de funções que têm infinitas descontinuidades próximas do intervalo de aproximação. A aproximação polinomial é geralmente inaceitável nessa situação.

Supondo que r é uma função racional de grau $N = n + m$ da forma

$$r(x) = \frac{p(x)}{q(x)} = \frac{p_0 + p_1x + \dots + p_nx^n}{q_0 + q_1x + \dots + q_mx^m}$$

que é usada para aproximar a função f em um intervalo fechado I contendo o zero. Para r ser definido no zero é necessário que $q_0 \neq 0$. Podemos assumir que $q_0 = 1$, se esse não for o caso simplesmente trocamos $p(x)$ por $p(x)/q_0$ e $q(x)/q_0$. Conseqüentemente seriam $N + 1$ parâmetros $q_1, q_2, \dots, q_m, p_0, p_1, \dots, p_n$ disponíveis para a aproximação de f por r .

A função $r(x)$ pode ser determinada de muitas formas, por exemplo minimizando alguma das normas das diferenças da função a ser aproximada e a função racional, ou de maneira que $r(x)$ interpole alguns pontos de $f(x)$, ou usando a técnica de Padé. Neste trabalho veremos duas formas de fazê-lo.

3.2 Frações Contínuas

Interpolação em certos pontos é uma base para selecionar uma aproximação racional. As principais ferramentas a usar são as frações contínuas e as diferenças recíprocas. As frações contínuas a serem usadas são da forma:

$$y(x) = y_1 + \frac{x - x_1}{\rho_1 + \frac{x - x_2}{\rho_2 - y_1 + \frac{x - x_3}{\rho_3 - \rho_1 + \frac{x - x_4}{\rho_4 - \rho_2}}}} \quad (3.1)$$

que, se necessário, podem ser continuadas. Não é difícil verificar que esta fração particular pode ser escrita como o quociente de dois polinômios quadráticos, ou seja uma função racional. Os coeficientes ρ são determinados de forma a haver interpolação. No caso presente obtém-se

$$\rho_1 = \frac{x_2 - x_1}{y_2 - y_1}$$

$$\rho_2 - y_1 = \frac{\frac{x_3 - x_2}{x_3 - x_1} - \frac{x_2 - x_1}{y_3 - y_1}}{y_2 - y_1}$$

com expressões semelhantes para ρ_3 e ρ_4 .

Exemplo 3.2.1 Determinar a função racional $y(x) = \frac{1}{a+bx}$ tal que $y(1) = 1$ e $y(3) = \frac{1}{2}$.

Solução: Por substituição, obtém-se que $a + b = 1$ e $a + 3b = 2$, ou seja, $a = b = \frac{1}{2}$. Assim a fração pretendida é $y(x) = \frac{2}{1+x}$.

Este problema muito simples explica o fato de que a determinação de uma função racional por interpolação corresponde a resolver um sistema de equações lineares em que as incógnitas são os coeficientes da função que estamos construindo.

Exemplo 3.2.2 Usar uma aproximação racional para interpolar $\tan(1,565)$ a partir dos dados da tabela das diferenças recíprocas, abaixo:

x	$\tan(x)$
1,53	24,498
	0,0012558
1,54	32,461
	-0,033
	0,0006403
	2,7279
1,55	48,078
	-0,022
	-0,4167
	0,0002245
	1,7145
1,56	92,631
	-0,0045
	0,0000086
1,57	12255,8

Solução: Nesta tabela estão indicadas as diferenças recíprocas até a quarta ordem. A interpolação conforme a equação 3.1 é feita por:

$$\tan(1,565) \simeq 24,498 + \frac{1,565 - 1,53}{0,001258 + \frac{1,565 - 1,54}{-24,531 + \frac{1,565 - 1,55}{2,7266 + \frac{1,565 - 1,56}{-0,3837}}}}$$

que resulta $\tan(1,565) = 172,552$ sendo que o correto é $172,5211218$. Este resultado está quase perfeito, o que é notável tendo em conta que estamos bastante próximos de um dos pólos ($\frac{\pi}{2}$) da função tangente.

3.3 Aproximação de Padé

O outro método de escolher uma função racional para uma aproximação desejada é conhecido como Aproximação de Padé. Antes porém vejamos alguns conceitos necessários, para um melhor entendimento do método.

Definição 3.3.1 A solução p de $f(x) = 0$, é dita uma raiz de multiplicidade m de $f(x)$, se $f(x)$ pode ser escrita como $f(x) = (x - p)^m q(x)$, para $x \neq p$, onde $\lim_{x \rightarrow p} q(x) \neq 0$.

Em essência, $q(x)$ representa aquela parte de $f(x)$ que não contribui para a raiz de f .

Teorema 3.3.1 Seja $f \in C^m[a, b]$, f tem uma raiz de multiplicidade m no ponto p se e somente se $0 = f^0(p) = f^1(p) = f^2(p) = \dots = f^{m-1}(p)$, mas $f^m(p) \neq 0$

Demonstração: (Ver Apêndice F).

Definição 3.3.2 A técnica de Aproximação de Padé escolhe $N + 1$ parâmetros de tal modo que $f^{(k)}(0) = r^{(k)}(0)$ para cada $k = 0, 1, 2, \dots, N$, onde f é a função que se deseja aproximar e r é uma função racional de grau N .

A aproximação de Padé é a extensão do Polinômio de Taylor para Funções Racionais. De fato, quando $n = N$ e $m = 0$, a aproximação de Padé é a n -ésima expansão do Polinômio de Taylor no ponto zero ou seja a n -ésima expansão do Polinômio de Maclaurin.

Considere a diferença

$$f(x) - r(x) = f(x) - \frac{p(x)}{q(x)} = \frac{f(x)q(x) - p(x)}{q(x)}$$

$$f(x) - r(x) = \frac{f(x) \sum_{i=0}^{\infty} q_i x^i - \sum_{i=0}^n p_i x^i}{q(x)}$$

e suponha que f tenha a expansão da Série de Maclaurin $f(x) = \sum_{i=0}^{\infty} a_i x^i$.

Então

$$f(x) - r(x) = \frac{\sum_{i=0}^{\infty} a_i x^i \sum_{i=0}^{\infty} q_i x^i - \sum_{i=0}^n p_i x^i}{q(x)} \quad (3.2)$$

O objetivo é escolher as constantes q_1, q_2, \dots, q_m e p_0, p_1, \dots, p_n de tal modo que $f^{(k)}(0) - r^{(k)}(0) = 0$ para cada $k = 0, 1, \dots, N$.

Isso equivale a $f - r$ ter uma raiz de multiplicidade $N + 1$ no ponto zero. Como conseqüência nos escolhemos q_1, q_2, \dots, q_m e p_0, p_1, \dots, p_n de tal modo que o numerador do membro direito da equação 3.2, ou seja,

$$(a_0 + a_1 x + a_2 x^2 + \dots)(1 + q_1 x + \dots + q_m x^m) - (p_0 + p_1 x + \dots + p_n x^n)$$

não tenha termos de grau menor ou igual a N . Para simplificar a notação, definimos $p_{n+1} = p_{n+2} = \dots = p_N = 0$ e $q_{m+1} = q_{m+2} = \dots = q_N = 0$. O coeficiente de x^k é $\left(\sum_{i=0}^k a_i q_{k-i}\right) - p_k$ e a função racional para a Aproximação de Padé, resulta da solução de $N + 1$ equações lineares

$$\sum_{i=0}^k a_i q_{k-i} = p_k, \quad k = 0, 1, \dots, N$$

em $N + 1$ incógnitas $q_1, q_2, \dots, q_m, p_0, p_1, \dots, p_n$.

Exemplo 3.3.1 A expansão da série de Maclaurin para e^{-x} é:

$$\sum_{i=0}^{\infty} \frac{(-1)^i}{i!} x^i$$

Para achar a aproximação de Padé para e^{-x} de grau 5 com $n = 3$ e $m = 2$ requer que encontremos p_0, p_1, p_2, p_3, q_1 e q_2 de modo que os coeficientes de x^k para

$k = 0, 1, 2, 3, 4$ e 5 sejam zero na expressão

$$(1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \dots)(1 + q_1x + q_2x^2) - (p_0 + p_1x + p_2x^2 + p_3x^3)$$

Expandindo e reagrupando os termos, temos:

$$\begin{aligned} x^5 : -\frac{1}{120} + \frac{1}{24}q_1 - \frac{1}{6}q_2 &= 0; & x^2 : \frac{1}{2} - q_1 + q_2 &= p_2 ; \\ x^4 : \frac{1}{24} - \frac{1}{6}q_1 + \frac{1}{2}q_2 &= 0; & x^1 : -1 + q_1 &= p_1 ; \\ x^3 : -\frac{1}{6} + \frac{1}{12}q_1 - q_2 &= 0; & x^0 : 1 &= p_0 ; \end{aligned}$$

a solução do sistema acima é:

$$p_0 = 1, p_1 = -\frac{3}{5}, p_2 = \frac{3}{20}, p_3 = -\frac{1}{60}q_1 = \frac{2}{5} \text{ e } q_2 = \frac{1}{20}$$

assim a aproximação de Padé é:

$$r(x) = \frac{1 - \frac{3}{5}x + \frac{3}{20}x^2 - \frac{1}{60}x^3}{1 + \frac{2}{5}x + \frac{1}{20}x^2}$$

Listando os valores de $r(x)$ e $P_5(x)$, o quinto Polinômio de Maclaurin ($P_5(x) = 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{24} - \frac{x^5}{120}$), a aproximação de Padé é claramente superior nesse exemplo:

Veja como ficam os gráficos de $f(x)$, de sua aproximação racional $S_3(x)$ e de $P_5(x)$ na figura 3.1.

x	e^{-x}	$P_5(x)$	$ e^{-x} - P_5(x) $	$r(x)$	$ e^{-x} - r(x) $
0.2	0.81873075	0.81873067	8.64×10^{-8}	0.81873075	7.55×10^{-9}
0.4	0.67032005	0.67031467	5.38×10^{-6}	0.67031963	4.11×10^{-7}
0.6	0.54881164	0.54875200	5.96×10^{-5}	0.54880763	4.00×10^{-6}
0.8	0.44932896	0.44900267	3.26×10^{-4}	0.44930966	1.93×10^{-5}
1.0	0.36787944	0.36666667	1.21×10^{-3}	0.36781609	6.33×10^{-5}

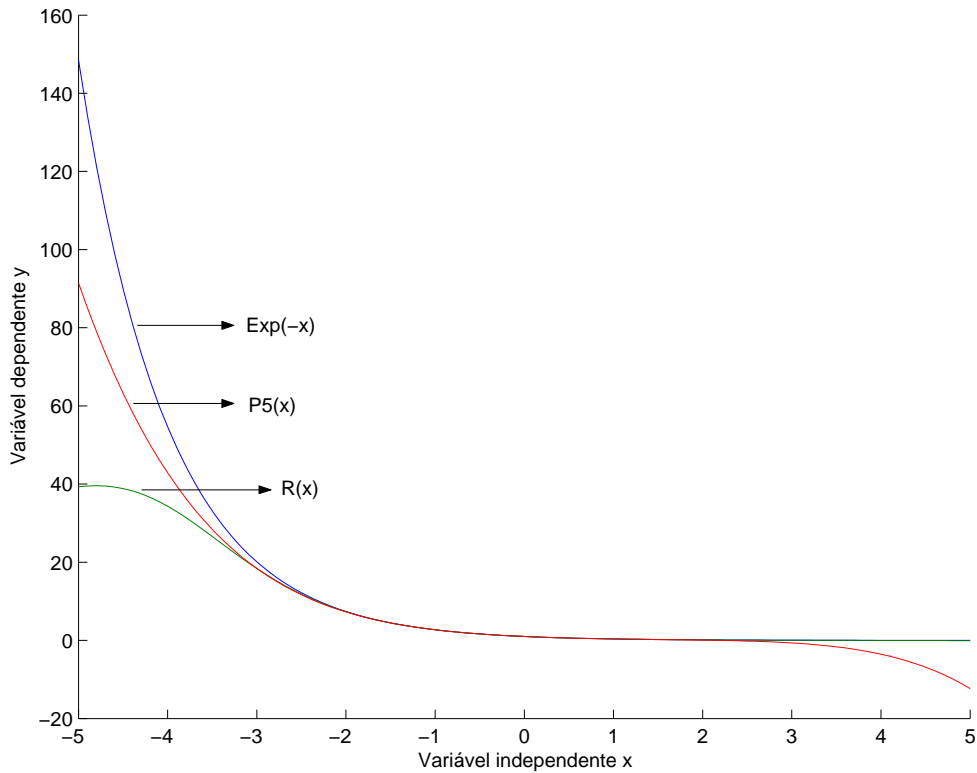


Figura 3.1: Aproximação racional de e^{-x} comparada com a aproximação de Maclaurin

A tabela é interessante para comparar os erros nos dois métodos, outro aspecto importante é o número de operações aritméticas requeridas para calcular valores de $P_5(x)$ e $r(x)$ no exemplo 3.3.1. Usando multiplicações sucessivas, $P_5(x)$ pode ser expressa como:

$$P_5(x) = 1 - x(1 - x(\frac{1}{2} - x(\frac{1}{6} - x(\frac{1}{24} - \frac{1}{120}x))))$$

Assumindo que os coeficientes de $1, x, x^2, x^3, x^4$ e x^5 são representados como decimais, para um simples cálculo de $P_5(x)$ na forma de multiplicações sucessivas, são necessárias cinco multiplicações e cinco adições/subtrações.

Usando multiplicações sucessivas, $r(x)$ é expressa como:

$$r(x) = \frac{1 - x\left(\frac{3}{5} - x\left(\frac{3}{20} - \frac{1}{60}x\right)\right)}{1 + x\left(\frac{2}{5} + \frac{1}{20}x\right)}$$

deste modo para um simples cálculo de $r(x)$ são necessárias cinco multiplicações, cinco adições/subtrações e uma divisão. Assim o esforço computacional parece ser menor na aproximação polinomial. Contudo, podemos reescrever $r(x)$ como uma divisão contínua (Frações contínuas):

$$\begin{aligned} r(x) &= \frac{1 - x\left(\frac{3}{5} - x\left(\frac{3}{20} - \frac{1}{60}x\right)\right)}{1 + x\left(\frac{2}{5} + \frac{1}{20}x\right)} \\ &= \frac{-\frac{1}{3}x^3 + 3x^2 - 12x + 20}{x^2 + 8x + 20} \\ &= -\frac{1}{3}x + \frac{17}{3} + \frac{-\frac{152}{3}x - \frac{280}{3}}{x^2 + 8x + 20} \\ &= -\frac{1}{3}x + \frac{17}{3} + \frac{-\frac{152}{3}}{\frac{x^2 + 8x + 20}{x + \frac{35}{19}}} \end{aligned}$$

ou

$$r(x) = -\frac{1}{3}x + \frac{17}{3} + \frac{-\frac{152}{3}}{x + \frac{117}{19} + \frac{\frac{3125}{361}}{x + \frac{35}{19}}}$$

Escrevendo dessa forma, um simples cálculo de $r(x)$ requer uma multiplicação, cinco adições/subtrações e duas divisões. Se a quantidade de computação requerida para a divisão é aproximadamente o mesmo que para multiplicação, o esforço computacional requerido para avaliação de $P_5(x)$ excede significativamente o esforço requerido para avaliar $r(x)$. Expressando a aproximação função racional nesta forma, ela é chamada

aproximação Frações - Contínuas. Essa é uma técnica clássica de aproximação de corrente interesse porque essa representação é eficiente computacionalmente.

Embora a aproximação por funções racionais no exemplo 3.3.1 dê resultados superiores a aproximação polinomial de mesmo grau, a aproximação tem uma larga variação na exatidão, a aproximação em 0.2 tem uma exatidão dentro da ordem de 8×10^{-9} , enquanto em $x = 1.0$ a aproximação e a função concordam dentro da ordem de 7×10^{-5} . Essa variação na exatidão era esperada, porque a aproximação de Padé é baseada em uma representação de Taylor de e^{-x} , e a representação de Taylor tem uma larga variação de exatidão em $[0.2; 1.0]$.

3.3.1 Aproximação por Funções Racionais Chebyshev

Para obter mais uniformidade na exatidão da aproximação funções racionais, usamos a classe dos Polinômios de Chebyshev, uma classe que exibe um comportamento mais uniforme.

O método geral de Aproximação Função Racional Chebyshev procede-se da mesma maneira que a Aproximação de Padé, exceto que cada termo x^k na Aproximação de Padé é trocado pelo k -ésimo polinômio de Chebyshev $T_k(x)$.

Supondo que queiramos aproximar a função f por uma função racional de grau N , a escrevemos na forma:

$$r(x) = \frac{\sum_{k=0}^n p_k T_k(x)}{\sum_{k=0}^m q_k T_k(x)},$$

onde $N = m + n$ e $q_0 = 1$. Escrevendo $f(x)$ como uma série envolvendo polinômios de Chebyshev:

$$f(x) - r(x) = \sum_{k=0}^{\infty} a_k T_k(x) - \frac{\sum_{k=0}^n p_k T_k(x)}{\sum_{k=0}^m q_k T_k(x)}$$

ou

$$f(x) - r(x) = \frac{\sum_{k=0}^{\infty} a_k T_k(x) \sum_{k=0}^m q_k T_k(x) - \sum_{k=0}^n p_k T_k(x)}{\sum_{k=0}^m q_k T_k(x)}.$$

Os coeficientes $q_1, q_2, \dots, q_m, p_0, p_1, \dots, p_n$ são escolhidos de tal modo que o numerador da fração acima, lado direito, tenha zero como coeficiente para $T_k(x)$ quando $k = 0, 1, 2, \dots, N$.

Isso implica que

$$\left(a_0 T_0(x) + a_1 T_1(x) + \dots \right) \left(T_0(x) + q_1 T_1(x) + \dots + q_m T_m(x) \right) - \left(p_0 T_0(x) + \dots + p_n T_n(x) \right)$$

não tem termos de grau menor ou igual a N .

Surgem dois problemas com o procedimento de Aproximação Racional de Chebyshev que fazem este aparentemente mais difícil de implementar que o método de Padé.

Um ocorre por causa do produto dos Polinômios de Chebyshev. Esse problema é resolvido fazendo uso da relação

$$T_i(x)T_j(x) = \frac{1}{2}[T_{i+j}(x) + T_{|i-j|}(x)]$$

O outro problema envolve a computação da Série de Chebyshev para $f(x)$. Na teoria, isso não é difícil, visto que se

$$f(x) = \sum_{k=0}^{\infty} a_k T_k(x)$$

Então a ortogonalidade dos Polinômios de Chebyshev implicam

$$a_0 = \frac{1}{\pi} \int_{-1}^1 \frac{f(x)}{\sqrt{1-x^2}} dx$$

e

$$a_k = \frac{2}{\pi} \int_{-1}^1 \frac{f(x)T_k(x)}{\sqrt{1-x^2}} dx$$

quando $k \geq 1$.

Embora, essas integrais possam raramente ser avaliadas de forma fechada, uma integração numérica para cada avaliação resolve este problema.

Exemplo 3.3.2 Os primeiros cinco termos da expansão de Chebyshev para e^{-x} são:

$$\begin{aligned}\tilde{P}_5(x) = & 1,266066T_0(x) - 1,130318T_1(x) + 0,271495T_2(x) \\ & - 0,044337T_3(x) + 0,005474T_4(x) - 0,000543T_5(x)\end{aligned}$$

Para determinar a Aproximação Racional de Chebyshev de grau 5 com $n = 3$ e $m = 2$ precisamos escolher q_1, q_2, p_0, p_1, p_2 e p_3 de modo tal que para $k = 0, 1, 2, 3, 4$ e 5 , os coeficientes de $T_k(x)$ sejam zero na expansão:

$$\tilde{P}_5 = \left(T_0(x) + q_1T_1(x) + q_2T_2(x) \right) - \left(p_0T_0(x) + p_1T_1(x) + p_2T_2(x) + p_3T_3(x) \right)$$

usando a relação produto para os Polinômios de Chebyshev e agrupando os termos, temos as equações:

$$T_0 : 1,266066 - 0,565159q_1 + 0,1357485q_2 = p_0$$

$$T_1 : -1,130318 + 1,401814q_1 - 0,583275q_2 = p_1$$

$$T_2 : 0,271495 - 0,58732q_1 + 1,268803q_2 = p_2$$

$$T_3 : -0,044337 + 0,138485q_1 - 0,565431q_2 = p_3$$

$$T_4 : 0,005474 - 0,022440q_1 + 0,1357485q_2 = 0$$

$$T_5 : -0,000543 + 0,002737q_1 - 0,0022169q_2 = 0$$

A solução desse sistema produz a função racional

$$r_T(x) = \frac{1,055265T_0(x) - 0,613016T_1(x) + 0,077478T_2(x) - 0,004506T_3(x)}{T_0(x) + 0,378331T_1(x) + 0,022216T_2(x)}$$

Convertendo o Polinômio de Chebyshev para potências de x temos:

$$r_T(x) = \frac{0,977787 - 0,599499x + 0,154956x^2 - 0,018022x^3}{0,977784 + 0,378331x + 0,044432x^2}$$

Listando e comparando os valores obtemos que a aproximação $r(x)$ é superior a $r_T(x)$ para $x = 0.2$ e $x = 0.4$, mas que o erro máximo para $r(x)$ é $6,33 \times 10^{-5}$ comparado com $9,13 \times 10^{-6}$ para $r_T(x)$.

O método de Chebyshev não produz a melhor Aproximação Função Racional no sentido da aproximação cujo máximo erro é mínimo.

Capítulo 4

Aproximação Polinomial

Trigonométrica

Funções trigonométricas são usadas para aproximar funções que têm um comportamento periódico, estas funções têm a propriedade que para alguma constante T (o período), $f(x + T) = f(x)$ para todo x . Mediante uma transformação de coordenadas podemos transformar um problema geral no caso em que $T = 2\pi$ e assim restringimos a aproximação para o intervalo padrão $[-\pi, \pi]$.

Definição 4.0.3 *Para cada inteiro positivo n , a escolha \mathcal{T}_n dos polinômios trigonométricos de grau menor que ou igual a n é a escolha de todas as combinações lineares de $\{\phi_0, \phi_1, \dots, \phi_{2n-1}\}$ onde*

$$\phi_0(x) = \frac{1}{\sqrt{2\pi}},$$

$$\phi_k(x) = \frac{1}{\sqrt{\pi}} \cos kx,$$

para $k = 1, 2, \dots, n$

e

$$\phi_{n+k}(x) = \frac{1}{\sqrt{\pi}} \text{sen } kx,$$

para cada $k = 1, 2, \dots, n - 1$.

O conjunto $\{\phi_0, \phi_1, \dots, \phi_{2n-1}\}$ é ortonormal em $[-\pi, \pi]$ com respeito a função peso $\omega \equiv 1$. A demonstração desta propriedade é semelhante a dos Polinômios de Chebyshev em $[-1, 1]$.

Para, $k \neq j$ e $j \neq 0$,

$$\int_{-\pi}^{\pi} \phi_{n+k}(x)\phi_j(x)dx = \int_{-\pi}^{\pi} \frac{1}{\sqrt{\pi}} \text{sen } kx \frac{1}{\sqrt{\pi}} \cos jx dx = \frac{1}{\pi} \int_{-\pi}^{\pi} \text{sen } kx \cos jx dx$$

Usando a identidade trigonométrica:

$$\text{sen } kx \cos jx = \frac{1}{2} \text{sen } (k + j)x + \frac{1}{2} \text{sen } (k - j)x,$$

obtemos:

$$\begin{aligned} \int_{-\pi}^{\pi} \phi_{n+k}(x)\phi_j(x)dx &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\text{sen } (k + j)x + \text{sen } (k - j)x \right] dx = \\ &= \frac{1}{2\pi} \left[-\frac{\cos(k + j)x}{k + j} - \frac{\cos(k - j)x}{k - j} \right]_{-\pi}^{\pi} = 0, \end{aligned}$$

visto que $\cos(k + j)\pi = \cos(k + j)(-\pi)$ e $\cos(k - j)\pi = \cos(k - j)(-\pi)$. O resultado anterior também vale quando $k = j$, nesse caso $\text{sen}(k - j)x = \text{sen}0 = 0$.

O procedimento para mostrar a ortogonalidade para as outras possibilidades de $\{\phi_0, \phi_1, \dots, \phi_{2n-1}\}$ é semelhante ao anterior utilizando-se as seguintes identidades trigonométricas:

$$\begin{aligned} \text{sen } j \cos k &= \frac{1}{2} \left[\text{sen}(j + k) + \text{sen}(j - k) \right] \\ \text{sen } j \text{sen } k &= \frac{1}{2} \left[\cos(j - k) - \cos(j + k) \right] \\ \cos j \cos k &= \frac{1}{2} \left[\cos(j + k) + \cos(j - k) \right] \end{aligned}$$

para converter os produtos em somas.

Definição 4.0.4 Dado $f \in \mathcal{C}[-\pi, \pi]$, a aproximação mínimos quadrados por funções \mathcal{T}_n é definida por:

$$S_n(x) = \sum_{k=0}^{2n-1} a_k \phi_k(x), \text{ onde } a_k = \int_{-\pi}^{\pi} f(x) \phi_k(x) dx$$

para cada $k = 0, 1, 2, \dots, 2n - 1$.

O limite de S_n com $n \rightarrow \infty$, é chamado de Série de Fourier de $f(x)$.

Exemplo 4.0.3 Determine o polinômio trigonométrico de \mathcal{T}_n que aproxima $f(x) = |x|$ para $-\pi < x < \pi$.

Este problema requer que encontremos:

$$\begin{aligned} a_0 &= \int_{-\pi}^{\pi} |x| \frac{1}{\sqrt{2\pi}} dx = -\frac{1}{\sqrt{2\pi}} \int_{-\pi}^0 x dx + \frac{1}{\sqrt{2\pi}} \int_0^{\pi} x dx = \\ &= \frac{2}{\sqrt{2\pi}} \int_0^{\pi} x dx = \frac{\sqrt{2}\pi^2}{2\sqrt{\pi}}. \end{aligned}$$

$$a_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} |x| \cos kx dx = \frac{2}{\sqrt{\pi}} \int_0^{\pi} x \cos kx dx = \frac{2}{\sqrt{\pi} k^2} [(-1)^k - 1]$$

Para cada $k = 1, 2, \dots, n$, e os coeficientes a_{n+k} . Os coeficientes a_{n+k} na expansão de Fourier são comumente denotados por b_k , isto é, $a_{n+k} = b_k$ para $k = 1, 2, \dots, n - 1$.

No nosso exemplo temos:

$$b_k = \frac{1}{\sqrt{\pi}} \int_{-\pi}^{\pi} |x| \text{sen } kx dx = 0 \text{ para cada } k = 1, 2, \dots, n - 1,$$

visto que o integrando é uma função ímpar.

O polinômio trigonométrico de \mathcal{T}_n aproximando f é, portanto:

$$S_n = \frac{\pi}{2} + \frac{2}{\pi} \sum_{k=0}^n \frac{(-1)^k - 1}{k^2} \cos kx.$$

Os primeiros polinômios trigonométricos para $f(x) = |x|$ são:

$$S_0 = \frac{\pi}{2}$$

$$S_1 = \frac{\pi}{2} - \frac{4}{\pi} \cos x$$

$$S_2 = S_1$$

$$S_3 = \frac{\pi}{2} - \frac{4}{\pi} \cos x - \frac{4}{9\pi} \cos 3x$$

Veja-os na figura abaixo:

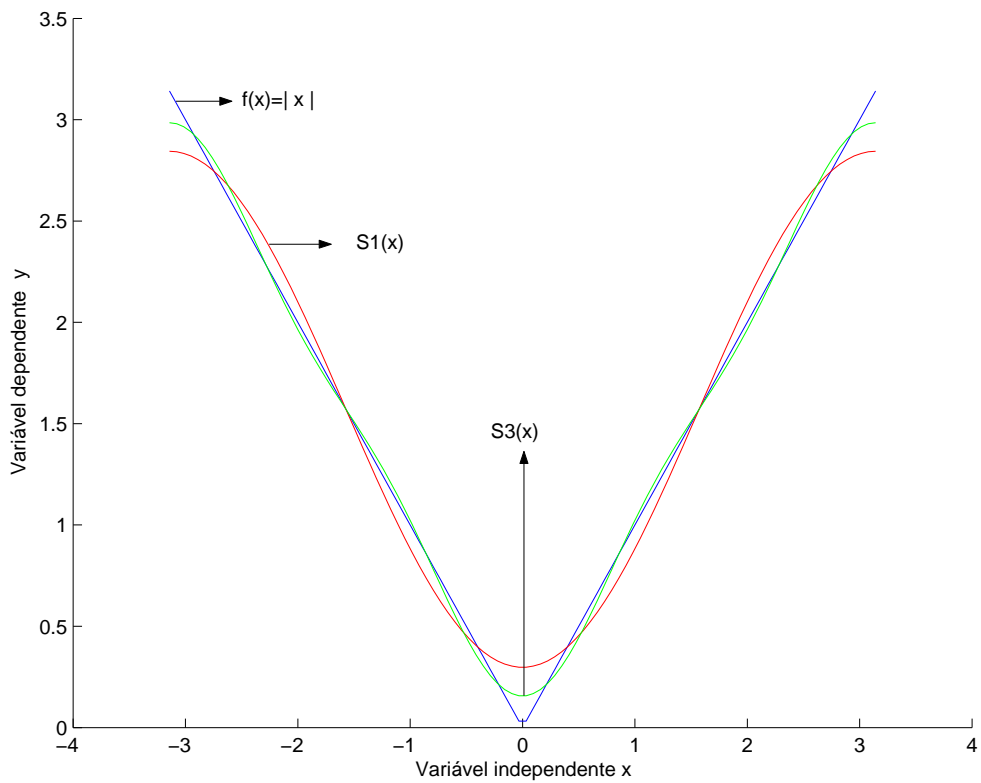


Figura 4.1: Polinômios trigonométricos

A *Série de Fourier* para f é:

$$S(x) = \lim_{n \rightarrow \infty} S_n(x) = \frac{\pi}{2} + \frac{2}{\pi} \sum_{k=0}^{\infty} \frac{(-1)^k - 1}{k^2} \cos kx$$

Visto que $|\cos kx| \leq 1$ para cada k e x , a série converge e $S(x)$ existe para todo número real x .

Há uma analogia discreta para Série de Fourier que é útil para a aproximação por quadrados mínimos e interpolação de grande quantidade de dados igualmente espaçados. É dado um conjunto $\{(x_j, y_j)\}_{j=0}^{2m-1}$ com um número $2m$ par de pontos numa partição de um intervalo fechado. Por conveniência, assumimos que o intervalo seja $[-\pi, \pi]$ e então:

$$x_j = -\pi + \left(\frac{j}{m}\right)\pi \text{ para cada } j = 0, 1, \dots, 2m - 1$$

Se esse não for o caso, uma simples transformação poderá ser usada para transformar os dados para essa forma.

Para um fixado $n < m$, considere o conjunto de \hat{T}_n consistindo de todas as combinações lineares de $\{\hat{\phi}_0, \hat{\phi}_k, \hat{\phi}_{k+n}\}$, onde

$$\hat{\phi}_0(x) = \frac{1}{2},$$

$$\hat{\phi}_k(x) = \cos kx, \text{ para cada } k = 1, 2, \dots, n,$$

e

$$\hat{\phi}_{k+n}(x) = \sin kx, \text{ para cada } k = 1, 2, \dots, n - 1.$$

A meta é determinar a combinação linear dessas funções para a qual

$$E(S_n) = \sum_{j=0}^{2m-1} \{y_j - S_n(x_j)\}^2$$

é minimizado. Isto é, desejamos minimizar o erro total

$$E(S_n) = \sum_{j=0}^{2m-1} \left(y_j - \left[\frac{a_0}{2} + a_n \cos kx_j + \sum_{k=1}^{n-1} (a_k \cos kx_j + a_{n+k} \sin kx_j) \right] \right)^2.$$

A determinação das constantes é simplificada pelo fato de que o conjunto é ortogonal com respeito a soma sobre os pontos igualmente espaçados $\{x_j\}_{j=0}^{2m-1}$ em $[-\pi, \pi]$.

E isso resulta que, para cada $k \neq l$

$$\sum_{j=0}^{2m-1} \hat{\phi}_k(x_j) \hat{\phi}_l(x_j) = 0.$$

A ortogonalidade segue do fato que se r e m são inteiros positivos com $r < 2m$, temos:

$$\sum_{j=0}^{2m-1} \cos rx_j = 0 \text{ e } \sum_{j=0}^{2m-1} \sin rx_j = 0$$

Para obter as constantes a_k para $k = 0, 1, \dots, n$ e $b_k \equiv a_{n+k}$ para $k = 1, 2, \dots, n-1$ no somatório:

$$S_n(x) = \frac{a_0}{2} + a_n \cos nx + \sum_{k=1}^{n-1} (a_k \cos kx + b_k \sin kx),$$

minimizamos a soma dos Quadrados Mínimos

$$E(a_0, \dots, a_n, b_1, \dots, b_{n-1}) = \sum_{j=0}^{2m-1} [y_j - S_n(x_j)]^2$$

fixando igual a zeros as derivadas parciais de E com respeito aos a'_k s e b'_k s. Isso implica que

$$a_k = \frac{1}{m} \sum_{j=0}^{2m-1} y_j \cos kx_j, \text{ para cada } k = 0, 1, \dots, n.$$

e

$$b_k = \frac{1}{m} \sum_{j=0}^{2m-1} y_j \sin kx_j, \text{ para cada } k = 1, 2, \dots, n-1.$$

Exemplo 4.0.4 *Seja $f(x) = x^4 - 3x^3 + 2x^2 - \tan x(x-2)$. Encontrar a aproximação dos Mínimos Quadrados para os dados $\{(x_j, y_j)\}_{j=0}^9$, onde $x_j = \frac{j}{5}$ e $y_j = f(x_j)$. Requer-se que os pontos $\{(x_j, y_j)\}_{j=0}^9$ primeiramente sejam transformados de $[0, 2]$ para $[-\pi, \pi]$. A transformação linear requerida é $z_j = \pi(x_j - 1)$ e a transformação dos dados é da forma*

$$\{(z_j, f(1 + \frac{z_j}{\pi}))\}_{j=0}^9,$$

conseqüentemente o Polinômio Trigonométrico dos Mínimos Quadrados é:

$$S_3(z) = \frac{a_0}{2} + a_3 \cos 3z + \sum_{k=1}^2 (a_k \cos kz + b_k \operatorname{sen} kz)$$

onde

$$a_k = \frac{1}{5} \sum_{j=0}^9 f\left(1 + \frac{z_j}{\pi}\right) \cos kz_j, \text{ para } k=0,1,2 \text{ e } 3$$

e

$$b_k = \frac{1}{5} \sum_{j=0}^9 f\left(1 + \frac{z_j}{\pi}\right) \operatorname{sen} kz_j, \text{ para } k=1 \text{ e } 2.$$

avaliando os somatórios produz-se a aproximação

$$S_3(z) = 0,76201 + 0,77177 \cos z + 0,017432 \cos 2z + 0,0065673 \cos 3z - \\ 0,38676 \operatorname{sen} z + 0,047816 \operatorname{sen} 2z$$

fazendo a conversão para voltar a variável x , temos:

$$S_3(x) = 0,76201 + 0,77177 \cos \pi(x-1) + 0,017432 \cos 2\pi(x-1) + \\ 0,0065673 \cos 3\pi(x-1) - 0,38676 \operatorname{sen} \pi(x-1) + 0,047816 \operatorname{sen} 2\pi(x-1)$$

a tabela abaixo lista os valores de $f(x)$ e $S_3(x)$

x	$f(x)$	$S_3(x)$	$ f(x) - S_3(x) $
0.125	0.26440	0.24060	2.38×10^{-2}
0.375	0.84081	0.85154	1.07×10^{-2}
0.625	1.36150	1.36248	9.74×10^{-4}
0.875	1.61282	1.60406	8.75×10^{-3}
1.125	1.36672	1.37566	8.94×10^{-3}
1.375	0.71697	0.71545	1.52×10^{-3}
1.625	0.07909	0.06929	9.80×10^{-3}
1.875	-0.14576	-0.12302	2.27×10^{-2}

Mas veja também o resultado da aproximação no gráfico:

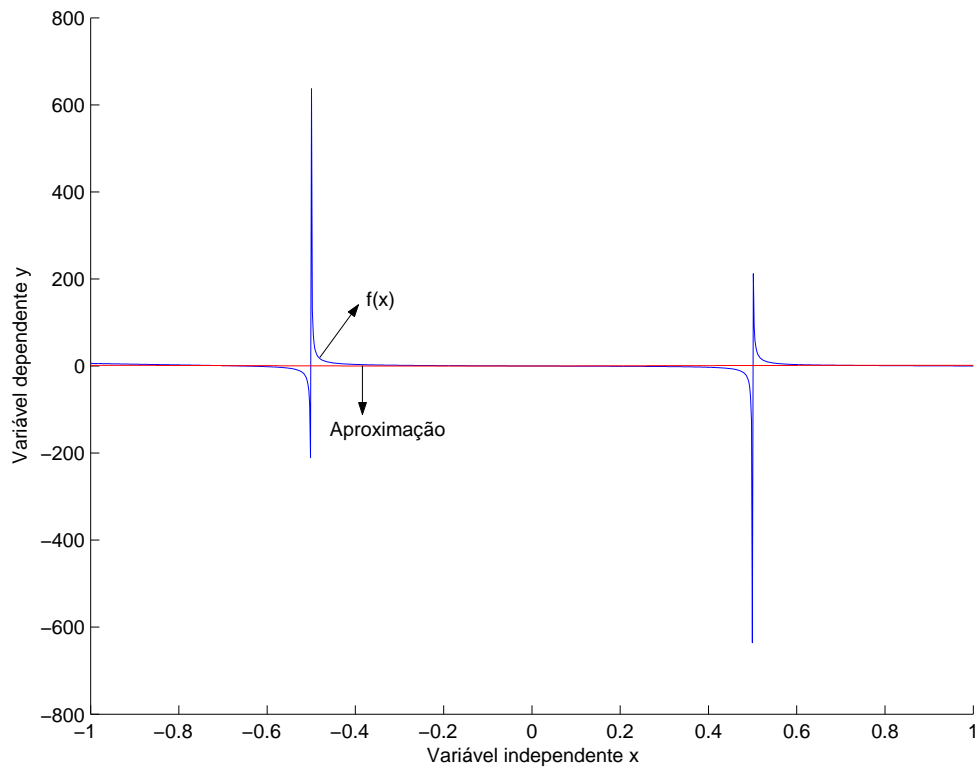


Figura 4.2: Aproximação trigonométrica de $f(x) = x^4 - 3x^3 + 2x^2 - \tan(x(x - 2))$

Conclusão

O presente trabalho introduziu vários métodos para a aproximação de funções e grande parte da teoria desses métodos. A partir desse estudo não podemos afirmar que existe dentre os métodos estudados um que seja mais eficiente, mas existem sim características nas funções que fazem com que a aproximação desta em um determinado intervalo seja melhor por um determinado método. Alguns exemplos ao longo deste trabalho ilustraram as características de cada método.

Uma continuação deste trabalho poderia fazer uma associação entre os métodos apresentados e também com a Aproximação Polinomial Minimax ou de Chebyshev, inclusive usando outras funções além de polinômios, o que restringiria bastante o tipo de função a ser aproximada, mas poderia melhorar a qualidade das aproximações. Outra possibilidade seria aprofundar o estudo dos polinômios ortogonais tendo em vista sua grande utilidade na matemática em áreas como a Teoria de Números, Equações Diferenciais, Frações Contínuas, Estabilidade Numérica e Algoritmos Rápidos e ainda estudar os aspectos computacionais em softwares como Matlab e Maple.

E finalmente, este trabalho foi uma ótima oportunidade de relacionar diversos assuntos abordados durante a graduação e de observar como a teoria e a prática estão intimamente ligadas e são aspectos muito importantes na matemática. Assim o objetivo desta disciplina foi atingido plenamente.

Referências Bibliográficas

- [1] BURDEN, R. L. and FAIRES, J. D. *Analysis Numerical*, New York, Fourth Edition, 1989.
- [2] SCHEID, F. *Análise Numérica*, 2^a. Edição, portugal, Makron Books, 1991.
- [3] CUNHA, C. *Métodos Numéricos para Engenharia e Ciências Aplicadas*, Campinas, São Paulo, Editora da Unicamp, 1993.
- [4] LAY, D. C. *Álgebra Linear e suas aplicações*, Rio de Janeiro, Rio de Janeiro, LTC Editora, 1999.
- [5] SPERANDIO, D. , MENDES, J. T. e MONKEN E SILVA, L. H. *Cálculo Numérico, Características Matemáticas e Computacionais dos Métodos Numéricos*, São Paulo, 2003.
- [6] MACHADO, M. J. S. *Ajuste de Curvas pelo Método dos Quadrados Mínimos usando Splines*, trabalho de Conclusão de Curso, UFSC, Florianópolis ,2003.

Apêndice A

Norma de Vetores e Matrizes

A definição de norma a seguir nos fornece uma maneira de medir o comprimento de um vetor. Existem várias maneiras de efetuar uma medida da grandeza de um vetor ou matriz. Introdz-se, a seguir, o conceito geral de norma de vetores e depois de matrizes.

Definição A.0.5 *Seja V um espaço vetorial sobre F . Uma norma em V é uma função real de V em \mathbb{R} , tal que $X \in V \rightarrow \|X\| \in \mathbb{R}$, para quaisquer que sejam $X, Y \in V$ e para todo escalar $c \in F$, tem-se:*

- (i) $\|X\| \geq 0$ (positividade); $\|X\| = 0$, se e somente se, $X = \mathbf{0}$ (separação);
- (ii) $\|cX\| = |c| \cdot \|X\|$ (homogeneidade), e
- (iii) $\|X + Y\| \leq \|X\| + \|Y\|$ (desigualdade triangular).

As normas mais usadas são as seguintes:

- **Norma um:** $\|X\|_1 = \sum_{i=1}^n |x_i|$,
- **Norma Euclidiana:** $\|X\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}}$ e
- **Norma do Máximo:** $\|X\|_\infty = \max_{1 \leq i \leq n} |x_i|$.

Essas três normas são casos particulares da norma:

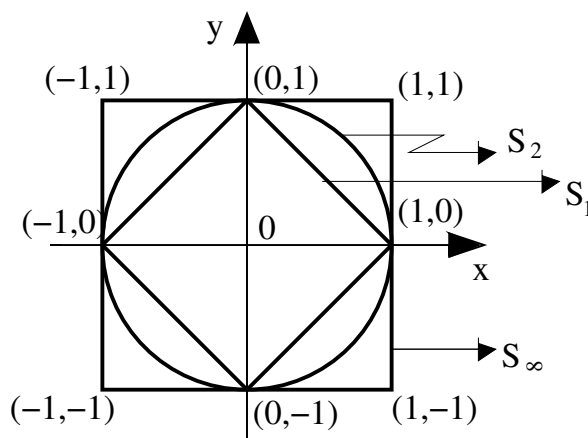
$$\|X\|_p = (|x_i|)^{\frac{1}{p}}$$

onde $p = 1$ para a norma um, $p = 2$ para a norma euclidiana e $p \rightarrow \infty$ para a norma do máximo.

Exemplo A.0.5 Considere o vetor $X = (1, 0, -1, 2)$; então, $\|X\|_1 = 4$; $\|X\|_2 = \sqrt{6}$; $\|X\|_\infty = 2$. Com o objetivo de fornecer uma representação geométrica para essas normas, considere o conjunto

$$S_p = \{x \in \mathbb{R}^2; \|X\|_p = 1\}, p = 1, 2, \infty.$$

A figura abaixo ilustra as bolas unitárias, ou seja, bolas de centro 0 e raio 1, que correspondem às normas definidas em \mathbb{R}^2 nos casos notáveis $p = 1, p = 2, p \rightarrow \infty$.



O quadrado de vértices $(1,0), (0,1), (-1,0), (0,-1)$ corresponde a S_1 , ou seja, à bola unitária dada por $\|X\|_1 \leq 1$. O disco de centro na origem e raio 1 corresponde à bola unitária $\|X\|_2 \leq 1$, S_2 , e o quadrado de vértices $(1,1), (-1,1), (-1,-1), (1,-1)$ corresponde a S_∞ , ou seja, à bola unitária dada por $\|X\|_\infty \leq 1$.

As normas euclidiana, um e do máximo para vetores no \mathbb{R}^n se relacionam da seguinte maneira:

$$\|X\|_\infty \leq \|X\|_1 \leq n \cdot \|X\|_\infty,$$

$$\|X\|_\infty \leq \|X\|_2 \leq n^{\frac{1}{2}} \cdot \|X\|_\infty ,$$

que indica a equivalência entre normas.

Como se observa, a norma euclidiana fornece o comprimento de um vetor. Logo, a norma um e a do máximo fornecem também alguma medida de comprimento de vetor.

Definição A.0.6 Uma seqüência de vetores $\{X_1, X_2, \dots, X_n, \dots\}$ em F^n é dita convergente para um vetor X se, e somente se, $\lim_{n \rightarrow \infty} \|X_n - X\| = 0$.

Vejamos a definição de norma de matriz:

Definição A.0.7 A norma de uma matriz $A \in C[m, n]$ é uma função real $A \in C[m, n] \rightarrow \|A\| \in \mathbb{R}$, tal que quaisquer $A, B \in C[m, n]$ e para todo escalar $c \in F$ as seguintes condições são satisfeitas:

- (i) $\|A\| \geq 0$ e $\|A\| = 0$, se, e somente se, $A = 0$;
- (ii) $\|cA\| = |c| \cdot \|A\|$;
- (iii) $\|A + B\| \leq \|A\| + \|B\|$, e (iv) $\|AB\| \leq \|A\| \cdot \|B\|$.

Definição A.0.8 Uma norma matricial é dita associada a uma norma vetorial se, para toda matriz $A \in C[n, n]$ e todo vetor $X \in C^n$ vale

$$\|AX\| \leq \|A\| \cdot \|X\|.$$

Para qualquer norma de vetor, existe uma norma matricial associada. De fato, escolhida uma norma $\|X\|$ vetorial, a norma matricial a ela associada pode ser definida por

$$\|A\| = \max_{X \neq 0} \frac{\|AX\|}{\|X\|}.$$

As normas matriciais usuais são:

- **Norma Euclidiana:** $\|A\|_E = \left[\sum_{j=1}^n \sum_{i=1}^n |a_{ij}|^2 \right]^{\frac{1}{2}}$;
- **Norma Um:** $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$, e

- **Norma do Máximo:** $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$.

Definição A.0.9 *Uma seqüência de matrizes $\{A_n\}_{n=1}^\infty$ converge para uma matriz A se, e somente se, $\lim_{n \rightarrow \infty} \|A_n - A\| = 0$.*

Apêndice B

Condicionamento de Matrizes

A definição de condicionamento de matrizes surge naturalmente ao fazer uma análise do erro dos métodos numéricos que permitem resolver $Ax = b$, examinando a estabilidade da solução x relativa a uma pequena perturbação do membro direito b . Seja A uma matriz não-singular, seja x a solução do sistema $Ax = b$, seja \hat{x} solução do problema perturbado $A\hat{x} = \hat{b}$. Restando estes dois sistemas de equações, tem-se:

$$A(x - \hat{x}) = b - \hat{b}$$

$$x - \hat{x} = A^{-1}(b - \hat{b})$$

Logo das propriedades da norma matricial se obtém:

$$\|x - \hat{x}\| = \|A^{-1}(b - \hat{b})\| \leq \|A^{-1}\| \cdot \|b - \hat{b}\|$$

e dividindo por $\|x\|$ se tem que

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|b - \hat{b}\|}{\|x\|} = \|A\| \cdot \|A^{-1}\| \frac{\|b - \hat{b}\|}{\|A\| \cdot \|x\|}$$

Logo a desigualdade $\|b\| = \|Ax\| \leq \|A\| \cdot \|x\|$, obtemos

$$\frac{\|x - \hat{x}\|}{\|x\|} \leq \|A\| \cdot \|A^{-1}\| \frac{\|b - \hat{b}\|}{\|b\|}$$

Mas para uma escolha adequada de b e \hat{b} , esta última desigualdade pode chegar a ser uma igualdade. O número

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

se denomina condição de A . Quando este número é muito grande, a solução de $Ax = b$ pode ser extremamente sensível a pequenas perturbações de b , e se diz que o sistema está ***mal condicionado***. Inversamente quando o número de condicionamento é pequeno, se diz que o sistema está ***bem condicionado***.

Apêndice C

Processo de Gram-Schmidt

O processo de Gram-Schmidt é um algoritmo simples para produzir uma base ortogonal ou ortonormal para qualquer subespaço não-nulo de \mathbb{R}^n .

Veja este problema:

Seja W o subespaço gerado por x_1 e x_2 , ou seja $W = \text{Span}\{x_1, x_2\}$, onde

$$x_1 = \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix} \text{ e } x_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}.$$

Construir uma base ortogonal $\{v_1, v_2\}$ para W .

Solução: Veja a figura que ilustra o subespaço W , junto com x_1 , x_2 e a projeção p de x_2 sobre x_1 . A componente de x_2 ortogonal a x_1 é $x_2 - p$, que está em W , pois é formado por x_2 e um múltiplo de x_1 .

Sejam

$$v_1 = x_1$$
$$v_2 = x_2 - p = x_2 - \frac{x_2 \cdot x_1}{x_1 \cdot x_1} x_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} - \frac{15}{45} \begin{pmatrix} 3 \\ 6 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$$

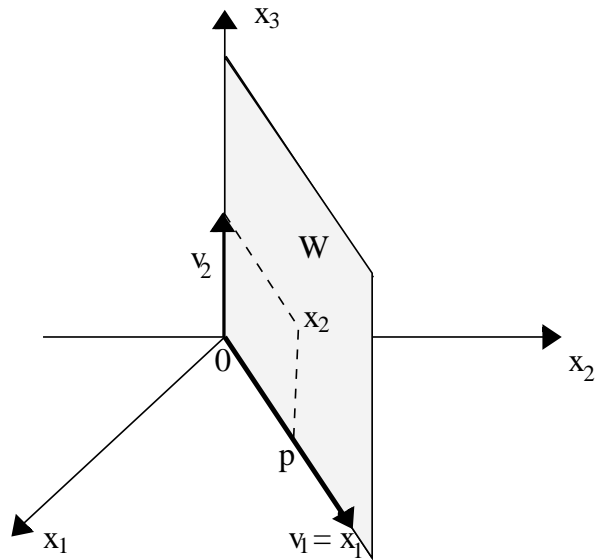


Figura C.1: Processo de Gram-Schmidt

Então, $\{v_1, v_2\}$ é um conjunto ortogonal de vetores não-nulos de W . Como $\dim W = 2$, $\{v_1, v_2\}$ é uma base para W .

Agora veja este problema que ilustra de maneira mais completa o que acontecerá no processo de Gram-Schmidt.

Sejam $x_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$, $x_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ e $x_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix}$. É fácil ver que $\{x_1, x_2, x_3\}$ é um

conjunto linearmente independente, logo forma uma base para um subespaço W de \mathbb{R}^4 .

Como construir uma base ortogonal para W ?

Solução:

Passo 1: Sejam $v_1 = x_1$ e $W_1 = \text{Span}\{x_1\} = \text{Span}\{v_1\}$.

Passo 2: Seja v_2 o vetor obtido subtraindo-se de x_2 sua projeção sobre o subespaço W_1 , isto é,

$$\begin{aligned} v_2 &= x_2 - \text{proj}_W x_2 = \\ &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 = \end{aligned}$$

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{3}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}$$

Como no problema apresentado anteriormente, v_2 é a componente de x_2 ortogonal a x_1 e $\{v_1, v_2\}$ é uma base ortogonal para o subespaço W_2 gerado por x_1 e x_2 .

Passo 2'(Opcional): Se apropriado, multiplique v_2 por um escalar para simplificar cálculos posteriores. Como v_2 tem elementos contendo frações, é conveniente multiplicá-lo por 4 e substituir v_1, v_2 pela base

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \text{ e } v_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Passo 3: Seja v_3 o vetor obtido subtraindo-se de x_3 sua projeção sobre o subespaço W_2 . Use a base $\{v_1, v_2\}$ para calcular a projeção sobre W_2 :

$$\begin{aligned} \text{proj}_{W_2} x_3 &= \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 + \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 = \\ &= \frac{2}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{2}{12} \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ \frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \end{aligned}$$

Então v_3 é a componente de x_3 ortogonal a W_2 , isto é,

$$\begin{aligned} v_3 &= x_3 - \text{proj}_{W_2} x_3 = \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{2}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{pmatrix} \end{aligned}$$

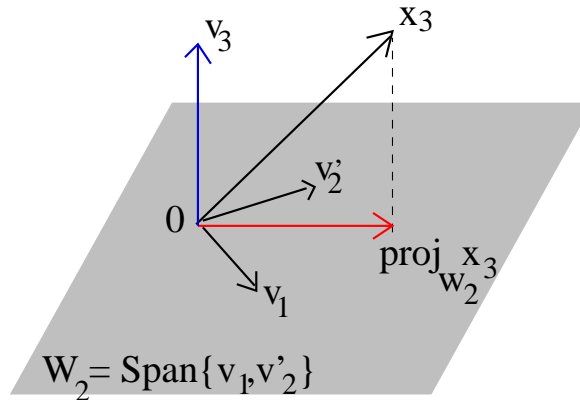


Figura C.2: Processo de Gram-Schmidt

Veja a figura acima para um diagrama dessa construção. Note que v_3 pertence a W , já que tanto x_3 como $\text{proj}_{W_2} x_3$ estão em W . Então $\{v_1, v'_2, v_3\}$ é um conjunto ortogonal de vetores não-nulos e portanto, um conjunto linearmente independente, em W . Note que W tem dimensão três, já que tem uma base com três vetores. Logo $\{v_1, v'_2, v_3\}$ é uma base ortogonal para W .

Assim podemos definir genericamente o processo de Gram-Schmidt como: dada uma base $\{x_1, \dots, x_p\}$ para um subespaço W de \mathbb{R}^n , defina

$$\begin{aligned}
 v_1 &= x_1 \\
 v_2 &= x_2 - \frac{x_2 \cdot v_1}{v_1 \cdot v_1} v_1 \\
 v_3 &= x_3 - \frac{x_3 \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_3 \cdot v_2}{v_2 \cdot v_2} v_2 \\
 &\quad \cdot \\
 &\quad \cdot \\
 &\quad \cdot \\
 v_p &= x_p - \frac{x_p \cdot v_1}{v_1 \cdot v_1} v_1 - \frac{x_p \cdot v_2}{v_2 \cdot v_2} v_2 - \dots - \frac{x_p \cdot v_{p-1}}{v_{p-1} \cdot v_{p-1}} v_{p-1}
 \end{aligned}$$

Então $\{v_1, \dots, v_p\}$ é uma base ortogonal para W .

Além disso, $\text{Span}\{v_1, \dots, v_k\} = \text{Span}\{x_1, \dots, x_k\}$ para $0 \leq k \leq p$.

O processo de Gram-Schmidt mostra que qualquer subespaço não-nulo W de \mathbb{R}^n tem uma base ortogonal, já que ele tem sempre base $\{x_1, \dots, x_p\}$ e o processo de Gram-Schmidt depende apenas da existência de projeções ortogonais sobre subespaços de W que já têm bases ortogonais.

Apêndice D

Série de Taylor

Uma função f a uma variável, contínua e infinitamente derivável, pode ser representada por uma série de potências da forma:

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

onde $f^{(0)}(a) = f(a)$, e $f^{(n)}(a)$ é a derivada de ordem n de f no ponto a . A equação acima denomina-se Série de Taylor da função f em torno do ponto $x = a$. Quando $a = 0$, a série recebe o nome de Maclaurin.

O emprego da Série de Taylor para representar f está limitado aos casos em que ela é convergente. Pela teoria das séries de potências, a série de Taylor é convergente para os valores de x que satisfazem a desigualdade

$$|x - a| < r,$$

onde r é o raio de convergência da série e é dado por $r = \lim_{n \rightarrow \infty} (n + 1) \frac{|f^{(n)}(a)|}{|f^{(n+1)}(a)|}$.

Se $h = x - a$, então a Série de Taylor escreve-se assim:

$$f(a + h) = \sum_{n=0}^{\infty} h^n \frac{f^{(n)}(a)}{n!}.$$

Nas aplicações da Série de Taylor, torna-se impossível computar todos os seus termos. O que se faz é considerar apenas um número finito deles. Se a série é truncada

após o n -ésimo termo, tem-se a aproximação:

$$f(a+h) \approx f(a) + hf'(a) + \frac{h^2}{2!}f''(a) + \dots + \frac{h^{n-1}}{(n-1)!}f^{(n-1)}(a)$$

Nessa aproximação comete-se um erro de truncamento, $R_n(x)$.

Este erro é dado por:

$$R_n(x) = \frac{f^{(n)}(\xi)(x-a)^n}{n!}, \quad a < \xi < x.$$

Esta é a fórmula de Lagrange para o erro de truncamento. Como ξ não é conhecido explicitamente, a fórmula de Lagrange pode ser usada para delimitar o erro, ou seja,

$$|R_n(x)| \leq \frac{M}{n!}(x-a)^n,$$

onde

$$M = \max |f^{(n)}(t)|, \quad a \leq t \leq x.$$

Exemplo D.0.6 *Desenvolver a função f definida por $f(x) = e^x$ em Série de Taylor em torno do ponto $x = 0$. Calcule e^{-1} usando, da série obtida, cinco termos, e delimite o erro cometido.*

Nesse caso,

$$f(x) = e^x, \quad a = 0 \quad \text{e} \quad f(0) = f'(0) = f''(0) = \dots = f^{(n)}(0) = 1.$$

Portanto,

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + \dots$$

e o raio de convergência da série é infinito. Então

$$e^{-1} \approx 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} = 0,375.$$

Delimitação do erro:

$$|R_5(-1)| \leq \frac{M}{n!}|(-1)|.$$

Como $M = \max |e^t|$, $-1 \leq t \leq 0$, então $M = 1$, assim:

$$|R_5(-1)| \leq 0,008333.$$

Portanto, o valor aproximado para e^{-1} obtido pela série com cinco termos tem pelo menos uma casa decimal correta.

Apêndice E

Teorema da Multiplicidade

Teorema E.0.2 *Seja $f \in \mathcal{C}^m[a, b]$, f tem uma raiz de multiplicidade m no ponto p se e somente se $0 = f^0(p) = f^1(p) = f^2(p) = \dots = f^{m-1}(p)$, mas $f^m(p) \neq 0$*

Prova:

(\rightarrow) *Se f tem uma raiz de multiplicidade m em p então $f(x)$ pode ser escrito como:*

$$f(x) = (x - p)^m q(x) \text{ onde, } \lim_{x \rightarrow p} q(x) \neq 0$$

e assim derivando $f(x)$ temos:

$$f^1(x) = m(x - p)^{m-1}q(x) + (x - p)^m q^1(x)$$

$$f^2(x) = m(m-1)(x-p)^{m-2}q(x) + m(x-p)^{m-1}q^1(x) + m(x-p)^{m-1}q^1(x) + (x-p)^m q^2(x)$$

$$f^3(x) = m(m-1)(m-2)(x-p)^{m-3}q(x) + 3m(m-1)(x-p)^{m-2}q^1(x) + \\ + 3m(x-p)^{m-1}q^2(x) + (x-p)^m q^3(x)$$

$$f^4(x) = m(m-1)(m-2)(m-3)(x-p)^{m-4}q(x) + 4m(m-1)(m-2)(x-p)^{m-3}q^1(x) + \\ + 6m(m-1)(x-p)^{m-2}q^2(x) + 4m(x-p)^{m-1}q^3(x) + (x-p)^m q^4(x)$$

assim até a $(m-1)$ -ésima derivada teremos em todas as parcelas desta soma um termo $(x-p)$, a partir da m -ésima derivada teremos pelo menos uma parcela desta soma sem o termo $(x-p)$, isso implica que $0 = f^0(p) = f^1(p) = f^2(p) = \dots = f^{m-1}(p)$ e

$f^m(p) \neq 0$.

(\leftarrow) Se $0 = f^0(p) = f^1(p) = f^2(p) = \dots = f^{m-1}(p)$ e $f^m(p) \neq 0$,

se expandirmos f como uma Série de Taylor de grau $(m - 1)$, sobre o ponto p teremos:

$$\begin{aligned} f(x) &= f(p) + f^1(p)(x-p) + f^2(p)(x-p)^2 + \dots + f^{m-1}(p)(x-p)^{m-1} + f^m(p)(x-p)^m = \\ &= f^m(p)(x-p)^m \end{aligned}$$

Tomando $q(x) = f^m(p)$, então teremos:

$$f(x) = q(x)(x-p)^m \text{ onde } \lim_{x \rightarrow p} q(x) \neq 0$$

ou seja f tem uma raiz de multiplicidade m .