

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
PROMEC/INSTITUTO DE MATEMÁTICA

MÉTODOS MATRICIAIS

Julio Cesar Ruiz Claeysen

Porto Alegre, novembro de 2005

ÍNDICE

1	Introdução	1
1.1	Problemas Centrais	1
2	Matrizes e Vetores: Operações Básicas	14
2.1	Visualização Gráfica de Vetores e Matrizes	16
2.2	Adição e Subtração de Matrizes	19
2.3	Multiplicação Matricial	20
2.3.1	Potências de uma Matriz Quadrada e Polinômios Matriciais	26
2.3.2	Matriz Transposta	28
2.3.3	Multiplicação Matricial através de Somas	30
2.3.4	Matrizes Simétricas	31
2.3.5	Conjugada de uma Matriz	31
2.3.6	Matrizes Hermitianas	32
2.4	Submatrizes	33
2.5	Matrizes Bloco	33
2.6	Alguns Tipos de Matrizes	35
2.6.1	Matrizes Triangulares	35
2.6.2	Matrizes Diagonais	36
2.6.3	Matrizes Tridiagonais e Matrizes Banda	37
2.6.4	Matrizes Estocásticas	37
2.6.5	Matrizes Booleanas e de Permutação	38
2.6.6	Matrizes Circulantes e de Toeplitz	39
2.6.7	Matrizes Esparsas e Cheias	40
2.7	O Problema da Divisão Matricial	40
2.8	Inversa de uma matriz quadrada	41
2.8.1	Determinante de uma Matriz Quadrada	42
2.8.2	Propriedades dos Determinantes	44
2.8.3	A Identidade de Cramer	45
2.8.4	Fórmula Analítica para a Inversa de uma Matriz	45
3	Sistemas de Equações Algébricas Lineares	48
3.1	Introdução	48
3.2	Sistemas Lineares Regulares	49
3.2.1	O Método De Cramer	49
3.2.2	O Método da Eliminação	50
3.2.3	Fatorização Matricial LU	52
3.3	Complexidade e Condicionamento Computacional	55
3.3.1	Exemplos Sobre Condicionamento	57
3.3.2	Fatorização L D U	61
3.3.3	Fatorização LDL^t	61
3.3.4	Fatorização LL^t de Cholesky	62
3.4	Sistemas Lineares Singulares	65

3.4.1	Posto e Nulidade	67
3.5	Sistemas Lineares Retangulares	70
3.6	Soluções Não Nulas de Sistemas Homogêneos	72
3.7	Geometria Matricial	76
3.7.1	Subespaços Lineares, Bases e Dimensão	80
4	O Método Espectral	84
4.1	Autovalores e Autovetores	84
4.1.1	Polinômio característico	86
4.1.2	Identidade de Cayley-Hamilton	91
4.2	Vetores Ortogonais	93
4.2.1	Processo de Ortogonalização	96
4.3	Ortogonalidade do Sistema Adjunto Homogêneo	98
4.4	Matrizes Ortogonais	99
4.5	Fatorização Espectral de Matrizes Simétricas	101
4.5.1	O Método Espectral	103
4.5.2	O Caso Singular	105
4.5.3	Fatorização e Diagonalização	105
4.5.4	Potências de uma Matriz simétrica	106
4.5.5	Desacoplamento	107
4.6	Extensão do Método Espectral	108
4.6.1	Biortogonalidade da Matriz Adjunta	108
4.6.2	Diagonalização de Matrizes Não Defeituosas	109
4.7	Formas Quadráticas	111
4.7.1	Quociente de Rayleigh	114
4.7.2	Formas Definidas e Modos Normais	115
4.8	Decomposição de Matrizes em Valores Singulares	118
4.8.1	Interpretação Geométrica	122
5	Equações Diferenciais Matriciais	124
5.1	Um sistema de primeira ordem como modelo da hemodiálise	124
5.2	A equação $Cu' + Bu = F(t)$	127
5.2.1	Fórmula integral	128
5.2.2	O Método de Variação de Parâmetros de Lagrange	129
5.3	O método matricial espectral	130
5.3.1	Caso homogêneo	131
5.3.2	O caso não-homogêneo	134
5.4	Métodos não-espectrais	136
5.4.1	Método Polinomial	140
5.5	Comportamento assintótico	142
5.6	Integração de um modelo de rastreamento de aeronaves rígidas	142
5.7	Um modelo de segunda ordem para a suspensão de um carro	150
5.8	A equação $Mu'' + Cu' + Ku = F(t)$	153
5.8.1	O método de Cauchy	155
5.8.2	O método espectral	156
5.8.3	Caso homogêneo	157
5.8.4	O caso não-homogêneo	165
5.8.5	Métodos não-espectrais	166
5.8.6	A formulação de estado	171
5.9	Comportamento assintótico	173
5.10	Filtragem	173

6	O Método Dos Mínimos Quadrados	174
6.1	A Equação Matricial Normal	177
6.1.1	Resolução da Equação Normal	179
6.1.2	Inversa Generalizada de uma Matriz	182
6.2	A Fatorização QR	183
6.2.1	Matrizes de Householder	187
7	Métodos Iterativos e Equações em Diferenças	190
7.1	Introdução	190
7.2	Métodos de Jacobi e de Gauss-Seidel	190
7.2.1	Método de Gauss-Seidel	192
7.3	Equações em Diferenças	193
7.3.1	Equações Homogêneas	194
7.3.2	O Método Espectral: Soluções Exponenciais	194
7.3.3	Desacoplamento	196
7.4	Soluções Transientes e Estacionárias	198
7.4.1	Método Operacional	198
7.5	Método da Potência	199
8	Métodos Variacionais	201
8.0.1	Método de Rayleigh-Ritz	202

ÍNDICE DE FIGURAS

1.1	Discretização da placa	4
1.2	Grade Espacial	5
1.3	Ajuste de Curva por mínimos Quadrados	11
2.1	Gráfico de um vetor	17
2.2	Gráfico de um vetor	18
2.3	Gráfico de uma matriz	18
2.4	Matriz banda	37
3.1	Equilíbrio de Força Elástica	59
3.2	Representação Geométrica	76
3.3	Colinearidade e Não-colinearidade	77
4.1	Vetores ortogonais	93
4.2	Forma Quadrática	111
4.3	Forma Quadrática	113
4.4	Interpretação Geométrica	123
4.5	Caso Bidimensional	123
5.1	Diagrama da máquina de hemodiálise para o rim	126
5.2	Forças que agem em uma aeronave	143
5.3	Nomenclatura para movimentos longitudinais de uma aeronave	144
5.4	Nomenclatura para movimentos laterais de uma aeronave	144
5.5	Controle do ângulo de inclinação através dos ailerões	145
5.6	Ângulos de Euler	148
5.7	Dinâmica Lateral	148
5.8	Dinâmica Longitudinal	148
5.9	Modelo da suspensão da metade de um carro	151
5.10	Entrada tipo oscilatória nas rodas	153
5.11	Modos dos sistemas de discos	162
5.12	Componentes da resposta livre	163
5.13	Respostas $u_1(t)$ e $u_2(t)$	171
6.1	Método dos Mínimos Quadrados	174

CAPÍTULO 1

Introdução

1.1 Problemas Centrais

O objetivo principal dos métodos matriciais é a resolução de sistemas de equações lineares. Eles podem estar relacionados com problemas de natureza estática (*permanente*) ou dinâmica (*evolutiva*). Os métodos são analíticos ou computacionais. Como problemas centrais permanentes tem-se a resolução dos seguintes tipos de equações matriciais:

$$\mathbf{A}\mathbf{u} = \mathbf{b} \quad (1.1)$$

$$\mathbf{A}\mathbf{v} = \lambda\mathbf{v} \quad (1.2)$$

$$(\lambda^2\mathbf{M} + \lambda\mathbf{C} + \mathbf{K})\mathbf{v} = \mathbf{0} \quad (1.3)$$

$$\mathbf{A}^T\mathbf{A}\mathbf{u} = \mathbf{A}^t\mathbf{b} \quad (1.4)$$

Aqui os coeficientes \mathbf{A} , \mathbf{M} , \mathbf{C} e \mathbf{K} são matrizes, enquanto as incógnitas \mathbf{u} e \mathbf{v} e o termo não-homogêneo \mathbf{b} podem ser vetores ou matrizes.

Os problemas centrais evolutivos consistem na resolução dos seguintes tipos de equações matriciais:

$$\frac{d\mathbf{u}}{dt} + \mathbf{A}\mathbf{u} = \mathbf{f}(t) \quad (1.5)$$

$$\mathbf{M}\frac{d^2\mathbf{u}}{dt^2} + \mathbf{C}\frac{d\mathbf{u}}{dt} + \mathbf{K}\mathbf{u} = \mathbf{f}(t) \quad (1.6)$$

$$\mathbf{u}_{k+1} + \mathbf{A}\mathbf{u}_k = \mathbf{f}_k \quad (1.7)$$

$$\mathbf{M}\mathbf{u}_{k+2} + \mathbf{C}\mathbf{u}_{k+1} + \mathbf{K}\mathbf{u}_k = \mathbf{f}_k \quad (1.8)$$

onde as incógnitas e os termos não homogêneos dependem de uma variável t , que assume valores reais ou de uma variável k , a qual assume valores inteiros.

Como derivados desses problemas, tem-se sistemas lineares sujeitos a restrições e sistemas não-lineares, que serão descritos posteriormente.

Os vetores e as matrizes nas equações anteriores decorrem de dois processos :

- **DADOS OBTIDOS EXPERIMENTALMENTE (OU OBSERVADOS), QUE PERMITEM PROPOR UM MODELO MATEMÁTICO APROXIMADO;**
- **MODELOS MATEMÁTICOS OBTIDOS ATRAVÉS DE LEIS FENOMENOLÓGICAS E HIPÓTESES RELATIVAS AO PROBLEMA.**

Os dados geram valores que podem ser agrupados de maneira vetorial. Ainda, de modo mais geral, valores matriciais. Entretanto, os modelos são descritos por equações cujas soluções, por sua vez, geram valores. Devido a complexidade de um problema, os dados podem ser utilizados para propor um modelo matemático cuja solução deve ser validada. Um modelo equacionado através de leis deve ser tal que sua solução verifique com os dados disponíveis, dentro de uma certa tolerância. Um bom modelo depende da sua previsibilidade, isto é, que os resultados sejam aceitáveis ou verificáveis.

Alguns dos problemas centrais serão ilustrados através de exemplos a seguir.

Exemplo 1.1

Um problema típico, que provém da análise dimensional, ocorre no escoamento de um fluido. Aqui as variáveis são a velocidade V , a densidade ρ , o diâmetro D , a gravidade g , a viscosidade μ . Em termos das usuais unidades fundamentais para massa M , comprimento L e tempo T , tem-se:

$$\begin{array}{l} \text{Grandeza :} \\ \text{Dimensões :} \end{array} \quad \begin{array}{ccccc} V & \rho & D & g & \mu \\ LT^{-1} & ML^{-3} & L & LT^{-2} & ML^{-1}T^{-1} \end{array} \quad (1.9)$$

Deseja-se, quando for possível, formular produtos adimensionais da forma

$$V^a \rho^b D^c g^d \mu^e \quad (1.10)$$

e, determinar o maior número possível de produtos.

Diz-se que um produto é adimensional, quando, ao substituir cada grandeza pelas suas dimensões, como na tabela acima, a soma dos expoentes em cada dimensão é zero. Em outras palavras,

$$(LT^{-1})^a (ML^{-3})^b L^c (LT^{-2})^d (ML^{-1}T^{-1})^e = M^o L^o T^o. \quad (1.11)$$

Então, as três equações seguintes devem ser satisfeitas para as cinco incógnitas:

$$\begin{array}{l} \text{(Das potências de } M) \\ \text{(Das potências de } L) \\ \text{(Das potências de } T) \end{array} \quad \begin{array}{cccccc} b + & e & & & & = 0 \\ a - & 3b + & c + & d - & e & = 0 \\ -a & -2d & -e & & & = 0 \end{array} \quad (1.12)$$

Tem-se, portanto, o sistema linear homogêneo

$$\mathbf{Ax} = \mathbf{0},$$

onde

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & -3 & 1 & 1 & -1 \\ -1 & 0 & 0 & -2 & -1 \end{bmatrix} \quad (1.13)$$

é a matriz dos coeficientes do sistema,

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} \quad (1.14)$$

é a matriz coluna formada pelas incógnitas do sistema e $\mathbf{0}$ é matriz coluna formada por zeros.

Exemplo 1.2

Considere-se a situação em que uma placa homogênea retangular está exposta a uma taxa de geração de energia por unidade de volume (fonte de calor) dada por uma função dada $f(x, y)$. Em condições de regime permanente, a distribuição de temperatura $u(x, y)$ está descrita pela equação de Poisson

$$k[u_{xx} + u_{yy}] + f(x, y) = 0 \quad (1.15)$$

para os pontos interiores da placa. Aqui, a constante k denota a condutividade térmica característica do material da placa. Suponha-se que nos lados da placa a temperatura u assuma o valor $g(x, y)$, para um ponto do contorno (x, y) .

Neste exemplo pretende-se determinar a temperatura num número finito de pontos na placa, os quais estão situados sobre uma grade superposta. Isto pode ser realizado de maneira aproximada através de diversos métodos numéricos: diferenças finitas, volumes finitos, elementos finitos, etc...

O método utilizado e descrito a seguir é o das diferenças finitas. Considera-se uma grade discreta uniforme de pontos na placa

$$P = (x_i, y_j), \quad i = 1, 2, \dots, n - 1; \quad j = 1, 2, \dots, m - 1. \quad (1.16)$$

Para as derivadas parciais de segunda ordem são consideradas as aproximações centrais:

$$u_{xx}(x_i, y_j) \approx \frac{u(x_{i+1}, y_j) - 2u(x_i, y_j) + u(x_{i-1}, y_j))}{\Delta x^2} \quad (1.17)$$

$$u_{yy}(x_i, y_j) \approx \frac{u(x_i, y_{j+1}) - 2u(x_i, y_j) + u(x_i, y_{j-1}))}{\Delta y^2} \quad (1.18)$$

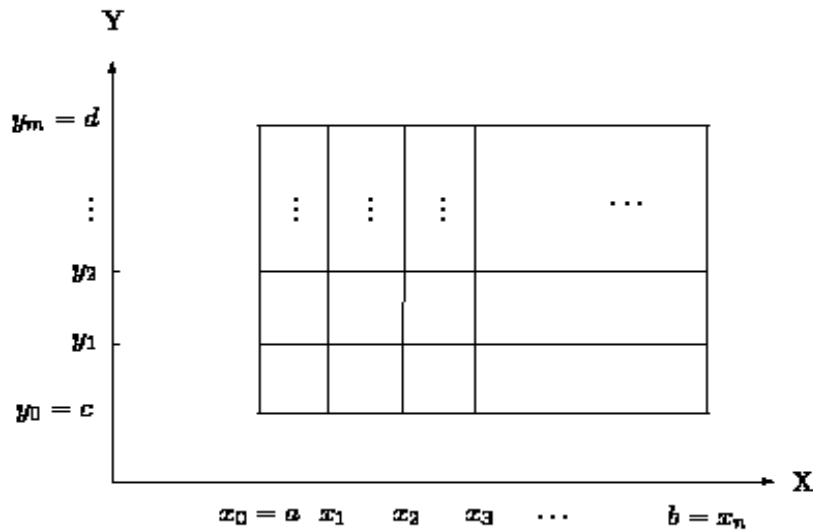


Figura 1.1 – Discretização da placa

Por simplicidade, suponha-se que $\Delta x^2 = \Delta y^2 = h^2$. Substituindo estas aproximações na equação (1.15), obtém-se

$$k[u(x_{i+1}, y_j) + u(x_{i-1}, y_j) - 4u(x_i, y_j) + u(x_i, y_{j+1}) + u(x_i, y_{j-1})] + h^2 f(x_i, y_j) \approx 0 \quad (1.19)$$

Do ponto de vista numérico, a equação (1.19) considera-se como uma igualdade, isto é

$$k[u_{i+1,j} + u_{i-1,j} - 4u_{i,j} + u_{i,j+1} + u_{i,j-1}] + h^2 f_{i,j} = 0 \quad (1.20)$$

onde $u_{i,j} = u(x_i, y_j)$ e $f_{i,j} = f(x_i, y_j)$ para $i = 1 : n - 1$, $j = 1 : m - 1$.¹ Espera-se que o valor $u_{i,j}$ obtida da equação (1.20) forneça uma boa aproximação para $u(x_i, y_j)$.

¹A variação de índices $i = 1, 2, 3, \dots, n - 1$ será abreviada por $i = 1 : (n - 1)$. Similarmente, $i = 1, p + 1, p + 2, \dots, p + n$ por $i = 1 : p : n$

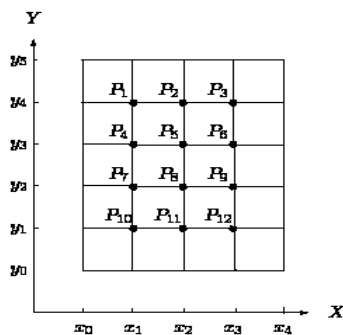


Figura 1.2 – Grade Espacial

Para cada ponto $\mathbf{P} = (x_i, y_j)$ da grade espacial, observe que a equação anterior envolve aproximações de u para pontos localizado simetricamente em torno de $\mathbf{P} = (x_i, y_j)$, isto é, $\mathbf{N} = (x_i, y_{j+1})$, $\mathbf{S} = (x_i, y_{j-1})$, $\mathbf{E} = (x_{i-1}, y_j)$ e $\mathbf{W} = (x_{i+1}, y_j)$.

$$\begin{array}{c} \mathbf{N} \\ | \\ \mathbf{E} - \mathbf{P} - \mathbf{W} \\ | \\ \mathbf{S} \end{array}$$

Deste modo, para cada ponto interior $\mathbf{P}=(x_i, y_j)$ da grade espacial a equação (1.20) pode ser escrita como

$$4\mathbf{u}(\mathbf{P}) - \mathbf{u}(\mathbf{N}) - \mathbf{u}(\mathbf{S}) - \mathbf{u}(\mathbf{E}) - \mathbf{u}(\mathbf{W}) = \frac{h^2}{k}\mathbf{f}(\mathbf{P}) \quad (1.21)$$

Observe-se que para pontos interiores \mathbf{P} que são adjacentes ao contorno, algum dos pontos \mathbf{N} , \mathbf{S} , \mathbf{E} , \mathbf{W} estarão no contorno e os valores de \mathbf{u} nesses pontos devem ser obtido das condições fornecidas.

Se a grade possui m pontos, há m equações desse tipo. As condições de contorno do problema fornecem os seguintes dados:

$$\begin{aligned} u_{0,j} &= u(x_0, y_j) = g(x_0, y_j), & j &= 1, 2, \dots, m \\ u_{n,j} &= u(x_n, y_j) = g(x_n, y_j), & j &= 1, 2, \dots, m \\ u_{i,0} &= u(x_i, y_0) = g(x_i, y_0), & i &= 1, 2, \dots, n-1 \\ u_{i,m} &= u(x_i, y_m) = g(x_i, y_m), & i &= 1, 2, \dots, n-1. \end{aligned} \quad (1.22)$$

Agora, suponha-se uma placa quadrada com lado de comprimento 0.5 metros. Dois lados adjacentes estão à temperatura fixa de 0° . Entretanto, a temperatura nos outros dois lados aumenta linearmente desde 0° , no vértice, até 100° , no vértice onde estes lados se

encontram. Considere-se que os lados com temperatura nula estão ao longo dos eixos x e y . Assim,

$$u(0, y) = 0, \quad u(x, 0) = 0, \quad u(x, 0.5) = 200x \quad e \quad u(0.5, y) = 200y, \quad (1.23)$$

para $0 < x < 0.5$ e $0 < y < 0.5$.

Escolha-se um espaçamento horizontal e vertical igual a $h = 0.125$ entre dois pontos vizinhos. Deste modo, $n = m = 4$. Definindo

$$u_k = u(\mathbf{P}_k) \quad (1.24)$$

tem-se as seguintes equações para as temperaturas nesses pontos:

$$\begin{aligned} P_1 : & \quad 4u_1 - u_2 - u_4 = u_{0,3} + u_{1,4} + h^2 f(x_1, y_3), \\ P_2 : & \quad 4u_2 - u_3 - u_1 - u_5 = u_{2,4} + h^2 f(x_2, y_3), \\ P_3 : & \quad 4u_3 - u_2 - u_6 = u_{4,3} + u_{3,4} + h^2 f(x_3, y_3), \\ P_4 : & \quad 4u_4 - u_5 - u_1 - u_7 = u_{0,2} + h^2 f(x_1, y_2), \\ P_5 : & \quad 4u_5 - u_6 - u_4 - u_2 - u_8 = h^2 f(x_2, y_2), \\ P_6 : & \quad 4u_6 - u_5 - u_3 - u_9 = u_{4,2} + h^2 f(x_3, y_2), \\ P_7 : & \quad 4u_7 - u_8 - u_4 = u_{0,1} + u_{1,0} + h^2 f(x_1, y_1), \\ P_8 : & \quad 4u_8 - u_9 - u_7 - u_5 = u_{2,0} + h^2 f(x_2, y_1), \\ P_9 : & \quad 4u_9 - u_8 - u_6 = u_{3,0} + u_{4,1} + h^2 f(x_3, y_1), \end{aligned} \quad (1.25)$$

onde os termos à direita das equações são obtidos das condições de contorno e da fonte f . De fato, tem-se que

$$\begin{aligned} u_{1,0} = u_{2,0} = u_{3,0} = u_{0,1} = u_{0,2} = u_{0,3} &= 0 \\ u_{1,4} = u_{4,1} = 25, \quad u_{2,4} = u_{4,2} = 50, \quad u_{3,4} = u_{4,3} &= 75. \end{aligned} \quad (1.26)$$

O sistema linear, $\mathbf{A}\mathbf{u} = \mathbf{b}$, associado a este problema, terá a forma

$$\begin{bmatrix} 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 4 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 4 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 4 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 4 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \\ u_6 \\ u_7 \\ u_8 \\ u_9 \end{bmatrix} = \begin{bmatrix} 25 \\ 50 \\ 150 \\ 0 \\ 0 \\ 50 \\ 0 \\ 0 \\ 25 \end{bmatrix} + \frac{(0.125)^2}{k} \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \\ f_6 \\ f_7 \\ f_8 \\ f_9 \end{bmatrix} \quad (1.27)$$

onde $f_k = f(\mathbf{P}_k)$. Neste exemplo, o vetor \mathbf{b} é a soma dos valores de u no contorno da placa com os valores da fonte multiplicados por uma constante igual ao quadrado do espaçamento e divididos pela condutividade térmica.

Exemplo 1.3

Considere-se a procura de soluções da forma

$$u = e^{i\omega t} \phi(x), \quad i = \sqrt{-1}$$

para a equação da onda

$$u_{tt} = u_{xx},$$

com as condições de contorno

$$u(0, t) = 0, \quad u(L, t) = 0.$$

Substituindo a primeira equação na segunda, decorre:

$$\phi''(x) + \omega^2 \phi(x) = 0$$

$$\phi(0) = 0, \quad \phi(L) = 0.$$

Aproximando

$$\phi''(x_i) \approx \frac{\phi(x_{i+1}) - 2\phi(x_i) + \phi(x_{i-1}))}{\Delta x^2}$$

e introduzindo as incógnitas $\phi_i \simeq \phi(x_i)$, obtém-se o esquema numérico

$$\frac{\phi_{i+1} - 2\phi_i + \phi_{i-1}}{\Delta x^2} = -\omega^2 \phi_i, \quad i = 1 : n,$$

onde $\phi_0 = \phi_{n+1} = 0$, em virtude das condições de contorno.

O sistema linear $\mathbf{A}\phi = \lambda\phi$, associado a este problema, tem a forma:

$$\begin{bmatrix} 2 & -1 & 0 & \cdots & \cdots & 0 \\ -1 & 2 & -1 & & & \\ 0 & -1 & 2 & -1 & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \\ \vdots & & & \ddots & \ddots & -1 \\ 0 & \cdots & & -1 & 2 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix} = \lambda \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_N \end{bmatrix},$$

onde $\lambda = -\omega^2 \Delta x^2$.

Exemplo 1.4

Considere-se uma viga de comprimento L apoiada em $x = 0$ e livre em $x = L$, com rigidez de flexão EI e massa m por unidade de comprimento l , que executa movimento vertical, com flexão u , sujeito a uma força externa distribuída f e a uma força axial P . Deseja-se determinar vibrações livres da forma $u = e^{t\lambda} \phi(x)$ onde a amplitude é aproximada por $\phi = c_1 \phi_1(x) + \dots + c_n \phi_n(x)$ para certas funções base $\phi_k(x)$, $k = 1 : n$ que satisfazem as condições de contorno da viga.

Pelo princípio de Hamilton, a funcional

$$I = \int_0^{t_f} (T - V + W) d\tau$$

onde T , V , W denotam a energia cinética, energia potencial e trabalho da força externa na viga, assume um valor estacionário, isto é

$$I' = \left. \frac{dI}{d\epsilon} \right|_{\epsilon=0} = \lim_{\epsilon \rightarrow 0} \frac{I(u + \epsilon v) - I(u)}{\epsilon} = 0,$$

quando u é o deslocamento da viga entre os instantes de tempo inicial 0 e tempo final t_f . Aqui $v = \delta u$ denota uma perturbação que é nula nesses instantes de tempo.

No caso da viga, tem-se

$$\begin{aligned} T &= \frac{1}{2} \int_0^L m \left(\frac{\partial u}{\partial t} \right)^2 dx \\ V &= \frac{1}{2} \int_0^L \left[EI \left(\frac{\partial^2 u}{\partial x^2} \right)^2 + P \left(\frac{\partial u}{\partial x} \right)^2 \right] dx \\ W &= \int_0^L f u dx, \end{aligned}$$

Decorre,

$$\begin{aligned}\frac{dI}{d\epsilon} &= \int_0^{t_f} \frac{d}{d\epsilon} \left[\frac{1}{2} \int_0^L m \left(\frac{\partial u}{\partial t} + \epsilon \frac{\partial v}{\partial t} \right)^2 - EI \left(\frac{\partial^2 u}{\partial x^2} + \epsilon \frac{\partial^2 v}{\partial x^2} \right)^2 - P \left(\frac{\partial u}{\partial x} + \epsilon \frac{\partial v}{\partial x} \right)^2 \right. \\ &\quad \left. + 2f(u + \epsilon v) dx \right] dt \\ \frac{dI}{d\epsilon} &= \int_0^{t_f} \left[\int_0^L m \left(\frac{\partial u}{\partial t} + \epsilon \frac{\partial v}{\partial t} \right) \frac{\partial v}{\partial t} - EI \left(\frac{\partial^2 u}{\partial x^2} + \epsilon \frac{\partial^2 v}{\partial x^2} \right) \frac{\partial^2 v}{\partial x^2} - P \left(\frac{\partial u}{\partial x} + \epsilon \frac{\partial v}{\partial x} \right) \frac{\partial v}{\partial x} \right. \\ &\quad \left. + f(u + \epsilon v) dx \right] dt\end{aligned}$$

Para $\epsilon = 0$, tem-se o valor estacionário

$$\int_0^{t_f} \left[\int_0^L m \frac{\partial u}{\partial t} \frac{\partial v}{\partial t} - EI \frac{\partial^2 u}{\partial x^2} \frac{\partial^2 v}{\partial x^2} - P \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + f v dx \right] dt = 0. \quad (1.28)$$

As vibrações livres aproximadas u podem ser escritas de maneira compacta

$$u = e^{t\lambda} (c_1 \phi_1(x) + \dots + c_n \phi_n(x)) = e^{t\lambda} c^T \Phi^T(x) = e^{t\lambda} \Phi(x) c$$

onde $\Phi = [\phi_1 \ \phi_2 \ \dots \ \phi_n]$ e c é o vetor coluna de componentes c_k , $k = 1 : n$ e c^T o vetor transposto. Segue,

$$\begin{aligned}\frac{\partial u}{\partial t} &= \lambda e^{t\lambda} \Phi(x) c \\ \frac{\partial u}{\partial x} &= e^{t\lambda} \Phi(x) c \\ \frac{\partial^2 u}{\partial x^2} &= e^{t\lambda} \Phi''(x) c.\end{aligned} \quad (1.29)$$

Suponha-se que a perturbação v é do mesmo tipo que u , isto é, $v = e^{t\lambda} \Phi(x) \Delta c$. Substituindo u, v em (1.28) com $f = 0$, que corresponde ao caso de vibrações livres, vem

$$\int_0^{t_f} \int_0^L e^{2t\lambda} c^T [m\lambda^2 \Phi^T(x) \Phi(x) - EI(\Phi''(x))^T \Phi''(x) - P(\Phi'(x))^T \Phi'(x)] \Delta c dx dt = 0.$$

Por simples integração no tempo, elimina-se a integração temporal. Assim,

$$\int_0^L c^T [m\lambda^2 \Phi^T(x) \Phi(x) - EI(\Phi''(x))^T \Phi''(x) - P(\Phi'(x))^T \Phi'(x)] \Delta c dx dt = 0.$$

ou,

$$\Delta c^T \left[\int_0^L (m\lambda^2 \Phi^T(x) \Phi(x) - EI(\Phi''(x))^T \Phi''(x) - P(\Phi'(x))^T \Phi'(x)) dx \right] = 0.$$

Como Δc é um vetor arbitrário, a expressão entre colchetes deve ser nula. Obtém-se a equação

$$[\lambda^2 \mathbf{M} + \mathbf{K}]c = 0 \quad (1.30)$$

onde M , C são as matrizes de componentes

$$m_{ij} = m \int_0^L \phi_i(x) \phi_j(x) dx$$

$$k_{ij} = - \int_0^L (EI \phi_i''(x) \phi_j''(x) + P \phi_i' \phi_j') dx$$

OBSERVAÇÃO

A partir do principio de Hamilton, obtém-se a equação diferencial que governa o movimento da viga. O procedimento é como segue. No termo correspondente a energia cinética, é trocada a ordem da integração e realizada uma integração por partes com respeito de t . Os dois termos correspondentes a energia potencial são integrados por partes com respeito de x , duas vezes o primeiro e uma vez o segundo termo. Decorre

$$0 = \int_0^L \int_0^{t_f} m \frac{\partial u}{\partial t} \frac{\partial v}{\partial t} dt dx - \int_0^{t_f} \int_0^L \left(EI \frac{\partial^2 u}{\partial x^2} \frac{\partial^2 v}{\partial x^2} - P \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + f v \right) dx dt$$

$$0 = \int_0^L \left[m \frac{\partial u}{\partial t} v \right]_0^{t_f} dx - \int_0^L \int_0^{t_f} m \frac{\partial^2 u}{\partial t^2} v dt dx - \int_0^{t_f} \left[EI \frac{\partial^2 u}{\partial x^2} \frac{\partial v}{\partial x} - \frac{\partial}{\partial x} \left(EI \frac{\partial^2 u}{\partial x^2} \right) v + P \frac{\partial u}{\partial x} v \right]_0^L dt$$

$$- \int_0^{t_f} \int_0^L \left[\frac{\partial^2}{\partial x^2} \left(EI \frac{\partial^2 u}{\partial x^2} \right) - \frac{\partial}{\partial x} \left(P \frac{\partial u}{\partial x} \right) \right] v dx dt + \int_0^{t_f} \int_0^L f v dx dt.$$

Agrupando termos e utilizando o fato que v é zero nos instantes 0 e t_f , vem

$$\int_0^{t_f} \int_0^L \left[m \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left(EI \frac{\partial^2 u}{\partial x^2} \right) - \frac{\partial}{\partial x} \left(P \frac{\partial u}{\partial x} \right) - f \right] v dx dt =$$

$$- \int_0^{t_f} \left[EI \frac{\partial^2 u}{\partial x^2} \frac{\partial v}{\partial x} - \frac{\partial}{\partial x} \left(EI \frac{\partial^2 u}{\partial x^2} \right) v + P \frac{\partial u}{\partial x} v \right]_0^L dt.$$

A perturbação v é arbitrária e pode ser escolhida de maneira conveniente. Por exemplo, se v e sua primeira derivada espacial são nulas em $x=0$ e L , segue que o integrando do termo a esquerda deve ser nulo. Assim,

$$m \frac{\partial^2 u}{\partial t^2} + \frac{\partial^2}{\partial x^2} \left(EI \frac{\partial^2 u}{\partial x^2} \right) - \frac{\partial}{\partial x} \left(P \frac{\partial u}{\partial x} \right) - f = 0, \quad 0 < x < L. \quad (1.31)$$

Por outro lado, a validade desta equação implica a nulidade do termo a direita. Novamente, por ser v arbitrário, segue que o integrando deve ser nulo. Isto é o caso se

$$\left(- \frac{\partial}{\partial x} \left(EI \frac{\partial^2 u}{\partial x^2} \right) + P \frac{\partial u}{\partial x} \right) v \Big|_0^L = 0, \quad EI \frac{\partial^2 u}{\partial x^2} \frac{\partial v}{\partial x} \Big|_0^L = 0. \quad (1.32)$$

Pelo tipo de viga (apoiada-livre), tem-se as condições de contorno

$$u = 0, \quad EI u_{xx} = 0 \quad em \quad x = 0 \quad (1.33)$$

$$-(EI u_{xx})_x + P u_x = 0, \quad EI u_{xx} = 0 \quad em \quad x = L \quad (1.34)$$

e verifica-se a nulidade em (1.32). Pois, é suficiente escolher v tal que $v = 0$ em $x = 0$.

Exemplo 1.5

Suponha-se que, como resultado de medições no laboratório ou de algo similar, foi obtida uma coleção de valores y_1, y_2, \dots, y_n correspondentes a uma coleção de alguma outra grandeza, x_1, x_2, \dots, x_n . Observa-se que graficando estes valores, tem-se pontos em torno de alguma curva.

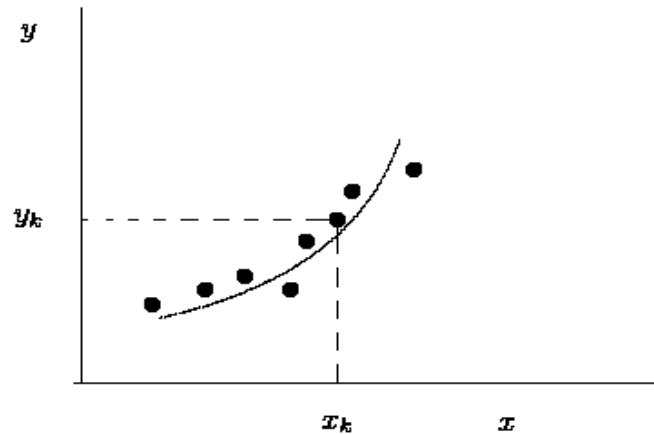


Figura 1.3 – Ajuste de Curva por mínimos Quadrados

Suponha-se, como simples exemplo, que a curva é da forma

$$y = a + bx + cx^2$$

Equacionando esta fórmula com as medições, tem-se o sistema

$$\begin{aligned} y_1 &= a + bx_1 + cx_1^2 \\ y_2 &= a + bx_2 + cx_2^2 \\ y_3 &= a + bx_3 + cx_3^2 \\ &\vdots \\ y_n &= a + bx_n + cx_n^2. \end{aligned}$$

Matricialmente, $\mathbf{y} = \mathbf{A}\mathbf{u}$, onde

$$\mathbf{A} = \begin{bmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \\ \vdots & \vdots & \vdots \\ 1 & x_n & x_n^2 \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} \quad \text{e} \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}.$$

O problema consiste em determinar a matriz coluna \mathbf{u} , formada pelos parâmetros da curva y , a partir da equação $\mathbf{A}\mathbf{u} = \mathbf{b}$. Usualmente, o número das medições é muito maior do que o número de parâmetros envolvidos na curva de ajuste y . O sistema é assim ambíguo ou, mais precisamente, *sobredeterminado*. Multiplicando ambos os membros da equação $\mathbf{A}\mathbf{u} = \mathbf{y}$ pela matriz transposta de \mathbf{A} , obtida trocando as linhas pelas colunas da matriz \mathbf{A} , isto é,

$$\mathbf{A}^t = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ x_1 & x_2 & x_3 & \dots & x_n \\ x_1^2 & x_2^2 & x_3^2 & \dots & x_n^2 \end{bmatrix},$$

resulta um sistema da forma

$$\mathbf{A}^t \mathbf{A} \mathbf{u} = \mathbf{A}^t \mathbf{y},$$

em que o número de equações é igual ao número de incógnitas. A escolha deste procedimento é decorrência de uma técnica de resolução de sistemas sobredeterminados, a ser vista posteriormente, chamada de método dos mínimos quadrados: *obter os coeficientes, de modo tal que a soma dos quadrados dos erros entre o valor medido e o do valor previsto, pela fórmula da curva, tenha um valor mínimo*.

Exemplo 1.6

Segundo a figura, suponha-se que $x(t)$ e $y(t)$ denotam as temperaturas nos quartos A e B, respectivamente. A lei de resfriamento de Newton estabelece que a taxa de variação da temperatura num quarto é proporcional à diferença entre temperatura do quarto e a do meio adjacente.

Suponha-se que o quarto A é aquecido com uma lareira que gera 100.000,00 Btu/hora de calor e que tem uma capacidade de calor de 0.00020°F/Btu. Isto significa que, se o quarto A está selado e isolado, a temperatura no quarto A aumenta 0.20°F a cada mil Btu de calor gerado pela lareira. Como a lareira gera 100.000 Btu por hora, a temperatura no quarto A aumentará 20°F a cada hora. Se as constantes de troca de calor entre os quartos são as seguintes: entre o quarto A e o exterior é $\frac{1}{5}$, entre o quarto B e o exterior é de $\frac{1}{10}$, e entre o quarto A e quarto B é de $\frac{1}{2}$, então, a equação para a temperatura no quarto A é

$$\frac{dx}{dt} = -\frac{1}{5}[x(t) - 0] - \frac{1}{2}[x(t) - y(t)] + 20$$

Para o quarto B, tem-se

$$\frac{dy}{dt} = -\frac{1}{2}[y(t) - x(t)] - \frac{1}{10}[y(t) - 0]$$

Simplificando,

$$\frac{dx}{dt} = -\frac{7}{10}x(t) + \frac{1}{2}y(t) + 20$$

$$\frac{dy}{dt} = \frac{1}{2}x(t) - \frac{3}{5}y(t)$$

Matricialmente,

$$\frac{d}{dt} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} -\frac{1}{7} & \frac{1}{2} \\ \frac{1}{2} & -\frac{3}{5} \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} 20 \\ 0 \end{bmatrix}$$

Assim, decorre um sistema de equações diferenciais lineares

$$\frac{d\mathbf{u}}{dt} = \mathbf{A}\mathbf{u} + \mathbf{f}(t),$$

onde \mathbf{A} é a matriz dos coeficientes, \mathbf{u} o vetor formado pelas incógnitas $x = x(t)$, $y = y(t)$, e $\mathbf{f}(t)$ o vetor correspondente ao termo fonte, neste caso 20 e 0.

CAPÍTULO 2

Matrizes e Vetores: Operações Básicas

Nesta seção serão apresentados conceitos, definições e propriedades básicas relativas às matrizes.

As matrizes são uma peça fundamental para representar os mais variados sistemas de equações, permitindo compactá-los de modo a exibir suas características, as quais seriam difíceis de serem evidenciadas em outra forma. Além disso, convenientes analogias com equações mais simples podem ser utilizadas como um primeiro passo na abordagem de um sistema.

Uma matriz \mathbf{A} de ordem $m \times n$ é um arranjo de elementos dispostos numa formação de m linhas e de n colunas,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}. \quad (2.1)$$

O par de subíndices i e j , associados ao elemento a_{ij} , denotam respectivamente seus números de linha e coluna ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$). Uma matriz com o mesmo número n de linhas e de colunas é chamada de matriz quadrada de ordem n . Uma matriz não necessariamente quadrada é dita matriz retangular. Caso exista necessidade de enfatizar que os elementos de uma matriz são números, ela será dita *matriz numérica*.

De modo abreviado, uma matriz \mathbf{A} de ordem $m \times n$ é usualmente escrita na forma

$$\mathbf{A} = [a_{ij}]. \quad (2.2)$$

Por exemplo, a matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 8 & 16 & 32 \\ 8 & 16 & 32 & 64 \end{bmatrix} \quad (2.3)$$

pode ser denotada por $\mathbf{A} = [a_{ij}]$, com $a_{ij} = 2^{i+j}$. Neste exemplo, $i = 1, 2$ e $j = 1, 2, 3, 4$.

Os elementos de uma matriz podem ser funções ou, também, operadores diferenciais. Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 2t & t^2 & t - 4 \\ 8 & \text{sen}(t) & e^{2t} \end{bmatrix} \quad (2.4)$$

$$\mathbf{A} = \begin{bmatrix} \frac{d^2}{dt^2} & \frac{d}{dt} \\ 4\frac{d}{dt} & \frac{d^2}{dt^2} + 3\frac{d}{dt} \end{bmatrix} \quad (2.5)$$

A notação indicial abreviada $i = 1 : n$ será utilizada no lugar de $i = 1, 2, \dots, n$.

Um caso especial de matriz é um **vetor**, o qual é uma matriz com uma única coluna ou uma única linha. Assim,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad (2.6)$$

são vetores **coluna**. Entretanto

$$\mathbf{v} = [v_1 \quad v_2 \quad \cdots \quad v_n] \quad (2.7)$$

é um vetor **linha**. As matrizes coluna e as matrizes linha serão indistintamente referidas como vetores.

A *ordem* de um vetor é seu número de linhas ou colunas. Para denotar os elementos de um vetor, é requerido somente um subíndice. Assim, um elemento típico de um vetor coluna \mathbf{x} é seu i -ésimo elemento x_i que aparece na linha i . Os vetores são comumente denotados por letras minúsculas e as matrizes por letras maiúsculas.

Algumas vezes, é conveniente considerar linhas ou colunas específicas de uma matriz \mathbf{A} . Para isto, definem-se os vetores

$$\mathbf{a}^i = [a_{i1} \quad a_{i2} \quad \cdots \quad a_{in}] \quad (2.8)$$

e

$$\mathbf{a}_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix} \quad (2.9)$$

para denotar a linha i e coluna j da \mathbf{A} , respectivamente. Esta notação permite escrever uma matriz \mathbf{A} , convenientemente, em termos de suas linhas ou colunas, ou seja :

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \vdots \\ \mathbf{a}^n \end{bmatrix} = [\mathbf{a}_1 \quad \mathbf{a}_2 \quad \cdots \quad \mathbf{a}_n]. \quad (2.10)$$

As matrizes aparecem de maneira natural no contexto de **sistemas de equações algébricas lineares**. Por exemplo, ao sistema

$$\begin{aligned} 4x_1 - 2x_2 + x_3 &= 1 \\ x_1 + x_2 + 3x_3 &= 0 \\ -2x_1 + 5x_2 - 3x_3 &= -2 \end{aligned} \quad (2.11)$$

associa-se a matriz

$$\mathbf{A} = \begin{bmatrix} 4 & -2 & 1 \\ 1 & 1 & 3 \\ -2 & 5 & -3 \end{bmatrix} \quad (2.12)$$

formada pelos coeficientes do sistema, o vetor

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (2.13)$$

formado pelas incógnitas e o vetor

$$\mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \quad (2.14)$$

formado pelos termos não-homogêneos.

As matrizes também podem ser obtidas dos sistemas que provém de **substituições lineares ou transformações lineares** entre duas coleções de variáveis. Assim, ao sistema

$$\begin{aligned} y_1 &= 6x_1 + 3x_2 + x_3 \\ y_2 &= x_1 - x_2 + 2x_3 \\ y_3 &= 5x_1 + x_2 - x_3 \end{aligned} \quad (2.15)$$

são associados à matriz

$$\mathbf{A} = \begin{bmatrix} 6 & 3 & 1 \\ 1 & -1 & 2 \\ 5 & 1 & -1 \end{bmatrix} \quad (2.16)$$

formada pelos coeficientes, a variável independente é dada pelo vetor

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad (2.17)$$

e a variável dependente pelo vetor

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}. \quad (2.18)$$

Coube a Caley a idéia de associar matrizes a sistemas e a Sylvester a introdução do termo "matriz". As matrizes possuem, hoje em dia, um sentido bem mais amplo: podem servir para representar informações de maneira compacta ou armazenar dados das mais variadas aplicações.

2.1 Visualização Gráfica de Vetores e Matrizes

Os vetores numéricos podem ser convenientemente visualizados como gráficos de funções de variáveis discretas. Assim, o vetor

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad (2.19)$$

ou $\mathbf{u} = [u_1 \ u_2 \ \dots \ u_n]$ pode ser descrito pela tabela funcional

i	1	2	3	.	n
u_i	u_1	u_2	u_3	.	u_n

ou

i	u_i
1	u_1
2	u_2
3	u_3
.	.
n	u_n

na qual i é uma variável independente discreta e u_i é a correspondente variável dependente. Graficando estes valores, obtém-se a visualização geométrica do vetor . Por exemplo, o vetor

$$\mathbf{u} = \begin{bmatrix} 4 \\ 6 \\ 3 \\ 8 \\ 0 \end{bmatrix} \quad (2.20)$$

pode ser representado pela figura 2.1

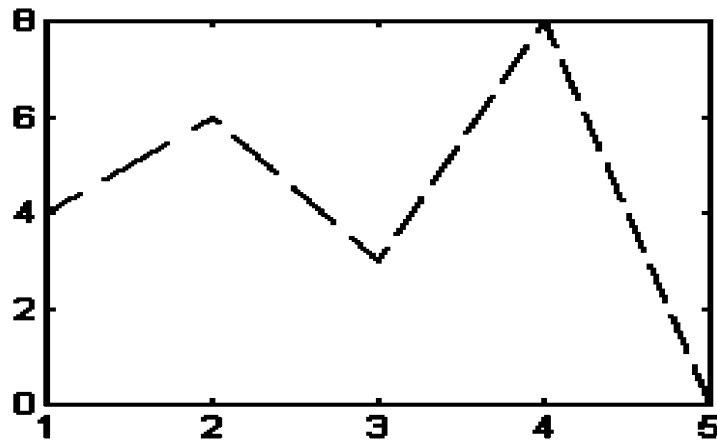


Figura 2.1 – Gráfico de um vetor

De outro lado, considerando o vetor

$$\mathbf{u} = [\text{sen}(\pi/6) \text{ sen}(2\pi/6) \text{ sen}(3\pi/6) \text{ sen}(4\pi/6) \text{ sen}(5\pi/6) \text{ sen}(6\pi/6)] \quad (2.21)$$

cujos elementos são obtidos da função $\text{sen}(\pi x/6)$ fazendo, sucessivamente, $x= 1, 2, 3, 4, 5$ e 6 obtém-se o gráfico

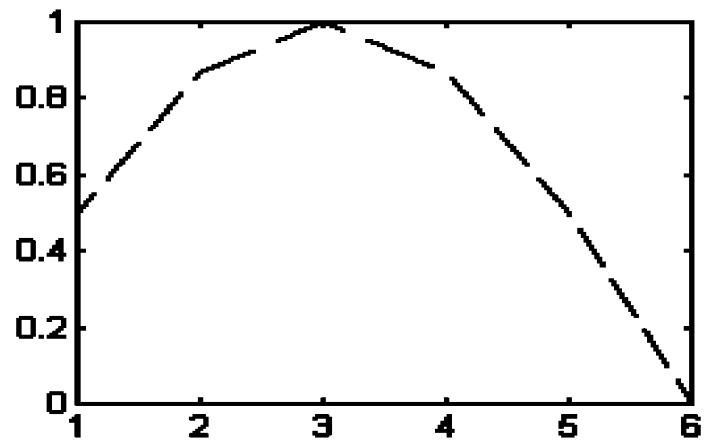


Figura 2.2 – Gráfico de um vetor

Analogamente aos vetores, uma matriz numérica pode ser representada através de uma tabela funcional

i/j	1	2	...	n
1	a_{11}	a_{12}	...	a_{1n}
2	a_{21}	a_{22}	...	a_{2n}
·	·	·	·	·
m	a_{m1}	a_{m2}	...	a_{mn}

com as variáveis independentes discretas i e j e uma variável dependente real a_{ij} . Graficando esta tabela, obtém-se uma valiosa representação geométrica das matrizes numéricas. Por exemplo, a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & -4 \\ -1 & -2 & -1 \end{bmatrix}, \quad (2.22)$$

é representada pelo gráfico,

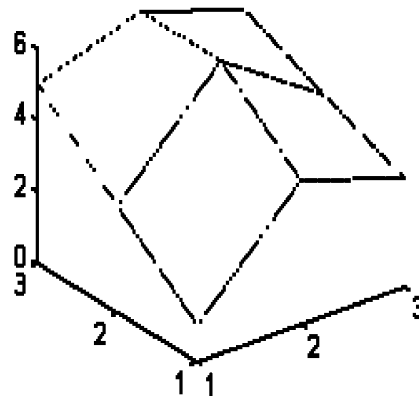


Figura 2.3 – Gráfico de uma matriz

ou, seja, uma superfície que une valores de uma função sobre um domínio discreto no plano.

2.2 Adição e Subtração de Matrizes

A álgebra matricial numérica está fundamentada em quatro operações: **adição**, **multiplicação por escalares**, **multiplicação matricial** e **transposição**. Todas estas operações provém do desejo de resolver sistemas de equações lineares e do manuseio com substituições lineares. Uma vez caracterizadas estas operações, as definições e propriedades são estabelecidas **algebricamente**, isto é, com ênfase nas regras operacionais e não na natureza dos elementos que compõem as matrizes.

Duas matrizes \mathbf{A} e \mathbf{B} com elementos a_{ij} e b_{ij} , respectivamente, são ditas iguais quando são da mesma ordem e todos seus correspondentes elementos são iguais, isto é, $a_{ij} = b_{ij}$ para cada par de índices i, j . Então, escreve-se $\mathbf{A} = \mathbf{B}$.

Matrizes \mathbf{A} e \mathbf{B} da mesma ordem $m \times n$ podem ser adicionadas ou subtraídas termo a termo. Por exemplo

$$\begin{bmatrix} 2 & 3 & -1 \\ 0 & 1 & 2 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 1 \\ 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 2+0 & 3+(-1) & -1+1 \\ 0+2 & 1+0 & 3+2 \end{bmatrix} = \begin{bmatrix} 2 & 2 & 0 \\ 2 & 1 & 5 \end{bmatrix}$$

Formalmente, a **soma** $\mathbf{C} = \mathbf{A} + \mathbf{B}$ e a **diferença** $\mathbf{D} = \mathbf{A} - \mathbf{B}$ são as matrizes com elementos

$$c_{ij} = a_{ij} + b_{ij}, \quad d_{ij} = a_{ij} - b_{ij}, \quad (2.23)$$

para cada $i = 1 : m ; j = 1 : n$, respectivamente.

Quando \mathbf{A} e \mathbf{B} não são da mesma ordem, $\mathbf{A} + \mathbf{B}$ ou $\mathbf{A} - \mathbf{B}$ não estão definidas.

Uma matriz, na qual todo elemento é zero, é chamada de **matriz nula**. Quando não há confusão em relação a sua ordem, a matriz nula escreve-se \mathbf{O} . Verifica-se que

$$\begin{aligned} \mathbf{A} + \mathbf{O} &= \mathbf{A} \\ \mathbf{A} + (-\mathbf{A}) &= \mathbf{O} \end{aligned} \quad (2.24)$$

onde \mathbf{O} é a matriz nula da mesma ordem que \mathbf{A} .

As regras algébricas usuais com números, tais como a comutatividade e associatividade, são válidas para o caso de matrizes. Por exemplo,

$$\begin{aligned} \mathbf{A} + \mathbf{B} &= \mathbf{B} + \mathbf{A} \\ \mathbf{A} + (\mathbf{B} + \mathbf{C}) &= (\mathbf{A} + \mathbf{B}) + \mathbf{C} \end{aligned} \quad (2.25)$$

O **produto** de uma matriz $\mathbf{A} = [a_{ij}]$ de ordem $m \times n$ por um número (ou escalar) α é a matriz $\alpha\mathbf{A}$, cujos elementos são:

$$b_{ij} = \alpha a_{ij}, \quad i = 1 : m ; j = 1 : n \quad (2.26)$$

O efeito de multiplicar uma matriz \mathbf{A} por um número α é chamado de **multiplicação escalar**. Por exemplo, se

$$\mathbf{A} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \end{bmatrix}, \quad \text{então} \quad \alpha\mathbf{A} = \begin{bmatrix} \alpha & 3\alpha \\ 2\alpha & 4\alpha \end{bmatrix}. \quad (2.27)$$

Este tipo de produto possui as mesmas regras da multiplicação com números. Por exemplo,

$$\begin{aligned}\alpha(\beta\mathbf{A}) &= (\alpha\beta)\mathbf{A} \\ \alpha(\mathbf{A} + \mathbf{B}) &= \alpha\mathbf{A} + \alpha\mathbf{B} \\ (\alpha + \beta)\mathbf{A} &= \alpha\mathbf{A} + \beta\mathbf{A} \\ \alpha\mathbf{A} &= \mathbf{A}\alpha\end{aligned}\tag{2.28}$$

2.3 Multiplicação Matricial

A multiplicação de matrizes por matrizes, ou **multiplicação matricial**, difere em importantes aspectos da multiplicação ordinária com números. Particularmente, no referente a comutatividade e a existência de divisores de zero.

Considere-se dois conjuntos de variáveis, relacionados por

$$\begin{aligned}y_1 &= a_{11}z_1 + a_{12}z_2 \\ y_2 &= a_{21}z_1 + a_{22}z_2 \\ y_3 &= a_{31}z_1 + a_{32}z_2\end{aligned}\tag{2.29}$$

ou abreviadamente, $\mathbf{y} = \mathbf{Az}$. Suponha-se que o conjunto de variáveis \mathbf{z} está relacionado a um terceiro conjunto \mathbf{x} através das relações

$$\begin{aligned}z_1 &= b_{11}x_1 + b_{21}x_2 \\ z_2 &= b_{21}x_1 + b_{22}x_2\end{aligned}\tag{2.30}$$

ou seja, $\mathbf{y} = \mathbf{Bx}$. É de interesse escrever $\mathbf{y} = \mathbf{Az} = \mathbf{A}(\mathbf{Bx}) = \mathbf{ABx}$.

Para tanto, substitui-se a segunda relação na primeira e agrupando-se os termos, o que resulta

$$\begin{aligned}y_1 &= (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 \\ y_2 &= (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2 \\ y_3 &= (a_{31}b_{11} + a_{32}b_{21})x_1 + (a_{32}b_{12} + a_{32}b_{22})x_2\end{aligned}\tag{2.31}$$

Este resultado é claramente da forma $\mathbf{y} = \mathbf{Cx}$ onde todos os elementos da matriz \mathbf{C} são somas de produtos de elementos de \mathbf{A} com elementos de \mathbf{B} . A regra de combinação é a seguinte:

$$c_{ij} = [a_{i1} \ a_{i2}] \begin{bmatrix} b_{1j} \\ b_{2j} \end{bmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j}\tag{2.32}$$

O subíndice i denota a localização da linha da matriz \mathbf{A} e o subíndice j o da coluna da matriz \mathbf{B} .

Duas matrizes \mathbf{A} e \mathbf{B} podem ser multiplicadas para formar o produto \mathbf{AB} somente quando o número de colunas de \mathbf{A} for igual ao número de linhas de \mathbf{B} . Matrizes que satisfazem esta condição são ditas **compatíveis**. Se \mathbf{A} tem ordem $m \times n$ e \mathbf{B} é de ordem $n \times p$, o produto \mathbf{AB} é definido pela matriz \mathbf{C} de ordem $m \times p$.

$$\mathbf{C} = [c_{ij}],\tag{2.33}$$

com elementos

$$c_{ij} = \sum_{k=1}^n a_{ik}b_{kj}, \quad i = 1 : m; \quad j = 1 : p.\tag{2.34}$$

Em outras palavras, o elemento c_{ij} é obtido multiplicando os elementos da i -ésima linha da primeira matriz pelos correspondentes elementos da j -ésima coluna da segunda e adicionando os produtos. Assim,

$$c_{ij} = [a_{i1} \ a_{i2} \ \cdots \ a_{in}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{nj} \end{bmatrix} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} \quad (2.35)$$

A notação introduzida por Einstein

$$c_{ij} = a_{ik}b_{kj} \quad (2.36)$$

é as vezes utilizada para escrever o produto de maneira abreviada. O termo $a_{ik}b_{kj}$, com o subíndice k repetido, significa a *soma* de todos tais termos, com k tomando cada um de seus possíveis valores. Assim, por exemplo,

$$a_{ik}b_{kj} = a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} \quad (2.37)$$

quando k varia de 1 a 3.

A multiplicação matricial pode ser escrita de maneira compacta como segue. Utilizando a notação com linhas ou colunas para matrizes, o produto

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \mathbf{a}^3 \\ \vdots \\ \mathbf{a}^m \end{bmatrix} [\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3 \ \cdots \ \mathbf{b}_p] \ , \quad (2.38)$$

é dado pela matriz

$$\mathbf{C} = \begin{bmatrix} \mathbf{a}^1\mathbf{b}_1 & \mathbf{a}^1\mathbf{b}_2 & \mathbf{a}^1\mathbf{b}_3 & \cdots & \mathbf{a}^1\mathbf{b}_p \\ \mathbf{a}^2\mathbf{b}_1 & \mathbf{a}^2\mathbf{b}_2 & \mathbf{a}^2\mathbf{b}_3 & \cdots & \mathbf{a}^2\mathbf{b}_p \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{a}^m\mathbf{b}_1 & \mathbf{a}^m\mathbf{b}_2 & \mathbf{a}^m\mathbf{b}_3 & \cdots & \mathbf{a}^m\mathbf{b}_p \end{bmatrix} . \quad (2.39)$$

Em particular, se \mathbf{A} é uma matriz $m \times p$ e \mathbf{b} um vetor coluna $p \times 1$, obtém-se que

$$\mathbf{Ab} = \begin{bmatrix} \mathbf{a}^1\mathbf{b} \\ \mathbf{a}^2\mathbf{b} \\ \mathbf{a}^3\mathbf{b} \\ \vdots \\ \mathbf{a}^m\mathbf{b} \end{bmatrix} . \quad (2.40)$$

Consequentemente,

$$\mathbf{C} = \mathbf{AB} = \mathbf{A}[\mathbf{b}_1 \ \mathbf{b}_2 \ \mathbf{b}_3 \ \cdots \ \mathbf{b}_p] = [\mathbf{Ab}_1 \ \mathbf{Ab}_2 \ \mathbf{Ab}_3 \ \cdots \ \mathbf{Ab}_p]. \quad (2.41)$$

Assim, o produto pode ser interpretado como a matriz \mathbf{A} atuando sobre cada uma das colunas da matriz \mathbf{B} .

Exemplo 2.7

Determinar o produto \mathbf{AB} para

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & 3 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 0 & -1 \\ 1 & 1 \\ 2 & 0 \\ 1 & 2 \end{bmatrix}.$$

Solução

Como \mathbf{A} possui quatro colunas e \mathbf{B} quatro linhas o produto está bem definido. Além disso, sabe-se que \mathbf{AB} será de ordem 2. Para determinar os elementos de \mathbf{AB} , calcula-se o produto de \mathbf{A} com cada uma das colunas de \mathbf{B}

$$\mathbf{Ab}_1 = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 7 \end{bmatrix}$$

$$\mathbf{Ab}_2 = \begin{bmatrix} 1 & 0 & -1 & 2 \\ 0 & 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 8 \end{bmatrix}$$

Portanto, o produto \mathbf{AB} é dado por

$$\mathbf{AB} = [\mathbf{Ab}_1 \quad \mathbf{Ab}_2] = \begin{bmatrix} 0 & 3 \\ 7 & 8 \end{bmatrix}.$$

Na multiplicação matricial, as matrizes \mathbf{A} , \mathbf{B} e \mathbf{C} são, em geral, de tamanhos diferentes.

Exemplo 2.8

a) Determinar o produto de duas matrizes \mathbf{A} e \mathbf{B} , onde

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 5 & -1 \\ 4 & 9 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 1 & 5 & 7 & 2 \\ 4 & 8 & 1 & 6 \end{bmatrix},$$

logo o produto \mathbf{AB} representada pela matrix \mathbf{C} é dada por

$$\mathbf{C} = \mathbf{AB} = \begin{bmatrix} 11 & 31 & 23 & 18 \\ 1 & 17 & 34 & 4 \\ 40 & 92 & 37 & 62 \end{bmatrix}.$$

b) Determinar o produto da matriz \mathbf{A} pelo vetor \mathbf{b} , onde

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 5 & -1 \\ 4 & 9 \end{bmatrix} \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 4 \end{bmatrix},$$

então o produto de \mathbf{A} por \mathbf{b} resultará na vetor \mathbf{c} dada por

$$\mathbf{c} = \mathbf{Ab} = \begin{bmatrix} 11 \\ 1 \\ 40 \end{bmatrix}.$$

Observe-se que \mathbf{BA} e \mathbf{bA} não podem ser definidos nestes exemplos.

Exemplo 2.9

Calcular \mathbf{ab} e \mathbf{ba} para as matrizes linha e coluna

$$\mathbf{a} = [1 \ 2 \ 3] \quad \text{e} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix}.$$

Solução

$$\mathbf{ab} = [1 \ 2 \ 3] \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} = 1 + 0 + 3 \times 4 = 13,$$

$$\mathbf{ba} = \begin{bmatrix} 1 \\ 0 \\ 4 \end{bmatrix} [1 \ 2 \ 3] = \begin{bmatrix} 1 \times 1 & 1 \times 2 & 1 \times 3 \\ 0 \times 1 & 0 \times 2 & 0 \times 3 \\ 4 \times 1 & 4 \times 2 & 4 \times 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 4 & 8 & 12 \end{bmatrix}.$$

A ordem dos fatores, na multiplicação matricial, é crucial em dois aspectos :

1. A troca de posição dos fatores \mathbf{A} e \mathbf{B} , na multiplicação matricial, pode ocasionar uma indefinição. Os produtos \mathbf{AB} e \mathbf{BA} estão simultaneamente bem definidos, apenas em duas situações:
 - (a) \mathbf{A} , \mathbf{B} são matrizes quadradas de mesma ordem;
 - (b) \mathbf{A} é de ordem $m \times p$ e \mathbf{B} é de ordem $p \times m$.
2. Ainda que os produtos \mathbf{AB} e \mathbf{BA} estejam definidos, em geral, tem-se que $\mathbf{AB} \neq \mathbf{BA}$. Por exemplo, considerando

$$\mathbf{A} = \begin{bmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \\ 2 & 3 \end{bmatrix}, \quad (2.42)$$

resulta

$$\mathbf{AB} = \begin{bmatrix} 10 & 13 \\ 11 & 14 \end{bmatrix} \neq \mathbf{BA} = \begin{bmatrix} 7 & 3 & 8 \\ 11 & 4 & 9 \\ 12 & 5 & 13 \end{bmatrix}. \quad (2.43)$$

Esta não-comutatividade, na multiplicação matricial, também pode ocorrer com matrizes que possuam produtos \mathbf{AB} e \mathbf{BA} da mesma ordem. Por exemplo, para as matrizes

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad (2.44)$$

tem-se que

$$\mathbf{AB} = \begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}, \quad \mathbf{BA} = \begin{bmatrix} -1 & -1 \\ 0 & 1 \end{bmatrix}. \quad (2.45)$$

e, novamente, $\mathbf{AB} \neq \mathbf{BA}$.

Quando $\mathbf{AB} = \mathbf{BA}$, as matrizes \mathbf{A} e \mathbf{B} são ditas **matrizes comutativas**.

De acordo com a definição de multiplicação matricial, o produto de uma matriz \mathbf{A} de ordem $m \times n$ por um vetor coluna \mathbf{x} de ordem $n \times 1$ é um vetor coluna de ordem $m \times 1$, cujos elementos são dados por

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \quad (2.46)$$

ou de maneira abreviada:

$$\sum_{k=1}^n a_{ik}x_k, \quad i = 1 : m. \quad (2.47)$$

Assim,

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}^1 \\ \mathbf{a}^2 \\ \vdots \\ \mathbf{a}^m \end{bmatrix} \mathbf{x} = \begin{bmatrix} \mathbf{a}^1 \mathbf{x} \\ \mathbf{a}^2 \mathbf{x} \\ \vdots \\ \mathbf{a}^m \mathbf{x} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^n a_{1k}x_k \\ \sum_{k=1}^n a_{2k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{bmatrix} \quad (2.48)$$

Introduzindo os vetores $n \times 1$

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \quad (2.49)$$

pode-se obter a j -ésima coluna \mathbf{a}_j de uma matriz \mathbf{A} $n \times n$ pela simples multiplicação com \mathbf{e}_j , isto é,

$$\mathbf{Ae}_j = \mathbf{a}_j \quad (2.50)$$

Utilizando a definição de igualdade de matrizes e multiplicação matricial, decorre que o sistema de equações algébricas lineares

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
\cdots &= \cdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned} \tag{2.51}$$

pode ser escrito como uma simples equação matricial

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \tag{2.52}$$

ou

$$\mathbf{Ax} = \mathbf{b}. \tag{2.53}$$

Similarmente, um sistema de substituições lineares ou das variáveis (u_1, u_2, \dots, u_n) para as variáveis (v_1, v_2, \dots, v_m) , tal como

$$\begin{aligned}
v_1 &= a_{11}u_1 + a_{12}u_2 + \cdots + a_{1n}u_n \\
v_2 &= a_{21}u_1 + a_{22}u_2 + \cdots + a_{2n}u_n \\
\vdots & \\
v_m &= a_{m1}u_1 + a_{m2}u_2 + \cdots + a_{mn}u_n
\end{aligned} \tag{2.54}$$

é escrito na forma matricial compacta

$$\mathbf{v} = \mathbf{Au}, \tag{2.55}$$

onde a matriz $\mathbf{A} = [a_{ij}]$, formada pelos coeficientes da substituição, caracteriza a transformação. As variáveis independente e dependente são os vetores

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}, \quad \mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}. \tag{2.56}$$

Por exemplo, a transformação no plano cartesiano

$$\begin{aligned}
x' &= x + y \\
y' &= y
\end{aligned} \tag{2.57}$$

pode ser escrita na forma matricial

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} \tag{2.58}$$

ou, simplesmente,

$$\mathbf{v} = \mathbf{Au}, \text{ onde } \mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{2.59}$$

Em particular, o vetor \mathbf{u} com elementos $x = 2$ e $y = 1$ é transformado no vetor \mathbf{v} com elementos $x' = 3$ e $y' = 1$.

Em \mathbf{AB} , \mathbf{A} é dito que "pré-multiplica" \mathbf{B} , e \mathbf{B} é dito que "pós-multiplica" \mathbf{A} . Produtos de três ou mais matrizes podem existir, desde que, em sequência, sejam conformes. Neste caso, a lei associativa é obedecida, isto é,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}) , \quad (2.60)$$

Similarmente, tem-se a validade da lei distributiva

$$(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC} , \quad \mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}. \quad (2.61)$$

A **matriz unidade** ou **matriz identidade** \mathbf{I} de ordem n é a matriz cujos elementos da diagonal são iguais à unidade e os elementos fora dela são zero.

$$\mathbf{I} = \text{diag}[1 \ 1 \ 1 \ \dots \ 1] = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix} \quad (2.62)$$

Quando a ordem é evidente, a matriz identidade é denotada por \mathbf{I} . A matriz identidade tem algumas propriedades da unidade 1. Por exemplo, para \mathbf{A} uma matriz quadrada de ordem n e \mathbf{x} um vetor $n \times 1$, tem-se

$$\mathbf{AI} = \mathbf{IA} = \mathbf{A} , \quad \text{e} \quad \mathbf{Ix} = \mathbf{x} . \quad (2.63)$$

2.3.1 Potências de uma Matriz Quadrada e Polinômios Matriciais

Potências de uma Matriz Quadrada

Uma matriz pode ser multiplicada por ela mesma se, e somente se, é uma matriz quadrada e, neste caso, a notação indicial pode ser convenientemente utilizada. Assim, se \mathbf{A} é uma matriz quadrada de ordem n , define-se

$$\mathbf{A}^0 = \mathbf{I} , \quad \mathbf{A}^r = \mathbf{A}^{r-1}\mathbf{A} , \quad r \geq 1. \quad (2.64)$$

como sendo a r -ésima potência de uma matriz. Em virtude da lei associativa, tem-se

$$\mathbf{A}^1 = \mathbf{A}, \quad \mathbf{A}^2 = \mathbf{A.A}, \quad \mathbf{A}^3 = \mathbf{AAA}, \quad \mathbf{A}^4 = \mathbf{AAAA} , \quad (2.65)$$

e assim sucessivamente.

A propriedade dos expoentes é válida:

$$\mathbf{A}^{r+s} = \mathbf{A}^r \mathbf{A}^s , \quad (2.66)$$

para r e s inteiros não-negativos.

Polinômios Matriciais

A partir de um polinômio de grau m na variável z ,

$$p(z) = a_m z^m + a_{m-1} z^{m-1} + \dots + a_1 z + a_0, \quad (2.67)$$

com coeficientes numéricos, substituindo a variável z por uma matriz quadrada \mathbf{A} e o escalar a_o pela matriz $a_o\mathbf{I}$, obtém-se o **polinômio matricial**

$$p(\mathbf{A}) = a_m\mathbf{A}^m + a_{m-1}\mathbf{A}^{m-1} + \cdots + a_1\mathbf{A} + a_o\mathbf{I}. \quad (2.68)$$

É fácil ver que a igualdade de polinômios e as operações de soma e multiplicação

$$p(z) = h(z), \quad p(z) + q(z) = h(z) \quad \text{e} \quad p(z)q(z) = g(z), \quad (2.69)$$

implicam os análogos matriciais,

$$p(\mathbf{A}) = h(\mathbf{A}), \quad p(\mathbf{A}) + q(\mathbf{A}) = h(\mathbf{A}) \quad \text{e} \quad p(\mathbf{A})q(\mathbf{A}) = g(\mathbf{A}), \quad (2.70)$$

pois os coeficientes das potências de z , na forma expandida de $p = h$, $p + q$ ou pq podem ser identificados, termo a termo, com os coeficientes das correspondentes potências de \mathbf{A} .

Como consequência, a adição e a multiplicação (para expressões polinomiais em uma mesma matriz \mathbf{A}) são comutativas. Decorre, portanto, uma perfeita analogia , entre a álgebra dos polinômios escalares em uma variável e a álgebra dos polinômios matriciais. Deste modo, uma importante propriedade dos polinômios escalares pode ser diretamente estendida ao caso matricial:

Se z_1, z_2, \dots, z_m são as m raízes de $p(z) = 0$, tem-se

$$p(z) = a_m(z - z_1)(z - z_2)\cdots(z - z_m). \quad (2.71)$$

O correspondente polinômio matricial pode ser escrito na forma fatorizada, isto é,

$$p(\mathbf{A}) = a_m(\mathbf{A} - z_1\mathbf{I})(\mathbf{A} - z_2\mathbf{I})\cdots(\mathbf{A} - z_m\mathbf{I}). \quad (2.72)$$

Exemplo 2.10

$$\text{Seja } \mathbf{A} = \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix} \quad \text{e suponha-se } p(z) = 3z^2 - 9z + 6,$$

então

$$p(\mathbf{A}) = 3 \begin{bmatrix} 5 & -8 \\ 2 & -3 \end{bmatrix} - 9 \begin{bmatrix} 3 & -4 \\ 1 & -1 \end{bmatrix} + 6 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix}.$$

Por outro lado, se $p(z) = 3z^2 - 6z + 6 = 3(z - 1)(z - 2)$ tem-se

$$p(\mathbf{A}) = 3(\mathbf{A} - \mathbf{I})(\mathbf{A} - 2\mathbf{I}) = 3 \begin{bmatrix} 2 & -4 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 1 & -4 \\ 1 & -3 \end{bmatrix} = \begin{bmatrix} -6 & 12 \\ -3 & 6 \end{bmatrix}$$

Deve-se observar que, também, $p(\mathbf{A}) = 3(\mathbf{A} - 2\mathbf{I})(\mathbf{A} - \mathbf{I})$ devido à comutatividade dos fatores.

Exemplo 2.11

Estabelecer que

$$I - \mathbf{A}^k = (I - \mathbf{A})(I + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1}) = (I + \mathbf{A} + \mathbf{A}^2 + \cdots + \mathbf{A}^{k-1})(I - \mathbf{A}).$$

Solução

Decorre da fatorização polinomial

$$1 - z^k = (1 - z)(1 + z + z^2 + \cdots + z^{k-1})$$

e da comutatividade das potências de \mathbf{A} .

A conhecida fórmula do binômio de Newton ,

$$(\mathbf{A} + \mathbf{B})^m = \sum_{k=0}^m \binom{m}{k} \mathbf{A}^{m-k} \mathbf{B}^k ,$$

também pode ser utilizada para matrizes quadradas \mathbf{A} e \mathbf{B} da mesma ordem desde que comutem, $\mathbf{AB} = \mathbf{BA}$. A fórmula é estabelecida por indução, de modo análogo ao da álgebra elementar.

Exemplo 2.12

Calcular \mathbf{A}^{10} onde \mathbf{A} é a matriz

$$\mathbf{A} = \begin{bmatrix} \alpha & 1 & 0 \\ 0 & \alpha & 1 \\ 0 & 0 & \alpha \end{bmatrix} .$$

Solução

A matriz pode ser escrita na forma

$$\mathbf{A} = \mathbf{D} + \mathbf{N} , \quad \mathbf{D} = \begin{bmatrix} \alpha & & \\ & \alpha & \\ & & \alpha \end{bmatrix} , \quad \mathbf{N} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} .$$

Observe-se que \mathbf{D} comuta com \mathbf{N} e que $\mathbf{N}^3 = 0$, ou seja \mathbf{N} é nilpotente e, portanto, $\mathbf{N}^k = 0$ para $k \geq 3$. Assim, do binômio de Newton

$$\mathbf{A}^{10} = \mathbf{D}^{10} + 10\mathbf{D}^9\mathbf{N} + 45\mathbf{N}^2 = \begin{bmatrix} \alpha^{10} & \alpha^9 & 45 \\ 0 & \alpha^{10} & \alpha^9 \\ 0 & 0 & \alpha^{10} \end{bmatrix} .$$

2.3.2 Matriz Transposta

O sistema linear

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 &= b_2 \end{aligned} \tag{2.73}$$

é escrito na forma matricial $\mathbf{Ax} = \mathbf{b}$ considerando as incógnitas e termos não-homogêneos como vetores *coluna*, isto é

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad b = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{2.74}$$

e

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \tag{2.75}$$

Porém, pode-se igualmente considerar vetores *linha* para as incógnitas e os termos não-homogêneos. Nesta situação, o sistema é representado por $\mathbf{x}^* \mathbf{A}^* = \mathbf{b}^*$; onde

$$\mathbf{x}^* = [x_1 \quad x_2 \quad x_3] \quad \mathbf{b}^* = [b_1 \quad b_2] \tag{2.76}$$

e

$$\mathbf{A}^* = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \\ a_{13} & a_{23} \end{bmatrix} \quad (2.77)$$

Observa-se que \mathbf{x}^* , \mathbf{b}^* e \mathbf{A}^* são matrizes formadas a partir de \mathbf{x} , \mathbf{b} e \mathbf{A} , respectivamente, trocando suas linhas pelas suas colunas.

Em geral, se numa matriz $m \times n$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad (2.78)$$

as linhas são permutadas pelas colunas, obtém-se uma matriz de ordem $n \times m$ denominada a **matriz transposta** de \mathbf{A} :

$$\mathbf{A}^t = \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \cdots & \cdots & \cdots & \cdots \\ a_{1n} & a_{2n} & \cdots & a_{mn} \end{bmatrix}. \quad (2.79)$$

Em particular, a transposta de um vetor linha

$$\mathbf{v} = [v_1 \ v_2 \ \cdots \ v_n] \quad (2.80)$$

é o vetor coluna

$$\mathbf{v}^t = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad (2.81)$$

e vice-versa.

Dada uma matriz $\mathbf{A} = [a_{ij}]$ de ordem $n \times m$ define-se sua *matriz transposta* como sendo a matriz $\mathbf{A}^* = [a_{ji}]$ de ordem $m \times n$. As notações \mathbf{A}' ou \mathbf{A}^t também são utilizadas para denotar a transposta de \mathbf{A} .

As transpostas dos vetores coluna definidos anteriormente \mathbf{e}_j $n \times 1$ são os vetores linha $n \times 1$

$$\mathbf{e}^1 = [1 \ 0 \ \cdots \ 0], \quad \mathbf{e}^2 = [0 \ 1 \ \cdots \ 0], \quad \dots, \quad \mathbf{e}^n = [0 \ 0 \ \cdots \ 0 \ 1] \quad (2.82)$$

Os elementos de uma matriz \mathbf{A} podem ser obtidos pré-multiplicando- pelo vetor linha \mathbf{e}^i e pós-multiplicando pelo vetor coluna \mathbf{e}_j , ou seja

$$a_{ij} = \mathbf{e}^i \mathbf{A} \mathbf{e}_j \quad (2.83)$$

Uma matriz e sua transposta são da mesma ordem unicamente quando o número de linhas é igual ao número de colunas, ou seja, matrizes quadradas. Por outro lado, os produtos $\mathbf{A}^* \mathbf{A}$ e $\mathbf{A} \mathbf{A}^*$ estão sempre bem definidos para qualquer matriz \mathbf{A} .

Se $\mathbf{C} = \mathbf{A} \mathbf{B}$, então

$$\mathbf{C}^* = \mathbf{B}^* \mathbf{A}^* \quad (2.84)$$

isto é, a transposição reverte a ordem dos fatores!

Com relação ao exemplo 2.2.8, em geral, o produto de um vetor coluna \mathbf{x} de ordem n e um vetor linha \mathbf{y}^t de ordem n , é uma matriz quadrada de ordem n

$$\mathbf{xy}^t = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} [y_1 \ y_2 \ \cdots \ y_n] = \begin{bmatrix} x_1y_1 & x_1y_2 & \cdots & x_1y_n \\ x_2y_1 & x_2y_2 & \cdots & x_2y_n \\ \vdots & \vdots & \vdots & \vdots \\ x_ny_1 & x_ny_2 & \cdots & x_ny_n \end{bmatrix}. \quad (2.85)$$

Entretanto, o produto de \mathbf{x}^t e \mathbf{y} é um escalar

$$\mathbf{x}^t\mathbf{y} = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + x_2y_2 + \cdots + x_ny_n. \quad (2.86)$$

2.3.3 Multiplicação Matricial através de Somas

A maneira usual da multiplicação matricial, "linhas por colunas", pode ser convenientemente modificada para uma regra de "soma de colunas vezes linhas". Para isto, deve-se observar que a i -ésima linha do produto \mathbf{AB} é escrita na forma

$$\begin{aligned} [\sum_{k=1}^p a_{ik}b_{k1} \quad \sum_{k=1}^p a_{ik}b_{k2} \quad \cdots \quad \sum_{k=1}^p a_{ik}b_{kp}] &= \sum_{k=1}^p a_{ik}[b_{k1} \quad b_{k2} \quad \cdots \quad b_{kp}] \\ &= \sum_{k=1}^p a_{ik}\mathbf{b}^k \end{aligned} \quad (2.87)$$

Assim, obtém-se que

$$\mathbf{AB} = \begin{bmatrix} \sum_{k=1}^p a_{1k} \mathbf{b}^k \\ \sum_{k=1}^p a_{2k} \mathbf{b}^k \\ \vdots \\ \sum_{k=1}^p a_{mk} \mathbf{b}^k \end{bmatrix} = \sum_{k=1}^p \begin{bmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{bmatrix} \mathbf{b}^k = \sum_{k=1}^p \mathbf{a}_k \mathbf{b}^k, \quad (2.88)$$

isto é, a soma do produto das colunas de \mathbf{A} com as correspondentes linhas de \mathbf{B} . Observe-se que $\mathbf{a}_k \mathbf{b}^k$ é uma matriz de ordem $m \times n$.

Exemplo 2.13

Realizar o produto das matrizes \mathbf{A} e \mathbf{B}

$$\text{Se } \mathbf{A} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix},$$

então

$$\begin{aligned} \mathbf{AB} &= \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} [2 \ 4] + \begin{bmatrix} 0 \\ 1 \end{bmatrix} [0 \ 3] \\ &= \begin{bmatrix} 2 & 4 \\ 4 & 8 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 4 \\ 4 & 11 \end{bmatrix}. \end{aligned}$$

2.3.4 Matrizes Simétricas

Uma matriz \mathbf{A} é dita **simétrica** quando coincide com sua transposta:

$$\mathbf{A}^t = \mathbf{A}. \quad (2.89)$$

Segue-se que uma matriz simétrica é quadrada e seus elementos simétricos em relação à diagonal principal são iguais, ou seja,

$$a_{ji} = a_{ij} \quad \text{para} \quad i \neq j \quad (2.90)$$

Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 2 & 4 \\ 4 & 1 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 3 & 5 \\ 0 & 5 & 4 \end{bmatrix}$$

são matrizes simétricas. Para sistemas lineares, cuja matriz de coeficientes é simétrica, podemos indistintamente escrever $\mathbf{Ax} = \mathbf{b}$ com \mathbf{x} , \mathbf{b} vetores coluna ou $\mathbf{x}\mathbf{A} = \mathbf{b}$, com \mathbf{x} , \mathbf{b} vetores linha, ou seja, a matriz é a mesma!

As matrizes simétricas são muito importantes na teoria e na prática. Elas podem, também, ser caracterizadas pela seguinte propriedade:

Uma matriz \mathbf{A} quadrada de ordem n é simétrica se, e somente se

$$\mathbf{x}^t \mathbf{A} \mathbf{y} = \mathbf{y}^t \mathbf{A} \mathbf{x} \quad (2.91)$$

para quaisquer vetores coluna \mathbf{x} , \mathbf{y} de ordem $n \times 1$.

De fato, se \mathbf{A} é simétrica, então da relação numérica

$$\mathbf{x}^t \mathbf{A} \mathbf{y} = (\mathbf{x}^t \mathbf{A} \mathbf{y})^t \quad (2.92)$$

obtém-se que $\mathbf{x}^t \mathbf{A} \mathbf{y} = \mathbf{y}^t \mathbf{A}^t \mathbf{x} = \mathbf{y}^t \mathbf{A} \mathbf{x}$. Por outro lado, se essa relação é válida, então é suficiente escolher $\mathbf{x} = \mathbf{e}^i$, $\mathbf{y} = \mathbf{e}_j$ para concluir que $a_{ij} = a_{ji}$.

É importante salientar que

Para qualquer matriz \mathbf{A} de ordem $m \times n$, a matriz

$$\mathbf{K} = \mathbf{A}^t \mathbf{A} \quad (2.93)$$

é sempre uma matriz simétrica.

Pois, utilizando a propriedade relativa à transposição do produto,

$$\mathbf{K}^t = (\mathbf{A}^t \mathbf{A})^t = \mathbf{A}^t (\mathbf{A}^t)^t = \mathbf{A}^t \mathbf{A} = \mathbf{K}. \quad (2.94)$$

Outras propriedades das matrizes simétricas serão estabelecidas futuramente.

2.3.5 Conjugada de uma Matriz

Sejam a e b números reais e $i = \sqrt{-1}$. Define-se $z = a + ib$ como sendo um **número complexo**.

Se $z = a + bi$, seu conjugado é definido e denotado por $\bar{z} = a - bi$. Se $z = a + bi$ e $w = \bar{z} = a - bi$, então $\overline{\bar{w}} = \bar{\bar{z}} = \overline{a - bi} = a + bi$, isto é, o conjugado do conjugado de um número complexo z é o próprio z .

Se $z = a + bi$ e $w = c + di$, então

$$1. z + w = (a + c) + (b + d)i$$

$$\overline{z + w} = (a + c) - (b + d)i = \bar{z} + \bar{w},$$

isto é, o conjugado da soma de dois números complexos é a soma de seus conjugados.

$$2. zw = (ac - bd) + (ad + bc)i$$

$$\overline{zw} = (ac - bd) - (ad + bc)i = (a - bi)(c - di) = \bar{z} \bar{w},$$

isto é, o conjugado do produto de dois números complexos é o produto de seus conjugados.

Dada uma matriz \mathbf{A} com números complexos como elementos, a matriz obtida de \mathbf{A} pela substituição de cada elemento por seu conjugado é chamada de **matriz conjugada de A** e denotada por $\overline{\mathbf{A}}$ (leia-se \mathbf{A} conjugada). Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 1 + 2i & i \\ 3 & 2 - 3i \end{bmatrix} \quad \text{e} \quad \overline{\mathbf{A}} = \begin{bmatrix} 1 - 2i & -i \\ 3 & 2 + 3i \end{bmatrix}$$

Utilizando as propriedades 1 e 2 dos números complexos, pode ser estabelecido que

$$\overline{(\mathbf{A} + \mathbf{B})} = \overline{\mathbf{A}} + \overline{\mathbf{B}} \quad \text{e} \quad \overline{(\mathbf{A}\mathbf{B})} = \overline{\mathbf{A}} \overline{\mathbf{B}}.$$

2.3.6 Matrizes Hermitianas

Uma matriz \mathbf{A} é dita **matriz Hermitiana** quando coincide com a conjugada da sua transposta:

$$\mathbf{A} = \overline{(\mathbf{A}^t)}. \quad (2.95)$$

Segue-se que uma matriz hermitiana é quadrada e seus elementos, simétricos em relação à diagonal, são conjugados, ou seja,

$$a_{ij} = \overline{a_{ji}}, \quad \text{para } i \neq j, \quad (2.96)$$

e os elementos na diagonal são números reais:

$$a_{ii} = \overline{a_{ii}}. \quad (2.97)$$

Por exemplo,

$$\mathbf{A} = \begin{bmatrix} 1 & 2 - i \\ 2 + i & 5 \end{bmatrix}, \quad \mathbf{A}^t = \begin{bmatrix} 1 & 2 + i \\ 2 - i & 5 \end{bmatrix}, \quad \overline{\mathbf{A}^t} = \begin{bmatrix} 1 & 2 - i \\ 2 + i & 5 \end{bmatrix},$$

portanto \mathbf{A} é hermitiana.

Usualmente, para cada matriz \mathbf{A} com números complexos denota-se por \mathbf{A}^* a matriz conjugada da transposta de \mathbf{A} , chamada de **matriz adjunta**. Assim

$$\mathbf{A}^* = \overline{(\mathbf{A}^t)}. \quad (2.98)$$

Certamente, uma matriz é Hermitiana quando $\mathbf{A} = \mathbf{A}^*$.

2.4 Submatrizes

Uma **submatriz** de uma matriz \mathbf{A} é uma matriz obtida de \mathbf{A} , suprimindo-se certas colunas ou linhas, ou é a própria matriz \mathbf{A} .

Exemplo 2.14

Determinar todas as submatrizes da seguinte matriz 2×3

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & -2 \end{bmatrix}$$

Tem-se submatrizes de ordens 2×3 , 2×2 , 2×1 , 1×2 , 1×1 e 1×3 . Sendo,

- 2×3 : \mathbf{A} ;
- 2×2 : $\begin{bmatrix} 2 & 1 \\ 0 & 3 \end{bmatrix}$, $\begin{bmatrix} 2 & -1 \\ 0 & -2 \end{bmatrix}$ e $\begin{bmatrix} 1 & -1 \\ 3 & -2 \end{bmatrix}$;
- 2×1 : $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 1 \\ 3 \end{bmatrix}$ e $\begin{bmatrix} -1 \\ -2 \end{bmatrix}$;
- 1×2 : $[2 \ 1]$, $[0 \ 3]$, $[2 \ -1]$, $[0 \ -2]$, $[1 \ -1]$, $[3 \ -2]$;
- 1×1 : $[2]$, $[1]$, $[-1]$, $[0]$, $[3]$, $[-2]$.
- 1×3 : $[2 \ 1 \ -1]$, $[0 \ 3 \ -2]$.

2.5 Matrizes Bloco

Matrizes, cujos elementos são matrizes, são denominadas **matrizes bloco**. Além de suas inerentes características matriciais, as matrizes bloco servem como um artifício operacional, que permite indicar aspectos especiais de uma matriz ou particioná-la.

Por exemplo,

$$\mathbf{A} = \left[\begin{array}{cc|cc} 1 & 2 & -1 & 0 \\ 3 & 4 & 0 & 1 \\ \hline -1 & 0 & 3 & 5 \\ 5 & 6 & 2 & 1 \\ \hline -7 & 8 & -1 & 3 \end{array} \right],$$

é uma matriz bloco. De outro lado, a seguinte matriz de ordem 10.

$$\mathbf{A} = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

pode ser escrita na forma de matriz bloco

$$\mathbf{A} = \left[\begin{array}{ccccc|ccccc} -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{array} \right].$$

Equivalentemente,

$$\mathbf{A} = \begin{bmatrix} \mathbf{T} & \mathbf{I} \\ \mathbf{0} & \mathbf{T} \end{bmatrix},$$

com

$$\mathbf{T} = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}, \quad \mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

A matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

pode ser representada pela matriz bloco

$$\mathbf{A} = \begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix}$$

onde

$$\mathbf{P} = \begin{bmatrix} 1 & 2 \\ 4 & 5 \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}, \quad \mathbf{R} = [7 \ 8] \quad \text{e} \quad \mathbf{S} = [9].$$

Similarmente, o vetor coluna

$$\mathbf{u} = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 9 \\ -1 \\ 2 \\ 4 \\ -7 \end{bmatrix},$$

pode ser denotado pelo vetor bloco

$$\mathbf{u} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix}$$

com

$$\mathbf{a} = \begin{bmatrix} 3 \\ 1 \\ 0 \\ 9 \end{bmatrix}, \quad \mathbf{b} = [-1] \quad \text{e} \quad \mathbf{c} = \begin{bmatrix} 2 \\ 4 \\ -7 \end{bmatrix}.$$

Certamente, a partição de matrizes em matrizes bloco pode ser feita de várias maneiras. É necessário, entretanto, observar as ordens dos blocos de tal modo, que as posições dos elementos sejam preservadas, bem como, seja mantida a ordem da matriz original. Assim, uma forma alternativa de escrever a matriz anterior em blocos é

$$\mathbf{P} = [1 \ 2], \quad \mathbf{Q} = \begin{bmatrix} 4 & 5 \\ 7 & 8 \end{bmatrix}, \quad \mathbf{R} = [3] \quad \text{e} \quad \mathbf{S} = \begin{bmatrix} 6 \\ 9 \end{bmatrix}.$$

As matrizes bloco decorrem, frequentemente, do particionamento de sistemas. Por exemplo, suponha-se que num sistema de n equações com n incógnitas, $\mathbf{Ax} = \mathbf{b}$, são separadas as primeiras $n - r$ equações e, logo, a seguir, separam-se as primeiras r incógnitas por uma linha de divisão vertical. Denote-se a coluna das primeiras r incógnitas por \mathbf{x}_1 e as restantes pela coluna \mathbf{x}_2 . Similarmente para \mathbf{b} . Isto pode ser exibido esquematicamente como

$$\begin{bmatrix} \mathbf{P} & \mathbf{Q} \\ \mathbf{R} & \mathbf{S} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{bmatrix}$$

Aqui as colunas $\mathbf{x}_1, \mathbf{b}_1$ são $r \times 1$, entretanto, \mathbf{x}_2 e \mathbf{b}_2 são $(n - r) \times 1$. \mathbf{P} e \mathbf{S} são matrizes quadradas de ordem r e $n - r$, respectivamente, e as matrizes \mathbf{Q} e \mathbf{R} são retangulares.

2.6 Alguns Tipos de Matrizes

Existem vários tipos de matrizes que possuem uma "estrutura" ou "padrão", ou seja, uma forma determinada. A seguir, alguns tipos serão considerados.

2.6.1 Matrizes Triangulares

Uma matriz quadrada \mathbf{A} , cujos elementos $a_{ij} = 0$ para $i > j$, é chamada de **matriz triangular superior**; uma matriz quadrada \mathbf{A} , cujos elementos $a_{ij} = 0$ para $i < j$, é chamada de **matriz triangular inferior**.

Assim,

$$\mathbf{U} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix} \quad (2.99)$$

é triangular superior e

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & \cdots & 0 \\ l_{21} & l_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \cdots & l_{nn} \end{bmatrix} \quad (2.100)$$

é triangular inferior. Certamente, a transposta de \mathbf{U} é uma matriz do tipo \mathbf{L} e vice-versa.

Deve ser observado que a soma e o produto \mathbf{AB} de duas matrizes triangulares superiores (inferiores) é uma matriz triangular superior(inferior), pois no caso do produto, para $i > j$,

$$\underbrace{[0 \ 0 \ 0 \ \cdots \ 0]}_i \ [a_{ii} \ \cdots \ a_{in}] \begin{bmatrix} b_{1j} \\ b_{2j} \\ \vdots \\ b_{jj} \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = 0. \quad (2.101)$$

Por exemplo,

$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} \begin{bmatrix} 3 & 2 & 1 \\ 0 & 5 & 4 \\ 0 & 0 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 12 & 27 \\ 0 & 20 & 46 \\ 0 & 0 & 36 \end{bmatrix}$$

Para o caso de matrizes triangulares inferiores, pode-se proceder de maneira análoga ou utilizar a propriedade da transposta do produto com matrizes triangulares superiores.

2.6.2 Matrizes Diagonais

A matriz quadrada

$$\mathbf{D} = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}, \quad (2.102)$$

a qual é simultaneamente triangular superior e inferior, é chamada de **matriz diagonal**. Será escrita de maneira abreviada como

$$D = \text{diag}[d_{11} \ d_{22} \ \cdots \ d_{nn}]. \quad (2.103)$$

O produto \mathbf{DA} de uma matriz diagonal \mathbf{D} por uma matriz \mathbf{A}

$$\begin{bmatrix} \alpha_1 & 0 & \cdots & 0 \\ 0 & \alpha_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \alpha_n \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} \alpha_1 a_{11} & \alpha_1 a_{12} & \cdots & \alpha_1 a_{1n} \\ \alpha_2 a_{21} & \alpha_2 a_{22} & \cdots & \alpha_2 a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n a_{n1} & \alpha_n a_{n2} & \cdots & \alpha_n a_{nn} \end{bmatrix} \quad (2.104)$$

é obtido multiplicando a primeira linha de \mathbf{A} por α_1 , a segunda linha de \mathbf{A} por α_2 , e assim, sucessivamente. No caso do produto \mathbf{AD} , a multiplicação é com as colunas em vez das linhas de \mathbf{A}

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \beta_1 & 0 & \cdots & 0 \\ 0 & \beta_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \beta_n \end{bmatrix} = \begin{bmatrix} \beta_1 a_{11} & \beta_2 a_{12} & \cdots & \beta_n a_{1n} \\ \beta_1 a_{21} & \beta_2 a_{22} & \cdots & \beta_n a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_1 a_{n1} & \beta_2 a_{n2} & \cdots & \beta_n a_{nn} \end{bmatrix} \quad (2.105)$$

2.6.3 Matrizes Tridiagonais e Matrizes Banda

Das mais variadas aplicações surgem matrizes que possuem um determinado padrão de elementos que são zeros. Dentre estas, destacam-se as **matrizes tridiagonais**

$$\begin{bmatrix} a_1 & b_1 & 0 & & \cdots & 0 \\ c_1 & a_2 & b_2 & 0 & & \cdots & 0 \\ 0 & c_2 & a_3 & b_3 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & & \cdots & c_{n-2} & a_{n-1} & b_{n-1} \\ 0 & & \cdots & 0 & c_{n-1} & a_n \end{bmatrix}. \quad (2.106)$$

Mais geralmente, uma **matriz banda** de ordem n e de largura $2k + 1$

$$a_{ij} = 0 \quad \text{para} \quad i - j > k, \quad (2.107)$$

para algum inteiro k não negativo entre 0 e $n - 1$. Numa matriz banda todos os elementos fora da diagonal e das primeiras k sub-diagonais, acima ou abaixo da diagonal, são nulos. A forma de tal matriz é ilustrada pela figura 2.4

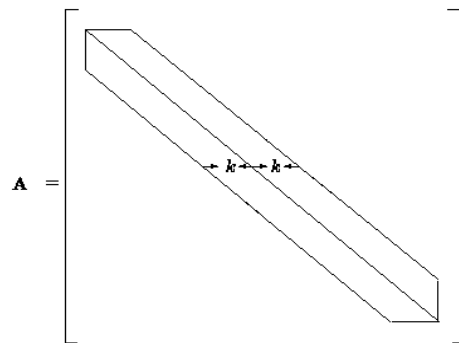


Figura 2.4 – Matriz banda

2.6.4 Matrizes Estocásticas

Para matrizes, cujos elementos são números reais, tem-se **matrizes não-negativas**, nas quais todos os elementos são números não-negativos. Quando todos os elementos são estritamente positivos, uma matriz não-negativa é dita **matriz positiva**. Similarmente, são definidas as matrizes não-positivas e matrizes negativas.

Uma classe importante de matrizes não-negativas, é constituída pelas **matrizes estocásticas**

$$\mathbf{P} = [p_{ij}] ,$$

em que os p_{ij} são números entre 0 e 1 e satisfazem a propriedade

$$\sum_{i=1}^n p_{ij} = 1$$

para cada coluna da matriz \mathbf{P} . Quando numa matriz estocástica, a propriedade é também válida para as linhas

$$\sum_{j=1}^n p_{ij} = 1,$$

ela é dita *duplamente* estocástica.

2.6.5 Matrizes Booleanas e de Permutação

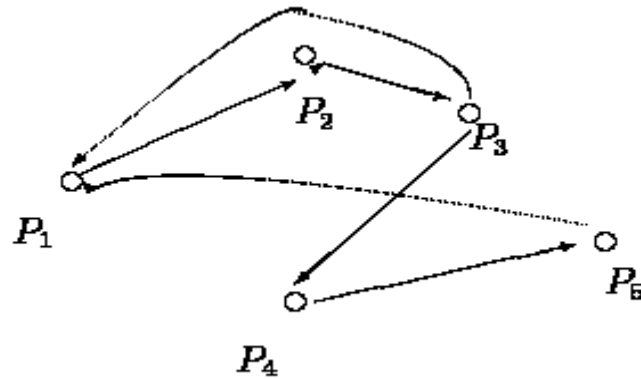
Matrizes numéricas em que todos seus elementos são 0 ou 1, são chamadas de **matrizes booleanas**. Dois casos particulares de tais matrizes são:

- **Matrizes de Permutação**. São obtidas através da troca de linhas na matriz identidade \mathbf{I} . Por exemplo,

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

é uma matriz de permutação, obtida da identidade \mathbf{I}_4 trocando-se a primeira linha com a terceira e esta com a segunda linha. As matrizes de permutação possuem as seguintes propriedades:

1. $\mathbf{P}\mathbf{P}^t = \mathbf{P}^t\mathbf{P} = \mathbf{I}$;
 2. O produto de duas matrizes de permutação é uma matriz de permutação;
 3. A transposta de uma matriz de permutação é também uma matriz de permutação.
- **Matrizes de Incidência**. A estrutura de zeros e uns obedece à conexão dos vértices, através de lados, num grafo. Mais precisamente, quando dois vértices p_i e p_j num grafo estão conectados através de um lado, escrevemos $a_{ij} = 1$, caso contrário fazemos $a_{ij} = 0$. Por exemplo, para o grafo



tem-se associada a matriz de incidência:

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Quando associa-se um sentido de entrada ou de saída de um lado, em relação a um vértice, obtém-se matrizes de incidência lado-nó . Por exemplo, para o grafo está associada a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 1 & 0 & -1 \end{bmatrix}$$

Os lados correspondem às linhas da matriz; os nós às colunas. Cada linha indica o nó que deixa, por -1 e o nó que entra, por $+1$.

2.6.6 Matrizes Circulantes e de Toeplitz

Matrizes, nas quais cada linha, a partir da segunda, é obtida da linha precedente, através do deslocamento de cada um de seus elementos uma coluna à direita, com o último elemento trocado ciclicamente com o primeiro elemento, são chamadas de **matrizes circulares** ou simplesmente circulares.

Estas matrizes têm a forma geral

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_n \\ c_n & c_1 & c_2 & \cdots & c_{n-1} \\ c_{n-1} & c_n & c_1 & \cdots & c_{n-2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ c_2 & c_3 & c_4 & \cdots & c_1 \end{bmatrix} \quad (2.108)$$

Observe-se que cada subdiagonal contém os mesmos elementos. Matrizes com esta propriedade, são chamadas **matrizes Toeplitz** . Por exemplo,

$$\begin{bmatrix} 1 & 2 & 0 & 9 & 7 \\ 5 & 1 & 2 & 0 & 9 \\ 0 & 5 & 1 & 2 & 0 \\ 4 & 0 & 5 & 1 & 2 \\ 3 & 4 & 0 & 5 & 1 \end{bmatrix}$$

é uma matriz que não é circulante, porém possui a estrutura Toeplitz.

2.6.7 Matrizes Esparsas e Cheias

Uma matriz quadrada é dita **matriz esparsa**, quando a maioria dos seus elementos são nulos. Caso contrário, a matriz é dita **cheia**, isto é, a maior parte de seus elementos são não nulos.

2.7 O Problema da Divisão Matricial

Da álgebra elementar, são conhecidos os seguintes postulados básicos:

1. Lei comutativa da adição : $a + b = b + a$;
2. Lei associativa da adição : $(a + b) + c = a + (b + c)$;
3. Lei comutativa da multiplicação : $ab = ba$;
4. Lei associativa da multiplicação : $(ab)c = a(bc)$;
5. Lei distributiva da multiplicação : $(a + b)c = ac + bc$; $c(a + b) = ca + cb$;
6. A não fatorização do zero: se $ab = 0$, então $a = 0$ ou $b = 0$;
7. Se $a \neq 0$, a equação $ax = 1$ tem única solução para x , a qual é denotada por $x = 1/a$ ou a^{-1} . O número a^{-1} é chamado de *inverso* ou *recíproco* de a , caracterizado pela propriedade de ser o único número b que satisfaz $ab = ba = 1$. Para $a \neq 0$ a equação $ax = b$ possui a única solução $x = a^{-1}b$ que é, usualmente, denotada por b/a . O número b/a é chamado de o quociente de b por a e a operação de formar quocientes é conhecida como **divisão**.

Na álgebra das matrizes, os postulados acima 1, 2, 4 e 5 permanecem válidos, enquanto que os postulados 3, 6 e 7 não. Como já foi visto, a multiplicação matricial não é comutativa e, para o sexto postulado, considere-se as matrizes

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} \quad \text{e} \quad \mathbf{B} = \begin{bmatrix} 3 & 7 \\ 0 & 0 \end{bmatrix}.$$

Note-se que

$$\mathbf{AB} = 0, \tag{2.109}$$

sem que \mathbf{A} ou \mathbf{B} seja a matriz nula. O fato que a matriz nula pode possuir fatores não nulos ocasiona profunda diferença entre a álgebra das matrizes e a álgebra elementar. Em particular, a lei do cancelamento: $ca = cb, c \neq 0$ implica $a = b$, também não é válida. Por exemplo, para as matrizes

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 3 & 4 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 2 & 5 \\ 3 & 4 \end{bmatrix} \quad \text{e} \quad \mathbf{C} = \begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix},$$

tem-se que

$$\mathbf{CA} = \mathbf{CB} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix},$$

com $\mathbf{C} \neq 0$, porém, $\mathbf{A} \neq \mathbf{B}$. Por outro lado, pode-se ter a matriz nula como a potência de uma matriz não nula

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad \mathbf{A}^2 = 0. \quad (2.110)$$

Este tipo de matrizes são chamadas de matrizes nilpotentes.

Diz-se que as matrizes \mathbf{A} e \mathbf{B} são **divisores de zero**, se

$$\mathbf{AB} = \mathbf{0}, \mathbf{A} \neq \mathbf{0} \quad \text{e} \quad \mathbf{B} \neq \mathbf{0}. \quad (2.111)$$

A existência de divisores de zero na álgebra das matrizes dificulta o processo da **divisão**. Na álgebra ordinária, a grandeza

$$a^{-1} = 1/a \quad (2.112)$$

existe para qualquer valor de a exceto $a = 0$ uma vez que não podemos dividir por zero. Porém, se o zero tem fatores a situação é diferente. Para uma matriz \mathbf{A} , a operação

$$\mathbf{A}^{-1} = 1/\mathbf{A} \quad (2.113)$$

perde seu significado não somente quando $\mathbf{A} = \mathbf{0}$, mas, também, quando \mathbf{A} possui um divisor de zero.

2.8 Inversa de uma matriz quadrada

A resolução de sistemas de equações algébricas lineares, pelo método de Cramer, permite introduzir o conceito de "recíproco" ou de "inversa" de uma matriz quadrada. Para isto, considere-se o seguinte sistema de equações lineares simultâneas

$$\begin{aligned} ax + by &= f \\ cx + dy &= g \end{aligned} \quad (2.114)$$

e denote-se

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad (2.115)$$

a matriz dos coeficientes do sistema. Pelo método de Cramer, tem-se que

$$x = \frac{\det \begin{bmatrix} f & b \\ g & d \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{(df - bg)}{\det(\mathbf{A})} \quad (2.116)$$

$$y = \frac{\det \begin{bmatrix} a & f \\ c & g \end{bmatrix}}{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}} = \frac{ag - cf}{\det(\mathbf{A})} \quad (2.117)$$

uma vez que o determinante do sistema cumpra

$$\det(\mathbf{A}) = ad - bc \neq 0. \quad (2.118)$$

Matricialmente, tem-se a equação

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix} \quad (2.119)$$

ou, simplesmente, $\mathbf{Ax} = \mathbf{b}$. O método de Cramer fornece a solução

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} = (\det \mathbf{A})^{-1} \begin{bmatrix} df - bg \\ ag - cf \end{bmatrix} = (\det \mathbf{A})^{-1} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} f \\ g \end{bmatrix}. \quad (2.120)$$

Formalmente, escreve-se

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b} \quad (2.121)$$

onde

$$\mathbf{A}^{-1} = \frac{1}{(ad - cb)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \quad (2.122)$$

Pode ser facilmente verificado que a matriz \mathbf{A}^{-1} satisfaz a propriedade recíproca ou da inversa

$$\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}. \quad (2.123)$$

Em geral, uma matriz quadrada \mathbf{A} de ordem n é dita que possui **inversa** ou que é **não singular** quando existir uma matriz quadrada \mathbf{B} de ordem n que verifica

$$\mathbf{AB} = \mathbf{BA} = \mathbf{I} \quad (2.124)$$

Neste caso, a matriz \mathbf{B} é denotada por \mathbf{A}^{-1} e é referida como a inversa de \mathbf{A} .

A caracterização de matrizes quadradas de ordem arbitrária que possuem inversa pode ser obtida com o uso de determinantes.

2.8.1 Determinante de uma Matriz Quadrada

A definição de **determinante** de uma matriz quadrada \mathbf{A} , denotado por $\det(\mathbf{A})$ ou $|\mathbf{A}|$, pode ser dada de diversas maneiras. Dentre elas, a seguinte definição por indução, permite calcular o determinante através de determinantes de matrizes de menor ordem:

1. Se $\mathbf{A} = [a]$ é uma matriz 1×1 , então $\det(\mathbf{A}) = a$; se \mathbf{A} é uma matriz 2×2

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \quad (2.125)$$

então $\det(\mathbf{A}) = ad - bc$;

2. O **menor** M_{ij} é o determinante da submatriz de ordem $n - 1$, obtida de \mathbf{A} de ordem n , pela supressão da i -ésima linha e da j -ésima coluna ;

3. O **cofator** A_{ij} , associado a M_{ij} , é definido como

$$A_{ij} = (-1)^{i+j} M_{ij}; \quad (2.126)$$

4. O **determinante** de uma matriz quadrada \mathbf{A} de ordem n , $n > 1$, é dado por

$$\det(\mathbf{A}) = \sum_{j=1}^n a_{ij} A_{ij}, \quad \text{para qualquer } i = 1, 2, \dots, n \quad (2.127)$$

ou, equivalentemente,

$$\det(\mathbf{A}) = \sum_{i=1}^n a_{ij} A_{ij}, \quad \text{para qualquer } j = 1, 2, \dots, n. \quad (2.128)$$

Exemplo 2.15

Calcular o determinante da matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 0 \\ 3 & -1 & 1 \\ 4 & 6 & -2 \end{bmatrix}.$$

Solução

Utilizando a expansão em cofatores com $i = 3$,

$$\begin{aligned} \det(\mathbf{A}) &= a_{31}A_{31} + a_{32}A_{32} + a_{33}A_{33} \\ &= 4A_{31} + 6A_{32} - 6A_{33} \\ &= 4(-1)^{3+1} \det \begin{bmatrix} 0 & 0 \\ -1 & 1 \end{bmatrix} + 6(-1)^{3+2} \det \begin{bmatrix} 2 & 0 \\ 3 & 1 \end{bmatrix} \\ &\quad - 2(-1)^{3+3} \det \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix} \\ &= -8. \end{aligned}$$

Suponha-se que $\mathbf{A} = [a_{ij}]$ é de ordem n . Define-se a **matriz de cofatores** de \mathbf{A} como sendo

$$\text{cof}(\mathbf{A}) = [\mathbf{A}_{ij}], \quad (2.129)$$

onde \mathbf{A}_{ij} é o cofator do menor M_{ij} da matriz \mathbf{A} . A **matriz adjugada** ou **matriz adjunta** de \mathbf{A} , escrita como $\text{adj}(\mathbf{A})$, é definida como a transposta de $\text{cof}(\mathbf{A})$, ou seja,

$$\text{adj}(\mathbf{A}) = (\text{cof}(\mathbf{A}))^t, \quad (2.130)$$

Deve ser observado que, na literatura, é utilizada frequentemente a denominação “matriz adjunta” para a matriz adjugada.

Exemplo 2.16

Determinar $\text{cof}(\mathbf{A})$ e $\text{adj}(\mathbf{A})$ para a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}.$$

Solução

Tem-se que

$$\begin{aligned} \text{cof} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} &= \begin{bmatrix} \begin{vmatrix} 3 & 2 \\ 2 & 2 \end{vmatrix} & - \begin{vmatrix} 2 & 2 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 2 & 3 \\ 1 & 2 \end{vmatrix} \\ - \begin{vmatrix} 2 & 3 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 1 & 2 \end{vmatrix} & - \begin{vmatrix} 1 & 2 \\ 1 & 2 \end{vmatrix} \\ \begin{vmatrix} 2 & 3 \\ 3 & 2 \end{vmatrix} & - \begin{vmatrix} 1 & 3 \\ 2 & 2 \end{vmatrix} & \begin{vmatrix} 1 & 2 \\ 2 & 3 \end{vmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 2 & -2 & 1 \\ 2 & -1 & 0 \\ -5 & 4 & -1 \end{bmatrix} \end{aligned}$$

Segue-se que

$$\text{adj}(\mathbf{A}) = \begin{bmatrix} 2 & 2 & -5 \\ -2 & -1 & 4 \\ 1 & 0 & -1 \end{bmatrix}.$$

Exemplo 2.17

Achar a matriz adjugada de

$$\mathbf{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

Solução

Tem-se

$$\text{adj}(\mathbf{A}) = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

2.8.2 Propriedades dos Determinantes

A seguir, serão enunciadas algumas das conhecidas propriedades elementares dos determinantes:

1. Se todo elemento de uma linha (coluna) de uma matriz quadrada \mathbf{A} é zero, então $\det(\mathbf{A}) = 0$;
2. Se \mathbf{A} é uma matriz quadrada, então $\det(\mathbf{A}^t) = \det \mathbf{A}$, isto é, para qualquer propriedade relacionada às linhas de um determinante tem-se uma correspondente propriedade em termos das colunas e vice-versa;
3. O determinante de uma matriz \mathbf{A} muda de sinal quando permutam-se duas linhas ou duas colunas de \mathbf{A} ;
4. Se \mathbf{B} é formada de uma matriz quadrada \mathbf{A} pela multiplicação de todos os elementos de uma linha ou coluna de \mathbf{A} por um escalar k , então $\det(\mathbf{B}) = k \det(\mathbf{A})$;

5. O determinante do produto de matrizes quadradas é igual ao produto dos determinantes dessas matrizes, assim,

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}); \quad (2.131)$$

6. Se uma matriz \mathbf{A} possui duas linhas ou colunas proporcionais, então $\det(\mathbf{A}) = 0$.
7. Se, numa matriz \mathbf{A} , uma linha é adicionada a outra, o determinante da matriz resultante é igual ao da matriz \mathbf{A} . Pois em virtude da propriedade 2, pode-se adicionar uma coluna sem alterar o valor do determinante.

2.8.3 A Identidade de Cramer

A matriz adjugada possui a seguinte importante propriedade chamada de **identidade de Cramer**

$$\text{adj}(\mathbf{A})\mathbf{A} = \mathbf{A}\text{adj}(\mathbf{A}) = (\det(\mathbf{A}))\mathbf{I}. \quad (2.132)$$

Esta identidade é um adequado refinamento do conhecido método de Cramer para resolver sistemas lineares algébricos.

Exemplo 2.18

Verificar a identidade de Cramer para a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix}$$

Solução

Tem-se que

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 2 \\ 1 & 2 & 2 \end{bmatrix} \begin{bmatrix} 2 & 2 & -5 \\ -2 & -1 & 4 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -\mathbf{I}.$$

2.8.4 Fórmula Analítica para a Inversa de uma Matriz

Considere-se a equação matricial

$$\mathbf{Ax} = \mathbf{b}, \quad (2.133)$$

com \mathbf{A} uma matriz quadrada de ordem n e \mathbf{x}, \mathbf{b} vetores $n \times 1$. Multiplicando a esquerda ambos os membros da equação pela matriz adjugada de \mathbf{A} e utilizando a identidade de Cramer, resulta

$$\text{adj}(\mathbf{A})\mathbf{Ax} = \det(\mathbf{A})\mathbf{x} = \text{adj}(\mathbf{A})\mathbf{b}. \quad (2.134)$$

Uma matriz quadrada \mathbf{A} é dita **matriz singular** quando $\det(\mathbf{A}) = 0$ e **não-singular** quando $\det(\mathbf{A}) \neq 0$.

Para uma matriz não-singular \mathbf{A} , segue-se que a solução da equação matricial $\mathbf{Ax} = \mathbf{b}$ é dada por

$$\mathbf{x} = [\text{adj}(\mathbf{A})/\det(\mathbf{A})]\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}, \quad (2.135)$$

onde

$$\mathbf{A}^{-1} = [\text{adj}(\mathbf{A})/\det(\mathbf{A})] \quad (2.136)$$

verifica

$$\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I} \quad (2.137)$$

e portanto é a inversa da matriz \mathbf{A} .

Se \mathbf{A} não for uma matriz quadrada, ou se for uma matriz quadrada singular, então \mathbf{A}^{-1} não estará definida e será necessário estender o conceito de inversão.

Para matrizes quadradas não-singulares, as seguintes regras são válidas

$$(\mathbf{A}\mathbf{B})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1};$$

$$\det(\mathbf{A})\det(\mathbf{A}^{-1}) = 1; \quad (2.138)$$

$$(\mathbf{A}^t)^{-1} = (\mathbf{A}^{-1})^t.$$

De fato,

$$(\mathbf{A}\mathbf{B})(\mathbf{B}^{-1}\mathbf{A}^{-1}) = \mathbf{A}(\mathbf{B}\mathbf{B}^{-1})\mathbf{A}^{-1} = \mathbf{A}\mathbf{I}\mathbf{A}^{-1} = \mathbf{A}\mathbf{A}^{-1} = \mathbf{I} \quad (2.139)$$

e

$$(\mathbf{B}^{-1}\mathbf{A}^{-1})\mathbf{A}\mathbf{B} = \mathbf{B}^{-1}(\mathbf{A}\mathbf{A}^{-1})\mathbf{B} = \mathbf{B}^{-1}\mathbf{I}\mathbf{B} = \mathbf{B}^{-1}\mathbf{B} = \mathbf{I}. \quad (2.140)$$

A segunda propriedade decorre da regra para o determinante de um produto e de que a matriz identidade tem determinante igual a um.

Da relação $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ e, utilizando transposição, obtém-se:

$$(\mathbf{A}^{-1})^t\mathbf{A}^t = (\mathbf{A}\mathbf{A}^{-1})^t = \mathbf{I}^t = \mathbf{I}$$

$$\mathbf{A}^t(\mathbf{A}^{-1})^t = (\mathbf{A}^{-1}\mathbf{A})^t = \mathbf{I}^t = \mathbf{I}. \quad (2.141)$$

Exemplo 2.19

Para \mathbf{A} não singular, as equações matriciais

$$\mathbf{A}\mathbf{X} = \mathbf{B} \quad \text{e} \quad \mathbf{Y}\mathbf{A} = \mathbf{B}$$

podem ser resolvidas multiplicando ambos os membros por \mathbf{A}^{-1} :

$$\mathbf{X} = \mathbf{A}^{-1}\mathbf{B} \quad \text{e} \quad \mathbf{Y} = \mathbf{B}\mathbf{A}^{-1},$$

ou seja, a lei do corte pode ser utilizada.

As leis dos expoentes para inteiros negativos é estendida ao caso de matrizes não-singulares. Dada uma matriz quadrada não-singular \mathbf{A} e p um inteiro positivo, define-se

$$\mathbf{A}^{-p} = (\mathbf{A}^{-1})^p.$$

Decorre que para uma matriz quadrada não-singular \mathbf{A} e qualquer par de inteiros p e q

$$\mathbf{A}^p \mathbf{A}^q = \mathbf{A}^{p+q}$$

$$(\mathbf{A}^p)^r = \mathbf{A}^{pq} .$$

Resumindo:

1. *Os resultados obtidos até agora podem ser resumidos dizendo-se que, na maioria dos aspectos, a álgebra matricial assemelha-se à álgebra dos números, sendo que a matriz unidade e a matriz nula desempenham o papel de 1 e 0, respectivamente. Os pontos vitais nos quais as duas álgebras diferem, é que a multiplicação matricial não é comutativa e que a lei da divisão (e conseqüentemente a lei do corte) não é válida para qualquer matriz.*
2. *A matriz quadrada \mathbf{A} possui inversa somente se $\det(\mathbf{A}) \neq 0$. Neste caso,*

$$\mathbf{A}^{-1} = [\text{adj}(\mathbf{A})/\det(\mathbf{A})].$$

CAPÍTULO 3

Sistemas de Equações Algébricas Lineares

3.1 Introdução

O problema central da álgebra das matrizes é a resolução de equações lineares simultâneas

$$\begin{array}{cccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array} \quad (3.1)$$

ou, de maneira compacta,

$$\mathbf{Ax} = \mathbf{b} , \quad (3.2)$$

onde \mathbf{A} é uma matriz de ordem $m \times n$ e \mathbf{b} um vetor coluna de ordem $m \times 1$ formados pelos coeficientes do sistema e pelos termos não-homogêneos, respectivamente. O vetor coluna \mathbf{x} , de ordem $n \times 1$, tem como elementos as incógnitas do sistema.

Nas mais variadas aplicações, tem-se os seguintes casos:

1. O número de equações é igual ao número de incógnitas, isto é, \mathbf{A} é uma matriz quadrada de ordem n e \mathbf{x} um vetor $n \times 1$;
2. O número de equações é menor do que o número de incógnitas, ou seja, \mathbf{A} é uma matriz $m \times n$ e \mathbf{x} um vetor $n \times 1$ com $m < n$;
3. O número de equações é maior do que o número de incógnitas, daí que \mathbf{A} é uma matriz $m \times n$ e \mathbf{x} um vetor $n \times 1$ com $m > n$.

Estas três situações podem ser abordadas através dos métodos gerais:

1. Método de Cramer;
2. Eliminação Gaussiana;
3. Mínimos Quadrados.

Estes métodos introduzem importantes conceitos para a álgebra matricial.

3.2 Sistemas Lineares Regulares

Dentre os casos, acima mencionados, o mais simples e importante é quando o número de variáveis é igual ao número de equações. Os métodos, que serão descritos a seguir, fornecem subsídios úteis à resolução de sistemas em que o número de variáveis é diferente do número de equações.

Considere-se o seguinte sistema de n equações lineares com n incógnitas

$$\begin{array}{cccccc} a_{11}x_1 & +a_{12}x_2 & +\cdots & +a_{1n}x_n & = & b_1 \\ a_{21}x_1 & +a_{22}x_2 & +\cdots & +a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \\ a_{n1}x_1 & +a_{n2}x_2 & +\cdots & +a_{nn}x_n & = & b_n \end{array} \quad (3.3)$$

escrito, na forma matricial, como

$$\mathbf{Ax} = \mathbf{b}. \quad (3.4)$$

3.2.1 O Método De Cramer

Suponha-se \mathbf{A} uma matriz com determinante não nulo, isto é, que o sistema linear dado é **regular**. Neste caso, a solução pode ser obtida através da identidade de Cramer.

Multiplica-se ambos os membros da equação $\mathbf{Ax} = \mathbf{b}$, pela matriz adjugada de \mathbf{A} . Assim,

$$adj(\mathbf{A})\mathbf{Ax} = adj(\mathbf{A})\mathbf{b} \quad (3.5)$$

e, utilizando a identidade de Cramer, decorre

$$det(\mathbf{A})\mathbf{x} = adj(\mathbf{A})\mathbf{b}. \quad (3.6)$$

Como \mathbf{A} é uma matriz com determinante não nulo, então, a variável \mathbf{x} pode ser isolada facilmente, isto é,

$$\mathbf{x} = \frac{adj(\mathbf{A})\mathbf{b}}{det(\mathbf{A})}. \quad (3.7)$$

Outra alternativa, em termos das componentes de \mathbf{x} , é utilizar a expansão em cofatores, obtendo-se

$$x_k = \frac{det[\mathbf{a}_1 \ \mathbf{a}_2 \ \cdots \ \mathbf{a}_{k-1} \ \mathbf{b} \ \mathbf{a}_{k+1} \ \cdots \ \mathbf{a}_n]}{det(\mathbf{A})}, \quad k = 1 : n.$$

onde a k -ésima coluna \mathbf{a}_k da matriz \mathbf{A} é substituída pelo dado \mathbf{b} . Estas fórmulas são conhecidas como a **regra de Cramer** para sistemas de equações lineares numéricas.

Conclui-se que um sistema de n equações lineares numéricas com n incógnitas e com a condição que o determinante do sistema é não nulo ($det(\mathbf{A}) \neq 0$) possui para cada termo não-homogêneo \mathbf{b} , uma solução bem determinada por 3.7.

Exemplo 3.20

Resolver, pelo método de Cramer, o sistema

$$\begin{array}{rcl} x + y + 2z & = & 9 \\ x + 2y + z & = & 8 \\ 2x + y + z & = & 7 \end{array}$$

Solução

Tem-se

$$\det(\mathbf{A}) = \det \begin{bmatrix} 1 & 1 & 2 \\ 1 & 2 & 1 \\ 2 & 1 & 1 \end{bmatrix} = -4 \quad \text{e} \quad \text{adj}(\mathbf{A}) = \begin{bmatrix} 1 & 1 & 3 \\ 4 & -3 & 1 \\ -3 & 1 & 1 \end{bmatrix}$$

Assim,

$$\mathbf{x} = (1/-4) \begin{bmatrix} 1 & 1 & 3 \\ 4 & -3 & 1 \\ -3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 9 \\ 8 \\ 7 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}.$$

3.2.2 O Método da Eliminação

Outra maneira de resolver sistemas de equações lineares numéricas é a través da chamada **condensação pivotal** ou **método da eliminação**, cujo nome é associado ao de Gauss. O método consiste em eliminar variáveis do sistema, de modo sucessivo, até obter um sistema modificado em que a matriz associada é triangular superior. É conveniente ilustrar com exemplos simples o método da eliminação.

Exemplo 3.21

$$\begin{aligned} 10x_1 + x_2 - 5x_3 &= 1 \\ -20x_1 + 3x_2 + 20x_3 &= 2 \\ 5x_1 + 3x_2 + 5x_3 &= 6 \end{aligned}$$

A primeira equação, quando multiplicada pelo fator 20/10 e adicionada à segunda, tem o seu coeficiente em x_1 eliminado. Similarmente, subtraindo da terceira equação a primeira equação multiplicada por 5/10, o seu coeficiente em x_1 é também eliminado. Tem-se, portanto, o sistema de equações

$$\begin{aligned} 10x_1 + x_2 - 5x_3 &= 1 \\ 5x_2 + 10x_3 &= 4 \\ 2.5x_2 + 7.5x_3 &= 5.5 \end{aligned} \tag{3.8}$$

Subtraindo da terceira equação, no sistema 3.8, a segunda equação multiplicada por 2.5/5, elimina-se o seu coeficiente em x_3 . Resulta, assim, o sistema modificado

$$\begin{aligned} 10x_1 + x_2 - 5x_3 &= 1 \\ 5x_2 + 10x_3 &= 4 \\ 2.5x_3 &= 3.5 \end{aligned} \tag{3.9}$$

Deste modo, o sistema original reduz-se a um sistema dito na forma triangular superior $\mathbf{U}\mathbf{x} = \mathbf{c}$, onde

$$\mathbf{U} = \begin{bmatrix} 10 & 1 & -5 \\ 0 & 5 & 10 \\ 0 & 0 & 2.5 \end{bmatrix}$$

Os valores das incógnitas são obtidas pela **retrosubstituição**

$$\begin{aligned} x_3 &= \frac{3.5}{\mathbf{2.5}} = 1.4 \\ x_2 &= \frac{4 - 10x_3}{\mathbf{5}} = -2 \\ x_1 &= \frac{1 - x_2 + 5x_3}{\mathbf{10}} = 1 \end{aligned} \tag{3.10}$$

Este é um processo automático, no qual, desde que certos **elementos pivotais** sejam não nulos, o número de equações do sistema é irrelevante. Nas equações 5.8, 3.8 e 3.9 os elementos pivotais são enfatizados em negrito e aparecem como denominadores dos fatores utilizados em cada passo da eliminação. Isto também acontece na retrosubstituição 3.10. Do ponto de vista matricial, o método da eliminação tem os seguintes objetivos:

1. Transformar o sistema $\mathbf{Ax} = \mathbf{b}$ no sistema $\mathbf{Ux} = \mathbf{c}$ em que \mathbf{U} é uma matriz triangular superior.
2. Resolver o sistema $\mathbf{Ux} = \mathbf{c}$ por retrosubstituição.

Permutações e Pivôs Nulos

No exemplo anterior, o método da eliminação foi realizado supondo que os pivôs, em cada passo, eram não nulos. Considere-se, agora, o seguinte sistema.

Exemplo 3.22

Resolver o seguinte sistema de equações

$$\begin{aligned} 3x_1 - x_2 + 0x_3 &= 2 \\ 6x_1 - 2x_2 + x_3 &= -1 \\ -3x_1 + 3x_2 + x_3 &= 0. \end{aligned}$$

Eliminando x_1 na segunda e terceira equações, obtém-se o sistema modificado

$$\begin{aligned} 3x_1 - x_2 + 0x_3 &= 2 \\ 0x_2 + x_3 &= -5 \\ 2x_2 + x_3 &= 2. \end{aligned}$$

Para eliminar x_2 da terceira equação, nenhum múltiplo da segunda equação pode ser utilizado, pois o pivô para este passo é nulo. Contudo, com uma simples troca na ordem das equações

$$\begin{aligned} 3x_1 - x_2 + 0x_3 &= 2 \\ 2x_2 + x_3 &= 2 \\ x_3 &= -5, \end{aligned}$$

O sistema final modificado pode ser imediatamente resolvido por retrosubstituição. De fato, com essa troca, os elementos pivotais finais são 3, 2 e 1. Assim

$$\begin{aligned} x_3 &= \frac{-5}{\mathbf{1}} = -5 \\ x_2 &= \frac{2 - x_3}{\mathbf{2}} = \frac{7}{2} \\ x_1 &= \frac{2 + x_2}{\mathbf{3}} = \frac{11}{6}. \end{aligned}$$

Um zero, na localização do que seria um elemento pivotal, acarreta duas possibilidades: o problema pode ser facilmente resolvido, com uma troca de equações, ou pode ser sério.

Por exemplo, no sistema

$$\begin{array}{rccccrc} x_1 & +2x_2 & +3x_3 & +4x_4 & = & 2 & \\ & & 5x_3 & +6x_4 & = & -1 & \\ & & 7x_3 & +6x_4 & = & 0 & \\ & cx_2 & +7x_3 & +8x_4 & = & 1 & \end{array} \quad (3.11)$$

deve ser observado que o primeiro elemento pivotal é 1, porém a sua utilização não é necessária, pois as outras equações não apresentam a variável x_1 . Como o coeficiente da variável x_2 , na segunda equação, é zero e não pode ser utilizado para eliminar x_2 das equações seguintes, faz-se a troca da segunda com a quarta equação. Assim,

$$\begin{array}{rccccrc} x_1 & +2x_2 & +3x_3 & +4x_4 & = & 2 & \\ & cx_2 & +7x_3 & +8x_4 & = & 1 & \\ & & 5x_3 & +6x_4 & = & -1 & \\ & & 7x_3 & +6x_4 & = & 0 & . \end{array} \quad (3.12)$$

Agora, utilizando o coeficiente 5 da terceira equação como elemento pivotal para eliminar o coeficiente de x_3 , na última equação, tem-se

$$\begin{array}{rccccrc} x_1 & +2x_2 & +3x_3 & +4x_4 & = & 2 & \\ & cx_2 & +7x_3 & +8x_4 & = & 1 & \\ & & 5x_3 & +6x_4 & = & -1 & \\ & & & -12x_4 & = & 7 & . \end{array} \quad (3.13)$$

Aqui, os elementos pivotais são 1, c , 5 e -12 . Se c é não nulo, então resolve-se por retrosubstituição. Quando c for nulo, o problema é sério: a matriz do sistema dado é singular. Neste caso, a primeira e a segunda colunas seriam proporcionais e o determinante do sistema seria zero. De outro lado, a retrosubstituição não seria automática e faz-se necessário, portanto, uma maior análise do sistema, o qual será feito nas próximas seções.

3.2.3 Fatorização Matricial LU

Deve-se ressaltar, no primeiro exemplo, a relação existente entre o valor dado \mathbf{b} com o valor transformado \mathbf{c} , ou seja, como as operações realizadas com as linhas da matriz \mathbf{A} , para obter as linhas da matriz triangular superior \mathbf{U} , transformam \mathbf{b} em \mathbf{c} . Em primeiro lugar, b_1 não foi alterado. Logo, multiplica-se b_1 pelo fator $(20/10)$ e adiciona-se a b_2 . Similarmente, subtrai-se, de b_3 , o elemento b_1 multiplicado por $(5/10)$. Assim,

$$\begin{array}{rcccl} 1 \cdot b_1 & = & 1 & = & c_1 \\ (20/10) \cdot b_1 + 1 \cdot b_2 & = & 4 & = & c_2 \\ -(5/10) \cdot b_1 + 1 \cdot b_3 & = & 5.5 & & \end{array} \quad (3.14)$$

No segundo passo, os dois primeiros novos elementos, c_1 e c_2 , permanecem inalterados. Então, multiplica-se $c_2 = 4$ pelo fator $(2.5/5)$ e subtrai-se de 5.5, obtendo-se o terceiro elemento c_3 . Sendo $b_1 = c_1$, decorre

$$\begin{array}{rcccl} 1 \cdot b_1 & = & 1 & = & c_1 \\ (20/10) \cdot c_1 + 1 \cdot b_2 & = & 4 & = & c_2 \\ -(5/10) \cdot c_1 + 1 \cdot b_3 - (2.5/5) \cdot c_2 & = & 5.5 - (2.5/5)4 & = & 2 = c_3 \end{array} \quad (3.15)$$

ou, equivalentemente,

$$\begin{aligned} c_1 &= b_1 \\ -(20/10) \cdot c_1 + c_2 &= b_2 \\ (5/10) \cdot c_1 + (2.5/10) \cdot c_2 + c_3 &= b_3 . \end{aligned} \quad (3.16)$$

Na forma matricial, este sistema é escrito como

$$\begin{bmatrix} 1 & 0 & 0 \\ -20/10 & 1 & 0 \\ 5/10 & 2.5/5 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} , \quad (3.17)$$

ou seja,

$$\mathbf{Lc} = \mathbf{b} \quad (3.18)$$

onde

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ -20/10 & 1 & 0 \\ 5/10 & 2.5/5 & 1 \end{bmatrix}$$

Os fatores utilizados na eliminação são denominados **multiplicadores**. \mathbf{L} é uma matriz triangular inferior, cujas colunas são, precisamente, os multiplicadores utilizados na eliminação. A resolução do sistema triangular inferior $\mathbf{Lc} = \mathbf{b}$ é feita pela **substituição ascendente** 3.16.

Deste modo, a eliminação consiste, intrinsecamente, em partir o sistema $\mathbf{Ax} = \mathbf{b}$ em dois sistemas triangulares $\mathbf{Lc} = \mathbf{b}$ e, a seguir, $\mathbf{Ux} = \mathbf{c}$. As matrizes \mathbf{L} e \mathbf{U} provém da eliminação no seguinte sentido \mathbf{L} é formada pelos multiplicadores e \mathbf{U} pelos resultados sobre os elementos da matriz \mathbf{A} . O que é mais importante, tem-se a seguinte **fatorização matricial**

$$\mathbf{A} = \mathbf{LU} , \quad (3.19)$$

conhecida como a redução de Doolittle .

As matrizes triangulares possuem várias propriedades que facilitam a sua operacionalidade, dentre elas: o determinante de uma matriz triangular é igual ao produto dos elementos da diagonal; o produto de duas matrizes triangulares superiores ou inferiores é novamente uma matriz triangular da mesma estrutura; uma matriz triangular não singular é facilmente invertida e sua inversa possui uma estrutura idêntica. Como foi visto nos exemplos, a resolução de sistemas triangulares de equações lineares é um processo extremamente simples.

Exemplo 3.23

Determinar a fatorização \mathbf{LU} da matriz tridiagonal

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

Solução

O primeiro pivô é $a_{11} = 2$, aqui utilizado somente para eliminar o elemento $a_{21} = 1$. Pois, neste exemplo $a_{31} = a_{41} = 0$. Assim, subtraindo da segunda linha o múltiplo $l_{21} = 1/2$ da primeira linha, tem-se

$$\mathbf{A}^{(2)} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \text{e} \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & * & 1 & 0 \\ 0 & * & * & 1 \end{bmatrix}.$$

O segundo pivô é $a_{22}^{(2)} = 3/2$, sendo utilizado para eliminar apenas o elemento $a_{32}^{(2)} = 1$, uma vez que $a_{42}^{(2)} = 0$. O múltiplo é $l_{32} = 2/3$ e a subtração produz um zero na linha 3 e coluna 2

$$\mathbf{A}^{(3)} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & 0 \\ 0 & 0 & 4/3 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} \quad \text{e} \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 2/3 & 1 & 0 \\ 0 & 0 & * & 1 \end{bmatrix}.$$

O passo seguinte utiliza o terceiro pivô $a_{33}^{(3)} = 4/3$, e multiplica-o por $l_{43} = 3/4$, para eliminar o último elemento não nulo situado abaixo da diagonal. A subtração, $2 - 3/4 = 5/4$, fornece o quarto pivô e resulta a forma triangular final

$$\mathbf{A}^{(4)} = \mathbf{U} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & 0 \\ 0 & 0 & 4/3 & 1 \\ 0 & 0 & 0 & 5/4 \end{bmatrix} \quad \text{e} \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 2/3 & 1 & 0 \\ 0 & 0 & 3/4 & 1 \end{bmatrix}.$$

Os elementos na diagonal de \mathbf{U} são os pivôs, os quais são desconhecidos a priori para a matriz \mathbf{A} . São determinados durante o processo da eliminação.

É importante ressaltar o efeito da eliminação sobre os elementos nulos da matriz \mathbf{A} tridiagonal. Neste exemplo, um elemento nulo, situado na coluna do pivô e abaixo dele, não precisa ser eliminado. Segue-se, portanto, que $l_{31} = l_{41} = l_{42} = 0$. Similarmente, os elementos nulos, acima da banda tridiagonal, são preservados em \mathbf{U} . Esta é uma característica da eliminação relacionada às matrizes banda no processo da eliminação, os elementos nulos acima e abaixo da banda continuam nulos.

No caso da eliminação realizar troca de equações, a fatorização \mathbf{LU} sofre uma pequena alteração. Com relação ao segundo exemplo, introduz-se a matriz de permutação

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

que ocasiona a troca de linhas

$$\mathbf{PA} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3 & -1 & 0 \\ 6 & -2 & 1 \\ -3 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 0 \\ -3 & 3 & 1 \\ 6 & -2 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 0 \\ -3 & 3 & 1 \\ 6 & -2 & 1 \end{bmatrix}$$

A matriz \mathbf{P} produz o mesmo efeito sobre o termo \mathbf{b} , trocar -5 com 2 . Assim, considerando o sistema inicial

$$\mathbf{PAx} = \mathbf{Pb} \tag{3.20}$$

a eliminação prossegue sem dificuldade.

É natural perguntar-se, para que classe de matrizes a eliminação gaussiana não requer troca de linhas?. Em geral, a eliminação gaussiana prossegue sem troca de linhas, unicamente, se o produto

$$\det_1(\mathbf{A})\det_2(\mathbf{A})\cdots\det_n(\mathbf{A}) \quad (3.21)$$

é não-nulo. Aqui, $\det_i(\mathbf{A})$ denota o menor principal de ordem i da matriz \mathbf{A} . Assim, far-se-á necessária uma troca de linhas, uma vez que aconteça $\det_p(\mathbf{A}) = 0$ para algum p . Esta é uma condição teórica e na prática equivale a resolver o sistema. Entretanto, existem classes de matrizes para as quais não há necessidade da troca de linhas. Em particular, a classe de matrizes positivas definidas as quais, inclusive, possuem um tipo especial de fatorização.

3.3 Complexidade e Condicionamento Computacional

Para implementar os métodos apresentados, deve-se ter em conta os seguintes aspectos :

1. A complexidade computacional do método, isto é, o número de operações a serem executadas;
2. Identificação de fenômenos numéricos decorrentes da aproximação dos números reais (“precisão infinita”), por números que são realmente utilizados nos cálculos (“precisão finita”);

Os procedimentos computacionais são surpreendentemente diferentes dos utilizados na teoria. Exemplificando, suponha-se que se deseja obter uma *solução numérica*, para um sistema de 50 equações com 50 incógnitas. Que método utilizar: Cramer ou Eliminação?. Para determinar o tempo de computação do método, é necessário conhecer quantas operações das várias classes serão realizadas.

Considere-se, em primeiro lugar, a redução do sistema $\mathbf{Ax} = \mathbf{b}$ para o sistema $\mathbf{Ux} = \mathbf{c}$, \mathbf{A} e \mathbf{U} de ordem $n \times n$, através da eliminação Gaussiana. Sejam

$$M_n = \text{ o número de multiplicações ou divisões}$$

e

$$S_n = \text{ o número de adições ou subtrações .}$$

Pode ser estabelecido que um sistema $\mathbf{Ax} = \mathbf{b}$, que pode ser resolvido por eliminação Gaussiana, sem troca de linhas, requer

$$\alpha_n = M_n + m_n = n^2 + (n - 1)n(n + 1)/3 \text{ multiplicações ou divisões}$$

e

$$\beta_n = S_n + s_n = n(n - 1) + (n - 1)n(2n - 1)/6 \text{ adições ou subtrações .}$$

Observe-se que α_n e β_n são menores do que $n^2 + n^3/3$. Portanto, para n grande, a complexidade computacional da eliminação gaussiana é da ordem de

$$n^3/3 \text{ operações.} \quad (3.22)$$

Para se ter uma idéia desta complexidade, defina-se *1mflop* como um milhão de operações com números de ponto flutuante (número no computador) e considere-se os seguintes valores de n :

n	n^3	<i>megaflops</i>
10	1000	0.001
50	125000	0.125
100	1000000	1
500	125000000	125
800	512000000	512
1000	10^9	10^3
10000	10^{12}	10^6
100000	10^{15}	10^{12}

Num computador com uma velocidade de 300 megaflops/segundo, o tempo para resolver o sistema $\mathbf{Ax} = \mathbf{b}$ é, aproximadamente, de

n	<i>tempo (segundos)</i>
10	0.000003
50	0.0416
100	0.0033
500	0.4166
800	1.7066
1000	3.3333
10000	3333.33 (≈ 1 hora)
100000	33333.3 (≈ 9.2 horas).

Na contagem das operações envolvidas, pode-se evitar multiplicações por uns e zeros e subtrações de zeros. Isto reduz o trabalho total de operações, na eliminação, para uma ordem de n^3 operações.

Retornando, agora, ao hipotético sistema inicial de 50 equações e 50 incógnitas, considere-se o método de Cramer. Cada variável x_i é expressa como o quociente de dois determinantes de ordem 50×50 . Escrevendo o denominador $\det(\mathbf{A})$ em termos de cofatores relativos a primeira linha, tem-se

$$\det(\mathbf{A}) = a_{11}\det(\mathbf{A}_{11}) - a_{12}\det(\mathbf{A}_{12}) + \cdots - a_{1,50}\det(\mathbf{A}_{1,50}). \quad (3.23)$$

Cada cofator $(-1)^{i+j}\det(\mathbf{A}_{ij})$ é um determinante de ordem 49×49 . Se os cofatores fossem conhecidos, ter-se-ia 50 multiplicações. Porém, os cofatores 49×49 podem ser expandidos segundo novos cofatores e, assim, sucessivamente. A completa expansão do $\det(\mathbf{A})$ requer $50 \times 49 \times \cdots \times 2 \times 1$ multiplicações, ou seja $50!$, aproximadamente 3.04140932^{64} . Este processo de cálculo levaria anos num computador. Portanto, o método de Cramer é considerado, atualmente, como um método essencialmente teórico. A grande vantagem, sobre a eliminação, é que o método de Cramer "desacopla" as incógnitas, através de uma fórmula algébrica, ou seja, não é necessário conhecer os valores das outras incógnitas, para obter o valor de uma determinada incógnita.

Para n muito grande, a contagem de operações para o método da eliminação, parece ser desalentadora, uma vez que n^3 será extremamente grande. Não deve ser esquecido, porém, que a própria multiplicação de matrizes de ordem $n \times n$ requer uma ordem aproximada de n^3 operações. Deste modo, a complexidade computacional para formar \mathbf{A}^{-1} é da mesma ordem que no cálculo de \mathbf{A}^2 ou $\mathbf{A}^t\mathbf{A}$. O cálculo da inversa \mathbf{A}^{-1} de uma matriz \mathbf{A} não-singular de ordem n equivale a resolver n sistemas da forma

$$\mathbf{Ax} = \mathbf{e}_i, \quad (3.24)$$

onde $\mathbf{e}_i = [0 \ 0 \ \dots \ 0 \ 1 \ 0 \ \dots \ 0]^t$, com 1 no i -ésimo elemento para $i = 1, 2, \dots, n$. Neste caso, tem-se que

$$\mathbf{Av}_1 = \mathbf{e}_1, \quad \mathbf{Av}_2 = \mathbf{e}_2, \quad \dots, \quad \mathbf{Av}_n = \mathbf{e}_n. \quad (3.25)$$

Assim,

$$\mathbf{A}[\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n] = [\mathbf{e}_1 \ \mathbf{e}_2 \ \dots \ \mathbf{e}_n] = \mathbf{I} \quad (3.26)$$

e, portanto,

$$\mathbf{A}^{-1} = [\mathbf{v}_1 \ \dots \ \mathbf{v}_n]. \quad (3.27)$$

Quando as matrizes possuem estrutura, o que é frequente nas aplicações, o tempo de cômputo reduz-se consideravelmente. Por exemplo, se \mathbf{A} é uma matriz tridiagonal, o número de operações diminui drasticamente. Pois, em cada passo, somente na linha abaixo do pivô é realizada a eliminação. Incluída a retrossubstituição, a eliminação Gaussiana com matrizes tridiagonais é da ordem de

$$4n \text{ operações.} \quad (3.28)$$

Ou seja, que sistemas tridiagonais podem ser resolvidos quase que instantaneamente. Por exemplo, para $n = 100.000$, um computador com velocidade de 300 Mflops/seg, levaria, aproximadamente, 0.00033 segundos! Em geral, para matrizes banda de comprimento w a complexidade computacional de multiplicações e divisões é da ordem de

$$w(w-1)(3n-2w+1)/3 \quad (3.29)$$

operações. Observa-se que, para $w = n$, ou seja, no caso em que a matriz é cheia, se recupera o valor

$$\alpha_n \approx (n-1)n(n+1)/3.$$

Se a matriz banda for simétrica, o número de operações é reduzido para

$$(w-1)[3(w+2)n-2w(w+1)]. \quad (3.30)$$

3.3.1 Exemplos Sobre Condicionamento

Outro problema prático é o do "condicionamento do sistema". Pequenas variações nos dados podem ocasionar resultados muito diferentes. Nesta situação, o sistema é dito *mal condicionado*.

Exemplo 3.24

Considere-se os sistemas

$$1) \ x + y = 1, \quad x + 1.01y = 1$$

$$2) \ x + y = 1, \quad x + 1.01y = 1.01$$

$$3) \ x + y = 1, \quad 0.99x + y = 1.01$$

O primeiro possui a solução $x = 1, y = 0$. O segundo, $x = 0, y = 1$. Uma variação de 0.01 no membro da direita da segunda equação do sistema ocasiona uma modificação da solução com uma outra ordem de grandeza. Finalmente, a solução do terceiro sistema é

$x = -1, y = -2$, de modo que do primeiro sistema ao terceiro, uma variação de 0.01 em cada um dos coeficientes produz uma mudança na solução com uma outra ordem de grandeza.

Ao examinar as matrizes inversas de cada sistema,

$$\begin{bmatrix} 101 & -101 \\ -100 & 100 \end{bmatrix}, \quad \begin{bmatrix} 101 & -100 \\ -100 & 100 \end{bmatrix}, \quad \begin{bmatrix} 100 & -100 \\ -99 & 100 \end{bmatrix}$$

observa-se que os elementos são grandes, quando comparados com os do sistema dado. Em geral, se \mathbf{x} e $\mathbf{x} + \Delta\mathbf{x}$ são as respectivas soluções dos sistemas

$$\mathbf{Ax} = \mathbf{b} \quad , \quad \mathbf{A}(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b} .$$

Então, fazendo a diferença das duas equações, obtém-se

$$\mathbf{A}(\Delta\mathbf{x}) = \Delta\mathbf{b} .$$

Observa-se que, para uma pequena variação do dado $\Delta\mathbf{b}$, a variação do valor da solução $\Delta\mathbf{x}$ vai depender do comportamento da inversa da matriz \mathbf{A} ao multiplicar pela variação dos dados. Em particular, se as componentes de \mathbf{A} são pequenas e as da inversa são, em comparação, muito grandes.

Na prática, variações nos coeficientes podem estar ligadas ao arredondamento de cifras significativas. Pois, os coeficientes podem ser determinados experimentalmente, ou assumido que, para um propósito de desenho, eles estão dentro de uma certa tolerância. Se este não for o caso, muito pouco pode ser feito para melhorar a situação, uma vez que as equações foram formuladas. Em alguns problemas, uma reformulação com outro embasamento pode melhorar o condicionamento.

Exemplo 3.25

Considere-se o sistema

$$\frac{1}{2}x_1 + \frac{1}{3}x_2 + \frac{1}{4}x_3 = 1$$

$$\frac{1}{3}x_1 + \frac{1}{4}x_2 + \frac{1}{5}x_3 = \frac{1}{2}$$

$$\frac{1}{4}x_1 + \frac{1}{5}x_2 + \frac{1}{6}x_3 = \frac{1}{3}$$

Que possui solução exata

$$x = 12, \quad y = -30, \quad z = 20$$

Esses números são grandes quando comparados com os do sistema, o que sugere que os elementos da matriz inversa devem ser grandes. Isto somente pode acontecer se o determinante do sistema for pequeno. De fato, o determinante é $1/43.200 \doteq 0.00023$ e a matriz inversa é

$$\begin{bmatrix} 72 & -240 & 180 \\ -240 & 900 & -720 \\ 180 & -720 & 600 \end{bmatrix} .$$

Arredondando os coeficientes do sistema para duas casa decimais, tem-se

$$0.50x_1 + 0.33x_2 + 0.25x_3 = 1$$

$$0.33x_1 + 0.25x_2 + 0.20x_3 = 0.50$$

$$0.25x_1 + 0.20x_2 + 0.17x_3 = 0.33$$

A solução exata do sistema arredondado é

$$x = \frac{605}{112} \doteq 5.40, \quad y = -\frac{625}{112} \doteq -5.58, \quad z = \frac{9}{16} \doteq 0.56.$$

Estes números têm pouca semelhança com os da solução do sistema original. O valor do determinante do sistema arredondado é 0.000112, ou seja, quase cinco vezes o do sistema original. Além disso, os elementos da matriz inversa do sistema arredondado, com duas casas decimais, diferem da original por fatores da ordem de 5.

$$\begin{bmatrix} 22.32 & -54.46 & 31.25 \\ -54.46 & 200.89 & -156.25 \\ 31.25 & -156.25 & 143.75 \end{bmatrix}$$

Geometricamente, as equações dos sistemas dados representam planos quase paralelos, portanto de interseção muito longe! Pode ser observado que esse fenômeno numérico se apresenta quando o determinante do sistema possuir um valor menor do que se esperaria para os valores dos elementos da matriz.

Exemplo 3.26

Um sistema, constituído por três molas em série, fixadas em dois apoios rígidos, sujeito a duas forças de tensão p_1 e p_2

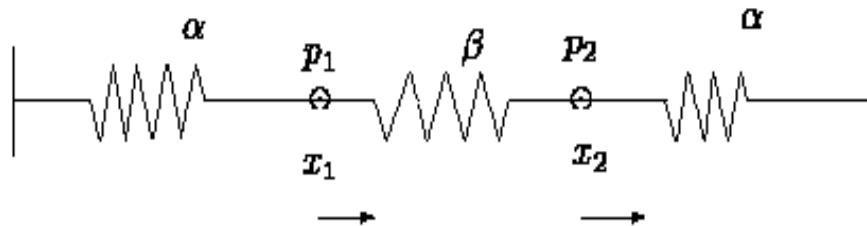


Figura 3.1 – Equilíbrio de Força Elástica

é representado pelas equações de equilíbrio elástico, descritas pelo sistema linear

$$(\lambda + \beta)x_1 - \beta x_2 = p_1$$

$$-\beta x_1 + (\lambda + \beta)x_2 = p_2 .$$

A matriz do sistema e sua inversa são dadas por

$$\mathbf{A} = \begin{bmatrix} \lambda + \beta & \beta \\ -\beta & \lambda + \beta \end{bmatrix} \quad \text{e} \quad \mathbf{A}^{-1} = \frac{1}{\lambda + 2\beta} \begin{bmatrix} 1 + \frac{\beta}{\lambda} & \frac{\beta}{\lambda} \\ \frac{\beta}{\lambda} & 1 + \frac{\beta}{\lambda} \end{bmatrix} .$$

O condicionamento deste problema depende dos parâmetros λ e β . Em particular, o sistema é mal condicionado, quando a natureza das molas é tal que o quociente $\frac{\beta}{\lambda}$ é grande.

Por exemplo, uma mola central muito “dura” e duas molas muito “macias” nas extremidades fixas às paredes.

Um sistema será dito *bem condicionado* quando um pequeno erro nos dados não ocasiona um erro relativo grande na solução. Em variadas estimativas destes erros, com o uso do conceito de norma matricial, a ser definido posteriormente, encontra-se o chamado *número de condicionamento*

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

onde $\|\cdot\|$ denota a norma da matriz A . Assume-se que esta norma verifica a propriedade

$$\|\mathbf{AB}\| \leq \|\mathbf{A}\| \|\mathbf{B}\|.$$

Por exemplo, $\|\mathbf{A}\| = \max_{j=1:n} \sum_{i=1}^m |a_{ij}|$ para uma matriz $\mathbf{A} = [a_{ij}]$. Em particular, $\|\mathbf{I}\|=1$ para a matriz identidade.

Da relação $\mathbf{AA}^{-1} = \mathbf{I}$, e da propriedade acima, segue que $\text{cond}(A) \geq 1$. Deve salientar-se que o valor deste número, dependerá da definição de norma utilizada.

3.3.2 Fatorização L D U

A fatorização \mathbf{LU} não apresenta simetria, no seguinte aspecto: enquanto \mathbf{U} possui pivôs na sua diagonal principal, os elementos diagonais de \mathbf{L} são unitários. Este fato pode ser revertido extraindo de \mathbf{U} uma matriz diagonal \mathbf{D} composta inteiramente pelos pivôs. Por exemplo,

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 0 & -4 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & & \\ & -1 & \\ & & -4 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 1/2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}. \end{aligned}$$

Resulta, deste modo, que a decomposição triangular $\mathbf{A} = \mathbf{LU}$ pode ser escrita na forma

$$\mathbf{A} = \mathbf{LDU}, \quad (3.31)$$

onde \mathbf{L} é triangular inferior e \mathbf{U} é triangular superior, ambas com elementos diagonais unitários. \mathbf{D} é a matriz diagonal formada pelos pivôs. Observe-se que a notação utilizada, ainda que possa gerar confusão, consiste em denotar a matriz triangular superior com elementos diagonais unitários pela mesma letra \mathbf{U} .

A decomposição $\mathbf{A} = \mathbf{LDU}$ permite resolver um sistema $\mathbf{Ax} = \mathbf{b}$ em três passos:

1. achar \mathbf{c} de $\mathbf{Lc} = \mathbf{b}$
2. achar \mathbf{z} de $\mathbf{Dz} = \mathbf{c}$
3. achar \mathbf{x} de $\mathbf{Ux} = \mathbf{z}$

A decomposição \mathbf{LDU} é única e permite flexibilidade para o cálculo prático das matrizes \mathbf{L} e \mathbf{U} . Por exemplo, isto acontece com a **redução de Crout**, na qual considera-se \mathbf{LD} como sendo uma matriz triangular inferior L (sem elementos unitários na sua diagonal) de modo que $\mathbf{A} = \mathbf{LU}$ com \mathbf{U} triangular superior e elementos unitários na sua diagonal. Esta redução permite calcular alternadamente as colunas de L e de \mathbf{U} .

3.3.3 Fatorização \mathbf{LDL}^t

Para matrizes simétricas, a decomposição \mathbf{LDU} pode ser escrita numa forma mais simples. Por exemplo, a matriz tridiagonal simétrica

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

possui a fatorização $\mathbf{A} = \mathbf{LU}$, onde

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 2/3 & 1 & 0 \\ 0 & 0 & 3/4 & 1 \end{bmatrix} \quad \text{e} \quad \mathbf{U} = \begin{bmatrix} 2 & 1 & 0 & 0 \\ 0 & 3/2 & 1 & 0 \\ 0 & 0 & 3/4 & 1 \\ 0 & 0 & 0 & 5/4 \end{bmatrix}.$$

Escrevendo

$$\mathbf{U} = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 5/4 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 0 & 0 \\ 0 & 1 & 2/3 & 0 \\ 0 & 0 & 1 & 3/4 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

decorre que

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 1 & 0 & 0 \\ 0 & 2/3 & 1 & 0 \\ 0 & 0 & 3/4 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 5/4 \end{bmatrix} \begin{bmatrix} 1 & 1/2 & 0 & 0 \\ 0 & 1 & 2/3 & 0 \\ 0 & 0 & 1 & 3/4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Assim, para matrizes simétricas a decomposição LDU resulta da forma

$$\mathbf{A} = \mathbf{LDL}^t. \quad (3.32)$$

Deve-se observar que se uma matriz possui essa fatorização, então ela deve ser simétrica. Pois, $\mathbf{A}^t = (\mathbf{LDL}^t)^t = (\mathbf{L}^t)^t \mathbf{D}^t \mathbf{L}^t = \mathbf{LDL}^t = \mathbf{A}$.

3.3.4 Fatorização \mathbf{LL}^t de Cholesky

Matrizes simétricas que possuem a fatorização \mathbf{LDL}^t com $\mathbf{D} > 0$, ou seja, que todos os pivôs são positivos são chamadas *matrizes positivas definidas*. Para este tipo de matrizes a fatorização pode ser ainda mais simples.

Com relação ao exemplo anterior, a matriz \mathbf{D} pode ser escrita na forma $\sqrt{\mathbf{D}}\sqrt{\mathbf{D}}$. Assim

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 3/2 & 0 & 0 \\ 0 & 0 & 3/4 & 0 \\ 0 & 0 & 0 & 5/4 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{3/2} & 0 & 0 \\ 0 & 0 & \sqrt{3/4} & 0 \\ 0 & 0 & 0 & \sqrt{5/4} \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{3/2} & 0 & 0 \\ 0 & 0 & \sqrt{3/4} & 0 \\ 0 & 0 & 0 & \sqrt{5/4} \end{bmatrix}$$

Denotando $L = \mathbf{L}\sqrt{\mathbf{D}}$, decorre que $L^t = \sqrt{\mathbf{D}}\mathbf{L}^t$. Deste modo, obtém-se a fatorização de Cholesky

$$\mathbf{A} = \mathbf{LL}^t,$$

onde L é uma matriz triangular inferior. No exemplo,

$$\begin{bmatrix} 2 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \\ 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 \\ \sqrt{2}/2 & \sqrt{3}/2 & 0 & 0 \\ 0 & \sqrt{2}/3 & 2/\sqrt{3} & 0 \\ 0 & 0 & \sqrt{3}/2 & \sqrt{5}/2 \end{bmatrix} \begin{bmatrix} \sqrt{2} & \sqrt{2}/2 & 0 & 0 \\ 0 & \sqrt{3}/2 & \sqrt{2}/3 & 0 \\ 0 & 0 & 2/\sqrt{3} & \sqrt{3}/2 \\ 0 & 0 & 0 & \sqrt{5}/2 \end{bmatrix}.$$

A obtenção da matriz L , na fatorização $\mathbf{A} = \mathbf{LL}^t$, é realizada como segue

$$\begin{aligned}
l_{11} &= a_{11} \\
l_{ii} &= \left[a_{ii} - \sum_{j=1}^{i-1} l_{ij}^2 \right]^{1/2}, \quad i = 2, \dots, n \\
l_{ij} &= (1/l_{jj}) \left[a_{ij} - \sum_{k=1}^{j-1} l_{jk} l_{ik} \right]; \quad j = 1, 2, \dots, n, \quad i = j+1, \dots, n \\
l_{ij} &= 0, \quad i < j.
\end{aligned} \tag{3.33}$$

Exemplo 3.27

Resolver o sistema

$$\begin{bmatrix} 4 & -1 & 0 & 0 \\ -1 & 4 & -1 & 0 \\ 0 & -1 & 4 & -1 \\ 0 & 0 & -1 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

utilizando o método de Cholesky.

Solução

Tem-se

$$\mathbf{L} = \begin{bmatrix} l_{11} & 0 & 0 & 0 \\ l_{21} & l_{22} & 0 & 0 \\ l_{31} & l_{32} & l_{33} & 0 \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix}.$$

Assim

$$\begin{aligned}
l_{11} &= 2 \\
l_{21} &= -1/2 & l_{22} &= (15/4)^{1/2} \\
l_{31} &= 0 & l_{32} &= -(4/15)^{1/2} & l_{33} &= (56/15)^{1/2} \\
l_{41} &= 0 & l_{42} &= 0 & l_{43} &= -(15/56)^{1/2} & l_{44} &= (209/56)^{1/2}.
\end{aligned}$$

A resolução de

$$\mathbf{L}\mathbf{L}^t\mathbf{x} = \mathbf{b},$$

equivale a resolver dois sistemas triangulares

$$\begin{aligned}
\mathbf{L}\mathbf{z} &= \mathbf{b} \\
\mathbf{L}^t\mathbf{x} &= \mathbf{z}
\end{aligned}$$

Primeiro para \mathbf{z} e, após, para \mathbf{x} .

Tem-se que

$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ -(1/2) & (15/4)^{1/2} & 0 & 0 \\ 0 & -(4/15)^{1/2} & (56/15)^{1/2} & 0 \\ 0 & 0 & -(15/56)^{1/2} & (209/56)^{1/2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

possui a solução

$$z_1 = 1/2, \quad z_2 = 1/(60)^{1/2}, \quad z_3 = 1/(840)^{1/2}, \quad z_4 = 1/(11704)^{1/2}$$

A solução do segundo sistema triangular é por retrossubstituição ascendente. Assim,

$$\begin{bmatrix} 2 & -1/2 & 0 & 0 \\ 0 & (15/4)^{1/2} & -(4/15)^{1/2} & 0 \\ 0 & 0 & (56/15)^{1/2} & -(15/56)^{1/2} \\ 0 & 0 & 0 & (209)^{1/2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1/2 \\ 1/(60)^{1/2} \\ 1/(840)^{1/2} \\ 1/(11704)^{1/2} \end{bmatrix}$$

possui a solução

$$x_4 = 1/209, \quad x_3 = 4/209, \quad x_2 = 15/209 \quad \text{e} \quad x_1 = 56/209.$$

3.4 Sistemas Lineares Singulares

Os sistemas de equações algébricas lineares

$$\mathbf{Ax} = \mathbf{b} ,$$

em que a matriz dos coeficientes \mathbf{A} possui determinante nulo, são chamados de *sistemas lineares singulares*. Este tipo de sistemas apresenta dois problemas: o sistema pode não ter solução ou, se possuir solução, esta não é determinada de maneira única. Ou seja, um sistema singular pode ser *inconsistente* (sem solução, no sentido usual) ou *consistente e subdeterminado* (tem solução, porém não necessariamente única).

Considere-se o seguinte sistema de equações lineares

$$\begin{aligned} x_1 + x_2 + \quad + x_4 &= b_1 \\ 2x_1 + 2x_2 + \quad + 2x_4 &= b_2 \\ \quad x_2 + x_3 &= b_3 \\ 2x_1 + 3x_2 + x_3 + 4x_4 &= b_4 \end{aligned} \tag{3.34}$$

Este sistema é singular, uma vez que o determinante do sistema é nulo, pois a segunda linha é um múltiplo da primeira.

O primeiro passo, na eliminação Gaussiana, produz o sistema

$$\begin{aligned} x_1 + x_2 + \quad + x_4 &= b_1 \\ 0x_1 + 0x_2 + 0x_3 + 0x_4 &= b_2 - 2b_1 \\ \quad x_2 + x_3 &= b_3 \\ \quad x_2 + x_3 + 2x_4 &= b_4 - 2b_1. \end{aligned}$$

Eliminando x_2 da última equação e permutando as linhas, decorre

$$\begin{aligned} x_1 + x_2 + \quad x_4 &= c_1 = b_1 \\ \quad x_2 + x_3 &= c_2 = b_3 \\ 2x_4 &= c_3 = b_4 - b_3 - 2b_1 \\ 0 &= c_4 = b_2 - 2b_1. \end{aligned} \tag{3.35}$$

Em termos matriciais

$$\mathbf{Ux} = \mathbf{c} , \tag{3.36}$$

onde

$$\mathbf{U} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} b_1 \\ b_3 \\ b_4 - b_3 - 2b_1 \\ b_2 - 2b_1 \end{bmatrix}.$$

Observe-se que, num sistema singular, pelo menos uma linha da matriz triangular \mathbf{U} deve ser nula, isto é, tem-se a presença de pivôs nulos. As linhas nula impõem restrições sobre o dado \mathbf{b} . No exemplo, o dado \mathbf{b} deve ser tal que $c_4 = b_2 - 2b_1 = 0$. *Se esta condição não é verificada, isto é, $b_2 - 2b_1 \neq 0$, então o sistema é inconsistente e não possui solução para o dado \mathbf{b} .* Satisfeita essa condição sobre o dado \mathbf{b} , o sistema $\mathbf{Ux} = \mathbf{c}$ pode ser resolvido por retrosubstituição para as *variáveis básicas* x_1, x_2 e x_4 , que correspondem aos pivôs não nulos, deixando-se x_3 , que corresponde a um pivô nulo, como uma *variável livre*. Assim,

$$\begin{aligned} x_4 &= (b_4 - b_3 - 2b_1)/2 \\ x_3 &= a \\ x_2 &= b_3 - a \\ x_1 &= b_1 - x_2 - x_4 = 2b_1 - (3/2)b_3 - b_4/2 + a , \end{aligned} \tag{3.37}$$

onde $x_3 = a$ é uma constante arbitrária. Deste modo, tem-se que a solução do sistema não é única, ou seja, não é bem determinada, pois para qualquer valor de a tem-se uma solução. Decorre que

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = a \begin{bmatrix} 1 \\ -1 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2b_1 - (3/2)b_3 - b_4/2 \\ b_3 \\ 0 \\ (b_4 - b_3 - 2b_1)/2 \end{bmatrix} .$$

Um critério teórico para saber a priori se um sistema singular é **inconsistente**, isto é, não possui solução no sentido usual, pode ser obtido da identidade de Cramer

$$\det(\mathbf{A})\mathbf{x} = \text{adj}(\mathbf{A})\mathbf{b} ,$$

Um sistema com $\det(\mathbf{A})=0$ será consistente se

$$0 = \text{adj}(\mathbf{A})\mathbf{b} , \quad (3.38)$$

Se o dado \mathbf{b} não satisfizer a condição 3.38, então o sistema para esse dado não poderá ter uma solução usual \mathbf{x} , isto é, o sistema é inconsistente com o dado \mathbf{b} . Por exemplo,

$$\begin{aligned} 4x_1 + 3x_2 &= 0 \\ -8x_1 - 6x_2 &= 1 \end{aligned}$$

não possui solução no sentido usual, pois a condição (8) não é satisfeita:

$$\text{adj} \begin{bmatrix} 4 & 3 \\ -8 & -6 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -6 & -3 \\ 8 & 4 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -3 \\ 4 \end{bmatrix} \neq 0 .$$

A não existência de solução clássica deste sistema deve-se ao fato que suas equações são incompatíveis. Pois, multiplicando a primeira equação por -2, obtém-se

$$-8x_1 - 6x_2 = 0 ,$$

a qual não é compatível com a segunda equação do sistema.

Se no sistema anterior, o dado \mathbf{b} for mudado, isto é,

$$\begin{aligned} 4x_1 + 3x_2 &= 1 \\ -8x_1 - 6x_2 &= -2 \end{aligned} ,$$

obtém-se um sistema consistente. Certamente, a segunda equação é redundante, pois é o dobro da primeira, e não acrescenta informação alguma para a obtenção da solução. Fazendo $x_2 = a$, obtém-se da primeira equação $x_1 = (1 - 3a)/4$. Assim,

$$\mathbf{x} = \begin{bmatrix} (1 - 3a)/4 \\ a \end{bmatrix}$$

satisfaz o sistema para qualquer valor da constante a . Ou seja, a solução não é bem determinada.

De este exemplo, conclui-se que um sistema singular pode ser consistente ou inconsistente em relação ao dado \mathbf{b} .

Em termos geométricos, o caso inconsistente, para o sistema acima, é representado como a situação de duas retas paralelas que não se cortam e, portanto, não há solução. No caso consistente, tem-se somente uma reta, então todos os pontos sobre ela fornecem uma solução do sistema, isto é, uma infinidade de soluções.

3.4.1 Posto e Nulidade

Para sistematizar as observações da seção anterior, é necessário introduzir os conceitos de **posto** e de **nulidade** de uma matriz \mathbf{A} . O posto pode ser definido com o auxílio de determinantes ou com a matriz \mathbf{U} que resulta da eliminação. Estas definições são equivalentes.

Uma matriz \mathbf{A} possui *posto* r quando uma das seguintes condições equivalentes se verifica

- r é a ordem da maior submatriz quadrada de \mathbf{A} que é não-singular. Isto é, \mathbf{A} possui ao menos um determinante de ordem r que é não nula, porém todos os determinantes de maior ordem que r se anulam.
- r é o número linhas não nulas da matriz triangular superior \mathbf{U} obtida por eliminação da matriz \mathbf{A} . Equivalentemente, r é o número de pivôs não nulos.

Uma matriz é dita de posto zero quando todos seus elementos são nulos, isto é, a matriz nula. Se \mathbf{A} é uma matriz quadrada de ordem n , a diferença

$$s = n - r$$

é chamada de *nulidade*. A nulidade de uma matriz é o número de pivôs nulos.

Matrizes quadradas não singulares de ordem n , possuem posto $r = n$ e nulidade $s = 0$.

Exemplo 3.28

A matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 \\ 4 & 8 & 4 \\ 0 & 1 & 2 \end{bmatrix}$$

possui posto 2, uma vez que, em particular, a submatriz

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$$

tem determinante não nulo e a única submatriz de ordem 3 de \mathbf{A} é a própria matriz \mathbf{A} que possui determinante nulo.

Por outro lado, por eliminação, obtém-se a matriz

$$\mathbf{U} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

cujo posto também é 2.

Para resolver um sistema singular $\mathbf{Ax} = \mathbf{b}$, procede-se da seguinte maneira:

1. A eliminação gaussiana é utilizada para obter o sistema equivalente $\mathbf{Ux} = \mathbf{c}$. Assume-se que com troca de linhas, se necessária, as linhas não nulas de \mathbf{U} aparecem em primeiro lugar e as nulas por último.
2. Determina-se o posto r e a nulidade s .

3. Os elementos de \mathbf{c} são examinados segundo duas opções:

- (a) Se ao menos um dos elementos c_{r+1}, \dots, c_n for não nulo, o sistema é *inconsistente* e portanto, o sistema não tem solução.
- (b) Se todos os elementos c_{r+1}, \dots, c_n forem nulos, então, o sistema é *consistente*, Definem-se como variáveis básicas as r variáveis que correspondem aos pivôs não-nulos e são definidas as outras $s = n - r$ como variáveis livres.

4. Com a opção 3b, a resolução de $\mathbf{U}\mathbf{x} = \mathbf{c}$ é feita por retrossubstituição, sendo as variáveis básicas obtidas em termos das variáveis livres.

Em resumo, *para sistemas singulares consistentes, todas as soluções de $\mathbf{U}\mathbf{x} = \mathbf{c}$ são obtidas atribuindo-se valores arbitrários às $s = n - r$ variáveis livres e resolvendo por retrossubstituição para as r variáveis básicas.*

Exemplo 3.29

Resolver o sistema, cuja forma reduzida $\mathbf{U}\mathbf{x} = \mathbf{0}$ é dada por

$$\begin{array}{rcccccc} x_1 & - & x_2 & + & 2x_3 & + & x_4 & + & x_5 & & = & c_1 \\ & & & & & & 3x_4 & + & 5x_5 & + & x_6 & = & c_2 \\ & & & & & & & & & & 4x_6 & = & c_3 \\ & & & & & & & & & & 0 & = & c_4 \\ & & & & & & & & & & 0 & = & c_5 \\ & & & & & & & & & & 0 & = & c_6 \end{array}$$

Solução

Tem-se duas situações. Se ao menos um dos valores c_4, c_5, c_6 for diferente de zero, o sistema será inconsistente e não possuirá solução no sentido usual. Se $c_4 = c_5 = c_6 = 0$, então o sistema será consistente e terá várias soluções. Para esta última situação, as variáveis básicas serão x_1, x_4 e x_6 , correspondentes aos pivôs não nulos 1, 3 e 4 respectivamente e as variáveis livres serão x_2, x_3 e x_5 . Fazendo-se

$$x_2 = \alpha, \quad x_3 = \beta \quad \text{e} \quad x_5 = \omega,$$

e por retrossubstituição, decorre

$$\begin{array}{l} x_6 = c_4/4 \\ x_4 = (c_2 - 5\omega - c_3/4) \\ x_1 = (c_1 - c_2/3 + c_3/12) + \alpha - 2\beta + 2\omega/3. \end{array}$$

Deste modo, obtém-se a solução geral

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} (c_1 - c_2/3 + c_3/12) + \alpha - 2\beta + 2\omega/3 \\ \alpha \\ \beta \\ c_2 - 5\omega - c_3/4 \\ \omega \\ c_4/4 \end{bmatrix}.$$

ou

$$\mathbf{x} = \begin{bmatrix} c_1 - c_2/3 + c_3/12 \\ 0 \\ 0 \\ c_2 - c_3/4 \\ 0 \\ c_4/4 \end{bmatrix} + \alpha \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \omega \begin{bmatrix} 2/3 \\ 0 \\ 0 \\ -5 \\ 1 \\ 0 \end{bmatrix}.$$

Decorre que a solução é da forma

$$\mathbf{x} = \mathbf{x}_p + \alpha \mathbf{h}_1 + \beta \mathbf{h}_2 + \omega \mathbf{h}_3.$$

Ou seja, é a superposição de uma solução particular \mathbf{x}_p ($\alpha = \beta = \omega = 0$) com múltiplos das soluções $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ da equação homogênea $\mathbf{U}\mathbf{x} = \mathbf{0}$, pois

$$\mathbf{U}(\mathbf{x} - \mathbf{x}_p) = \mathbf{U}\mathbf{x} - \mathbf{U}\mathbf{x}_p = \mathbf{b} - \mathbf{b} = \mathbf{0}$$

e, portanto,

$$\alpha \mathbf{U}\mathbf{h}_1 + \beta \mathbf{U}\mathbf{h}_2 + \omega \mathbf{U}\mathbf{h}_3 = \mathbf{0}.$$

Escolhendo $\alpha = 1, \beta = \omega = 0$, obtém-se que $\mathbf{U}\mathbf{h}_1 = \mathbf{0}$. Similarmente, com \mathbf{h}_2 e \mathbf{h}_3 para adequados valores de α, β e ω .

Em geral, considerando que o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ e o seu reduzido $\mathbf{U}\mathbf{x} = \mathbf{c}$ possuem as mesmas soluções, tem-se que a solução de um sistema consistente $\mathbf{A}\mathbf{x} = \mathbf{b}$, com \mathbf{A} de ordem n e de posto r , é da forma

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{h}_2 + \cdots + \alpha_s \mathbf{h}_s \quad (3.39)$$

onde $s = n - r$ é a nulidade da matriz \mathbf{A} . Aqui \mathbf{x}_p é uma solução particular do sistema, $\mathbf{h}_1, \cdots, \mathbf{h}_s$ são soluções da equação homogênea $\mathbf{A}\mathbf{x} = \mathbf{0}$ e $\alpha_1, \cdots, \alpha_s$ são parâmetros arbitrários.

Historicamente, a resolução de sistemas singulares com o uso de determinantes precede ao desenvolvimento sistemático da eliminação. Suponha-se que o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$, com \mathbf{A} uma matriz quadrada de ordem n , possui posto r menor do que n . Também, que as últimas $n-r$ equações são proporcionais ou combinações lineares das primeiras r equações, isto é, são redundantes e, portanto, descartáveis. Então, as equações a serem resolvidas são

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1r}x_r &= b_1 - a_{1,r+1}x_{r+1} - \cdots - a_{1n}x_n \\ a_{22}x_2 + \cdots + a_{2r}x_r &= b_2 - a_{2,r+1}x_{r+1} - \cdots - a_{2n}x_n \\ &\vdots \\ a_{rr}x_1 + \cdots + a_{rr}x_r &= b_r - a_{r,r+1}x_{r+1} - \cdots - a_{rn}x_n, \end{aligned}$$

pois as outras $n - r$ equações não acrescentam informação adicional sobre as incógnitas x_1, \cdots, x_r , uma vez que podem ser descartadas do sistema.

Escrevendo-se este último sistema na forma

$$\mathbf{C}\mathbf{v} = \mathbf{d}, \quad (3.40)$$

onde \mathbf{C} é a matriz não singular de ordem r , com elementos a_{ij} , e \mathbf{d} o vetor coluna $r \times 1$, com elementos $b_i - a_{i,r+1}x_{r+1} - \cdots - a_{in}x_n$, decorre que

$$\mathbf{v} = \frac{\text{adj}(\mathbf{C})}{\det(\mathbf{C})} \mathbf{d}.$$

Assim, a solução do sistema singular $\mathbf{A}\mathbf{x} = \mathbf{b}$ é dada por

$$\mathbf{x} = \begin{bmatrix} v \\ x_{r+1} \\ \vdots \\ x_n \end{bmatrix},$$

onde x_{r+1}, \cdots, x_n são $n - r$ constantes que podem ser arbitrárias.

Por este motivo, um sistema singular consistente é referido na literatura como um *sistema subdeterminado*, pois as equações do sistema não são suficientes para determinar uma única solução.

Exemplo 3.30

Resolver o sistema de equações lineares

$$\begin{aligned}x_1 - 2x_2 + 3x_3 &= 4 \\x_1 + x_2 + 2x_3 &= 5 \\2x_1 - x_2 + 5x_3 &= 9\end{aligned}$$

Solução

É fácil ver que o posto da matriz dos coeficientes é igual a dois e, portanto, o sistema é singular. Assim, considere-se o subsistema 2×2 de posto 2

$$\begin{aligned}x_1 - 2x_2 &= 4 - 3x_3 \\x_1 + x_2 &= 5 - 2x_3\end{aligned}$$

ou

$$\begin{bmatrix} 1 & -2 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 - 3x_3 \\ 5 - 2x_3 \end{bmatrix} .$$

Tem-se

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = (1/3) \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 4 - 3x_3 \\ 5 - 2x_3 \end{bmatrix} = (1/3) \begin{bmatrix} 14 - 7x_3 \\ 1 + x_3 \end{bmatrix} .$$

Portanto, a solução geral é dada por

$$\mathbf{x} = \begin{bmatrix} (14 - 7a)/3 \\ (1 + a)/3 \\ a \end{bmatrix} ,$$

onde $x_3 = a$ é uma constante arbitrária.

3.5 Sistemas Lineares Retangulares

O tratamento geral de sistemas de equações algébricas lineares com coeficientes numéricos

$$\mathbf{Ax} = \mathbf{b} ,$$

em que \mathbf{A} é de ordem $m \times n$, \mathbf{x} de ordem $n \times 1$ e \mathbf{b} de ordem $m \times 1$ é inteiramente análogo a do caso de sistemas singulares.

Exemplo 3.31

Resolver o sistema de equações

$$\begin{array}{rcccccc} & & & 3x_4 & + & 5x_5 & + & 2x_6 & = & b_1 \\ x_1 & - & x_2 & + & 2x_3 & + & x_4 & + & x_5 & = & b_2 \\ 2x_1 & - & 2x_2 & + & 4x_3 & + & 5x_4 & + & 7x_5 & + & 6x_6 & = & b_3 \\ 3x_1 & - & 3x_2 & + & 6x_3 & + & 6x_4 & + & 8x_5 & + & 2x_6 & = & b_4 \end{array} .$$

Solução

A eliminação gaussiana é realizada, no primeiro passo, permutam-se a primeira e segunda equações e, após, subtraindo-se múltiplos desta nova primeira equação das outras. O múltiplo a ser subtraído da terceira equação é 2 e da quarta equação é 3, obtendo-se o sistema

$$\begin{array}{rcccccc} x_1 & - & x_2 & + & 2x_3 & + & x_4 & + & x_5 & = & b_2 \\ & & & & & & 3x_4 & + & 5x_5 & + & 2x_6 & = & b_1 \\ & & & & & & 3x_4 & + & 5x_5 & + & 6x_6 & = & b_3 - 2b_2 \\ & & & & & & 3x_4 & + & 5x_5 & + & 2x_6 & = & b_4 - 3b_2 \end{array} .$$

Completa-se o processo da eliminação, subtraindo-se a segunda equação da terceira e quarta equações (o múltiplo é 1, em ambos casos). Assim,

$$\begin{array}{rcccccc} x_1 & - & x_2 & + & 2x_3 & + & x_4 & + & x_5 & = & b_2 \\ & & & & & & 3x_4 & + & 5x_5 & + & 2x_6 & = & b_1 \\ & & & & & & & & & & 4x_6 & = & b_3 - 2b_2 - b_1 \\ & & & & & & & & & & 0 & = & b_4 - 3b_2 - b_1 \end{array} .$$

Tem-se as variáveis básicas x_1, x_4, x_6 , correspondentes aos pivôs não nulos, e as variáveis livres x_2, x_3, x_5 , associadas aos pivôs nulos. Por retrosubstituição, decorre

$$\begin{array}{l} x_6 = (b_3 - 2b_2 - b_1)/4 \\ x_4 = (b_1 - 5x_5 - 2x_6)/3 = (3b_1 + 4b_2 - 2b_3 - 5u)/3 \\ x_1 = b_2 + x_2 - 2x_3 - x_4 - x_5, \end{array}$$

onde

$$\begin{array}{l} x_5 = u \\ x_3 = v \\ x_2 = w \end{array}$$

são constantes arbitrárias. Para que este procedimento seja inteiramente válido, deve-se assumir, a priori, que os elementos do dado \mathbf{b} satisfazem à relação de consistência

$$b_4 - 3b_2 - b_1 = 0 ,$$

a qual decorre da última equação do sistema final reduzido. Neste caso, segue que

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix} = \begin{bmatrix} b_2 \\ 0 \\ 0 \\ b_1 \\ 0 \\ b_3 - 2b_2 - b_1 \end{bmatrix} + w \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + v \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + u \begin{bmatrix} 1 \\ 0 \\ 0 \\ -5 \\ 1 \\ 0 \end{bmatrix} .$$

Matricialmente, a eliminação Gaussiana, incluída a troca de linhas, transformou a matriz do sistema original

$$\mathbf{A} = \begin{bmatrix} 0 & 0 & 0 & 3 & 5 & 2 \\ 1 & -1 & 2 & 1 & 1 & 0 \\ 2 & -2 & 4 & 5 & 7 & 6 \\ 3 & -3 & 6 & 6 & 8 & 2 \end{bmatrix} ,$$

na matriz escada

$$\mathbf{U} = \begin{bmatrix} 1 & -1 & 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 3 & 5 & 2 \\ 0 & 0 & 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} .$$

Para matrizes retangulares \mathbf{A} de ordem $m \times n$, a definição de posto r e nulidade s é a mesma dos sistemas singulares, isto é, r é o número de pivôs não nulos e $s = n - r$. Através de determinantes, r é a ordem de uma submatriz com determinante não nulo, porém, todos os determinantes das submatrizes de ordem maior que r são nulos. Ambas definições fornecem o mesmo valor. Observe-se que a nulidade é o número de variáveis livres e que r é o número de variáveis básicas.

Em geral, o procedimento é o seguinte

- É realizada a eliminação, incluindo troca de linhas, obtendo-se um sistema $\mathbf{U}\mathbf{x} = \mathbf{c}$.
- É identificado o posto r e a nulidade s .
- O sistema $\mathbf{U}\mathbf{x} = \mathbf{c}$ é dito **inconsistente** se, ao menos, algum dos elementos c_{r+1}, \dots, c_n é diferente de zero;
- O sistema $\mathbf{U}\mathbf{x} = \mathbf{c}$ é **consistente** se $c_{r+1} = c_{r+2} = \dots = c_n = 0$.
- Para sistemas consistentes, são atribuídos valores arbitrários às s variáveis livres, correspondentes aos pivôs nulos, e resolve-se por retrossubstituição para as r variáveis básicas, que correspondem aos pivôs não-nulos.

A rigor, a resolução de um sistema retangular consistente equivale a extrair o maior sub-sistema não-singular de ordem $r \times r$ e invertê-lo, considerando as outras $s = n - r$ variáveis como parâmetros. Mais precisamente,

A solução geral de um sistema retangular $\mathbf{A}\mathbf{x} = \mathbf{b}$, \mathbf{A} $m \times n$, para o qual o sistema reduzido $\mathbf{U}\mathbf{x} = \mathbf{c}$ é consistente, tem a forma

$$\mathbf{x} = \mathbf{x}_p + \alpha_1 \mathbf{h}_1 + \dots + \alpha_s \mathbf{h}_s .$$

Aqui \mathbf{x}_p é uma solução particular ($\mathbf{A}\mathbf{x}_p = \mathbf{b}$ ou $\mathbf{U}\mathbf{x}_p = \mathbf{c}$), $\mathbf{h}_1, \dots, \mathbf{h}_s$ são soluções do sistema homogêneo associado ($\mathbf{A}\mathbf{x} = \mathbf{0}$ ou $\mathbf{U}\mathbf{x} = \mathbf{0}$), $\alpha_1, \dots, \alpha_s$ são parâmetros livres e s é a nulidade do sistema, ou seja, $s = n - r$, onde r é o posto de \mathbf{A} .

3.6 Soluções Não Nulas de Sistemas Homogêneos

Um sistema homogêneo

$$\mathbf{A}\mathbf{x} = \mathbf{0} , \tag{3.41}$$

com \mathbf{A} uma matriz quadrada de ordem n , é sempre consistente, pois, a condição de consistência $\text{adj}(\mathbf{A})\mathbf{b} = 0$ é satisfeita para $\mathbf{b} = \mathbf{0}$. Em particular, $\mathbf{x} = \mathbf{0}$ é sempre uma solução. É de interesse determinar quando o sistema homogêneo possui soluções não nulas.

- Se $\det(\mathbf{A}) \neq 0$, ou seja, a matriz é não singular, então $\mathbf{A}\mathbf{x} = \mathbf{0}$ possui unicamente a solução nula. De fato,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{0} = \mathbf{0};$$

- Se $\det(\mathbf{A}) = 0$, ou seja, a matriz é singular, então $\mathbf{Ax} = \mathbf{0}$ possui uma infinidade de soluções não nulas que dependem de s parâmetros, onde s é a nulidade.

Pois, se \mathbf{A} é singular de ordem n , então o posto e a nulidade verificam $r < n$ e $s = n - r > 0$.

Resumindo:

O sistema $\mathbf{Ax}=\mathbf{0}$, \mathbf{A} matriz $n \times n$, possui soluções não nulas, se e somente se $\det\mathbf{A} = 0$. Equivalentemente, o posto de \mathbf{A} é menor que n .

Exemplo 3.32

Determinar valores de ω para os quais o problema de contorno

$$\begin{aligned}\ddot{\phi}(x) + \omega^2\phi(x) &= 0 \\ \phi(0) = 0, \quad \phi(L) &= 0\end{aligned}$$

possui como soluções funções não identicamente nulas.

Solução

A equação dada tem a solução geral

$$\phi(x) = c_1\cos\omega x + c_2\sen\omega x.$$

Para determinar as constantes c_1, c_2 devem-se verificar as condições dadas, isto é,

$$0 = \phi(0) = c_1\cos 0 + c_2\sen 0,$$

$$0 = \phi(L) = c_1\cos\omega L + c_2\sen\omega L.$$

Matricialmente, tem-se o sistema homogêneo

$$\begin{bmatrix} 1 & 0 \\ \cos\omega L & \sen\omega L \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Este sistema possui a solução nula $c_1 = c_2 = 0$ que não é de interesse, pois $\phi(x)$ seria uma função identicamente nula. Para obter uma solução não nula desse sistema, seu determinante, denotado por $\Delta(\omega)$, deve ser nulo. Assim,

$$\Delta(\omega) = \sen\omega L = 0$$

Decorre que

$$\omega = \frac{n\pi}{L}$$

e, para estes valores, o sistema é dado por

$$\begin{bmatrix} 1 & 0 \\ \cos\frac{n\pi}{L} & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Por eliminação gaussiana, resulta

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

é a matriz dos coeficientes do sistema,

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix}$$

é a matriz coluna formada pelas incógnitas do sistema e $\mathbf{0}$ é a matriz coluna formada por zeros.

Realizando a eliminação gaussiana, com troca de linhas, obtém-se

$$\begin{bmatrix} 1 & -3 & 1 & 1 & -1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

A matriz deste sistema 3×5 possui posto 3. Tem-se as variáveis básicas a , b , c e as variáveis livres d , e . Assim,

$$\begin{aligned} c &= d - e, \\ b &= -e, \\ a &= 3b - c - d + e = -3e - d + e - d + e = -e - 2d \end{aligned}$$

Segue que

$$\begin{bmatrix} a \\ b \\ c \\ d \\ e \end{bmatrix} = \begin{bmatrix} -e - 2d \\ -e \\ d - e \\ d \\ e \end{bmatrix} = e \begin{bmatrix} -1 \\ -1 \\ -1 \\ 0 \\ 1 \end{bmatrix} + d \begin{bmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$$

Para $e = -1$ e $d = 0$, decorre o produto

$$\frac{\rho V D}{\mu}$$

conhecido como o *número de Reynolds*. Para $e = 0$ e $d = -1$ tem-se o *número de Froude*

$$\frac{V^2}{Lg}$$

Com a solução obtida, observa-se que existe uma infinidade de combinações possíveis. Esta é uma característica dos sistemas retangulares homogêneos com mais variáveis do que equações.

3.7 Geometria Matricial

A manipulação algébrica com matrizes é extremamente valiosa, porque fornece, simultaneamente, uma notação compacta e um procedimento sistemático para a resolução de sistemas de equações lineares. Embora que uma matriz seja, basicamente, uma ferramenta algébrica, ganha-se muito, se as operações estritamente algébricas são complementadas por uma descrição geométrica.

Os vetores no plano são usualmente representados como segmentos de reta desde a origem (ponto correspondente ao vetor nulo) até o ponto (x_1, x_2) que se considera como o vetor dado \mathbf{x}

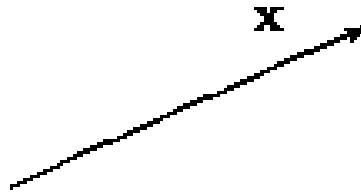


Figura 3.2 – Representação Geométrica

Considere-se a coleção de combinações lineares de dois vetores no plano \mathbf{x} e \mathbf{y}

$$\mathbf{u} = \alpha\mathbf{x} + \beta\mathbf{y}$$

onde α e β são escalares . Equivalentemente, para \mathbf{x} e \mathbf{y} considerados vetores coluna, tem-se matricialmente

$$\mathbf{u} = \alpha\mathbf{x} + \beta\mathbf{y} = \mathbf{A}\mathbf{c},$$

onde

$$\mathbf{A} = [\mathbf{x} \ \mathbf{y}] = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}.$$

Com relação à figura,

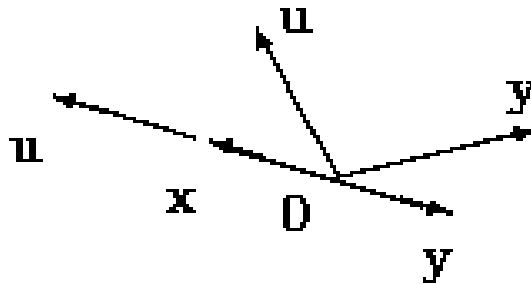


Figura 3.3 – Colinearidade e Não-colinearidade

observe-se que se os vetores \mathbf{x} e \mathbf{y} não forem colineares, então qualquer vetor \mathbf{u} , no plano, poderá escrito de maneira única na forma de combinação linear de \mathbf{x} e \mathbf{y} , para certos escalares α e β , que dependem do vetor \mathbf{u} , isto é, $\mathbf{c} = \mathbf{A}^{-1}\mathbf{u}$. Entretanto, se os vetores \mathbf{x} e \mathbf{y} forem colineares, isto é,

$$\mathbf{y} = k\mathbf{x},$$

para um certo coeficiente de proporcionalidade k , a matriz \mathbf{A} será singular porque seu determinante é nulo. Um vetor qualquer \mathbf{u} poderá, ou não, ser escrito como uma combinação linear de \mathbf{x} e \mathbf{y} , segundo esteja na mesma direção da reta L que contém \mathbf{x} e \mathbf{y} ("consistente") ou fora da reta L ("inconsistente"). Pois, a combinação linear $\mathbf{u} = \alpha\mathbf{x} + \beta\mathbf{y}$ pode ser escrita

$$\mathbf{u} = (\alpha + k\beta)\mathbf{x} = \alpha_1\mathbf{x}.$$

Em particular, para $\mathbf{y} = -\mathbf{x}$, obtém-se $\mathbf{x} + \mathbf{y} = \mathbf{0}$. Ou seja, $\mathbf{0} = \mathbf{A}\mathbf{c}$ com $\mathbf{c} = [1, -1]^t$ uma solução não nula.

Resumindo, dois vetores no plano \mathbf{x} , \mathbf{y} são não colineares ou colineares, segundo o vetor $\mathbf{0}$ possa ser escrito somente como uma combinação linear nula ou não, respectivamente. Estes conceitos podem ser estendidos ao caso de vetores arbitrários.

Os vetores $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ de ordem $m \times 1$ são ditos *linearmente independentes*, se

$$\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \dots + \mathbf{a}_nx_n = \mathbf{0}.$$

for possível unicamente com $x_1 = x_2 = \dots = x_n = 0$. Eles são ditos *linearmente dependentes*, se

$$\mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \dots + \mathbf{a}_nx_n = \mathbf{0}.$$

é possível sem que todos os x_j sejam iguais a zero.

Para caracterizar quando uma coleção de vetores $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ é linearmente independente ou linearmente dependente, escreve-se a combinação linear como um sistema homogêneo

$$\mathbf{A}\mathbf{x} = \mathbf{0},$$

onde \mathbf{A} é a matriz cujas colunas são esses vetores, \mathbf{x} é um vetor coluna $n \times 1$ e $\mathbf{0}$ é o vetor nulo $m \times 1$. Mais precisamente,

$$\mathbf{Ax} = \begin{bmatrix} \mathbf{a}^1 & \mathbf{a}^2 & \dots & \mathbf{a}^n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \mathbf{a}_1x_1 + \mathbf{a}_2x_2 + \dots + \mathbf{a}_nx_n = \mathbf{0} .$$

Tem-se duas opções:

- O sistema $\mathbf{Ax} = \mathbf{0}$ possui somente a solução nula, ou seja, a combinação linear

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{0}$$

é possível unicamente se $x_1 = x_2 = \dots = x_n = 0$.

- O sistema $\mathbf{Ax} = \mathbf{0}$ possui soluções não nulas, ou seja

$$x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n = \mathbf{0} .$$

sem que todos os x_j sejam iguais a zero.

Na primeira situação, os vetores coluna $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ são *linearmente independentes*. Isto ocorre quando a nulidade da matriz \mathbf{A} é $s = 0$ e o posto é $r = n$.

Se há soluções não nulas, então os vetores coluna $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n$ são *linearmente dependentes*. Isto ocorre quando a nulidade é $s > 0$ e o posto é $r < n$.

Exemplo 3.34

Os vetores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} 2 \\ 3 \\ 6 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 4 \\ 8 \\ 2 \end{bmatrix} \quad \text{e} \quad \mathbf{v}_4 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

são linearmente dependentes ?

Solução

A matriz, cujas colunas são os vetores dados, é de ordem 3×4 . Ela possui posto 3 ($< n = 4$) e nulidade $s = 4 - 3 = 1 > 0$. Portanto, os vetores são linearmente dependentes.

Exemplo 3.35

Os vetores *padrão* de ordem $n \times 1$

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

são linearmente independentes. De fato,

$$\mathbf{0} = c_1\mathbf{e}_1 + c_2\mathbf{e}_2 + \dots + c_n\mathbf{e}_n = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix},$$

unicamente, se $c_1 = c_2 = \dots = c_n = 0$. Além disso, cada vetor coluna \mathbf{x} de ordem $n \times 1$ pode ser escrito na forma

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2 + \dots + x_n \mathbf{e}_n ,$$

isto é, como uma combinação linear dos vetores $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$.

A coleção de todos os vetores coluna com n elementos numéricos reais é, usualmente, denotada por \mathbf{E}_n .

Exemplo 3.36

Os vetores $\mathbf{v} = [1 \ 2 \ 3]^t$ e $\mathbf{w} = [3 \ 2 \ 1]^t$ são linearmente independentes ?

Solução

A relação $\alpha \mathbf{v} + \beta \mathbf{w} = \mathbf{0}$ equivale ao sistema

$$\begin{aligned} \alpha + 3\beta &= 0 \\ 2\alpha + 2\beta &= 0 \\ 3\alpha + \beta &= 0 , \end{aligned}$$

o qual possui solução não nula se, e somente se, o posto da matriz

$$\begin{bmatrix} 1 & 3 \\ 2 & 2 \\ 3 & 1 \end{bmatrix}$$

for menor que 2. Como o posto desta matriz é exatamente 2, decorre que os vetores dados são linearmente independentes.

Exemplo 3.37

Para um sistema $\mathbf{Ax} = \mathbf{b}$ com nulidade $s > 0$, as soluções não nulas são da forma

$$\mathbf{x} = c_1 \mathbf{h}_1 + c_2 \mathbf{h}_2 + \dots + c_s \mathbf{h}_s = \mathbf{Hc}$$

Os vetores $\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_s$ são linearmente independentes. Pois, $\mathbf{x} = \mathbf{0}$ ocorrerá somente no caso em que $c_1 = c_2 = \dots = c_s = 0$, isto é, a matriz \mathbf{H} de ordem $n \times s$ possui posto s .

Da discussão acima, decorre que

O posto r de uma matriz é igual ao máximo número de colunas de \mathbf{A} linearmente independentes. Se a análise for feita em termos da matriz transposta, então o posto r será igual ao número máximo de linhas linearmente independentes. Também, se \mathbf{U} é a matriz obtida de \mathbf{A} por eliminação gaussiana, então r será o número de colunas ou de linhas linearmente independentes, e, portanto, igual ao número de pivôs não nulos.

Com relação a um sistema não homogêneo $\mathbf{Ax} = \mathbf{b}$, dizer que \mathbf{x} é solução equivale a afirmar que o vetor \mathbf{b} pode ser escrito como *combinação linear* das colunas da matriz \mathbf{A} , isto é

$$\mathbf{b} = \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \dots + \mathbf{a}_n x_n .$$

Se \mathbf{b} não puder ser escrito nessa forma, para nenhuma coleção de escalares x_1, x_2, \dots, x_n , então o sistema $\mathbf{Ax} = \mathbf{b}$ não possuirá solução.

3.7.1 Subespaços Lineares, Bases e Dimensão

Em variadas aplicações, encontram-se coleções de vetores \mathbf{H} , que são descritas por combinações lineares. Por exemplo, a coleção de vetores da forma

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ 0 \\ c \\ 0 \end{bmatrix}$$

onde a , b , e c são parâmetros arbitrários, pode ser escrita

$$\mathbf{x} = a \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}.$$

Esta combinação linear, por sua vez, pode ser escrita na forma matricial $\mathbf{x} = \Phi \mathbf{c}$, onde Φ é a matriz, cujas colunas são os vetores da combinação e \mathbf{c} o vetor dos parâmetros, ou seja,

$$\mathbf{x} = \Phi \mathbf{c} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix}.$$

Uma coleção de vetores \mathbf{H} é denominada *espaço (ou subespaço) linear* quando é fechada com respeito a combinações lineares com elementos da coleção. Mais precisamente, dados os vetores \mathbf{x} , \mathbf{y} em \mathbf{H} e os escalares α , β , o vetor $\alpha \mathbf{x} + \beta \mathbf{y}$ é um elemento de \mathbf{H} .

No exemplo anterior, tem-se que

$$\alpha \mathbf{x} + \beta \mathbf{y} = \alpha \Phi \mathbf{c} + \beta \Phi \mathbf{d} = \Phi \mathbf{e},$$

onde $\mathbf{e} = \alpha \mathbf{c} + \beta \mathbf{d}$ é um elemento da mesma coleção. Portanto, a coleção desses vetores é um espaço.

Em algumas situações, os elementos da matriz Φ podem ser funções. Por exemplo, denote-se por H a coleção formada por todos os polinômios de grau não maior do que r . Qualquer polinômio de grau s

$$P(z) = c_0 + c_1 z + c_2 z^2 + \cdots + c_s z^r$$

pode ser escrito na forma

$$P(z) = [1 \quad z \quad z^2 \quad \cdots \quad z^r] \begin{bmatrix} c_0 \\ c_1 \\ \vdots \\ c_r \end{bmatrix} = \Phi \mathbf{c}.$$

Aqui, Φ é uma matriz linha cujos elementos são funções, logo H é um espaço linear.

Em muitas aplicações, tem-se espaços \mathbf{H} cujos elementos são da forma

$$\mathbf{x} = c_1 \mathbf{h}_1 + c_2 \mathbf{h}_2 + \cdots + c_s \mathbf{h}_k,$$

para determinados vetores \mathbf{h}_j , $j = 1 : k$.
Matricialmente,

$$\mathbf{x} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \cdots \quad \mathbf{h}_k] \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix} = \Phi \mathbf{c}.$$

Um espaço linear com elementos do tipo $\mathbf{x} = \Phi \mathbf{c}$ é dito de *dimensão* r quando a matriz Φ possui exatamente r colunas linearmente independentes. Neste caso, essas colunas são chamadas de *base*. Certamente, $r \leq k$ e r é o posto da matriz Φ , o qual caracteriza a dimensão deste tipo de espaço. A matriz Φ_r , formada pelas r colunas de Φ que são linearmente independentes, é chamada de uma *matriz base* espaço do \mathbf{H} .

Quando é fixada esta matriz base, os elementos do espaço \mathbf{H} podem ser re-escritos como

$$\mathbf{x} = \Phi_r \mathbf{x}_r$$

onde \mathbf{x}_r é um vetor coluna $r \times 1$. Deste modo, opera-se como se forem elementos do espaço \mathbf{E}_r dos vetores coluna $r \times 1$. Mais precisamente, se

$$\mathbf{x} = \Phi_r \mathbf{x}_r, \quad \mathbf{y} = \Phi_r \mathbf{y}_r$$

$$a\mathbf{x} + b\mathbf{y} = \Phi_r a\mathbf{x}_r + b\mathbf{y}_r = \Phi_r \mathbf{z}_r,$$

onde $\mathbf{z}_r = a\mathbf{x}_r + b\mathbf{y}_r$.

Em termos algébricos, isto significa dizer que existirá um *isomorfismo* entre um espaço \mathbf{H} de dimensão r e o espaço \mathbf{E}_r . Ou seja, que as operações são basicamente realizadas em termos dos vetores coordenadas \mathbf{x}_r .

Na prática, o conceito de dimensão refere-se ao número de parâmetros básicos (ou “graus de liberdade”) necessários para caracterizar a coleção. Por exemplo, a coleção de vetores linha

$$\mathbf{x} = a[1 \quad -2] + b[1 \quad 1] + c[2 \quad -1]$$

possui três parâmetros, porém, somente dois parâmetros básicos são necessários. Pois,

$$[2 \quad -1] = [1 \quad -2] + [1 \quad 1]$$

implica

$$\mathbf{x} = (a + c)[1 \quad -2] + (b + c)[1 \quad 1] = c_1[1 \quad 2] + c_2[1 \quad 1].$$

Por outro lado, para estabelecer que a dimensão do espaço \mathbf{H} , formado pelos polinômios de grau r é igual a r , deve-se, retomar o conceito de independência linear, pois, neste exemplo, os elementos base são as funções: $1, z, z^2, \dots, z^r$. Elas serão *linearmente independentes*, se a combinação linear

$$\alpha_0 + \alpha_1 z + \alpha_2 z^2 + \cdots + \alpha_s z^s = 0,$$

(que a rigor, é uma *relação funcional* válida para qualquer valor de z) possuir somente a solução

$$\alpha_0 = \alpha_1 = \cdots = \alpha_s = 0.$$

É um fato conhecido da álgebra elementar que um polinômio se anula identicamente se, e somente se, os seus coeficientes forem nulos. Portanto, como \mathbf{H} é um espaço gerado por r

elementos linearmente independentes, resulta que \mathbf{H} é um espaço de dimensão r .

Exemplo 3.38

A coleção \mathbf{H} de vetores \mathbf{x} da forma

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ 0 \\ c \\ 0 \end{bmatrix}$$

possui dimensão 3?

Solução Tem-se que a combinação linear

$$\mathbf{z} = \alpha\mathbf{x} + \beta\mathbf{y} = \alpha \begin{bmatrix} a \\ b \\ 0 \\ c \\ 0 \end{bmatrix} + \beta \begin{bmatrix} d \\ e \\ 0 \\ f \\ 0 \end{bmatrix} = \begin{bmatrix} \alpha a + \beta d \\ \alpha b + \beta e \\ 0 \\ \alpha c + \beta f \\ 0 \end{bmatrix}$$

é um vetor do mesmo tipo dos vetores dados, isto é, com a terceira e quinta componentes nulas, portanto, \mathbf{H} é um subespaço linear. Por outro lado, qualquer elemento \mathbf{x} de \mathbf{H} pode ser escrito como

$$\mathbf{x} = \begin{bmatrix} a \\ b \\ 0 \\ c \\ 0 \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} + c \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = a\mathbf{u}_1 + b\mathbf{u}_2 + c\mathbf{u}_3 ,$$

onde os vetores

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} , \quad \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \text{e} \quad \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

são linearmente independentes. Assim, \mathbf{H} é um subespaço de dimensão 3.

Exemplo 3.39

Suponha-se que o intervalo $[0, L]$ é particionado nos subintervalos $[0, x_1]$, $[x_1, x_2]$, $[x_2, L]$. Considere-se a coleção \mathbf{L} de funções contínuas que são lineares em cada subintervalo. Então \mathbf{L} é um subespaço das funções contínuas com dimensão 4.

Solução

Tem-se que a combinação linear de esse tipo de funções é uma função do mesmo tipo. Portanto, \mathbf{L} é um subespaço linear. Afirma-se que a dimensão é 4. De fato, em cada subintervalo são necessárias duas constantes para caracterizar a reta que descreve a função, isto é, $y = ax + b$. Sendo três subintervalos, serão necessárias no total seis constantes: a, b, c, d, e e f . Por outro lado, sendo a função contínua, deve-se ter as seguintes relações de continuidade nos extremos dos subintervalos interiores

$$\begin{aligned} ax_1 + b &= cx_1 + b \\ cx_2 + d &= ex_2 + f \end{aligned}$$

Matricialmente,

$$\begin{bmatrix} x_1 & 1 & -x_1 & -1 & 0 & 0 \\ 0 & 0 & x_2 & 1 & -x_2 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Como a matriz dos coeficientes possui posto 2, a nulidade será $s = 6 - 2 = 4$. Portanto, haverá somente 4 constantes independentes para caracterizar os elementos do subespaço \mathbf{L} .

OBSERVAÇÃO

Em cada sub-intervalo $[x_j, x_{j+1}]$, $j = 0 : 2$ com $x_0 = 0$, $x_3 = L$, um elemento $u(x)$ do subespaço \mathbf{L} , pode ser escrito em termos dos valores de u nos extremos do sub-intervalo como

$$u(x) = u(x_j)l_1(x) + u(x_{j+1})l_2(x), \quad x_j \leq x \leq x_{j+1},$$

onde $l_1(x)$, $l_2(x)$ são os elementos (*interpoladores locais*)

$$l_1(x) = \frac{x_{j+1} - x}{x_{j+1} - x_j}$$

$$l_2(x) = \frac{x - x_j}{x_{j+1} - x_j}.$$

Por outro lado, cada elemento de \mathbf{L} , pode ser escrito na forma

$$u(x) = u(x_1)\phi_1(x) + u(x_2)\phi_2(x) + u(x_3)\phi_3(x) + u(x_4)\phi_4(x)$$

onde

$$\phi_1(x) = \begin{cases} \frac{x_1 - x}{x_1 - x_0} & x_0 \leq x \leq x_1 \\ 0 & x_1 \leq x_3 \end{cases}$$

$$\phi_2(x) = \begin{cases} \frac{x - x_1}{x_1 - x_0} & x_0 \leq x \leq x_1 \\ \frac{x_2 - x}{x_2 - x_1} & x_1 \leq x \leq x_2 \\ 0 & x \geq x_2 \end{cases}$$

$$\phi_3(x) = \begin{cases} \frac{x - x_2}{x_3 - x_2} & x_2 \leq x \leq x_3 \\ \frac{x_3 - x}{x_3 - x_2} & x_2 \leq x \leq x_3 \end{cases}$$

$$\phi_4(x) = \begin{cases} 0 & x \leq x_2 \\ \frac{x_3 - x}{x_3 - x_2} & x_2 \leq x \leq x_3 \end{cases}$$

são os elementos de uma base (*global*) em \mathbf{L} .

CAPÍTULO 4

O Método Espectral

Este método é utilizado com sistemas lineares simétricos. Tem como base uma propriedade de *ortogonalidade* existente entre os *autovetores* das matrizes simétricas. A definição destes conceitos será feita a seguir.

4.1 Autovalores e Autovetores

A multiplicação de um vetor não nulo \mathbf{x} por uma matriz \mathbf{A} gera um novo vetor \mathbf{y} , que pode ser visualizado como uma transformação do vetor original \mathbf{x} . Existirá um vetor transformado, proporcional ao vetor original? Ou seja,

$$\mathbf{Ax} = \lambda \mathbf{x}, \quad (4.1)$$

para algum escalar λ ? Em caso afirmativo, diz-se que \mathbf{x} é um *autovetor* da matriz \mathbf{A} associado ao *autovalor* λ .

Os autovalores e autovetores aparecem de maneira natural na procura de soluções do tipo exponencial para variados sistemas de equações, nos quais as incógnitas são funções.

Exemplo 4.40

Determinar para que valores do escalar λ o sistema

$$\frac{du}{dt} - 4u + 3v = 0$$

$$\frac{dv}{dt} + u - v = 0$$

possui soluções não nulas do tipo

$$u = ae^{\lambda t}, \quad v = be^{\lambda t}$$

Solução

Substituindo no sistema, $u = ae^{\lambda t}$, $v = be^{\lambda t}$, decorre

$$(\lambda a - 4a + 3b)e^{\lambda t} = 0$$

$$(\lambda b + a - b)e^{\lambda t} = 0.$$

Simplificando, obtém-se o sistema

$$(\lambda - 4)a + 3b = 0$$

$$-a + (\lambda - 1)b = 0$$

Matricialmente,

$$\begin{bmatrix} \lambda - 4 & 3 \\ 1 & \lambda - 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

que é do tipo $(\lambda \mathbf{I} - \mathbf{A})\mathbf{x} = \mathbf{0}$ ou, equivalentemente, $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$.

O caso da solução $a = b = 0$ não é de interesse, pois fornece as soluções nulas $u = v = 0$. Para que este sistema homogêneo possua uma solução não nula, o determinante deve ser nulo. Assim,

$$\det \begin{bmatrix} \lambda - 4 & 3 \\ 1 & \lambda - 1 \end{bmatrix} = (\lambda - 4)(\lambda - 1) - 3 = \lambda^2 - 5\lambda + 1 = 0.$$

Decorre que o sistema dado possui soluções com expoentes $\lambda_1 = \frac{5+\sqrt{21}}{2}$ e $\lambda_2 = \frac{5-\sqrt{21}}{2}$.

Exemplo 4.41

Determinar os valores do escalar λ , para os quais o sistema

$$u_{k+1} + 6u_k + 3v_k = 0$$

$$v_{k+1} + u_k - 2v_k = 0$$

possui soluções não nulas do tipo

$$u_k = a\lambda^k, \quad v_k = b\lambda^k.$$

Solução

Substituindo u e v no sistema dado, decorre

$$(\lambda a - 6a + 3b)\lambda^k = 0$$

$$(\lambda b + a - 2b)\lambda^k = 0,$$

e simplificando,

$$(\lambda a - 6)a - 3b = 0$$

$$a + (\lambda - 2)b = 0$$

o qual é do tipo $\lambda \mathbf{I} - \mathbf{A}\mathbf{x} = \mathbf{0}$ ou, equivalentemente, $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$. Matricialmente,

$$\begin{bmatrix} \lambda - 6 & 3 \\ 1 & \lambda - 2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

O caso $a = b = 0$ não é de interesse, pois fornece soluções nulas. Para que este sistema homogêneo possua uma solução não nula, o determinante deve ser nulo. Assim,

$$\det \begin{bmatrix} \lambda - 6 & 3 \\ 1 & \lambda - 2 \end{bmatrix} = (\lambda - 6)(\lambda - 2) - 3 = \lambda^2 - 8\lambda + 9 = 0.$$

Decorre que o sistema dado de equações em diferenças possui soluções do tipo $c\lambda_k$ com bases exponenciais $\lambda_1 = 4 + \sqrt{7}$ e $\lambda_2 = 4 - \sqrt{7}$.

Exemplo 4.42

Para a transformação linear

$$x' = x + y$$

$$y' = 2y$$

descrita pela matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix},$$

tem-se que

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{e} \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

são autovetores de \mathbf{A} associados aos autovalores $\lambda_1 = 1$ e $\lambda_2 = 2$, respectivamente.

4.1.1 Polinômio característico

Para determinar os autovalores e autovetores de uma matriz deve-se observar que a relação

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}, \quad \text{com} \quad \mathbf{x} \neq \mathbf{0}, \quad (4.2)$$

pode ser escrita na forma de um sistema linear homogêneo

$$[\lambda\mathbf{I} - \mathbf{A}]\mathbf{x} = \mathbf{0}, \quad \mathbf{x} \neq \mathbf{0} \quad (4.3)$$

Para que este sistema possua soluções não nulas, seu determinante

$$\det[\lambda\mathbf{I} - \mathbf{A}] = \begin{bmatrix} \lambda - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & \lambda - a_{22} & \cdots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda - a_{nn} \end{bmatrix} \quad (4.4)$$

deve ser nulo. Este determinante é um polinômio de grau n em λ , denominado *polinômio característico* da matriz \mathbf{A} e denotado por $P(\lambda)$. Assim,

$$P(\lambda) = \det[\lambda\mathbf{I} - \mathbf{A}] = \lambda^n + b_1\lambda^{n-1} + b_2\lambda^{n-2} + \cdots + b_{n-1}\lambda + b_n, \quad (4.5)$$

e a equação

$$P(\lambda) = \det[\lambda\mathbf{I} - \mathbf{A}] = 0 \quad (4.6)$$

denomina-se a *equação característica* da matriz \mathbf{A} . Como esta equação possui n raízes, decorre que *os autovalores de uma matriz quadrada \mathbf{A} de ordem n são as n raízes do seu polinômio característico*. Se uma raiz ocorrer precisamente k vezes, o autovalor será dito de multiplicidade k .

A fatorização do polinômio característico

$$P(\lambda) = \det[\lambda\mathbf{I} - \mathbf{A}] = (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n),$$

em termos de suas raízes permite determinar se uma matriz é singular ou não. Pois, fazendo $\lambda = 0$, e lembrando que $\det(-\mathbf{A}) = (-1)^n \det(\mathbf{A})$, tem-se

$$\det(\mathbf{A}) = \lambda_1 \lambda_2 \cdots \lambda_n,$$

e decorre a seguinte caracterização:

Uma matriz A é dita singular se, e somente se, ao menos um dos seus autovalores for nulo; é não-singular, quando todos seus autovalores forem não nulos.

Em particular, quando uma matriz for diagonal ou triangular, é fácil concluir se é singular: basta que algum elemento da diagonal seja zero. De fato, o polinômio característico é, para este tipo de matrizes, o produto dos elementos da diagonal, isto é, dos autovalores.

Exemplo 4.43

Determinar os autovalores da matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & -1 & 1 \\ 0 & 0 & 1 \end{bmatrix} .$$

Solução

O polinômio característico de \mathbf{A} é

$$\det \begin{bmatrix} \lambda - 1 & 0 & -2 \\ 0 & \lambda + 1 & -1 \\ 0 & 0 & \lambda - 1 \end{bmatrix} = (\lambda - 1)(\lambda + 1)(\lambda - 1) = \lambda^3 - 2\lambda + 1.$$

Suas raízes são 1 e -1, esta última com multiplicidade 2.

Embora uma matriz $n \times n$ não necessariamente terá n autovalores distintos, *poderá ter* n autovetores linearmente independentes.

Exemplo 4.44

Determinar os autovetores associados aos autovalores da matriz

$$\mathbf{A} = \begin{bmatrix} 5 & 2 & 2 \\ 3 & 6 & 3 \\ 6 & 6 & 9 \end{bmatrix} .$$

Solução

Para a matriz dada,

$$\lambda \mathbf{I} - \mathbf{A} = \left[\lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} - \begin{bmatrix} 5 & 2 & 2 \\ 3 & 6 & 3 \\ 6 & 6 & 9 \end{bmatrix} \right] = \begin{bmatrix} \lambda - 5 & -2 & -2 \\ -3 & \lambda - 6 & -3 \\ -6 & -6 & \lambda - 9 \end{bmatrix} .$$

O determinante, desta última matriz, é dado por

$$\lambda^3 - 20\lambda^2 + 93\lambda - 126 = (\lambda - 3)^2(\lambda - 14) .$$

A equação característica de \mathbf{A} é $(\lambda - 3)^2(\lambda - 14) = 0$, a qual tem como soluções os autovalores $\lambda = 3$ de multiplicidade 2, e $\lambda = 14$ de multiplicidade um. Os autovetores, correspondentes ao autovalor $\lambda = 3$, são obtidos resolvendo o sistema singular $(3\mathbf{I} - \mathbf{A})\mathbf{x} = 0$, ou, equivalentemente

$$(\mathbf{A} - 3\mathbf{I})\mathbf{x} = 0 ,$$

ou seja,

$$\begin{bmatrix} 2 & 2 & 2 \\ 3 & 3 & 3 \\ 6 & 6 & 6 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Efetuada a eliminação, decorre o sistema

$$\begin{bmatrix} 2 & 2 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

ou, simplesmente, a equação

$$2x_1 + 2x_2 + 2x_3 = 0 ,$$

que possui a variável básica x_1 e as variáveis livres $x_2 = a$ e $x_3 = b$. Assim, o vetor coluna

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -a - b \\ a \\ b \end{bmatrix} ,$$

onde a e b são constantes arbitrárias, ambas não nulas simultaneamente, é a forma geral dos autovetores associados ao autovalor $\lambda = 3$. Mais precisamente,

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = a \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Observe-se que, em particular, cada coluna é um autovetor e que eles são linearmente independentes.

Quando $\lambda = 14$, se obtém $(\mathbf{A} - 14\mathbf{I})\mathbf{x} = 0$, isto é,

$$\begin{bmatrix} -9 & 2 & 2 \\ 3 & -8 & 3 \\ 6 & 6 & -5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Com as operações sobre as linhas, decorre

$$\begin{bmatrix} -9 & 2 & 2 \\ 0 & -22/3 & 11/3 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

As variáveis básicas são x_1 e x_2 e a variável livre é $x_3 = c$. Logo, o autovetor associado ao autovalor $\lambda = 14$ é dado por

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = c \begin{bmatrix} 1/3 \\ 1/2 \\ 1 \end{bmatrix} ,$$

sendo c uma constante arbitrária não nula.

Resumindo, para o autovalor duplo $\lambda = 3$ tem-se dois autovetores associados

$$\mathbf{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} , \quad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} ,$$

que são linearmente independentes. Para o autovalor simples tem-se o autovetor associado

$$\mathbf{v}_3 = \begin{bmatrix} 1/3 \\ 1/2 \\ 1 \end{bmatrix} .$$

Podem ser verificados que estes autovetores são linearmente independentes.

Exemplo 4.45

Determinar os autovetores associados aos autovalores da matriz simétrica

$$\mathbf{A} = \begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix}.$$

Solução

Para a matriz dada, a equação característica é o determinante da matriz

$$\begin{bmatrix} \lambda - 7 & 2 & -1 \\ 2 & \lambda - 10 & 23 \\ -1 & 2 & \lambda - 7 \end{bmatrix}.$$

Assim,

$$\lambda^3 - 24\lambda^2 - 180\lambda - 432 = 0$$

e segue que os autovalores são 6, 6 e 12. Para $\lambda = 6$, tem-se

$$\begin{bmatrix} 1 & -2 & -1 \\ 2 & -4 & 2 \\ -1 & 2 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Efetuando-se a eliminação, decorre a equação

$$x_1 - 2x_2 + x_3 = 0,$$

que possui a variável básica x_1 e as variáveis livres $x_2 = d$ e $x_3 = e$. Assim, o vetor coluna

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2d - e \\ d \\ e \end{bmatrix},$$

onde d e e são constantes arbitrárias, ambas não nulas simultaneamente, é a forma geral do autovetor associado ao autovalor $\lambda = 6$. Em particular, pode-se escolher $v_1 = [1 \ 0 \ -1]^t$ e $v_2 = [1 \ 1 \ 1]^t$. Quando $\lambda = 12$, se obtém que $v_3 = [1 \ -2 \ 1]^t$ é um autovetor associado ao autovalor $\lambda = 12$.

Matrizes Defeituosas

Em geral, uma matriz $n \times n$ não possui n autovetores linearmente independentes. A matriz \mathbf{A} de ordem $n \times n$ é dita *não defeituosa* quando possui n autovetores linearmente independentes. Caso contrário é dita *defeituosa*. Por exemplo, matrizes cujos autovalores são todos distintos são não defeituosas. Entretanto, matrizes com autovalores repetidos podem ser ou não ser defeituosas.

Exemplo 4.46

Obter os autovalores e autovetores da matriz

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}.$$

Solução

Tem-se o polinômio característico

$$p(\lambda) = \det[\lambda \mathbf{I} - \mathbf{A}] = \begin{vmatrix} \lambda & -2 & 0 \\ 0 & \lambda & -2 \\ 0 & 0 & \lambda - 1 \end{vmatrix} = \lambda^2(\lambda - 1).$$

Os autovalores são $\lambda_1 = \lambda_2 = 0$, $\lambda_3 = 1$. Para o autovalor duplo 0, a equação dos autovetores correspondentes é

$$\begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Efetuando a eliminação, vem

$$\begin{bmatrix} 0 & -2 & 0 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Decorre $-2x_2 = 0$, $-2x_3 = 0$. Assim, para $\lambda_1 = \lambda_2 = 0$, o autovetor correspondente é

$$\mathbf{v}_1 = a \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

onde a é uma constante arbitrária. Para $\lambda_3 = 1$, tem-se o sistema

$$\begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & -2 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Decorre $x_1 = 2x_2$, $x_2 = 2b$, $x_3 = b$, b uma constante. Assim,

$$\mathbf{v}_2 = b \begin{bmatrix} 4 \\ 2 \\ 1 \end{bmatrix}.$$

A matriz dada é 3×3 , porém somente há dois autovetores linearmente independentes. Portanto, a matriz é defeituosa.

Em geral, uma matriz é defeituosa quando possuir um autovalor λ de multiplicidade m tal que a nulidade s do sistema $[\lambda \mathbf{I} - \mathbf{A}]\mathbf{v} = \mathbf{0}$ é menor que m . Matrizes não defeituosas são aquelas em que a multiplicidade de cada autovalor λ coincide com a nulidade do correspondente sistema $[\lambda \mathbf{I} - \mathbf{A}]\mathbf{v} = \mathbf{0}$.

Autovalores e Autovetores Complexos

Deve-se salientar que se \mathbf{A} for uma matriz com elementos reais, então todos os coeficientes do polinômio característico serão reais, entretanto, algumas de suas raízes poderão ser complexas. De um resultado bem conhecido da álgebra elementar, as raízes complexas aparecem como complexas conjugadas. Os autovetores correspondentes aos autovalores reais terão seus elementos reais. Porém, os autovetores associados aos autovalores complexos podem ter elementos complexos.

Exemplo 4.47

A matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

possui os autovalores $\lambda_1 = 2 + i$, $\lambda_2 = 2 - i$, onde i é a unidade imaginária, e $\lambda_3 = 1$. Para $\lambda_1 = 2 + i$, tem-se

$$[\lambda_1 \mathbf{I} - \mathbf{A}] \mathbf{x} = \begin{bmatrix} i & -1 & 0 \\ 1 & i & 0 \\ 0 & 0 & 1+i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Efetuada a eliminação

$$\begin{bmatrix} i & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1+i \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} .$$

Decorre $x_3 = 0, x_2 = a, ix_1 = a$. Como $i \cdot i = -1$, obtém-se $x_1 = -ai$. Assim

$$\mathbf{v}_1 = a \begin{bmatrix} -i \\ 1 \\ 0 \end{bmatrix} .$$

Efetuada os cálculos para $\lambda_2 = 2 - i$, decorre

$$\mathbf{v}_2 = b \begin{bmatrix} i \\ 1 \\ 0 \end{bmatrix} .$$

Escolhendo $a = b = 1$, segue que *autovetores de matriz real correspondentes a autovalores complexos conjugados são complexos conjugados*. A rigor, para uma matriz real $\bar{\mathbf{A}} = \mathbf{A}$. Então,

$$\mathbf{A} \mathbf{v} = \lambda \mathbf{v} \quad \rightarrow \quad \mathbf{A} \bar{\mathbf{v}} = \bar{\lambda} \bar{\mathbf{v}} .$$

4.1.2 Identidade de Cayley-Hamilton

Um resultado na álgebra das matrizes, sem contrapartida na álgebra elementar, foi estabelecido por Hamilton para uma classe especial de matrizes e, após, enunciado na sua forma geral por Cayley que o considerou "evidente", omitindo sua demonstração. Este resultado tem profundas raízes na existência de divisores de zero na álgebra das matrizes e, de certo modo, "compensa" a falta de comutatividade com as matrizes.

Qualquer matriz quadrada \mathbf{A} , com elementos numéricos, satisfaz sua equação característica, isto é, $P(\mathbf{A}) = \mathbf{0}$, ou, na forma expandida

$$P(\mathbf{A}) = \mathbf{A}^n + b_1 \mathbf{A}^{n-1} + \dots + b_{n-1} \mathbf{A} + b_n \mathbf{I} = \mathbf{0} . \quad (4.7)$$

A prova deste importante resultado é mostrada a seguir.

Utilizando-se a identidade de Cramer com a matriz $(\lambda \mathbf{I} - \mathbf{A})$, vem

$$(\lambda \mathbf{I} - \mathbf{A}) \text{adj}(\lambda \mathbf{I} - \mathbf{A}) = \det(\lambda \mathbf{I} - \mathbf{A}) \mathbf{I} . \quad (4.8)$$

Os elementos de $\text{adj}(\lambda \mathbf{I} - \mathbf{A})$, para \mathbf{A} de ordem $n \times n$, são todos polinômios em λ de grau $n - 1$. Assim, pode-se escrever

$$\text{adj}(\lambda \mathbf{I} - \mathbf{A}) = \lambda^{n-1} \mathbf{C}_1 + \lambda^{n-2} \mathbf{C}_2 + \cdots + \lambda \mathbf{C}_{n-1} + \mathbf{C}_n, \quad (4.9)$$

onde cada \mathbf{C}_j é uma matriz de ordem $n \times n$. Multiplicando ambos os membros da expressão por $(\lambda \mathbf{I} - \mathbf{A})$ e igualando os coeficientes da mesma potência em λ , segue da identidade de Cramer

$$\begin{array}{rcl} \text{para } \lambda^{n-1} & : & \mathbf{C}_1 = \mathbf{I} \\ \text{para } \lambda^{n-2} & : & \mathbf{C}_2 - \mathbf{A}\mathbf{C}_1 = b_1 \mathbf{I} \\ & \vdots & \vdots \\ \text{para } \lambda & : & \mathbf{C}_n - \mathbf{A}\mathbf{C}_{n-1} = b_{n-1} \mathbf{I} \\ \text{para } 1 & : & -\mathbf{A}\mathbf{C}_n = b_n \mathbf{I}. \end{array}$$

Multiplicando, pela esquerda, a primeira equação por \mathbf{A}^n , a segunda por \mathbf{A}^{n-1} , etc., e adicionando os produtos resultantes, vem

$$\begin{aligned} \mathbf{A}^n \mathbf{C}_1 + \mathbf{A}^{n-1} (\mathbf{C}_2 - \mathbf{A}\mathbf{C}_1) + \cdots + \mathbf{A} (\mathbf{C}_n - \mathbf{A}\mathbf{C}_{n-1}) &= \\ = [\mathbf{A}^n + b_1 \mathbf{A}^{n-1} + \cdots + b_{n-1} \mathbf{A} + b_n \mathbf{I}] \mathbf{I}. \end{aligned} \quad (4.10)$$

Agora, como todos os termos do membro da esquerda se cancelam, tem-se

$$0 = \mathbf{P}(\mathbf{A}) \mathbf{I}, \quad (4.11)$$

isto é,

$$\mathbf{P}(\mathbf{A}) = 0. \quad (4.12)$$

A seguir, serão consideradas duas consequências da identidade de Cayley-Hamilton. Da relação $\mathbf{P}(\mathbf{A}) = 0$, vem

$$\mathbf{A}^n = -b_1 \mathbf{A}^{n-1} - b_2 \mathbf{A}^{n-2} - \cdots - b_{n-1} \mathbf{A} - b_n \mathbf{I} \quad (4.13)$$

e multiplicando por \mathbf{A} , segue

$$\begin{aligned} \mathbf{A}^{n+1} &= -b_1 \mathbf{A}^n - b_2 \mathbf{A}^{n-1} - \cdots - b_{n-1} \mathbf{A}^2 - b_n \mathbf{I} \\ &= -b_1 (-b_1 \mathbf{A}^{n-1} - \cdots - b_{n-1} \mathbf{A} - b_n \mathbf{I}) - b_2 \mathbf{A}^{n-1} - \cdots - b_n \mathbf{I}. \end{aligned} \quad (4.14)$$

Agrupando as potências em \mathbf{A} , observa-se que \mathbf{A}^{n+1} é uma combinação em $\mathbf{I}, \mathbf{A}, \dots, \mathbf{A}^{n-1}$. Em geral, por indução, segue a seguinte propriedade

Qualquer potência \mathbf{A}^k de uma matriz \mathbf{A} de ordem n pode ser expressa como uma combinação linear das potências $\mathbf{I}, \mathbf{A}, \mathbf{A}^2, \dots, \mathbf{A}^{n-1}$.

Esta propriedade, estende-se para potências negativas de uma matriz não singular. Pois, multiplicando $\mathbf{P}(\mathbf{A}) = 0$, à esquerda por \mathbf{A}^{-1} , decorre

$$\mathbf{A}^{-1} = \frac{-\mathbf{A}^{n-1} - b_1 \mathbf{A}^{n-2} - \cdots - b_{n-2} \mathbf{A} - b_{n-1} \mathbf{I}}{b_{n-1}}, \quad (4.15)$$

uma vez que $b_n = \det[-\mathbf{A}]$ é diferente de zero.

Este procedimento para calcular a inversa de uma matriz, tem um custo computacional muito elevado, porém pode ser utilizado com matrizes de pequeno porte, ou servir como ponto de partida para a obtenção de métodos mais eficientes.

4.2 Vetores Ortogonais

Os conceitos euclidianos de comprimento e de ortogonalidade, relacionados pelo Teorema de Pitágoras, são estendidos ao caso de vetores. Para um vetor \mathbf{x} , cujas componentes x_i são números reais, define-se

$$\|\mathbf{x}\| = [x_1^2 + x_2^2 + \cdots + x_n^2]^{1/2}, \quad (4.16)$$

como o comprimento ou norma (euclidiana) do vetor \mathbf{x} .

Dois vetores \mathbf{x} e \mathbf{y} são visualmente ortogonais

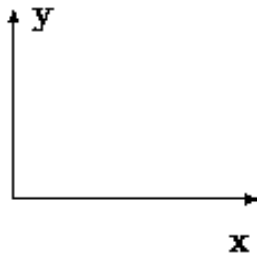


Figura 4.1 – Vetores ortogonais

desde que formem um triângulo retângulo. O teorema de Pitágoras, neste caso, forneceria a relação

$$\|\mathbf{x} - \mathbf{y}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2, \quad (4.17)$$

ou, equivalentemente,

$$\begin{aligned} (x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2 &= \\ &= (x_1^2 + \cdots + x_n^2) + (y_1^2 + \cdots + y_n^2). \end{aligned} \quad (4.18)$$

Tal relação é válida, unicamente, se

$$(x_1y_1 + x_2y_2 + \cdots + x_ny_n) = 0. \quad (4.19)$$

Observe-se que esta grandeza é idêntica a $\mathbf{x}^t\mathbf{y}$, isto é, ao produto da multiplicação de um vetor linha $1 \times n$ (o vetor \mathbf{x}^t) por uma matriz $n \times 1$ (o vetor coluna \mathbf{y})

$$\mathbf{x}^t\mathbf{y} = [x_1 \quad \cdots \quad x_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = x_1y_1 + \cdots + x_ny_n \quad (4.20)$$

ou, utilizando a notação para a soma,

$$\mathbf{x}^t\mathbf{y} = \sum_{i=1}^n x_iy_i. \quad (4.21)$$

Esta combinação aparece em toda discussão da geometria do espaço dos vetores coluna, cujos elementos são números reais e, às vezes, é chamada de produto escalar de dois vetores, denotado por $\langle \mathbf{u}, \mathbf{v} \rangle$ ou $\mathbf{u} \cdot \mathbf{v}$. Aqui, optar-se-á por denominá-la **produto interno** e por preservar a notação $\mathbf{x}^t\mathbf{y}$. Um **produto interno**, para vetores com elementos reais, será sempre caracterizado pelas seguintes propriedades básicas:

1. $\mathbf{x}^t(\mathbf{y} + \mathbf{z}) = \mathbf{x}^t\mathbf{y} + \mathbf{x}^t\mathbf{z}$;
2. $\mathbf{x}^t(\beta\mathbf{y}) = \beta\mathbf{x}^t\mathbf{y}$, β escalar real;
3. $\mathbf{x}^t\mathbf{y} = \mathbf{y}^t\mathbf{x}$;
4. $\mathbf{x}^t\mathbf{x} \geq 0$, $\mathbf{x}^t\mathbf{x} = 0$, somente, se $\mathbf{x} = 0$;
5. $|\mathbf{x}^t\mathbf{y}| \leq \|\mathbf{x}\| \|\mathbf{y}\|$.

A última propriedade é conhecida como a *desigualdade de Cauchy-Schwarz* e pode ser obtida a partir das 4 primeiras propriedades. Esta desigualdade permite definir ângulos entre vetores não nulos. Pois, elevando ao quadrado ambos membros da desigualdade e dividindo pelo termo à esquerda, tem-se

$$\left[\frac{|\mathbf{x}^t\mathbf{y}|}{\|\mathbf{x}\| \|\mathbf{y}\|} \right]^2 \leq 1 \quad (4.22)$$

ou, equivalentemente,

$$-1 \leq \frac{\mathbf{x}^t\mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \leq 1 . \quad (4.23)$$

Se α for um ângulo cuja medida em radianos varia entre 0 e π , então $\cos \alpha$ assume uma única cada valor entre -1 e 1 inclusive. Portanto,

$$\cos \alpha = \frac{\mathbf{x}^t\mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|}, \quad (4.24)$$

para algum ângulo $0 \leq \alpha \leq \pi$. Define-se α como sendo o *ângulo entre x e y* . Esta definição coincide com as fórmulas usuais da geometria analítica para o cosseno do ângulo entre dois vetores no plano ou espaço.

As noções de comprimento e de ortogonalidade estão associadas ao produto interno. Dir-se-á, que os vetores \mathbf{x} e \mathbf{y} são *ortogonais*, quando

$$\mathbf{x}^t\mathbf{y} = 0. \quad (4.25)$$

Por outro lado, a norma euclidiana de um vetor \mathbf{x} poderá ser escrita

$$\|\mathbf{x}\| = [\mathbf{x}^t\mathbf{x}]^{1/2}. \quad (4.26)$$

Dois vetores \mathbf{x} e \mathbf{y} são ditos *ortonormais*, quando são ortogonais e possuem norma unitária. Certamente, o ângulo entre dois vetores ortogonais é $\pi/2$.

Exemplo 4.48

Os seguintes vetores são mutuamente ortogonais:

1.

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \mathbf{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \dots \quad \mathbf{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix};$$

2.

$$\mathbf{v}_1 = \begin{bmatrix} \text{sen} \frac{\pi}{n+1} \\ \text{sen} \frac{2\pi}{n+1} \\ \vdots \\ \text{sen} \frac{n\pi}{n+1} \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} \text{sen} \frac{2\pi}{n+1} \\ \text{sen} \frac{4\pi}{n+1} \\ \vdots \\ \text{sen} \frac{2n\pi}{n+1} \end{bmatrix}, \quad \dots \quad \mathbf{v}_k = \begin{bmatrix} \text{sen} \frac{k\pi}{n+1} \\ \text{sen} \frac{2k\pi}{n+1} \\ \vdots \\ \text{sen} \frac{kn\pi}{n+1} \end{bmatrix}.$$

Solução

1. É imediato pela definição.

2. Tem-se que

$$\begin{aligned} \mathbf{v}_i^t \mathbf{v}_j &= \sum_{q=1}^n \text{sen} \frac{qin}{(n+1)} \text{sen} \frac{qjn}{(n+1)} \\ &= \frac{\sum_{q=1}^n \cos \frac{q(i-j)n}{(n+1)} - \sum_{q=1}^n \cos \frac{q(i+j)n}{(n+1)}}{2} \\ &= \frac{\text{sen} \frac{\frac{n}{2}(i-j)\pi}{(n+1)} \cos \frac{(n+1)(i-j)\pi}{(n+1)}}{\text{sen} \frac{(i-j)\pi}{(n+1)}} - \\ &\quad \frac{\text{sen} \frac{(n/2)(i+j)\pi}{(n+1)} \cos \frac{(n+1)(i+j)\pi}{(n+1)}}{\text{sen} \frac{(i+j)\pi}{(n+1)}} \\ &= 0 - 0 = 0. \end{aligned}$$

A importância do conceito da ortogonalidade se torna evidente na seguinte propriedade:

Sejam $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k$ vetores coluna $n \times 1$ mutuamente ortogonais, ou seja,

$$(\mathbf{v}_i)^t \mathbf{v}_j = \mathbf{0}, \quad \text{para } i \neq j. \quad (4.27)$$

Então, para um vetor particular \mathbf{x} com a representação

$$\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_k \mathbf{v}_k, \quad (4.28)$$

tem-se os seguintes valores para os coeficientes

$$c_i = \frac{\mathbf{x}^t \mathbf{v}_i}{\mathbf{v}_i^t \mathbf{v}_i} \quad (4.29)$$

e, além disso,

$$\|\mathbf{x}\|^2 = \mathbf{x}^t \mathbf{x} = c_1^2 \|\mathbf{v}_1\|^2 + \dots + c_k^2 \|\mathbf{v}_k\|^2. \quad (4.30)$$

Isto é uma decorrência do produto interno de \mathbf{x} com cada vetor \mathbf{v}_i

$$\mathbf{v}_i^t \mathbf{x} = c_1 \mathbf{v}_i^t \mathbf{v}_1 + \dots + c_k \mathbf{v}_i^t \mathbf{v}_k \quad (4.31)$$

e de utilizar a hipótese de ortogonalidade.

Desta propriedade, decorre, em particular, que *vetores ortogonais são linearmente independentes*. Pois, a combinação

$$\mathbf{0} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \cdots + c_k \mathbf{v}_k \quad (4.32)$$

possui os coeficientes

$$c_i = \frac{\mathbf{0}^t \mathbf{v}_i}{\mathbf{v}_i^t \mathbf{v}_i} = 0 \quad (4.33)$$

4.2.1 Processo de Ortogonalização

Dados p vetores arbitrários $m \times 1$, o posto r da matriz cujas colunas são esses vetores permite selecionar r deles que são linearmente independentes. A seguir é dado um procedimento que permite gerar r vetores mutuamente ortogonais.

Suponha-se que $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_r$ são r vetores de ordem m que são linearmente independentes. Defina-se

$$\begin{aligned} \mathbf{w}_1 &= \mathbf{v}_1 \\ \mathbf{w}_2 &= \mathbf{v}_2 - \frac{\mathbf{w}_1^t \mathbf{v}_2}{\mathbf{w}_1^t \mathbf{w}_1} \mathbf{w}_1 \\ \mathbf{w}_3 &= \mathbf{v}_3 - \frac{\mathbf{w}_2^t \mathbf{v}_3}{\mathbf{w}_2^t \mathbf{w}_2} \mathbf{w}_2 - \frac{\mathbf{w}_1^t \mathbf{v}_3}{\mathbf{w}_1^t \mathbf{w}_1} \mathbf{w}_1 \\ &\vdots \\ \mathbf{w}_k &= \mathbf{v}_k - \frac{\mathbf{w}_{k-1}^t \mathbf{v}_k}{\mathbf{w}_{k-1}^t \mathbf{w}_{k-1}} \mathbf{w}_{k-1} - \cdots - \frac{\mathbf{w}_1^t \mathbf{v}_k}{\mathbf{w}_1^t \mathbf{w}_1} \mathbf{w}_1 \end{aligned} \quad (4.34)$$

Então, os vetores

$$\mathbf{q}_k = \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|}, \quad k = 1 : r \quad (4.35)$$

são mutuamente ortonormais.

A prova deste resultado, conhecido como o *processo de ortogonalização de Gram-Schmidt*, é por construção:

1. Defina-se $\mathbf{w}_1 = \mathbf{v}_1$;
2. Considere-se $\mathbf{w}_2 = \mathbf{v}_2 + \alpha \mathbf{w}_1$. Como \mathbf{w}_2 deve ser ortogonal a \mathbf{w}_1

$$\mathbf{w}_1^t \mathbf{w}_2 = \mathbf{w}_1^t \mathbf{v}_2 + \alpha \mathbf{w}_1^t \mathbf{w}_1 = 0$$

e

$$\alpha = -\frac{\mathbf{w}_1^t \mathbf{v}_2}{\mathbf{w}_1^t \mathbf{w}_1}.$$

Assim,

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{w}_1^t \mathbf{v}_2}{\mathbf{w}_1^t \mathbf{w}_1} \mathbf{w}_1.$$

3. Considere-se $\mathbf{w}_3 = \mathbf{v}_3 + \alpha \mathbf{w}_2 + \beta \mathbf{w}_1$. Como $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3$ devem ser mutuamente ortogonais

$$\mathbf{w}_1^t \mathbf{w}_3 = \mathbf{w}_1^t \mathbf{v}_3 + \alpha \mathbf{w}_1^t \mathbf{w}_2 + \beta \mathbf{w}_1^t \mathbf{w}_1 = \mathbf{w}_1^t \mathbf{v}_3 + \beta \mathbf{w}_1^t \mathbf{w}_1 = 0$$

$$\mathbf{w}_2^t \mathbf{w}_3 = \mathbf{w}_2^t \mathbf{v}_3 + \alpha \mathbf{w}_2^t \mathbf{w}_2 + \beta \mathbf{w}_2^t \mathbf{w}_1 = \mathbf{w}_2^t \mathbf{v}_3 + \alpha \mathbf{w}_2^t \mathbf{w}_2 = 0.$$

Então

$$\alpha = -\frac{\mathbf{w}_2^t \mathbf{v}_3}{\mathbf{w}_2^t \mathbf{w}_2},$$

$$\beta = -\frac{\mathbf{w}_1^t \mathbf{v}_3}{\mathbf{w}_1^t \mathbf{w}_1}$$

e

$$\mathbf{w}_3 = \mathbf{v}_3 - \frac{\mathbf{w}_2^t \mathbf{v}_3}{\mathbf{w}_2^t \mathbf{w}_2} \mathbf{w}_2 - \frac{\mathbf{w}_1^t \mathbf{v}_3}{\mathbf{w}_1^t \mathbf{w}_1} \mathbf{w}_1.$$

4. Continua-se o processo até a obtenção de \mathbf{w}_r .

Utilizando o fato que $\|\mathbf{w}_j\|^2 = \mathbf{w}_j^t \mathbf{w}_j$, tem-se os seguintes vetores \mathbf{q}_k ortonormais

$$\mathbf{q}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}$$

$$\mathbf{w}_k = \mathbf{v}_k - \sum_{i=1}^{k-1} \mathbf{q}_i^t \mathbf{v}_k \mathbf{q}_i, \quad k = 2 : n$$

$$\mathbf{q}_k = \frac{\mathbf{w}_k}{\|\mathbf{w}_k\|}, \quad k = 2 : n.$$

Na prática, este resultado significa o seguinte. É conhecido que os elementos $\mathbf{x} = \Phi \mathbf{c}$, Φ uma matriz $m \times n$ com posto r e \mathbf{c} um vetor $n \times 1$, de espaço linear $d\mathbf{H}$, podem ser escritos $\mathbf{x} = \Phi_r \mathbf{x}_r$ com Φ_r uma matriz $m \times r$ com posto r e \mathbf{x}_r um vetor $r \times 1$. Com o processo de Gram-Schmidt, os vetores do espaço \mathbf{H} , poderá ser escritos

$$\mathbf{x} = \Psi \mathbf{c} \tag{4.36}$$

onde Ψ é uma matriz cujas colunas são mutuamente ortonormais.

Exemplo 4.49

Ortonormalizar os vetores

$$\mathbf{v}_1 = [1 \ 1 \ 1]^t, \quad \mathbf{v}_2 = [1 \ -2 \ 1]^t \text{ e } \mathbf{v}_3 = [1 \ 2 \ 3]^t.$$

Solução

$$\mathbf{w}_1 = \mathbf{v}_1 = [1 \ 1 \ 1]^t,$$

$$\mathbf{w}_2 = \mathbf{v}_2 - \frac{\mathbf{q}_1^t \mathbf{v}_2}{\mathbf{q}_1^t \mathbf{q}_1} \mathbf{q}_1 = [1 \ -2 \ 1]^t - \frac{0}{3} \mathbf{v}_1 = [1 \ -2 \ 1]^t,$$

$$\begin{aligned} \mathbf{w}_3 &= \mathbf{v}_3 - \frac{\mathbf{q}_2^t \mathbf{v}_3}{\mathbf{q}_2^t \mathbf{q}_2} \mathbf{q}_2 - \frac{\mathbf{q}_1^t \mathbf{v}_3}{\mathbf{q}_1^t \mathbf{q}_1} \mathbf{q}_1 \\ &= [1 \ 2 \ 3]^t - \frac{0}{6} \mathbf{q}_2 - \frac{6}{3} [1 \ 1 \ 1]^t = [-1 \ 0 \ 1]^t. \end{aligned}$$

Os vetores

$$\mathbf{q}_1 = \frac{\mathbf{q}_1}{\|\mathbf{q}_1\|} = \left[\frac{1}{\sqrt{3}} \ \frac{1}{\sqrt{3}} \ \frac{1}{\sqrt{3}} \right]^t,$$

$$\mathbf{q}_2 = \frac{\mathbf{q}_2}{\|\mathbf{q}_2\|} = \left[\frac{1}{\sqrt{6}} \ \frac{-2}{\sqrt{6}} \ \frac{1}{\sqrt{6}} \right]^t$$

e

$$\mathbf{q}_3 = \frac{\mathbf{q}_3}{\|\mathbf{q}_3\|} = \left[\frac{-1}{\sqrt{2}} \ 0 \ \frac{1}{\sqrt{2}} \right]^t$$

são ortonormais.

OBSERVAÇÕES

- O processo de Gram-Schmidt permite construir uma base ortogonal, no espaço n -dimensional, a partir de s vetores dados que sejam linearmente independentes, com $s < n$. De fato, como o espaço E_n tem dimensão n , devem existir $n - s$ vetores linearmente independentes aos s vetores dados. Assim, obtém-se n vetores linearmente independentes. A afirmação decorre da ortogonalização destes n vetores pelo processo de Gram-Schmidt.
- Em termos computacionais, para manter a ortogonalidade dos vetores calculados pelo método de Gram-Schmidt, é conveniente modificar o processo. Ao invés de deixar os vetores originais \mathbf{v}_k inalterados, estes vão sendo modificados passo a passo. Ao calcular w_k , os vetores $\mathbf{v}_k, \mathbf{v}_{k+1}, \dots, \mathbf{v}_n$ são modificados requerendo que sejam ortogonais a $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{k-1}$. Esta modificação consiste em retirar de cada vetor \mathbf{v}_k uma única parcela $\frac{\mathbf{w}_k^t \mathbf{v}_i}{\mathbf{w}_k^t \mathbf{w}_k} \mathbf{w}_k$ a cada passo.

$$\begin{aligned} \mathbf{q}_1 &= \mathbf{w}_1^{(0)} / \|\mathbf{w}_1^{(0)}\|, \quad \mathbf{w}_j^{(0)} = v_j, \quad j = 1 : n \\ \mathbf{w}_i^{(j)} &= \mathbf{w}_i^{(j-1)} - \mathbf{q}_j^t \mathbf{w}_i^{(j-1)} \mathbf{q}_j, \quad i = j + 1 : n \\ \mathbf{q}_j &= \mathbf{w}_j^{(j-1)} / \|\mathbf{w}_j^{(j-1)}\| \end{aligned}$$

Pode ser mostrado que os \mathbf{q}_j assim calculados são os memos do processo de Gram-Schmidt.

Exemplo 4.50

Ortogonalizar os vetores $\mathbf{v}_1 = [1 \ 1 \ 1]^t$, $\mathbf{v}_2 = [1 \ 1 \ -1]^t$, $\mathbf{v}_3 = [1 \ -1 \ 1]^t$ pelo processo de Gram-Schmidt modificado.

O primeiro vetor é

$$\mathbf{q}_1 = \mathbf{w}_1^{(0)} / \|\mathbf{w}_1^{(0)}\| = \frac{1}{\sqrt{3}} [1 \ 1 \ 1]^t.$$

Para obter os seguintes dois vetores, tem-se

$$\begin{aligned} \mathbf{w}_2^{(1)} &= \mathbf{v}_2 - \mathbf{q}_1^t \mathbf{v}_2 \mathbf{q}_1 = \frac{2}{3} [1 \ 1 \ -2]^t \\ \mathbf{w}_3^{(1)} &= \mathbf{v}_3 - \mathbf{q}_1^t \mathbf{v}_3 \mathbf{q}_1 = \frac{2}{3} [1 \ -2 \ 1]^t \\ \mathbf{q}_2 &= \mathbf{w}_2^{(1)} / \|\mathbf{w}_2^{(1)}\| = \frac{1}{\sqrt{3}} [1 \ 1 \ -2]^t \\ \mathbf{w}_3^{(2)} &= \mathbf{v}_3 - \mathbf{q}_2^t \mathbf{w}_3^{(1)} \mathbf{q}_2 = [1 \ -1 \ 0]^t \\ \mathbf{q}_3 &= \mathbf{w}_3^{(2)} / \|\mathbf{w}_3^{(2)}\| = \frac{1}{\sqrt{2}} [1 \ -1 \ 0]^t. \end{aligned}$$

4.3 Ortogonalidade do Sistema Adjunto Homogêneo

Considere-se o sistema

$$\mathbf{Ax} = \mathbf{b}, \quad (4.37)$$

onde \mathbf{A} é uma matriz de ordem $m \times n$. O sistema

$$\mathbf{A}^t \mathbf{y} = \mathbf{d} , \quad (4.38)$$

é referido como o *sistema adjunto* do sistema dado.

As soluções destes sistemas possuem uma propriedade que permite estabelecer a consistência do sistema original, dada por

$$\mathbf{y}^t \mathbf{b} - \mathbf{x}^t \mathbf{d} = \mathbf{0} . \quad (4.39)$$

De fato, considerando-se o produto interno da primeira equação com \mathbf{y} e o da segunda com \mathbf{x} , vem

$$\mathbf{y}^t \mathbf{A} \mathbf{x} = \mathbf{y}^t \mathbf{b} \quad \text{e} \quad \mathbf{x}^t \mathbf{A}^t \mathbf{y} = \mathbf{x}^t \mathbf{d} . \quad (4.40)$$

Sendo

$$\mathbf{y}^t \mathbf{A} \mathbf{x} = (\mathbf{y}^t \mathbf{A} \mathbf{x})^t = \mathbf{x}^t \mathbf{A}^t \mathbf{y} , \quad (4.41)$$

o resultado segue por simples subtração.

Do sistema adjunto homogêneo

$$\mathbf{A}^t \mathbf{y} = \mathbf{0} , \quad (4.42)$$

tem-se, em particular que

$$\mathbf{y}^t \mathbf{b} = \mathbf{0} , \quad \text{para cada } \mathbf{y} \text{ tal que } \mathbf{A}^t \mathbf{y} = \mathbf{0} . \quad (4.43)$$

Ou seja, se $\mathbf{A} \mathbf{x} = \mathbf{b}$ possuir solução, necessariamente, o dado \mathbf{b} deve ser ortogonal a cada solução do sistema adjunto homogêneo.

Esta caracterização de consistência pode ser resumida como segue:

Para que o sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$, \mathbf{A} de ordem $m \times n$, possua solução, é necessário que o dado \mathbf{b} seja ortogonal a cada solução do sistema homogêneo adjunto $\mathbf{A}^t \mathbf{y} = \mathbf{0}$.

4.4 Matrizes Ortogonais

Uma matriz quadrada \mathbf{A} de ordem $n \times n$ é dita *matriz ortogonal* quando suas n colunas são mutuamente ortonormais. Por exemplo, a matriz

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{10}} & 0 & \frac{3}{\sqrt{10}} \\ 0 & 1 & 0 \\ -\frac{3}{10} & 0 & \frac{1}{\sqrt{10}} \end{bmatrix}$$

tem três colunas mutuamente ortonormais.

As matrizes ortogonais possuem várias propriedades que refletem a ortogonalidade de suas colunas. Além disso, a classe das matrizes ortogonais é fechada em relação à multiplicação matricial. A seguir, são enunciadas as seguintes propriedades:

1. $\mathbf{A} \mathbf{A}^t = \mathbf{A}^t \mathbf{A} = \mathbf{I}$
2. $(\mathbf{A} \mathbf{x})^t \mathbf{A} \mathbf{y} = \mathbf{x}^t \mathbf{y}$
3. $|\det(\mathbf{A})| = 1$

4. O produto de matrizes ortogonais é uma matriz ortogonal.

A primeira propriedade estabelece que toda matriz ortogonal é não singular. Também, que a matriz inversa de uma matriz ortogonal é sua matriz transposta. Entretanto, a segunda propriedade diz que as matrizes ortogonais preservam o produto interno, em particular, o comprimento de um vetor. Essas propriedades decorrem da ortonormalidade das colunas. O valor unitário decorre de propriedades dos determinantes.

Exemplo 4.51

A matriz

$$\mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} & 0 \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

é ortogonal.

O vetor $\mathbf{x} = [a \ b \ c]^t$, tem comprimento

$$\|\mathbf{x}\| = \sqrt{a^2 + b^2 + c^2}.$$

Por outro lado,

$$\mathbf{Ax} = \mathbf{v} = \left[\frac{a}{\sqrt{3}} + \frac{b}{\sqrt{6}} - \frac{c}{\sqrt{2}} \quad \frac{a}{\sqrt{3}} - 2\frac{b}{\sqrt{6}} \quad \frac{a}{\sqrt{3}} + \frac{b}{\sqrt{6}} + \frac{c}{\sqrt{2}} \right]^t,$$

possui o mesmo comprimento do vetor \mathbf{x} .

4.5 Fatorização Espectral de Matrizes Simétricas

As matrizes simétricas com elementos reais possuem as seguintes propriedades fundamentais:

1. Os autovalores de uma matriz simétrica são reais;
2. Os autovetores de uma matriz simétrica associados a autovalores diferentes são ortogonais;
3. Se \mathbf{A} for uma matriz simétrica de ordem n , então a matriz \mathbf{A} possuirá n autovetores ortogonais. Ou seja, qualquer matriz simétrica será não defeituosa.
4. O sistema $\mathbf{Ax} = \mathbf{b}$, com \mathbf{A} uma matriz simétrica, possui solução, não necessariamente única, quando \mathbf{b} for ortogonal com qualquer solução do sistema homogêneo $\mathbf{Ay} = \mathbf{0}$.

De fato, da relação $\mathbf{Ax} = \lambda\mathbf{x}$, tomando os complexos conjugados, vem

$$\overline{\mathbf{Ax}} = \overline{\lambda\mathbf{x}} \quad \text{ou} \quad \mathbf{A}\overline{\mathbf{x}} = \overline{\lambda}\overline{\mathbf{x}}. \quad (4.44)$$

Formando-se os produtos

$$\overline{\mathbf{x}}^t \mathbf{Ax} = \lambda \overline{\mathbf{x}}^t \mathbf{x}, \quad \mathbf{x}^t \mathbf{A}\overline{\mathbf{x}} = \overline{\lambda} \mathbf{x}^t \overline{\mathbf{x}}, \quad (4.45)$$

e, pelo fato que \mathbf{A} é simétrica real, ou seja, $\overline{\mathbf{x}}^t \mathbf{Ax} = \mathbf{x}^t \mathbf{A}\overline{\mathbf{x}}$, decorre por subtração que

$$0 = (\lambda - \overline{\lambda}) \|\mathbf{x}\|^2. \quad (4.46)$$

Por ser um autovetor, \mathbf{x} é um vetor não-nulo. Assim, $\lambda = \overline{\lambda}$.

Das relações

$$\mathbf{Ax} = \lambda\mathbf{x}, \quad \mathbf{Ay} = \beta\mathbf{y}, \quad (4.47)$$

obtem-se que

$$\mathbf{y}^t \mathbf{Ax} = \lambda \mathbf{y}^t \mathbf{x}, \quad \mathbf{x}^t \mathbf{Ay} = \beta \mathbf{x}^t \mathbf{y}, \quad (4.48)$$

e utilizando o fato que a matriz é simétrica, decorre

$$0 = (\lambda - \beta) \mathbf{y}^t \mathbf{x}. \quad (4.49)$$

Por hipótese, $\lambda - \beta \neq 0$ e, portanto, \mathbf{x} e \mathbf{y} são ortogonais.

Exemplo 4.52

Determinar os autovetores associados aos autovalores da matriz simétrica e verificar sua ortogonalidade.

$$\mathbf{A} = \begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix}.$$

Solução

Para a matriz dada, a equação característica é o determinante da matriz

$$\begin{bmatrix} \lambda - 7 & 2 & -1 \\ 2 & \lambda - 10 & 23 \\ -1 & 2 & \lambda - 7 \end{bmatrix}$$

dado por

$$\lambda^3 - 24\lambda^2 - 180\lambda - 432$$

e os autovalores são 6, 6 e 12. Para $\lambda = 6$, tem-se

$$\begin{bmatrix} -1 & 2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Efetuando a eliminação, decorre a equação

$$x_1 - 2x_2 + x_3 = 0,$$

que possui a variável básica x_1 e as variáveis livres x_2 e x_3 . Assim, o vetor coluna

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2x_2 - x_3 \\ x_2 \\ x_3 \end{bmatrix},$$

onde x_2 e x_3 são constantes arbitrárias, ambas não nulas simultaneamente, é a forma geral do autovetor associado ao autovalor $\lambda = 6$. Em particular, pode-se escolher $\mathbf{v}_1 = [1 \ 0 \ -1]^t$ e $\mathbf{v}_2 = [1 \ 1 \ 1]^t$ que são ortogonais. Quando $\lambda = 12$, se obtém que $\mathbf{v}_3 = [1 \ -2 \ 1]^t$ é um autovetor associado ao autovalor $\lambda = 12$ e ortogonal a \mathbf{v}_1 e \mathbf{v}_2 .

A prova da terceira propriedade será realizada como ilustração para $n = 2$. O caso geral pode ser estabelecido por indução (veja-se a literatura) ou por outros argumentos, os quais não acrescentam informação prática substancial para o método espectral.

Suponha-se que \mathbf{A} é de ordem 2, que λ_1 é um autovalor de \mathbf{A} e \mathbf{q}_1 um autovetor associado de norma unitária. Será estabelecido que \mathbf{A} tem um segundo autovetor \mathbf{q}_2 , associado a um autovalor λ_2 , o qual é ortogonal a \mathbf{q}_1 . Para isto, considere-se a matriz

$$\mathbf{Q} = \begin{bmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{bmatrix} \quad (4.50)$$

onde $\mathbf{q}_1 = [q_{11} \ q_{21}]^t$ é o autovetor correspondente ao autovalor λ_1 . Procura-se obter a segunda coluna de \mathbf{Q} a ser obtida de maneira conveniente, escolhendo q_{12} e q_{22} , tais que

$$q_{12}q_{12} + q_{22}q_{22} = 1 \quad \text{e} \quad q_{12}q_{11} + q_{22}q_{21} = 0. \quad (4.51)$$

Por ser um autovetor, \mathbf{q}_1 é não nulo e ao menos um de seus elementos é diferente de zero. Suponha-se que $q_{11} \neq 0$, então

$$q_{22} = \beta, \quad q_{12} = -\beta \frac{q_{21}}{q_{11}} \quad (4.52)$$

satisfazem a segunda equação. Escolhe-se β de modo que a primeira equação seja satisfeita. A matriz \mathbf{Q} , por construção, é certamente ortogonal. Do produto

$$\mathbf{B} = \mathbf{Q}^t \mathbf{A} \mathbf{Q} = \mathbf{Q}^t [\mathbf{A} \mathbf{q}_1 \ \mathbf{A} \mathbf{q}_2] = \mathbf{Q}^t [\lambda_1 \mathbf{q}_1 \ \mathbf{A} \mathbf{q}_2], \quad (4.53)$$

decorre

$$\begin{aligned}\mathbf{Q}^t \mathbf{A} \mathbf{Q} &= \begin{bmatrix} q_{11} & q_{21} \\ q_{12} & q_{22} \end{bmatrix} \begin{bmatrix} \lambda_1 q_{11} & a_{11} q_{12} + a_{12} q_{22} \\ \lambda_1 q_{21} & a_{21} q_{12} + a_{22} q_{22} \end{bmatrix} \\ &= \begin{bmatrix} \lambda_1 (q_{11}^2 + q_{21}^2) & b_{12} \\ \lambda_1 (q_{11} q_{12} + q_{22} q_{21}) & b_{22} \end{bmatrix} = \begin{bmatrix} \lambda_1 & b_{21} \\ 0 & b_{22} \end{bmatrix}.\end{aligned}\quad (4.54)$$

Uma vez que \mathbf{B} é simétrica, pois $\mathbf{B}^t = \mathbf{Q}^t \mathbf{A}^t \mathbf{Q} = \mathbf{Q}^t \mathbf{A} \mathbf{Q} = \mathbf{B}$, obtém-se que $b_{12} = 0$. Assim,

$$\mathbf{Q}^t \mathbf{A} \mathbf{Q} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & b_{22} \end{bmatrix} = \mathbf{D}.\quad (4.55)$$

Desta relação segue que $\mathbf{A} \mathbf{Q} = \mathbf{Q} \mathbf{D}$, ou seja, que $\mathbf{A} \mathbf{q}_1 = \lambda_1 \mathbf{q}_1$ e $\mathbf{A} \mathbf{q}_2 = b_{22} \mathbf{q}_2$. Como b_{22} deve ser o outro autovalor λ_2 de \mathbf{A} e \mathbf{q}_2 seu autovetor associado, o qual é ortogonal ao autovetor \mathbf{q}_1 , associado a λ_1 , conclui-se que a matriz simétrica real \mathbf{A} de ordem 2×2 possui 2 autovetores ortogonais.

A prova da última propriedade decorre da ortogonalidade do sistema adjunto homogêneo.

4.5.1 O Método Espectral

Considere-se o sistema

$$\mathbf{A} \mathbf{x} = \mathbf{b},\quad (4.56)$$

onde \mathbf{A} é uma matriz simétrica de ordem n .

Este tipo de sistema pode ser resolvido com o uso de autovetores e autovalores. Suponha-se que $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$ são os n autovetores ortogonais associados aos autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ da matriz \mathbf{A} .

Escreva-se a solução \mathbf{x} na forma

$$\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n,\quad (4.57)$$

onde os coeficientes c_k devem ser determinados. Substituindo na equação e utilizando o fato que os \mathbf{v}_k são autovetores, isto é, $\mathbf{A} \mathbf{v}_k = \lambda_k \mathbf{v}_k$, vem

$$\mathbf{A}(c_1 \mathbf{v}_1 + \dots + c_n \mathbf{v}_n) = c_1 \mathbf{A} \mathbf{v}_1 + \dots + c_n \mathbf{A} \mathbf{v}_n = \mathbf{b}\quad (4.58)$$

$$c_1 \lambda_1 \mathbf{v}_1 + \dots + c_n \lambda_n \mathbf{v}_n = \mathbf{b},\quad (4.59)$$

Aplicando o produto interno de \mathbf{v}_k com a última expressão, tem-se membros com o vetor

$$\mathbf{v}_k^t [c_1 \lambda_1 \mathbf{v}_1 + \dots + c_n \lambda_n \mathbf{v}_n] = \mathbf{v}_k^t \mathbf{b}\quad (4.60)$$

e decorre pela ortogonalidade

$$c_k \lambda_k \mathbf{v}_k^t \mathbf{v}_k = \mathbf{v}_k^t \mathbf{b} = \mathbf{b}^t \mathbf{v}_k, \quad k = 1, 2, \dots, n.\quad (4.61)$$

Se a matriz simétrica \mathbf{A} for não singular, então todos seus autovalores serão diferentes de zero e os coeficientes c_k são dados por

$$c_k = \frac{\mathbf{b}^t \mathbf{v}_k}{\lambda_k \mathbf{v}_k^t \mathbf{v}_k}.\quad (4.62)$$

Deste modo, a solução do sistema $\mathbf{Ax} = \mathbf{b}$ será da forma

$$\mathbf{x} = \frac{\mathbf{b}^t \mathbf{v}_1}{\lambda_1 \mathbf{v}_1^t \mathbf{v}_1} \mathbf{v}_1 + \cdots + \frac{\mathbf{b}^t \mathbf{v}_n}{\lambda_n \mathbf{v}_n^t \mathbf{v}_n} \mathbf{v}_n, \quad (4.63)$$

Os vetores

$$\mathbf{q}_k = \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}, \quad k = 1 : n \quad (4.64)$$

são mutuamente ortonormais. Além disso

$$\frac{\mathbf{b}^t \mathbf{v}_k}{\lambda_k \mathbf{v}_k^t \mathbf{v}_k} \mathbf{v}_k = \mathbf{b}^t \frac{\frac{\mathbf{v}_k}{\|\mathbf{v}_k\|}}{\lambda_k} \frac{\mathbf{v}_k}{\|\mathbf{v}_k\|} = \frac{\mathbf{b}^t \mathbf{q}_k}{\lambda_k} \mathbf{q}_k \quad (4.65)$$

Assim,

$$\mathbf{x} = \frac{\mathbf{b}^t \mathbf{q}_1}{\lambda_1} \mathbf{q}_1 + \frac{\mathbf{b}^t \mathbf{q}_2}{\lambda_2} \mathbf{q}_2 + \cdots + \frac{\mathbf{b}^t \mathbf{q}_n}{\lambda_n} \mathbf{q}_n = \sum_{k=1}^n \frac{\mathbf{b}^t \mathbf{q}_k}{\lambda_k} \mathbf{q}_k \quad (4.66)$$

Exemplo 4.53

Resolver pelo método espectral o sistema $\mathbf{Ax} = \mathbf{b}$, onde

$$\mathbf{A} = \begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}.$$

Solução

Os autovalores são $\lambda_1 = 6 = \lambda_2$, $\lambda_3 = 12$ e os autovetores correspondentes $\mathbf{v}_1 = [1 \ 0 \ -1]^t$, $\mathbf{v}_2 = [1 \ 1 \ 1]^t$ e $\mathbf{v}_3 = [1 \ -2 \ 1]^t$. Escreve-se a solução na forma

$$\mathbf{x} = c_1 v_1 + c_2 v_2 + c_3 v_3,$$

onde v_1, v_2, v_3 são os autovetores de \mathbf{A} . Substituindo os dados acima, vem

$$\mathbf{Ax} = c_1 6v_1 + c_2 6v_2 + c_3 12v_3 = \mathbf{b}$$

Assim,

$$c_1 = \frac{(v_1)^t \mathbf{b}}{6(v_1)^t v_1} = -\frac{1}{6}$$

$$c_2 = \frac{(v_2)^t \mathbf{b}}{6(v_2)^t v_2} = \frac{1}{9}$$

$$c_3 = \frac{(v_3)^t \mathbf{b}}{6(v_3)^t v_3} = \frac{1}{9}.$$

Portanto, a solução é o vetor

$$\mathbf{x} = \begin{bmatrix} \frac{1}{18} \\ -\frac{1}{9} \\ \frac{7}{18} \end{bmatrix}.$$

O método acima descrito é chamado de *método espectral*. Num contexto funcional é, simplesmente, o método de Fourier para a resolução de equações diferenciais.

4.5.2 O Caso Singular

Se a matriz \mathbf{A} for singular, então para que $\mathbf{Ax} = \mathbf{b}$ seja consistente, o dado \mathbf{b} deverá ser ortogonal a cada solução do sistema adjunto $\mathbf{A}^t \mathbf{w} = \mathbf{0}$. Como \mathbf{A} é simétrica, isto implica que \mathbf{b} deverá ser ortogonal com cada autovetor correspondente a um autovalor nulo. Nesta situação, a solução não é única, e é dada por

$$\mathbf{x} = \sum_{\lambda_k \neq 0} \frac{\mathbf{v}_k^t \mathbf{b}}{\lambda_k \mathbf{v}_k^t \mathbf{v}_k} \mathbf{v}_k + \sum_{\lambda_k = 0} c_k \mathbf{v}_k, \quad (4.67)$$

onde os c_k são constantes arbitrárias.

4.5.3 Fatorização e Diagonalização

O método espectral pode ser formulado matricialmente como segue. Defina-se a matriz

$$\mathbf{Q} = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \cdots \quad \mathbf{q}_n], \quad (4.68)$$

onde $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ são autovetores ortogonais correspondentes aos autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$, e normalizados pela condição

$$\mathbf{q}_k^t \mathbf{q}_k = 1, \quad k = 1, 2, \dots, n. \quad (4.69)$$

Uma vez que as colunas de \mathbf{Q} são ortonormais, tem-se que $\mathbf{Q}^t \mathbf{Q} = \mathbf{I}$, ou seja, \mathbf{Q} é uma matriz ortogonal. As relações

$$\mathbf{A} \mathbf{q}_1 = \lambda_1 \mathbf{q}_1, \quad \mathbf{A} \mathbf{q}_2 = \lambda_2 \mathbf{q}_2, \quad \cdots, \quad \mathbf{A} \mathbf{q}_n = \lambda_n \mathbf{q}_n \quad (4.70)$$

podem ser escritas na forma

$$\mathbf{A} [\mathbf{q}_1 \quad \cdots \quad \mathbf{q}_n] = [\mathbf{A} \mathbf{q}_1 \quad \cdots \quad \mathbf{A} \mathbf{q}_n] = [\lambda_1 \mathbf{q}_1 \quad \cdots \quad \lambda_n \mathbf{q}_n]. \quad (4.71)$$

Definindo a matriz diagonal

$$\mathbf{D} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix},$$

decorre que

$$\mathbf{A} \mathbf{Q} = \mathbf{Q} \mathbf{D}. \quad (4.72)$$

Multiplicando esta igualdade à direita por \mathbf{Q}^t e utilizando o fato que $\mathbf{Q}^t \mathbf{Q} = \mathbf{I}$, obtém-se a importante *fatorização espectral* de uma matriz simétrica real:

$$\mathbf{A} = \mathbf{Q} \mathbf{D} \mathbf{Q}^t, \quad (4.73)$$

denominada *fatorização espectral* ou *diagonalização* de uma matriz simétrica. . A matriz diagonal \mathbf{D} , formada com os autovalores é dita *matriz espectral* e a matriz ortogonal \mathbf{Q} , formada com os autovetores, é dita *matriz modal*.

Exemplo 4.54

Determinar as matrizes modal e espectral para a matriz \mathbf{A} do exemplo anterior. Verificar a fatorização espectral.

Solução

Tem-se que

$$\mathbf{D} = \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{bmatrix}$$

e

$$\mathbf{Q} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \\ 0 & \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{6}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{6}} \end{bmatrix}.$$

Da multiplicação matricial, decorre que $\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^t$.

4.5.4 Potências de uma Matriz simétrica

A fatorização espectral de uma matriz simétrica permite calcular facilmente a inversa de uma matriz simétrica não singular e as potências de uma matriz simétrica qualquer.

Se \mathbf{A} for uma matriz simétrica não singular, o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ possuirá a solução $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$. Por outro lado, da fatorização espectral e sendo \mathbf{Q} é uma matriz ortogonal, tem-se

$$\mathbf{A}^{-1} = (\mathbf{Q}^t\mathbf{D}\mathbf{Q})^{-1} = \mathbf{Q}^t\mathbf{D}^{-1}\mathbf{Q} \quad (4.74)$$

Assim,

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b} = \mathbf{Q}^t\mathbf{D}^{-1}\mathbf{Q}\mathbf{b}. \quad (4.75)$$

Observe-se que o termo à direita é uma expressão compacta para a fórmula obtida pelo método espectral.

A fatorização espectral de uma matriz simétrica permite, também, calcular suas potências. Se $\mathbf{A} = \mathbf{Q}^t\mathbf{D}\mathbf{Q}$, então

$$\mathbf{A}^2 = \mathbf{A}.\mathbf{A} = (\mathbf{Q}^t\mathbf{D}\mathbf{Q})(\mathbf{Q}^t\mathbf{D}\mathbf{Q}) = \mathbf{Q}^t\mathbf{D}^2\mathbf{Q}. \quad (4.76)$$

Por indução,

$$\mathbf{A}^k = \mathbf{Q}\mathbf{D}^k\mathbf{Q}^t, \quad k = 0, 1, 2, \dots \quad (4.77)$$

Observe-se que, para \mathbf{A} simétrica não singular, esta fórmula é também válida para k negativo.

Exemplo 4.55

Calcular \mathbf{A}^{-2} para a matriz do exemplo anterior.

Solução

Tem-se que

$$\mathbf{A}^{-2} = \mathbf{Q}\mathbf{D}^{-2}\mathbf{Q}^t = \mathbf{Q} \begin{bmatrix} \frac{1}{36} & 0 & 0 \\ 0 & \frac{1}{36} & 0 \\ 0 & 0 & \frac{1}{144} \end{bmatrix} \mathbf{Q}^t.$$

4.5.5 Desacoplamento

A fatorização espectral

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^t \quad (4.78)$$

permite transformar a equação

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (4.79)$$

numa mais simples. De fato, substituindo \mathbf{A} pela sua fatorização

$$\mathbf{Q}\mathbf{D}\mathbf{Q}^t\mathbf{x} = \mathbf{b} \quad (4.80)$$

e multiplicando à esquerda por \mathbf{Q}^t , obtém-se

$$\mathbf{D}\mathbf{Q}^t\mathbf{x} = \mathbf{Q}^t\mathbf{b} \quad (4.81)$$

ou, simplesmente,

$$\mathbf{D}\mathbf{y} = \mathbf{d}, \quad (4.82)$$

onde

$$\mathbf{y} = \mathbf{Q}^t\mathbf{x}, \quad \mathbf{d} = \mathbf{Q}^t\mathbf{b}. \quad (4.83)$$

Assim,

$$\begin{aligned} \lambda_1 y_1 &= d_1 \\ \lambda_2 y_2 &= d_2 \\ &\vdots \\ \lambda_n y_n &= d_n. \end{aligned} \quad (4.84)$$

Neste sistema, cada variável y_k está isolada, ou seja, o sistema está desacoplado para as variáveis y_1, y_2, \dots, y_n . Se todos os autovalores de \mathbf{A} forem não-nulos, tem-se

$$y_i = (\lambda_i)^{-1}d_i, \quad i = 1, \dots, n, \quad (4.85)$$

ou

$$\mathbf{y} = \mathbf{D}^{-1}\mathbf{d}. \quad (4.86)$$

A solução do sistema original $\mathbf{A}\mathbf{x} = \mathbf{b}$ é a mesma que foi determinada anteriormente, ou seja,

$$\mathbf{x} = \mathbf{Q}\mathbf{y} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{b} = \mathbf{Q}\mathbf{D}^{-1}\mathbf{Q}^t\mathbf{b} = \mathbf{A}^{-1}\mathbf{b}. \quad (4.87)$$

Geometricamente, o método espectral é uma conveniente mudança de coordenadas para *desacoplar* um sistema simétrico.

Resumindo:

Se \mathbf{A} for uma matriz simétrica de ordem n , então \mathbf{A} possuirá somente autovalores reais e n autovetores ortogonais associados. Além disso, se $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$ forem estes autovetores normalizados e $\mathbf{D} = \text{diag}[\lambda_1 \ \dots \ \lambda_n]$, então

$$\mathbf{A} = \mathbf{Q}\mathbf{D}\mathbf{Q}^t, \quad (4.88)$$

onde $\mathbf{Q} = [\mathbf{q}_1 \ \dots \ \mathbf{q}_n]$ uma matriz ortogonal.

4.6 Extensão do Método Espectral

A classe das matrizes simétricas é restrita, contudo o método espectral pode ser estendido à uma classe mais ampla: as matrizes quadradas de ordem n que possuem exatamente n autovetores linearmente independentes. Tal extensão é feita de duas maneiras distintas que são especificadas a seguir.

4.6.1 Biortogonalidade da Matriz Adjunta

O primeiro procedimento consiste em introduzir a matriz transposta \mathbf{A}^t e seus n autovetores. Além destas matrizes possuírem os mesmos autovalores, tem-se a seguinte propriedade da **bi-ortogonalidade** :

Se \mathbf{v} é um autovetor de \mathbf{A} com autovalor λ e \mathbf{w} um autovetor de \mathbf{A}^t com diferente autovalor β , então \mathbf{v} e \mathbf{w} são ortogonais.

De fato,

$$\begin{aligned} Av &= \lambda v, & A^t w &= \beta w, \\ w^t Av &= \lambda w^t v, & v^t A^t w &= \beta v^t w \end{aligned} \quad (4.89)$$

$$0 = (\lambda - \beta)w^t v$$

Suponha-se que uma matriz quadrada real \mathbf{A} de ordem n possua n autovetores independentes \mathbf{v}_i e que sejam \mathbf{w}_j os n autovetores de \mathbf{A}^t . Considere-se a equação

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad (4.90)$$

e escreva-se

$$\mathbf{x} = c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n, \quad (4.91)$$

obtendo-se, na equação anterior

$$\mathbf{A}\mathbf{x} = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \dots + c_n \lambda_n \mathbf{v}_n = \mathbf{b}. \quad (4.92)$$

Pela ortogonalidade entre os autovetores de \mathbf{A}^t e os autovetores de \mathbf{A} , decorre

$$\begin{aligned} c_1 \lambda_1 \mathbf{w}_1^t \mathbf{v}_1 &= \mathbf{w}_1^t \mathbf{b} \\ c_2 \lambda_2 \mathbf{w}_2^t \mathbf{v}_2 &= \mathbf{w}_2^t \mathbf{b} \\ &\vdots \\ c_n \lambda_n \mathbf{w}_n^t \mathbf{v}_n &= \mathbf{w}_n^t \mathbf{b}. \end{aligned} \quad (4.93)$$

Destas relações, para o caso em que \mathbf{A} é não-singular, pode-se obter os coeficientes c_i e formar a solução \mathbf{x} . Observe-se que este processo é muito semelhante ao utilizado com matrizes simétricas, porém é mais complexo, pois requer o cálculo dos autovetores de \mathbf{A} e de \mathbf{A}^t . Mais precisamente:

Se a matriz \mathbf{A} é não singular, ou seja todos seus autovalores λ_k são diferentes de zero, decorre que a solução do sistema $\mathbf{Ax} = \mathbf{b}$, escreve-se

$$\mathbf{x} = \sum_{j=1}^n \frac{\mathbf{w}_j^t \mathbf{b}}{\mathbf{w}_j^t \mathbf{v}_j} \mathbf{v}_j \quad (4.94)$$

onde os \mathbf{v}_j são os autovetores de \mathbf{A} e os \mathbf{w}_j são os autovetores de \mathbf{A}^t .

4.6.2 Diagonalização de Matrizes Não Defeituosas

Além do desacoplamento produzido pela bi-ortogonalidade, a segunda maneira de estender o método espectral não requer o uso direto da matriz transposta e é obtida da maneira seguinte. Suponha-se que uma matriz quadrada \mathbf{A} de ordem n possua n autovetores independentes $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, associados aos autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ respectivamente. Agora, a equação

$$\mathbf{Ax} = \mathbf{b} \quad (4.95)$$

com

$$\begin{aligned} \mathbf{x} &= c_1 \mathbf{v}_1 + c_2 \mathbf{v}_2 + \dots + c_n \mathbf{v}_n, \\ \mathbf{b} &= d_1 \mathbf{v}_1 + d_2 \mathbf{v}_2 + \dots + d_n \mathbf{v}_n \end{aligned} \quad (4.96)$$

é equivalente à equação

$$\mathbf{Ax} = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \dots + c_n \lambda_n \mathbf{v}_n = d_1 \mathbf{v}_1 + \dots + d_n \mathbf{v}_n. \quad (4.97)$$

Pela independência linear dos autovetores de \mathbf{A} , obtém-se

$$\begin{aligned} c_1 \lambda_1 &= d_1 \\ c_2 \lambda_2 &= d_2 \\ &\vdots \\ c_n \lambda_n &= d_n. \end{aligned} \quad (4.98)$$

Se a matriz \mathbf{A} é não singular, ou seja, todos seus autovalores λ_k são diferentes de zero, decorre que a solução do sistema $\mathbf{Ax} = \mathbf{b}$ é

$$\mathbf{x} = \sum_{j=1}^n \frac{d_j}{\lambda_j} \mathbf{v}_j. \quad (4.99)$$

Deve ser observado que as relações

$$\mathbf{Av}_1 = \lambda_1 \mathbf{v}_1, \quad \mathbf{Av}_2 = \lambda_2 \mathbf{v}_2, \quad \dots, \quad \mathbf{Av}_n = \lambda_n \mathbf{v}_n \quad (4.100)$$

podem ser escritas na forma matricial

$$[\mathbf{Av}_1 \quad \dots \quad \mathbf{Av}_n] = [\lambda_1 \mathbf{v}_1 \quad \dots \quad \lambda_n \mathbf{v}_n] \quad (4.101)$$

ou, simplesmente,

$$\mathbf{A}[\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}. \quad (4.102)$$

Definindo as matrizes

$$\begin{aligned} \mathbf{V} &= [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n], \\ \mathbf{\Omega} &= \text{diag}[\lambda_1 \ \cdots \ \lambda_n] \end{aligned} \quad (4.103)$$

decorre

$$\mathbf{AV} = \mathbf{V}\mathbf{\Omega} \quad \text{ou} \quad \mathbf{A} = \mathbf{V}\mathbf{\Omega}\mathbf{V}^{-1}, \quad (4.104)$$

uma vez que a matriz \mathbf{V} é não singular, pois suas colunas são independentes, por hipótese.

Esta extensão do método espectral equivale a substituir na equação $\mathbf{Ax} = \mathbf{b}$, a fatorização $\mathbf{A} = \mathbf{V}\mathbf{\Omega}\mathbf{V}^{-1}$. Assim,

$$(\mathbf{V}\mathbf{\Omega}\mathbf{V}^{-1})\mathbf{x} = \mathbf{b} \quad (4.105)$$

e determinando a solução

$$\mathbf{x} = (\mathbf{V}\mathbf{\Omega}^{-1}\mathbf{V}^{-1})\mathbf{b}, \quad (4.106)$$

ou seja,

$$\mathbf{A}^{-1} = \mathbf{V}\mathbf{\Omega}^{-1}\mathbf{V}^{-1}. \quad (4.107)$$

Observe-se que com as substituições

$$\mathbf{x} = \mathbf{V}\mathbf{c}, \quad \mathbf{b} = \mathbf{V}\mathbf{d} \quad (4.108)$$

na equação $\mathbf{Ax} = \mathbf{b}$, obtém-se

$$(\mathbf{V}\mathbf{\Omega}\mathbf{V}^{-1})\mathbf{V}\mathbf{c} = \mathbf{V}\mathbf{d} \quad (4.109)$$

ou seja, o sistema desacoplado

$$\mathbf{\Omega}\mathbf{c} = \mathbf{d}.$$

4.7 Formas Quadráticas

Historicamente, num contexto mais geral do que para as matrizes, o método espectral é devido a Fourier. Na álgebra das matrizes, o estudo espectral foi desenvolvido por Sylvester em conexão com as formas quadráticas. Estas apresentações são complementares e valiosas na prática.

Considere-se um sistema composto por três massas conectadas em série com quatro molas, sendo duas delas engastadas como mostra a figura

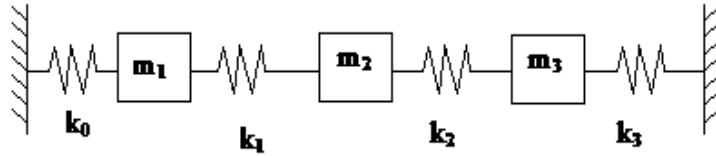


Figura 4.2 – Forma Quadrática

Se x_1, x_2, x_3 denotam os deslocamentos das massas vibrantes desde suas posições de equilíbrio, e se as molas obedecem a lei de Hooke, com constantes de rigidez k_i positivas, então da lei de Newton, aplicada a cada massa decorre

$$m_i \frac{d^2 x_i}{dt^2} = k_{i-1} x_{i-1} - (k_i + k_{i+1}) x_i + k_{i+1} x_{i+1}, \quad k = 1 : 3, \quad (4.110)$$

onde $x_0 = x_4 = 0$. Matricialmente, tem-se

$$\begin{bmatrix} m_1 & 0 & 0 \\ 0 & m_2 & 0 \\ 0 & 0 & m_3 \end{bmatrix} \begin{bmatrix} \ddot{x}_1 \\ \ddot{x}_2 \\ \ddot{x}_3 \end{bmatrix} + \begin{bmatrix} k_0 + k_1 & k_1 & 0 \\ -k_1 & k_1 + k_2 & -k_2 \\ 0 & -k_2 & k_2 + k_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (4.111)$$

ou

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{0}. \quad (4.112)$$

As matrizes \mathbf{M} , \mathbf{K} são chamadas *matriz de massa* e *matriz de rigidez* associadas a este sistema, respectivamente. Multiplicando a esquerda ambos os lados de (4.112) por $\dot{\mathbf{x}}^t$, vem

$$\dot{\mathbf{x}}^t \mathbf{M} \ddot{\mathbf{x}} + \dot{\mathbf{x}}^t \mathbf{K} \mathbf{x} = 0 \quad (4.113)$$

e substituindo

$$\dot{\mathbf{x}}^t \mathbf{M} \ddot{\mathbf{x}} = \frac{d}{dt} [\dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}}] \quad (4.114)$$

$$\dot{\mathbf{x}}^t \mathbf{K} \mathbf{x} = \frac{d}{dt} [\mathbf{x}^t \mathbf{K} \mathbf{x}] \quad (4.115)$$

decorre

$$\frac{d}{dt} \left[\frac{1}{2} \dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} + \frac{1}{2} \mathbf{x}^t \mathbf{K} \mathbf{x} \right] = 0 \quad (4.116)$$

isto é, a energia

$$\mathbf{E} = \frac{1}{2} \dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} + \frac{1}{2} \mathbf{x}^t \mathbf{K} \mathbf{x} \quad (4.117)$$

é preservada ao longo do movimento. O sistema (4.112) é dito *conservativo* e define-se a *energia cinética* e *energia potencial* do sistema como sendo as *formas quadráticas*

$$q_m = \frac{1}{2} \dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} \quad (4.118)$$

$$q_k = \frac{1}{2} \mathbf{x}^t \mathbf{K} \mathbf{x} \quad (4.119)$$

ou

$$\begin{aligned} q_m &= m_1 \dot{x}_1^2 + m_2 \dot{x}_2^2 + m_3 \dot{x}_3^2 \\ q_k &= (k_o + k_1) x_1^2 - 2k_1 x_1 x_2 + \\ &\quad + (k_1 + k_2) x_2^2 - 2k_2 x_2 x_3 + k_3 x_3^2, \end{aligned}$$

respectivamente.

Em geral, uma *forma quadrática homogênea em n variáveis* ou simplesmente *forma quadrática* q , é uma função real da forma

$$q = \sum k = 0n \sum j = 0n a_{ij} x_i x_j \quad (4.120)$$

A matriz quadrada $\mathbf{A} = [a_{ij}]$ de ordem n é chamada de matriz associada com a forma q e, vice-versa, diz-se que a forma q está associada à matriz \mathbf{A} . Matricialmente, pode-se escrever de maneira compacta

$$q = \mathbf{x}^t \mathbf{A} \mathbf{x}, \quad (4.121)$$

onde \mathbf{A} é a matriz associada e \mathbf{x} é o vetor coluna

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}. \quad (4.122)$$

Pode-se verificar que

$$q = \mathbf{x}^t \mathbf{A} \mathbf{x} = \frac{1}{2} \mathbf{x}^t (\mathbf{A} + \mathbf{A}^t) \mathbf{x} \quad (4.123)$$

Como a matriz $A = \mathbf{A} + \mathbf{A}^t$ é simétrica, será considerado a seguir que a matriz associada a quadrática é uma matriz simétrica. Por exemplo,

$$q = x^2 + xy + y^2 \quad (4.124)$$

pode ser escrita

$$q = [x \ y] \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad (4.125)$$

onde a matriz associada é simétrica.

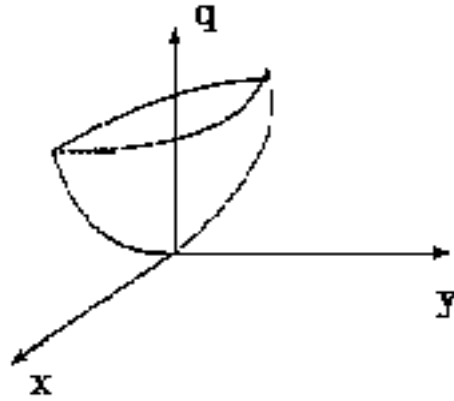


Figura 4.3 – Forma Quadrática

As formas quadráticas homogêneas mais simples são aquelas em que sua matriz associada é uma matriz diagonal

$$q = d_1x_1^2 + d_2x_2^2 + \dots + d_nx_n^2 = \mathbf{x}^t\mathbf{D}\mathbf{x}, \quad (4.126)$$

onde

$$\mathbf{D} = \begin{bmatrix} d_1 & 0 & \dots & 0 \\ 0 & d_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & d_n \end{bmatrix}.$$

Formas quadráticas deste tipo são ditas *diagonais* ou *desacopladas*.

A fatorização espectral de uma matriz simétrica permite reduzir uma forma quadrática a uma forma diagonal. Pois, substituindo $\mathbf{A} = \mathbf{Q}^t\mathbf{D}\mathbf{Q}$, decorre

$$q = \mathbf{x}^t\mathbf{Q}^t\mathbf{D}\mathbf{Q}\mathbf{x} = \mathbf{Q}\mathbf{x}^t\mathbf{D}\mathbf{Q}\mathbf{x}. \quad (4.127)$$

Introduzindo a mudança de variáveis

$$\mathbf{z} = \mathbf{Q}\mathbf{x} \quad (4.128)$$

obtém-se

$$q = \mathbf{z}^t \mathbf{D} \mathbf{z} = \lambda_1 z_1^2 + \lambda_2 z_2^2 + \dots + \lambda_n z_n^2, \quad (4.129)$$

onde os λ_k são os autovalores de \mathbf{A} .

A importância das formas diagonais reside no fato que elas representam combinações de quadrados e não possuem, portanto, termos de produtos cruzados. Por exemplo, as cônicas $cx^2 + dy^2 = 1$, cujo estudo é mais simples que o de $ax^2 + cxy + by^2 = 1$.

Observe-se que se a matriz \mathbf{A} for singular, então possuirá um autovalor nulo. Portanto, a coordenada de \mathbf{z} na direção do autovetor associado ao autovalor nulo estará ausente. Uma forma quadrática será dita *não degenerada* quando a matriz associada for não singular e será dita *degenerada* quando a matriz associada for singular.

Para uma matriz simétrica não singular, pode-se utilizar a fatorização $\mathbf{A} = \mathbf{L}^t \mathbf{D} \mathbf{L}$ para diagonalizar a forma quadrática. Mais precisamente,

$$q = \mathbf{x}^t \mathbf{L}^t \mathbf{D} \mathbf{L} \mathbf{x} = \mathbf{L} \mathbf{x}^t \mathbf{D} \mathbf{L} \mathbf{x} \quad (4.130)$$

$$q = \mathbf{z}^t \mathbf{D} \mathbf{z} = d_1 z_1^2 + d_2 z_2^2 + \dots + d_n z_n^2, \quad (4.131)$$

onde os d_k são os pivôs da matriz \mathbf{A} .

A diferença entre o uso da fatorização espectral e o da fatorização Gaussiana, além da não singularidade da matriz associada à forma, está no caráter das operações efetuadas para obtê-las. No método espectral, o processo é analítico (cálculo das raízes de um polinômio com um processo iterativo), entretanto, no método da eliminação o processo é algébrico (número finito de operações algébricas).

4.7.1 Quociente de Rayleigh

Se \mathbf{A} for uma matriz simétrica de ordem n , para cada autovetor, isto é,

$$\mathbf{A} \mathbf{v}_k = \lambda_k \mathbf{v}_k, \quad (4.132)$$

tem-se que

$$\mathbf{v}_k^t \mathbf{A} \mathbf{v}_k = \lambda_k \|\mathbf{v}_k\|^2 \quad (4.133)$$

ou

$$\lambda_k = \frac{\mathbf{v}_k^t \mathbf{A} \mathbf{v}_k}{\mathbf{v}_k^t \mathbf{v}_k}, \quad (4.134)$$

A expressão

$$R(\mathbf{v}) = \frac{\mathbf{v}^t \mathbf{A} \mathbf{v}}{\mathbf{v}^t \mathbf{v}} \quad (4.135)$$

é chamada de *quociente de Rayleigh*. Este quociente serve para caracterizar o maior ou o menor autovalor de uma matriz simétrica. Pois, substituindo a fatorização espectral $\mathbf{A} = \mathbf{Q}^t \mathbf{D} \mathbf{Q}$ no quociente e com a mudança de variáveis $\mathbf{y} = \mathbf{Q} \mathbf{x}$, obtém-se

$$R = \frac{\mathbf{v}^t \mathbf{Q}^t \mathbf{D} \mathbf{Q} \mathbf{v}}{\mathbf{v}^t \mathbf{v}} = \frac{\mathbf{y}^t \mathbf{D} \mathbf{y}}{\mathbf{y}^t \mathbf{y}}. \quad (4.136)$$

ou

$$R = \frac{\lambda_1 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2}{y_1^2 + y_2^2 + \cdots + y_n^2}. \quad (4.137)$$

Suponha-se que os autovalores estão na ordem decrescente

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n. \quad (4.138)$$

Mas, como

$$R = \frac{\lambda_1^2 y_1^2 + \lambda_2 y_2^2 + \cdots + \lambda_n y_n^2}{y_1^2 + y_2^2 + \cdots + y_n^2} \quad (4.139)$$

decorre que $R \leq \lambda_1^2$. Este valor é atingido quando $\mathbf{y} = \mathbf{Q}\mathbf{v}_1$. Deste modo é obtido o *Princípio de Rayleigh*

Seja \mathbf{A} uma matriz simétrica \mathbf{A} de ordem n . Suponha-se que o maior autovalor é λ_1 . então

$$\lambda_1 = \max R(\mathbf{v}) = \max \frac{\mathbf{v}^t \mathbf{A} \mathbf{v}}{\mathbf{v}^t \mathbf{v}}, \quad \mathbf{v} \neq \mathbf{0} \quad (4.140)$$

Além disso, R atinge valor máximo quando \mathbf{v} é um autovetor correspondente a λ_1 .

Observe que, se for realizada a minimização do quociente, será obtido o menor autovalor.

Este resultado não é válido para matrizes não simétricas ainda, com autovalores reais. Por exemplo,

Exemplo 4.56

Seja

$$\mathbf{A} = \begin{bmatrix} 1 & 94 \\ 0 & 2 \end{bmatrix}.$$

Esta matriz possui os autovalores $\lambda_1 = 2$ e $\lambda_2 = 1$. Porém, $\lambda_1 = 2$ não é o maior valor do quociente de Rayleigh. Pois, para o vetor $v_1 = v_2 = 1$

$$R = \frac{\mathbf{v}^t \mathbf{A} \mathbf{v}}{\mathbf{v}^t \mathbf{v}} = \frac{v_1^2 + 94v_1v_2 + 2v_2^2}{v_1^2 + v_2^2} = \frac{97}{2} \geq \lambda_1$$

O princípio de Rayleigh se aplica somente a matrizes simétricas.

4.7.2 Formas Definidas e Modos Normais

Uma matriz simétrica \mathbf{A} é dita *positiva definida* quando

$$\mathbf{x}^t \mathbf{A} \mathbf{x} > 0, \quad \mathbf{x} \neq \mathbf{0}. \quad (4.141)$$

Utilizando a fatorização de Cholesky e a fatorização espectral, tem-se as seguintes caracterizações das matrizes simétricas positivas definidas

1. Uma matriz simétrica é positiva definida se, e somente se, todos os pivôs são positivos.
2. Uma matriz simétrica é positiva definida se, e somente se, todos os autovalores são positivos

Este tipo de formas quadráticas aparecem frequentemente no estudo de sistemas mecânicos conservativos. Com relação ao exemplo das massas e molas no início desta seção, observa-se que

$$\mathbf{x}^t \mathbf{K} \mathbf{x} = k_o x_1^2 + k_1 (x_1 - x_2)^2 + k_2 (x_2 - x_3)^2 + k_3 x_3^2 \quad (4.142)$$

Como os k_i são positivos, tem-se $\mathbf{x}^t \mathbf{K} \mathbf{x} > 0$ exceto se $x_1 = x_2 = x_3 = 0$, ou seja, a matriz \mathbf{K} é positiva definida. Por outro lado, para a matriz \mathbf{M} tem-se

$$\dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} = m_1 \dot{x}_1^2 + m_2 \dot{x}_2^2 + m_3 \dot{x}_3^2 \quad (4.143)$$

Tem-se $\dot{\mathbf{x}}^t \mathbf{M} \dot{\mathbf{x}} > 0$, exceto se $\dot{x}_1 = \dot{x}_2 = \dots = \dot{x}_n = 0$ e segue que \mathbf{M} também é positiva definida.

Estas formas quadráticas podem ser diagonalizadas separadamente. Porém, as matrizes modais utilizadas não necessariamente são as mesmas para as matrizes \mathbf{M} e \mathbf{K} . A questão é a seguinte: *Será possível diagonalizar ambas as formas quadráticas de maneira simultânea, isto é, com uma mesma matriz modal \mathbf{Q} ?* A resposta é afirmativa.

Em primeiro lugar, relaciona-se \mathbf{M} e \mathbf{K} através de um problema de autovalor. Suponha-se que soluções oscilatórias do tipo

$$\mathbf{x} = e^{i\omega t} \mathbf{v} \quad (4.144)$$

devem ser determinadas, para certos valores de ω e vetores \mathbf{v} não nulos. Substituindo na equação

$$\dot{\mathbf{x}} = i\omega e^{i\omega t} \mathbf{v}, \quad \ddot{\mathbf{x}} = (i\omega)^2 e^{i\omega t} \mathbf{v} \quad (4.145)$$

decorre

$$e^{i\omega t} [-\omega^2 \mathbf{M} + \mathbf{K}] \mathbf{v} = \mathbf{0} \quad (4.146)$$

originando o problema *generalizado* de autovalor

$$\mathbf{K} \mathbf{v} = \omega^2 \mathbf{M} \mathbf{v}, \quad \mathbf{v} \neq \mathbf{0}. \quad (4.147)$$

As soluções deste problema são denominadas *modos*. Eles possuem uma certa propriedade de ortogonalidade. Suponha-se que \mathbf{v} , \mathbf{u} são soluções correspondentes a ω e γ , respectivamente, ou seja,

$$\mathbf{K} \mathbf{v} = \omega^2 \mathbf{M} \mathbf{v}, \quad \mathbf{v} \neq \mathbf{0}$$

$$\mathbf{K} \mathbf{u} = \gamma^2 \mathbf{M} \mathbf{u}, \quad \mathbf{u} \neq \mathbf{0}$$

Da simetria de \mathbf{K} e \mathbf{M} , vem

$$\begin{aligned} \mathbf{0} &= \mathbf{u}^t \mathbf{K} \mathbf{v} - \mathbf{v}^t \mathbf{K} \mathbf{u} \\ &= \omega^2 \mathbf{u}^t \mathbf{M} \mathbf{v} - \gamma^2 \mathbf{v}^t \mathbf{M} \mathbf{u} \\ &= [\omega^2 - \gamma^2] \mathbf{v}^t \mathbf{M} \mathbf{u}. \end{aligned}$$

Portanto, para $\omega \neq \gamma$,

$$\mathbf{u}^t \mathbf{M} \mathbf{v} = \mathbf{v}^t \mathbf{M} \mathbf{u} = \mathbf{0}$$

$$\mathbf{u}^t \mathbf{K} \mathbf{v} = \mathbf{v}^t \mathbf{K} \mathbf{u} = \mathbf{0}.$$

Como \mathbf{M} é simétrica e positiva definida, tem-se

$$\mathbf{M} = \mathbf{Q}^t \mathbf{D} \mathbf{Q} \quad (4.148)$$

onde \mathbf{D} é uma matriz diagonal cujos elementos diagonais, os autovalores de \mathbf{M} , são positivos e \mathbf{Q} uma matriz ortogonal. Sendo $\mathbf{Q}^t \mathbf{Q} = \mathbf{Q} \mathbf{Q}^t = \mathbf{I}$, segue

$$\mathbf{Q} \mathbf{M} \mathbf{Q}^t = \mathbf{D} \quad (4.149)$$

Escrevendo $\mathbf{D} = \sqrt{\mathbf{D}} \sqrt{\mathbf{D}}$, obtém-se

$$\left[\sqrt{\mathbf{D}} \mathbf{Q} \right] \mathbf{M} \left[\sqrt{\mathbf{D}} \mathbf{Q} \right]^t = \mathbf{I}. \quad (4.150)$$

Assim,

$$\mathbf{R} \mathbf{M} \mathbf{R}^t = \mathbf{I}, \quad \mathbf{R} = \sqrt{\mathbf{D}} \mathbf{Q} \quad (4.151)$$

Definindo

$$\mathbf{C} = \mathbf{R} \mathbf{K} \mathbf{R}^t \quad (4.152)$$

segue que a matriz \mathbf{C} é simétrica. Portanto,

$$\mathbf{C} = \mathbf{V} \mathbf{\Lambda} \mathbf{V}^t \quad (4.153)$$

onde \mathbf{V} é uma matriz ortogonal e $\mathbf{\Lambda}$ uma matriz simétrica. Decorre

$$\mathbf{V}^t \mathbf{C} \mathbf{V} = \mathbf{V}^t \mathbf{R} \mathbf{K} \mathbf{R}^t \mathbf{V} = \mathbf{\Lambda} \quad (4.154)$$

onde $\mathbf{\Lambda}$ é uma matriz diagonal. Defina-se

$$\mathbf{U} = \mathbf{R}^t \mathbf{V} = \sqrt{\mathbf{D}} \mathbf{Q} \mathbf{V} \quad (4.155)$$

Então,

$$\mathbf{U}^t \mathbf{K} \mathbf{U} = \mathbf{\Lambda}. \quad (4.156)$$

Similarmente, da relação $\mathbf{R} \mathbf{M} \mathbf{R}^t = \mathbf{I}$, segue

$$\mathbf{U}^t \mathbf{M} \mathbf{U} = \mathbf{I} \quad (4.157)$$

Resumindo:

Sejam \mathbf{M} , \mathbf{K} matrizes simétricas. Suponha-se que \mathbf{M} é positiva definida. Então, existe uma matriz não singular \mathbf{U} , tal que

$$\mathbf{U}^t \mathbf{M} \mathbf{U} = \mathbf{I}, \quad \mathbf{U}^t \mathbf{K} \mathbf{U} = \mathbf{\Lambda} \quad (4.158)$$

com $\mathbf{\Lambda}$ uma matriz diagonal cujos elementos diagonais ω_k^2 são não negativos, e positivos se \mathbf{K} for positiva definida. Além disso,

$$\mathbf{K} \mathbf{u}_j = \omega_j^2 \mathbf{M} \mathbf{u}_j \quad (4.159)$$

para cada coluna de \mathbf{U} . Matricialmente,

$$\mathbf{K} \mathbf{U} = \mathbf{M} \mathbf{U} \mathbf{\Lambda} \quad (4.160)$$

Com este resultado, tem-se as soluções complexas

$$e^{i\omega_k t} \mathbf{u}_k, \quad e^{-i\omega_k t} \mathbf{u}_k, \quad k = 1 : n, \quad (4.161)$$

as quais podem ser substituídas pelas soluções reais

$$\text{sen}(\omega_k t) \mathbf{u}_k, \quad \text{cos}(\omega_k t) \mathbf{u}_k. \quad (4.162)$$

Assim, a solução da equação $\mathbf{M}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{0}$ pode ser escrita

$$\mathbf{x}(t) = \sum_{k=1}^n [a_k \text{sen}(\omega_k t) + b_k \text{cos}(\omega_k t)] \mathbf{u}_k. \quad (4.163)$$

Para \mathbf{K} positiva definida, ou seja, os ω_k são positivos, as constantes c_k podem ser obtidas das condições iniciais

$$\mathbf{x}(0) = \mathbf{x}_o, \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_o. \quad (4.164)$$

Mais precisamente,

$$\sum_{k=1}^n b_k \mathbf{u}_k = \mathbf{x}_o \quad (4.165)$$

e

$$\sum_{k=1}^n a_k \omega_k \mathbf{u}_k = \dot{\mathbf{x}}_o. \quad (4.166)$$

Matricialmente,

$$[\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \begin{bmatrix} b_1 & \omega_1 a_1 \\ b_2 & \omega_2 a_2 \\ \vdots & \vdots \\ b_n & \omega_n a_n \end{bmatrix} = [\mathbf{x}_o \quad \dot{\mathbf{x}}_o]. \quad (4.167)$$

A matriz a esquerda é precisamente a matriz \mathbf{U} , a qual é não singular, e o sistema pode ser resolvido para a_k, b_k , isto é

$$\mathbf{b} = \mathbf{U}^{-1} \mathbf{x}_o \quad (4.168)$$

e

$$\mathbf{a} = \sqrt{\Lambda^{-1}} \mathbf{U}^{-1} \dot{\mathbf{x}}_o \quad (4.169)$$

onde $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_n]^t$, $\mathbf{a} = [a_1 \ a_2 \ \dots \ a_n]^t$ e $\sqrt{\Lambda^{-1}} = \text{diag}[\omega_1 \ \omega_2 \ \dots \ \omega_n]^t$.

4.8 Decomposição de Matrizes em Valores Singulares

Uma matriz \mathbf{A} de ordem $m \times n$ não admite necessariamente uma fatorização espectral. Porém, $\mathbf{A}^t \mathbf{A}$ e $\mathbf{A} \mathbf{A}^t$ são matrizes simétricas de ordem $n \times n$ e $m \times m$, respectivamente. Portanto, possuem autovalores reais e autovetores ortonormais. Mais precisamente, o problema de autovalor

$$\mathbf{A}^t \mathbf{A} \mathbf{v} = \lambda \mathbf{v}, \quad \mathbf{v} \neq \mathbf{0} \quad (4.170)$$

e elementos não-diagonais nulos, onde o número r de autovalores não nulos de $\mathbf{A}^t\mathbf{A}$ é o posto de \mathbf{A} .

Em particular, para matrizes quadradas a matriz $\mathbf{\Sigma}$ é diagonal, e se \mathbf{A} for não singular, então

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0.$$

Os números $\sigma_1 = \sqrt{\lambda_1}$, $\sigma_2 = \sqrt{\lambda_2}$, \dots , $\sigma_r = \sqrt{\lambda_r}$, isto é, as raízes positivas dos autovalores não nulos de $\mathbf{A}^t\mathbf{A}$, são chamados de *valores singulares* de \mathbf{A} e a fatorização $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t$ é dita *decomposição em valores singulares* de \mathbf{A} .

Exemplo 4.57

Construir a decomposição em valores singulares da matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 2 & 5 \\ 3 & 7 \end{bmatrix}.$$

Solução

Tem-se

$$\mathbf{A}^t\mathbf{A} = \begin{bmatrix} 14 & 33 \\ 33 & 48 \end{bmatrix}$$

e

$$\mathbf{A}\mathbf{A}^t = \begin{bmatrix} 5 & 12 & 17 \\ 12 & 29 & 41 \\ 17 & 41 & 58 \end{bmatrix}.$$

Os autovalores de $\mathbf{A}^t\mathbf{A}$ são, de maneira aproximada, $\lambda_1 = 91.9674$, $\lambda_2 = 0.0326$. Os correspondentes autovetores são

$$\mathbf{q}_1 = \begin{bmatrix} 0.3898 \\ 0.9209 \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} 0.9209 \\ -0.3898 \end{bmatrix}.$$

Para a matriz $\mathbf{A}\mathbf{A}^t$, tem-se os autovetores

$$\mathbf{u}_1 = \begin{bmatrix} 0.2327 \\ 0.5614 \\ 0.7941 \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} 0.5774 \\ 0.5774 \\ -0.5774 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -0.7826 \\ 0.5928 \\ -0.1848 \end{bmatrix}.$$

Assim,

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{91.9674} & 0 & 0 \\ 0 & \sqrt{0.0326} & 0 \end{bmatrix},$$

$$\mathbf{U} = \begin{bmatrix} 0.2327 & 0.5774 & -0.7826 \\ 0.5614 & 0.5774 & 0.5928 \\ 0.7941 & -0.5774 & -0.1848 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} 0.3898 & 0.9209 \\ 0.9209 & -0.3898 \end{bmatrix}.$$

Exemplo 4.58

Construir a decomposição em valores singulares da matriz

$$\mathbf{A} = \begin{bmatrix} 2 & 2 & -2 \\ 2 & 2 & -2 \\ -2 & -2 & 6 \end{bmatrix}.$$

Solução

A matriz

$$\mathbf{A}^t \mathbf{A} = \begin{bmatrix} 12 & 12 & -20 \\ 12 & 12 & -20 \\ -20 & -20 & 44 \end{bmatrix}$$

possui os autovalores 64, 4 e 0 com correspondentes autovetores ortonormais

$$\mathbf{q}_1 = \begin{bmatrix} -\frac{1}{\sqrt{6}} \\ -\frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad \mathbf{q}_3 = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}.$$

Assim,

$$\mathbf{V} = \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \end{bmatrix}$$

é a matriz espectral correspondente a $\mathbf{A}^t \mathbf{A}$. Similarmente, $\mathbf{A} \mathbf{A}^t$ possui a matriz modal dada por

$$\begin{bmatrix} \sqrt{64} & 0 \\ 0 & \sqrt{4} \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\mathbf{U} = \begin{bmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{3}} & 0 \end{bmatrix}.$$

Por outro lado,

$$\mathbf{\Sigma} = \begin{bmatrix} \sqrt{64} & 0 & 0 \\ 0 & \sqrt{4} & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Um cálculo direto mostra que

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} 8 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \mathbf{V}^t.$$

Deve-se observar que, no caso de matrizes simétricas, a decomposição em valores singulares coincide com a fatorização espectral, isto é, $\mathbf{U} = \mathbf{Q}$, $\mathbf{V}^t = \mathbf{Q}^t$ e $\mathbf{\Sigma} = \mathbf{D}$.

4.8.1 Interpretação Geométrica

A decomposição em valores singulares fornece uma importante caracterização do efeito geométrico de uma matriz como transformação. Para tanto, considere-se os vetores transformados

$$\mathbf{y} = \mathbf{A}\mathbf{x} , \quad (4.179)$$

onde \mathbf{A} é uma matriz de ordem $m \times n$. Com a mudança de variáveis

$$\mathbf{y} = \mathbf{U}\mathbf{w} , \quad \mathbf{x} = \mathbf{V}\mathbf{u}, \quad (4.180)$$

obtem-se

$$\begin{aligned} \mathbf{w} &= \mathbf{U}^t \mathbf{y} \\ &= \mathbf{U}^t \mathbf{A}\mathbf{x} \\ &= \mathbf{U}^t \mathbf{U} \mathbf{\Sigma} \mathbf{V}^t \mathbf{V}\mathbf{u} = \mathbf{\Sigma}\mathbf{u} \end{aligned}$$

ou, em termos das componentes,

$$\begin{aligned} w_1 &= \sigma_1 u_1 \\ w_2 &= \sigma_2 u_2 \\ &\vdots \\ w_r &= \sigma_r u_r \\ w_{r+1} &= 0 \\ &\vdots \\ w_m &= 0. \end{aligned}$$

Conseqüentemente,

$$\left(\frac{w_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{w_r}{\sigma_r}\right)^2 = u_1^2 + \cdots + u_r^2$$

e

$$w_{r+1} = \cdots = w_m = 0.$$

Tem-se dois casos:

1. $r = n$

Nesta situação, para u tal que $\|\mathbf{u}\|^2 = u_1^2 + \cdots + u_n^2 = 1$, obtém-se

$$\left(\frac{w_1}{\sigma_1}\right)^2 + \cdots + \left(\frac{w_n}{\sigma_n}\right)^2 = 1 .$$

Ou seja, a esfera unitária n -dimensional é transformada numa elipse n -dimensional. Aqui, \mathbf{A} e $\mathbf{\Sigma}$ são matrizes não singulares e as equações $\mathbf{A}\mathbf{x} = \mathbf{y}$ ou $\mathbf{\Sigma}\mathbf{u} = \mathbf{w}$ podem ser resolvidas de maneira única para qualquer dado.

2. $r < n$

A esfera n -dimensional é transformada no interior e contorno de uma elipse r -dimensional no hiperplano $w_{r+1} = \cdots = w_m = 0$. O sistema $\mathbf{A}\mathbf{x} = \mathbf{y}$ possui solução, somente, se o sistema $\mathbf{\Sigma}\mathbf{u} = \mathbf{w}$, possui solução. Este caso acontece unicamente quando \mathbf{w} está em qualquer direção da elipse r -dimensional, ou seja, $w_{r+1} = \cdots = w_m = 0$.

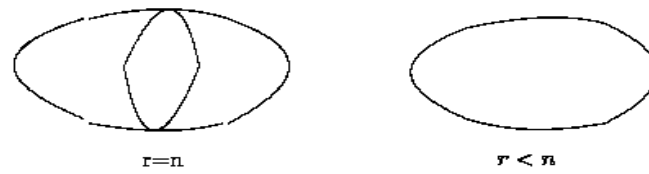


Figura 4.4 – Interpretação Geométrica

Para $n = 2$, por exemplo, tem-se as seguintes situações:

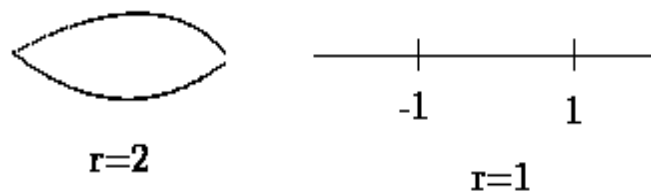


Figura 4.5 – Caso Bidimensional

Nesta situação, é claro que a equação $\Sigma \mathbf{u} = \mathbf{w}$ possui solução para qualquer \mathbf{w} . Entretanto, para $r = 1$ o sistema $\Sigma \mathbf{u} = \mathbf{w}$ possui solução, somente, se \mathbf{w} estiver sobre a linha . Em outras palavras, a existência de soluções está diretamente relacionada ao número de direções que a matriz \mathbf{A} preserva como transformação, isto é , o posto r é uma medida para avaliar o quanto a "solidez" é mantida (cheia) ou enfraquecida (comprimida ou distorcida).

Equações Diferenciais Matriciais

Neste capítulo serão considerados sistemas de equações diferenciais de primeira e segunda ordem, seguindo, no possível, a abordagem escalar. O estudo das equações diferenciais matriciais, abordado através da análise de modelos interessantes com aplicação em diversas áreas de interesse, fará uso dos três princípios das equações lineares:

1. PRINCÍPIO DA DECOMPOSIÇÃO

A solução de uma equação linear não-homogênea pode ser decomposta na soma de uma solução homogênea geral e uma solução não-homogênea particular.

2. PRINCÍPIO DA SUPERPOSIÇÃO

A combinação linear de soluções homogêneas é solução homogênea.

3. PRINCÍPIO DA REPRESENTAÇÃO

Existe uma solução fundamental que carrega toda a informação de uma equação diferencial linear

5.1 Um sistema de primeira ordem como modelo da hemodiálise

A principal função do rim é remover resíduos de produtos tais como uréia, creatinina e excesso de fluido do sangue. Quando os rins não estão funcionando corretamente, resíduos acumulam-se no sangue; quando níveis tóxicos são alcançados, a morte é certa. A remoção dos produtos tóxicos do metabolismo e a restauração do volume e da composição dos líquidos corporais em direção à normalidade pode ser realizada com a diálise com um rim artificial. Felizmente, a diálise nos rins remove os resíduos de produtos do sangue de pacientes cujos rins não estão funcionando adequadamente. Em certos tipos de insuficiência renal aguda, um rim artificial pode ser usado para fazer o paciente suportar até que os rins reassumam sua função. Milhares de pessoas com insuficiência renal irreversível, ou mesmo com remoção total dos rins, estão sendo mantidas por 15 a 20 anos pela diálise com rins artificiais [Keener].

O princípio básico do rim artificial é passar o sangue por delgados canais sanguíneos limitados por uma fina membrana. Do outro lado da membrana há um *líquido de diálise* para dentro do qual as substâncias indesejáveis do sangue passam por difusão.

A Fig. 5.9 mostra os componentes de um tipo de rim artificial no qual o sangue flui continuamente entre duas delgadas membranas de celofane; fora da membrana está o líquido da diálise. O celofane é suficientemente poroso para permitir que os componentes do plasma, exceto as proteínas plasmáticas, se difundam em ambas as direções: do plasma para o líquido da diálise ou do líquido da diálise de volta ao plasma. Se a concentração de uma substância for maior no plasma que no líquido da diálise, haverá transferência *efetiva* da substância do plasma para dentro do líquido da diálise.

A intensidade do movimento do soluto pela membrana de diálise depende

- do gradiente de concentração do soluto entre as duas soluções;
- da permeabilidade da membrana ao soluto;
- da área de superfície da membrana
- da quantidade de tempo em que o sangue e o líquido ficam em contato com a membrana.

Assim, a intensidade máxima de transferência de soluto ocorre inicialmente quando o gradiente de concentração é maior (quando a diálise é iniciada) e torna-se mais lenta à medida que o gradiente de concentração é dissipado. Num sistema de fluxo, como é o caso da "hemodiálise", na qual o sangue e o líquido da diálise fluem pelo rim artificial, a dissipação do gradiente de concentração pode ser reduzida e a difusão do soluto através da membrana pode ser otimizada aumentando-se o fluxo do sangue e/ou do líquido da diálise.

Durante o processo de hemodiálise, o sangue do paciente é bombeado, usualmente a uma taxa de 1 a 3 decilitros por minuto. O líquido da diálise contém algumas substâncias benéficas ao corpo, as quais são difusas para o sangue. O líquido da diálise flui em direção contrária à do sangue, usualmente a uma taxa de 2 a 6 decilitros por minuto. Os resíduos de produtos do sangue são difusos através da membrana a uma taxa proporcional a uma diferença de concentração dos resíduos de produtos no sangue e do líquido da diálise.

Seja $u(x)$ a concentração de resíduos no sangue e $v(x)$ a concentração de resíduos no líquido da diálise, onde x é a distância ao longo do dialisador; Q_D representa a taxa de fluxo do líquido da diálise através da máquina; e Q_S representa a taxa de fluxo do sangue através da máquina (dialisador).

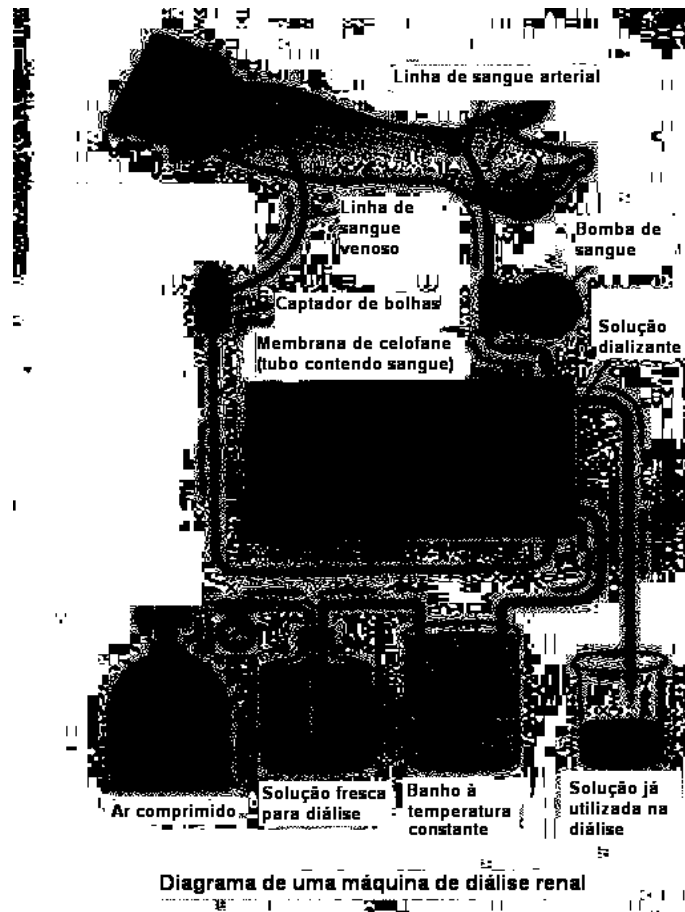


Figura 5.1 – Diagrama da máquina de hemodiálise para o rim

Então, tem-se o modelo

$$\begin{aligned} Q_S u' &= -k(u - v) + f(t) \\ -Q_D v' &= k(u - v) + g(t) \end{aligned} \quad (5.1)$$

onde k é a constante de proporcionalidade, $f(t)$ e $g(t)$ são fontes ("aditivos" na máquina ou no sangue).

Sejam, L o comprimento do dialisador, $u(0) = u_o$ a concentração inicial dos resíduos no sangue e $v(L) = 0$ a concentração inicial de resíduos no líquido de diálise. Então deve-se resolver o problema de valor inicial

$$\begin{aligned} Q_S u' + k(u - v) &= f(t) \\ -Q_D v' - k(u - v) &= g(t) \\ u(0) &= u_o \\ v(L) &= 0 \end{aligned} \quad (5.2)$$

O sistema de equações diferenciais lineares de 1ª ordem (5.2) pode ser escrito matricialmente como

$$C \frac{dU}{dt} + BU = F(t) \quad (5.3)$$

onde os coeficientes C , B são as matrizes

$$C = \begin{bmatrix} Q_S & 0 \\ 0 & -Q_D \end{bmatrix}, \quad B = \begin{bmatrix} k & -k \\ -k & k \end{bmatrix},$$

e U, F são os vetores coluna

$$U = \begin{bmatrix} u \\ v \end{bmatrix}, \quad dU/dt = U' = \begin{bmatrix} du/dt \\ dv/dt \end{bmatrix}, \quad F(t) = \begin{bmatrix} f(t) \\ g(t) \end{bmatrix}$$

5.2 A equação $Cu' + Bu = F(t)$

Um sistema de equações diferenciais lineares de 1ª ordem

$$c_{11} \frac{du_1}{dt} + c_{12} \frac{du_2}{dt} + \cdots + c_{1n} \frac{du_n}{dt} + b_{11}u_1 + b_{12}u_2 + \cdots + b_{1n}u_n = f_1(t) \quad (5.4)$$

$$c_{21} \frac{du_1}{dt} + c_{22} \frac{du_2}{dt} + \cdots + c_{2n} \frac{du_n}{dt} + b_{21}u_1 + b_{22}u_2 + \cdots + b_{2n}u_n = f_2(t) \quad (5.5)$$

$$\vdots \quad (5.6)$$

$$c_{n1} \frac{du_1}{dt} + c_{n2} \frac{du_2}{dt} + \cdots + c_{nn} \frac{du_n}{dt} + b_{n1}u_1 + b_{n2}u_2 + \cdots + b_{nn}u_n = f_n(t), \quad (5.7)$$

pode ser escrito matricialmente como

$$C \frac{dU}{dt} + BU = F(t) \quad (5.8)$$

onde os coeficientes $C = [c_{ij}]$, $B = [b_{ij}]$ são matrizes $n \times n$ e U, F são vetores coluna $n \times 1$. A matriz C será assumida não-singular.

Na literatura, é usual colocar (5.8) na forma normal

$$\frac{dU}{dt} = AU + f(t) \quad (5.9)$$

onde $A = -C^{-1}B$, $f = C^{-1}F$.

Uma forma integral da solução de (5.8) pode ser obtida de maneira análoga ao caso escalar em termos de uma solução fundamental $h(t)$. Por atuar sobre dados iniciais que, neste caso, são vetores, a solução $h(t)$ deverá ser uma matriz.

O cálculo desta solução $h(t)$ é usualmente realizado de maneira simbólica ou numérica. No cálculo simbólico a distinção fundamental entre uma ou outra técnica é o uso ou não de autovetores vinculados ao sistema (modos). Estas técnicas, podem ser agrupadas em três grandes métodos:

- Modal ou Espectral
- Não-Modal ou Não -Espectral
- Numérico

A seguir, a manipulação com matrizes deve ser realizada da maneira usual, como se fossem escalares. Somente deve-se ter atenção redobrada na ordem das matrizes num determinado produto. Pois, a propriedade da comutatividade nem sempre é válida para matrizes. Por outro lado, qualquer definição que envolva um processo limite, deve ser entendida como envolvendo o mesmo processo limite para cada componente. Por exemplo, para a matriz $f(t) = [f_{ij}(t)]$ de ordem $n \times n$, define-se

$$\frac{df}{dt} = \left[\frac{df_{ij}(t)}{dt} \right], \quad \int f(t)dt = \left[\int f_{ij}(t)dt \right], \quad (5.10)$$

desde que, respectivamente, cada componente seja diferenciável ou integrável.

5.2.1 Fórmulação integral

Considere-se o problema de valor inicial

$$C\dot{h}(t) + Bh(t) = 0, \quad (5.11)$$

onde 0 denota a matriz nula, com condição inicial

$$Ch(0) = I \quad (5.12)$$

A obtenção de $h(t)$ pode ser feita pelo método de Cauchy das séries de potência ou, série de Taylor, pois todas as derivadas de $h(t)$ em $t = 0$ podem ser obtidas diretamente da equação e, assim, serem utilizadas. Por exemplo, $C\dot{h}(0) + Bh(0) = 0$ implica $\dot{h}(0) = -(C^{-1}B)C^{-1}$, $C\ddot{h}(0) + B\dot{h}(0) = 0$ implica $\ddot{h}(0) = (C^{-1}B)^2C^{-1}$. Por indução, $h^k(0) = (-C^{-1}B)^kC^{-1}$. A expansão em série de Taylor de $h(t)$ pode ser escrita

$$h(t) = \sum_{k=0}^{\infty} \frac{h^k(0)}{k!} t^k = \sum_{k=0}^{\infty} \frac{(-C^{-1}B)^k t^k}{k!} C^{-1},$$

a qual, em analogia ao caso escalar, corresponde à expansão de Taylor da exponencial de uma matriz, isto é

$$h(t) = e^{-C^{-1}Bt}C^{-1} \quad (5.13)$$

vem a ser a solução do problema de valor inicial (5.11)-(5.12).

Para obter uma fórmula integral para a solução $u(t)$, fazemos atuar $h(t - \tau)$ como fator integrante. Para tanto, será utilizada a seguinte propriedade:

$$h'(t)C + h(t)B = 0, \quad h(0)C = I. \quad (5.14)$$

Esta propriedade segue por substituição direta de (5.13) em (5.11).

$$\boxed{h(t) \text{ é solução à esquerda e à direita da equação } Cu' + Bu = 0}$$

Então,

$$\begin{aligned} \int_0^t h(t - \tau)f(\tau)d\tau &= \int_0^t h(t - \tau)[Cu'(\tau) + Bu(\tau)]d\tau \\ &= h(t - \tau)Cu(\tau)|_0^t + \int_0^t [h'(t - \tau)C + h(t - \tau)B]u(\tau)d\tau \\ &= u(t) - h(t)Cu(0) \end{aligned}$$

Decorre a fórmula integral

$$u(t) = h(t)Cu(0) + \int_0^t h(t-\tau)F(\tau)d\tau, \quad h(t) = e^{-C^{-1}Bt}C^{-1}. \quad (5.15)$$

conhecida como *fórmula de variação de parâmetros*. Em particular, para condição inicial nula $u(0)=0$, tem-se que

$$u(t) = \int_0^t h(t-\tau)F(\tau)d\tau \quad (5.16)$$

é uma solução particular da equação não-homogênea (5.8). Ela é denominada de *resposta forçada*.

Por simples derivação e substituição, segue que, para τ considerado como parâmetro,

$$u(t) = h(t+\tau)$$

é solução da equação homogênea $C\dot{u} + Bu = 0$ com o valor inicial $u(0)=h(\tau)$. Da formula integral (5.15) com $F=0$, decorre que $h(t)$ satisfaz a *propriedade de semigrupo*

$$h(t+\tau) = h(t)Ch(\tau) = h(\tau)Ch(t) \quad (5.17)$$

Em particular,

$$h(t)^{-1} = Ch(-t)C \quad (5.18)$$

5.2.2 O Método de Variação de Parâmetros de Lagrange

Este método consiste em procurar a solução do mesmo tipo do que a homogênea, porém, permitindo que as constantes sejam funções. Assim, $u(t) = h(t)c(t)$, com $c(t)$ a ser determinada. Substituindo na equação, vem

$$C\dot{h}(t)c(t) + Ch(t)\dot{c}(t) + Bh(t)c(t) = Ah(t)c(t) + F(t).$$

Como $h(t)$ é solução homogênea, decorre que

$$Ch(t)\dot{c}(t) = F(t).$$

Assim, pela propriedade (5.18)

$$\dot{c}(t) = Ch(-t)CC^{-1}F(t) = Ch(-t)F(t).$$

Integrando, vem

$$c(t) = c(0) + \int_0^t Ch(-\tau)F(\tau)d\tau.$$

Então, utilizando a propriedade de semigrupo (5.17) e a condição $Ch(0) = I$, vem

$$\begin{aligned} u(t) &= h(t)c(0) + h(t) \int_0^t h(t-\tau)Ch(-\tau)F(\tau)d\tau \\ &= h(t)h(0)^{-1}u(0) + \int_0^t h(t)Ch(-\tau)F(\tau)d\tau \\ &= h(t)Cu(0) + \int_0^t h(t-\tau)F(\tau)d\tau. \end{aligned}$$

que não é outra coisa do que (5.15).

Propriedades da exponencial de uma matriz

Para a equação matricial

$$\frac{dU}{dt} = AU + f(t) \quad (5.19)$$

tem-se que $h(t)$ vem a ser a exponencial de uma matriz A :

$$h(t) = \sum_{k=0}^{\infty} \frac{h^k(0)}{k!} t^k = \sum_{k=0}^{\infty} \frac{A^k t^k}{k!} = e^{tA}. \quad (5.20)$$

A matriz exponencial e^{tA} possui as seguintes propriedades:

1. $e^{(t+\tau)A} = e^{tA}e^{\tau A}$,
2. $e^0 = I$,
3. $(e^{tA})^{-1} = e^{-tA}$.
4. A exponencial de uma matriz diagonal é uma outra matriz diagonal cujas componentes são as exponenciais das componentes da matriz diagonal dada. Assim

$$e \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} t = \begin{bmatrix} e^{\lambda_1 t} & 0 & \cdots & 0 \\ 0 & e^{\lambda_2 t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{\lambda_n t} \end{bmatrix}. \quad (5.21)$$

5. A solução da equação matricial na forma normal $u' = Au + f(t)$ é escrita na forma

$$u(t) = e^{tA}u(0) + \int_0^t e^{A(t-\tau)} f(\tau) d\tau. \quad (5.22)$$

5.3 O método matricial espectral

A seguir, apresenta-se o método espectral para a resolução de sistemas lineares de primeira ordem. Este método funciona bem para matrizes não-defeituosas, isto é, matrizes que geram bases de vetores formadas por seus autovetores. Dito de outro modo, se houver autovalores repetidos, então a multiplicidade geométrica de um autovalor (número de autovetores independente associados com esse autovalor) é igual a multiplicidade algébrica do autovalor (o número de vezes que é raiz da equação característica). Os seguintes classes de matrizes são particularmente não-defeituosas

- Matrizes com autovalores não repetidos
- Matrizes simétricas

5.3.1 Caso homogêneo

Considere-se a equação homogênea

$$Cu' + Bu = 0$$

O método de Euler de obter soluções do tipo exponencial é adaptado ao caso matricial considerando-se soluções do tipo

$$u = e^{\lambda t}v, \quad v \neq 0$$

Por substituição direta na equação homogênea, segue a equação algébrica ou *problema de autovalor*

$$[\lambda C + B]v = 0 \quad (5.23)$$

Esta equação possui solução v não-nula se o determinante do sistema é nulo, ou seja, λ deve ser uma raiz do polinômio característico

$$P(\lambda) = \det[\lambda C + B]. \quad (5.24)$$

Nesta situação, λ é dito *autovalor* e v um *autovetor* associado. Se os coeficientes C , B são $n \times n$, então, tem-se n autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ e os n autovetores correspondentes $\{v_1, v_2, \dots, v_n\}$. Pelo princípio da superposição linear e (5.21), tem-se que

$$u(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots + c_n e^{\lambda_n t} v_n = V e^{\Lambda t} c$$

é solução homogênea. Aqui c é o vetor cujas componentes são as constantes c_k , V é a matriz $n \times n$ cujas colunas são os autovetores v_k e

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_n \end{bmatrix}. \quad (5.25)$$

Para determinar as constantes c_k , utiliza-se o valor inicial $u_o = u(0)$ da solução. Assim,

$$u(0) = c_1 v_1 + c_2 v_2 + \dots + c_n v_n.$$

Tem-se o sistema algébrico linear $Vc = u_o$. Este sistema possui solução única $c = V^{-1}u_o$ para qualquer u_o dado, se V for não singular, ou seja, seu determinante é não nulo ou, também, suas n colunas são vetores linearmente independentes. Sob esta hipótese, tem-se que a solução da equação homogênea

$$Cu' + Bu = 0$$

é dada por

$$u(t) = \sum_{k=0}^n c_k e^{\lambda_k t} v_k = V e^{\Lambda t} V^{-1} u(0) \quad (5.26)$$

Como deve ser $u(t) = h(t)Cu(0)$, identifica-se $h(t)$ como sendo a matriz

$$h(t) = Ve^{\Lambda t}V^{-1}C^{-1} = \sum_{k=0}^n e^{\lambda_k t} v_k w_k^T C^{-1} \quad (5.27)$$

onde w^T denota o vetor transposto do vetor w , e w_k^T denota a k -ésima linha da matriz V^{-1} .

Exemplo 5.59

Resolver o problema de valor inicial

$$\begin{aligned} \dot{u} &= Au \\ u(0) &= u_o \end{aligned}$$

onde

$$A = \begin{bmatrix} 1 & 1 & -2 \\ -1 & 2 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \quad u_o = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

A equação característica com $C=I$ e $B=-A$ vem a ser

$$|\lambda I - A| = \begin{vmatrix} \lambda - 1 & -1 & 2 \\ 1 & \lambda - 2 & -1 \\ 0 & -1 & \lambda + 1 \end{vmatrix} = -\lambda^3 + 2\lambda^2 + \lambda - 2 = 0.$$

a qual é simplificada para $(\lambda + 1)(\lambda - 1)(\lambda - 2) = 0$. Então, os três autovalores são $\lambda_1 = -1, \lambda_2 = 1, \lambda_3 = 2$. Para determinar os correspondentes autovetores, substituímos sistematicamente $\lambda_1, \lambda_2, \lambda_3$ em $(\lambda I - A)w = 0$ para λ e resolvemos para w não nulo. Fazendo isto, vem

$$\lambda_1 = -1 \rightarrow$$

$$\begin{aligned} -2w_1 - w_2 + 2w_3 &= 0 \\ w_1 - 3w_2 - w_3 &= 0 \\ w_2 &= 0 \end{aligned}$$

$$\lambda_2 = 1 \rightarrow$$

$$\begin{aligned} -w_2 + 2w_3 &= 0 \\ w_1 - w_2 - w_3 &= 0 \\ -w_2 + 2w_3 &= 0 \end{aligned}$$

$$\lambda_3 = 2 \rightarrow$$

$$\begin{aligned} 2w_1 - w_2 + 2w_3 &= 0 \\ w_1 - w_3 &= 0 \\ -w_2 + 3w_3 &= 0 \end{aligned}$$

Resolvendo esses sistemas por eliminação gaussiana, obtém-se os correspondentes autovetores

$$\begin{aligned}\lambda_1 &= -1, & v_1 &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\ \lambda_2 &= 1, & v_2 &= \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} \\ \lambda_3 &= 2, & v_3 &= \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix}\end{aligned}$$

Assim, a solução general do problema é dada por

$$u(t) = c_1 e^{-t} v_1 + c_2 e^t + c_3 e^{2t}$$

Para determinar as constantes c_1, c_2, c_3 , utilizamos o valor inicial

$$u_o = c_1 v_1 + c_2 v - 2 + c_3 v_3 \rightarrow c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix}.$$

e obtém-se um sistema da forma $Vc = u_o$ cuja solução¹ é $c = V^{-1}u_o$. Assim,

$$\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ 3 & 2 & 1 \\ 1 & 3 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -1/6 & -1/3 & 7/6 \\ 1/2 & 0 & -1/2 \\ -1/3 & 1/3 & 1/3 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} = \begin{bmatrix} -5/2 \\ 3/2 \\ -1 \end{bmatrix}$$

A solução ao problema de valor inicial dado é

$$u(t) = -\frac{5}{2}e^{-t} \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} + \frac{3}{2}e^t \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix} - e^{2t} \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix},$$

ou,

$$u(t) = \begin{bmatrix} -\frac{15}{2}e^{-t} + \frac{9}{2}e^t - e^{2t} \\ -5e^{-t} + 3e^t - 3e^{-2t} \\ -\frac{5}{2}e^{-t} + \frac{3}{2}e^t - e^{2t} \end{bmatrix}.$$

¹A matriz V , cujas colunas são os autovetores, é inversível, pois os autovetores são linearmente independentes devido que correspondem a autovalores diferentes.

O caso simétrico

O método espectral aplica-se apenas à resolução de sistemas em que o problema de autovalor (5.23) é *não defeituoso*, isto é, que gere n autovetores linearmente independentes. Isto sempre ocorre quando C, B são matrizes simétricas que comutam. Em particular, quando $C = I$ é a matriz identidade e B é uma matriz simétrica. Nesta situação os autovetores são ortogonais e, além disso, a matriz V é ortogonal. Assim $V^{-1} = V^T$. Decorre a *expansão espectral da solução fundamental* $h(t)$ do sistema simétrico $u' + Bu = 0$, $B^T = B$:

$$(V\Lambda + B)V = 0, \quad C = I \quad (5.28)$$

$$h(t) = Ve^{\Lambda t}V^T = \sum_{k=0}^n e^{\lambda_k t} v_k v_k^T \quad (5.29)$$

5.3.2 O caso não-homogêneo

A solução da equação não-homogênea

$$Cu' + Bu = F(t) \quad (5.30)$$

pode ser obtida pelo método espectral com o uso do princípio da superposição linear. Procura-se a solução da forma

$$u(t) = u_h(t) + u_p(t). \quad (5.31)$$

onde a solução homogênea é escrita como combinação linear das n soluções do tipo exponencial $e^{\lambda t}v$, ou seja,

$$u_h(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \cdots + c_n e^{\lambda_n t} v_n. \quad (5.32)$$

Uma solução particular não-homogênea $u_p(t)$ pode ser decomposta em um somatório de n soluções particulares, sendo cada uma delas da mesma forma que as correspondentes componentes da entrada F , escrita como combinação linear dos autovetores. Em outras palavras, escreve-se em primeiro lugar o termo F como combinação linear da base de autovetores $\{v_1, v_2, \dots, v_n\}$ da seguinte maneira

$$F = VV^{-1}F = d_1 v_1 + d_2 v_2 + \cdots + d_n v_n, \quad (5.33)$$

onde V corresponde a matriz modal denotada por

$$V = [v_1 \ v_2 \ \cdots \ v_n] \quad \text{e} \quad V^{-1} = \begin{bmatrix} w_1^T \\ w_2^T \\ \vdots \\ w_n^T \end{bmatrix}. \quad (5.34)$$

Assim, $w_i^T v_j = \delta_{ij}$ para $i, j = 1 : n$ e $d_i = w_i^T f$ para cada $i = 1 : n$. Procura-se então uma solução particular da forma

$$u_p(t) = u_1^p(t) + u_2^p(t) + \cdots + u_n^p(t), \quad (5.35)$$

com u_j^p do mesmo tipo da componente $d_j(t)v_j$ de F , relativa ao autovetor v_j . Então as j -ésimas soluções particulares são tais que

$$u_j^p(t) = g_j(t)v_j. \quad (5.36)$$

Para determinar os valores dos coeficientes $g_j(t)$ resolve-se o sistema equivalente desacoplado,

$$\frac{dg_j}{dt}(t) = \lambda_j g_j(t) + d_j, \quad \text{para cada } j = 1 : n, \quad (5.37)$$

resultante da substituição da expressão (5.36) para u_j^p no sistema original (5.102). Uma solução particular não-homogênea para a equação (5.37) é dada por

$$g_j(t) = \int_0^t e^{\lambda_j(t-\tau)} d_j d\tau. \quad (5.38)$$

Para determinar as constantes da parte homogênea (5.32), utiliza-se o valor inicial da solução $u(t)$ em $t = 0$, resultando desta forma

$$c = V^{-1}(u_0 - u_p(0))$$

onde c corresponde ao vetor das constantes da parte homogênea $u_h(t)$.

Ao igual que no caso de sistemas simétricos homogêneos, o cálculo da inversa da matriz modal V simplifica-se para $V^{-1} = V^T$.

Exemplo 5.60

Resolver o problema de valor inicial

$$\begin{aligned} C\dot{u} + Bu &= F(t) \\ u(0) &= u_o \end{aligned}$$

onde

$$C = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 7/10 & -1/2 \\ -1 & 6/5 \end{bmatrix}, \quad F(t) = \begin{bmatrix} 20 \\ 0 \end{bmatrix}, \quad u_o = \begin{bmatrix} 70 \\ 70 \end{bmatrix},$$

Multiplicando a equação por C^{-1} , vem o sistema

$$\begin{aligned} \dot{u} &= Au + f(t), \\ A &= \begin{bmatrix} -7/10 & 1/2 \\ 1/2 & -3/5 \end{bmatrix}, \\ f(t) &= C^{-1}F(t) = \begin{bmatrix} 20 \\ 0 \end{bmatrix}. \end{aligned}$$

Os autovalores e autovetores da matriz A são

$$\begin{aligned} \lambda_1 &= -\frac{13 + \sqrt{101}}{20}, \quad v_1 = \begin{bmatrix} 1 \\ \frac{1 + \sqrt{101}}{10} \end{bmatrix}, \\ \lambda_2 &= -\frac{13 - \sqrt{101}}{20}, \quad v_2 = \begin{bmatrix} 1 \\ \frac{1 - \sqrt{101}}{10} \end{bmatrix}. \end{aligned}$$

Tem-se a matriz modal cujas colunas são os autovetores

$$V = \begin{bmatrix} 1 & 1 \\ \frac{1+\sqrt{101}}{10} & \frac{1-\sqrt{101}}{10} \end{bmatrix}$$

e como a matriz A é simétrica,

$$V^{-1} = V^T = \begin{bmatrix} 1 & \frac{1+\sqrt{101}}{10} \\ 1 & \frac{1-\sqrt{101}}{10} \end{bmatrix}$$

Tem-se a solução geral

$$u(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + u_p(t)$$

onde, neste caso, por ser $f(t) = b$ um vetor constante, a solução particular não-homogênea pode ser obtida pelo *método dos coeficientes a determinar*, isto é, supor que é do mesmo tipo do que $f(t)$. Como o termo não-homogêneo é não-constante, substitui-se $u_p(t) = d$ na equação dada $\dot{u}(t) = Au(t) + b$ e decorre o sistema $0 = Ad + b$. Sendo A não-singular, vem

$$d = -A^{-1}b = \begin{bmatrix} -60/17 & -50/17 \\ -50/17 & -70/17 \end{bmatrix} \begin{bmatrix} 20 \\ 0 \end{bmatrix} = \begin{bmatrix} -1200/17 \\ -1000/17 \end{bmatrix}.$$

Para determinar as constantes c_1 e c_2 , utiliza-se o valor inicial

$$u(0) = u_o = c_1 v_1 + c_2 v_2 + u_p(0) = Vc + d.$$

Decorre

$$\begin{aligned} c &= V^{-1}[u_o - d] = V^T[u_o - d] \\ &= \begin{bmatrix} 9/17 + 19/17\sqrt{101} \\ 9/17 - 19/17\sqrt{101} \end{bmatrix}. \end{aligned}$$

A solução do problema é

$$u(t) = \left(\frac{9 + 19\sqrt{101}}{17}\right)e^{-\left(\frac{13+\sqrt{101}}{20}\right)t} \begin{bmatrix} 1 \\ \frac{1+\sqrt{101}}{10} \end{bmatrix} + \left(\frac{9 - 19\sqrt{101}}{17}\right)e^{-\left(\frac{13-\sqrt{101}}{20}\right)t} \begin{bmatrix} 1 \\ \frac{1-\sqrt{101}}{10} \end{bmatrix} + \begin{bmatrix} -\frac{1200}{17} \\ -\frac{1000}{17} \end{bmatrix}.$$

5.4 Métodos não-espectrais

Estes métodos não utilizam autovetores. Serão descritos o método operacional da transformada de Laplace e o método polinomial.

- **Método Operacional**

Aplicando-se a transformada de Laplace em ambos os lados do sistema

$$C \frac{du}{dt} + Bu = F(t)$$

e utilizando-se a propriedade $L(u') = sL(u) - u(0)$ com a condição inicial dada tem-se que

$$(sC + B)U(s) - u(0) = F(s). \quad (5.39)$$

Considerando-se $\Delta(s) = sC + B$, que denota uma função matricial de primeira ordem, na variável s , resulta que

$$U(s) = H(s)u(0) + H(s)F(s), \quad (5.40)$$

onde $H(s) = (sC + B)^{-1}$.

A fim de retornar à variável original do problema, aplica-se a transformada inversa de Laplace.

Deve ser observado que $H(s)$ é o *fator de transferência* para entradas do tipo exponencial

$$F(t) = e^{st}v.$$

Pois, procurando soluções do mesmo tipo

$$u(t) = e^{st}w,$$

e substituindo no sistema (5.39), vem

$$(sC + B)w = v.$$

Assim,

$$w = H(s)v.$$

Exemplo 5.61

Resolver o sistema

$$\begin{aligned} \dot{u}_1 &= -2u_1 + u_2 + 1 \\ \dot{u}_2 &= u_1 - 2u_2 \end{aligned}$$

sujeito as condições iniciais

$$u_1(0) = 0 \quad u_2(0) = 1.$$

Aplicando a transformada de Laplace em cada equação do sistema, vem

$$\begin{aligned}(s+2)U_1(s) - U_2(s) &= \frac{1}{s} \\ -U_1(s) + (s+2)U_2(s) &= 1.\end{aligned}$$

Resolvendo este sistema algébrico, obtém-se

$$\begin{aligned}U_1(s) &= \frac{2}{s(s+3)} \\ U_2(s) &= \frac{s+1}{s(s+3)}\end{aligned}$$

Tomando a inversa da transformada de Laplace, decorre a solução

$$\begin{aligned}u_1(t) &= \frac{2}{3} - \frac{2}{3}e^{-3t} \\ u_2(t) &= \frac{1}{3} + \frac{2}{3}e^{-3t}\end{aligned}$$

Resposta Freqüência

Na prática, é muito importante a determinação de soluções diante termos não-homogêneos de natureza harmônica, isto é

$$C\dot{u} + Bu = \cos(\omega t)a + \operatorname{sen}(\omega t)b.$$

onde a, b são vetores coluna. Na forma normal, tem-se o sistema

$$\begin{aligned}\dot{u} &= Au(t) + f(t), \\ A &= -C^{-1}B, \\ f(t) &= \cos(\omega t)r + \operatorname{sen}(\omega t)w, \\ r &= C^{-1}a, \quad w = C^{-1}b.\end{aligned}$$

Procurando uma solução particular do mesmo tipo que $f(t)$

$$u_p(t) = \cos(\omega t)q + \operatorname{sen}(\omega t)d,$$

com q, d vetores coluna, e substituindo na equação, decorre o sistema

$$-\omega \operatorname{sen}(\omega t)q + \omega \cos(\omega t)d = \cos(\omega t)Aq + \operatorname{sen}(\omega t)Ad + \cos(\omega t)r + \operatorname{sen}(\omega t)w.$$

Igualando os termos em seno e cosseno, vem

$$\begin{aligned}\omega d - Aq &= r \\ -\omega q - Ad &= w\end{aligned}$$

ou, na forma de um sistema matricial bloco

$$\begin{bmatrix} -A & \omega I \\ -\omega I & -A \end{bmatrix} \begin{bmatrix} q \\ d \end{bmatrix} = \begin{bmatrix} r \\ w \end{bmatrix}.$$

Pré-multiplicando a esquerda pela matriz *adjunta*, vem

$$\begin{bmatrix} -A & -\omega I \\ \omega I & -A \end{bmatrix} \begin{bmatrix} -A & \omega I \\ -\omega I & -A \end{bmatrix} \begin{bmatrix} q \\ b \end{bmatrix} = \begin{bmatrix} -A & -\omega I \\ \omega I & -A \end{bmatrix} \begin{bmatrix} r \\ w \end{bmatrix},$$

vem

$$\begin{bmatrix} A^2 + \omega^2 I & 0 \\ 0 & A^2 + \omega^2 I \end{bmatrix} \begin{bmatrix} q \\ b \end{bmatrix} = \begin{bmatrix} -Ar - \omega w \\ \omega r - Aw \end{bmatrix}.$$

Deste modo, se $\lambda = i\omega$ não é autovalor da matriz A^2 , e portanto, $\det[\omega^2 I + A] \neq 0$, segue que

$$q = -(\omega^2 I + A)^{-1}(Ar + \omega w),$$

$$w = (\omega^2 I + A^2)^{-1}(\omega r - Aw).$$

A solução particular $u_p(t)$ a uma entrada harmônica vem a ser a resposta freqüência

$$u_p(t) = (\omega^2 I + A^2)^{-1} [-(Ar + \omega w)\cos(\omega t) + (\omega r - Aw)\sin(\omega t)] \quad (5.41)$$

Observa-se que as amplitudes r e w da entrada com freqüência ω , são modificadas pelo fator

$$H(i\omega) = (\omega^2 I + A)^{-1} \quad (5.42)$$

chamado de função freqüência.

Exemplo 5.62

Obter a resposta freqüência e função freqüência do sistema

$$C\dot{u} + Bu(t) = F(t),$$

onde

$$C = \begin{bmatrix} 3 & -4 \\ -4 & 8 \end{bmatrix}, \quad B = \begin{bmatrix} 5 & -3 \\ -3 & 7 \end{bmatrix}, \quad F = \begin{bmatrix} \cos(t) \\ \sin(2t) \end{bmatrix}$$

O sistema dado na forma normal ou de estado, está dado por $\dot{u} = Au + f(t)$ com

$$A = \begin{bmatrix} -\frac{7}{2} & -\frac{1}{2} \\ -\frac{11}{8} & -\frac{113}{100} \end{bmatrix}, \quad f = \begin{bmatrix} \cos(t) + \frac{1}{2}\sin(2t) \\ \frac{1}{2}\cos(t) + \frac{3}{8}\sin(2t) \end{bmatrix}$$

O termo não-homogêneo pode ser escrito na forma

$$f(t) = \cos(t) \begin{bmatrix} 1 \\ \frac{1}{2} \end{bmatrix} + \sin(2t) \begin{bmatrix} \frac{1}{2} \\ \frac{3}{8} \end{bmatrix}$$

Pelo princípio da superposição linear, a resposta frequência será

$$u_p(t) = u_1(t) + u_2(t)$$

onde $u_1(t)$ é a resposta à frequência $\omega_1 = 1$ e $u_2(t)$ é a resposta à frequência $\omega_2 = 2$. Observe-se que a amplitude w_1 do termo seno é nula para a primeira frequência e que a amplitude r_2 do termo cosseno é nula para a segunda frequência, decorre de (5.41) que

$$u_p(t) = (\omega_1^2 i + A^2)^{-1} [-Ar_1 \cos(\omega_1 t) + \omega r_1 \sin(\omega_1 t)] \\ + (\omega_2^2 i + A)^{-1} [-\omega w_2 \cos(\omega_2 t) - Aw_2 \sin(\omega_2 t)]$$

e com o uso de software simbólico Maple, tem-se a resposta aproximada

$$u_p(t) = \begin{bmatrix} 0.249 \cos(t) + .0679 \sin(t) + .0490 \cos(2. t) + .104 \sin(2. t) \\ .119 \cos(t) + .0230 \sin(t) + - .0737 \cos(2. t) + .0751 \sin(2. t) \end{bmatrix}$$

Aqui a função frequência é dada por

$$H(i\omega) = (i\omega + A)^{-1} = \begin{bmatrix} \frac{8i\omega + 7}{-8\omega^2 + 37i\omega + 26} & \frac{4i\omega + 3}{-8\omega^2 + 37i\omega + 26} \\ \frac{4i\omega + 3}{-8\omega^2 + 37i\omega + 26} & \frac{3i\omega + 5}{-8\omega^2 + 37i\omega + 26} \end{bmatrix}$$

Observe-se que na entrada original $F(t)$ tem-se as frequências *desacopladas*, enquanto em $f(t)$ estão acopladas devido a multiplicação pela inversa da matriz C . A resposta frequência reflete o acoplamento das frequências.

5.4.1 Método Polinomial

A função exponencial de uma matriz está incluída na classe de funções matriciais da forma

$$f(A) = \sum_{k=0}^{\infty} f^k(0) \frac{A^k}{k!}, \quad (5.43)$$

obtidas a partir de séries de potências convergentes do tipo

$$f(x) = \sum_{k=0}^{\infty} f^k(0) \frac{x^k}{k!}, \quad (5.44)$$

pela substituição da variável x pela variável matricial A de ordem n . Este tipo de funções podem ser calculadas com o uso do teorema de Cayley-Hamilton.

Da divisão

$$\frac{f(x)}{p(x)} = q(x) + \frac{r(x)}{p(x)}, \quad (5.45)$$

resulta que

$$f(x) = p(x)q(x) + r(x) \quad (5.46)$$

onde

$$p(x) = \det[Ix - A]$$

corresponde ao polinômio característico da matriz A de grau n e

$$r(x) = \sum_{k=0}^{n-1} a_{n-k}x^{n-k}, \quad (5.47)$$

um polinômio de grau $n - 1$. Substituindo $z=A$ em (5.46), segue que

$$f(A) = q(A)p(A) + r(A) = r(A) \quad (5.48)$$

uma vez que, pelo teorema de Cayley-Hamilton, $p(A) = 0$. Assim, o cálculo do valor da soma para $f(A)$ é reduzido a determinação do polinômio $r(A)$

$$f(A) = r(A) = a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \cdots + a_1A + a_0I, \quad (5.49)$$

onde os escalares $a_{n-1}, a_{n-2}, \dots, a_1, a_0$ são determinados conforme segue.

Para cada autovalor λ_j de A , tem-se de (5.46) que

$$f(\lambda_j) = r(\lambda_j), \quad (5.50)$$

pois $p(\lambda_j) = 0$.

Se λ_j é um autovalor duplo de A , considere-se

$$p(\lambda_j) = 0, \quad \dot{p}(\lambda_j) = 0, \quad \ddot{p}(\lambda_j) \neq 0. \quad (5.51)$$

Derivando-se (5.46) decorre

$$\dot{f}(\lambda_j) = p(\lambda_j)\dot{q}(\lambda_j) + \dot{p}(\lambda_j)q(\lambda_j) + \dot{r}(\lambda_j) = \dot{r}(\lambda_j).$$

Em geral, para cada autovalor λ de multiplicidade k , sendo $k \geq 1$, formula-se as seguintes equações, envolvendo as derivadas de $f(\lambda)$ e $r(\lambda)$ com respeito a λ :

$$\begin{aligned} \dot{f}(\lambda)|_{\lambda=\lambda_j} &= \dot{r}(\lambda)|_{\lambda=\lambda_j} \\ \ddot{f}(\lambda)|_{\lambda=\lambda_j} &= \ddot{r}(\lambda)|_{\lambda=\lambda_j} \\ \vdots &= \vdots \\ f^{k-1}(\lambda)|_{\lambda=\lambda_j} &= r^{k-1}(\lambda)|_{\lambda=\lambda_j} \end{aligned} \quad (5.52)$$

Resolvendo o sistema de n equações (5.52) para os s autovalores distintos λ_j com multiplicidade m_j , $j = 1 : m$, tais que $m_1 + m_2 + \cdots + m_s = n$; obtém-se, assim, os valores dos coeficientes a_k e desta forma tem-se a expressão para $f(A)$ através da equação (5.49).

5.5 Comportamento assintótico

A maneira como as soluções se comportam quando t cresce indefinidamente é de interesse na teoria e nas aplicações. Como foi observado, as equações lineares produzem soluções semelhantes aos termos não-homogêneos, porém acompanhadas de algumas deformações devido a interação do sistema com a excitação. Se todas as soluções do sistema homogêneo tendem para zero quando t vai para o infinito, então o sistema é dito *assintoticamente estável* e as soluções homogêneas são ditas *respostas transientes*. A solução particular não-homogênea é dita *resposta permanente*.

Como as soluções do sistema $Cu' + Bu = 0$ são da forma $u(t) = h(t)Au(0)$, é claro que o seu comportamento é o da solução fundamental $h(t)$. Com o uso do método espectral, para o caso do problema de autovalor $\lambda A + B$ ser não defeituoso, a representação

$$h(t) = Ve^{\Lambda t}V^{-1}A^{-1} = \sum_{k=0}^n e^{\lambda_k t} v_k w_k^T A^{-1}$$

mostra que o comportamento de $h(t)$ segue o dos termos exponenciais. Assim,

$$|e^{\lambda_k t}| = |e^{Re\lambda_k t} e^{Im\lambda_k t}| = e^{Re\lambda_k t}.$$

Decorre daí que $h(t)$ é *assintoticamente estável se, e somente se todos os autovalores possuem parte real negativa*.

No caso geral $u' = Au$, $A = -C^{-1}B$, uma matriz qualquer, pode ser utilizada a decomposição matricial de Jordan: existe uma matriz não-singular V tal que $A = VJV^{-1}$, $J = D + N$, $DN = ND$ com D uma matriz diagonal cujos elementos diagonais são os autovalores de C , e N uma matriz nilpotente ($N^m = 0$, para algum $m \leq n$). Então, $h(t) = Ve^{tD}e^{tN}V^{-1}$. Observa-se que o comportamento de $h(t)$ com (5.13) vem a ser o de e^{Dt} , isto é, da parte real dos autovalores como foi feito no caso não-defeituoso pois, por N ser uma matriz nilpotente a sua exponencial é reduzida a um polinômio.

5.6 Integração de um modelo de rastreo de aeronaves rígidas

A fim de ilustrar a aplicação dos diferentes métodos espectrais e não-espectrais na solução de sistemas lineares de primeira ordem, apresenta-se um modelo na área de dinâmica de controle de sistemas. [Lyshevski], [Bryson].

A dinâmica de muitos aviões pode ser modelada considerando-os como corpos rígidos que sofrem ação de forças do tipo gravitacional, aerodinâmica e de propulsão conforme mostra a Fig. 5.2.

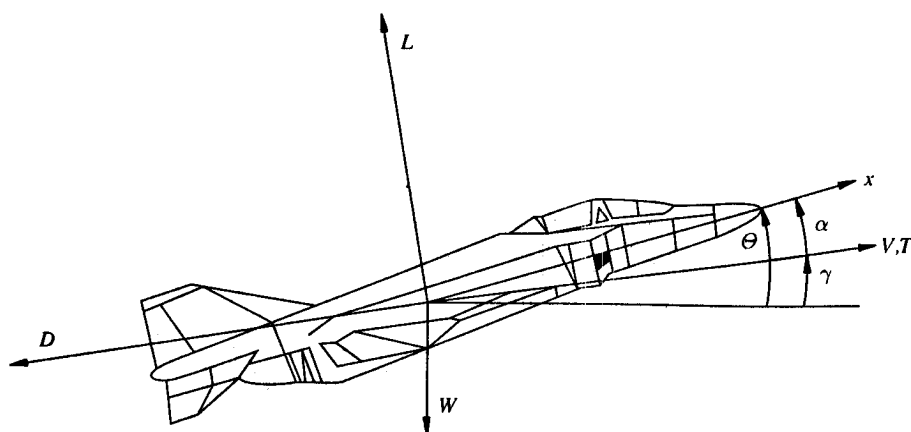


Figura 5.2 – Forças que agem em uma aeronave

Neste modelo, as forças atuantes são a elevação causada pela aerodinâmica das asas, que atua perpendicularmente à velocidade denotada por L ; o empuxo T produzido pelos motores do avião na direção do vetor velocidade; as forças de arraste, que correspondem a forças aerodinâmicas que se opõem ao empuxo, normalmente são proporcionais à velocidade dadas por D e o peso W , força devido a gravidade que atua perpendicularmente ao horizonte.

Os movimentos das aeronaves podem ser descritos fornecendo-se a posição, a velocidade do centro de massa, a orientação e a velocidade angular de um conjunto de eixos fixos no corpo em relação a um conjunto de eixos de referência. Usualmente é conveniente escolher eixos de referência fixos em relação à terra. Considerando-se que a velocidade dos aviões é pequena se comparada com a velocidade orbital, pode ser assumido que os eixos de referência fixos à terra podem ser aproximados como eixos inerciais.

As equações dinâmicas do movimento são formuladas em termos da aceleração inercial $\dot{\mathbf{v}}$ e da aceleração angular ω_B , cujas componentes da velocidade do centro de massa e velocidade angular em relação aos eixos de referência são denotadas por $\mathbf{v} = [u, v, w]$ e $\omega_B = [p, q, r]$, respectivamente.

As equações cinemáticas relacionam as taxas de localização do centro de massa do avião no sistema de eixos de referência $[x, y, z]$ com a velocidade do corpo e, as taxas dos ângulos de Euler $[\psi, \theta, \phi]$ com as componentes da velocidade angular.

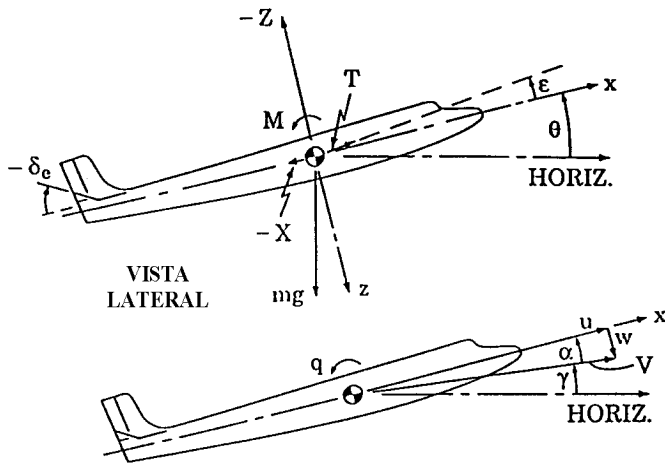


Figura 5.3 – Nomenclatura para movimentos longitudinais de uma aeronave

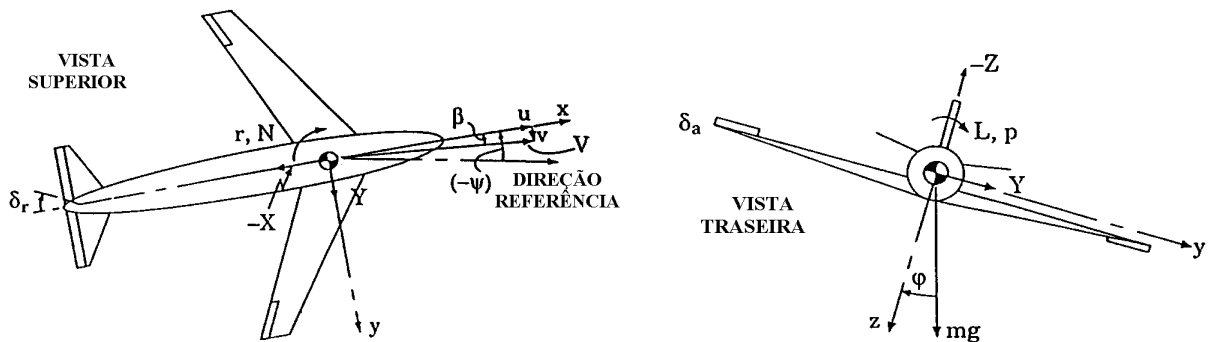


Figura 5.4 – Nomenclatura para movimentos laterais de uma aeronave

Os movimentos longitudinais e laterais de uma aeronave são descritos empregando-se, usualmente, eixos alinhados aos eixos do avião, conforme Figs. 5.3 e 5.4. Tais movimentos são monitorados pelo deslocamento de superfícies de controle sobre asas e cauda (ailerões, flaps e lemes) que podem trocar a inclinação, de modo que as forças aerodinâmicas mudam a orientação da ponta do avião. A Fig. 5.5 ilustra o comportamento do ângulo de inclinação sob a ação de diferentes deflexões nos ailerões.

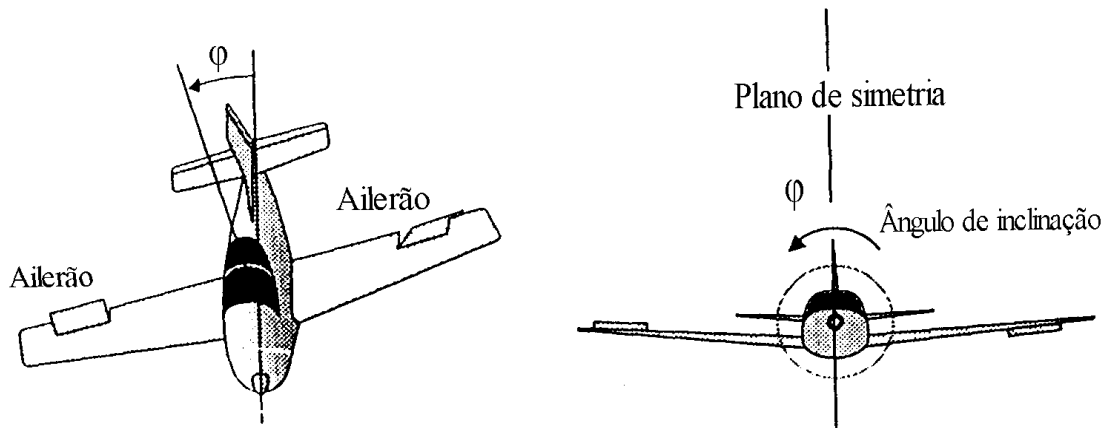


Figura 5.5 – Controle do ângulo de inclinação através dos ailerões

Com o objetivo de ilustrar problemas de controle de rastreamento ou de dinâmica longitudinal e lateral de aeronaves será considerado neste trabalho, conforme [Lyshevski, 2001], um modelo linearizado para um avião caça.

O vetor velocidade é denotado por v [m/s], sendo que o ângulo de ataque dado por α [rad], fornece a direção do vetor velocidade v , ou seja, corresponde ao ângulo formado entre o vetor velocidade do avião e o eixo x de orientação do seu corpo. O ângulo de inclinação longitudinal, formado entre o horizonte e o eixo x do avião denotado por θ [rad], o ângulo de giro em torno do eixo paralelo ao eixo normal do avião dado por φ [rad] e o ângulo de rolamento longitudinal dado por ψ [rad] são denominados na literatura de ângulos de Euler; sendo que as respectivas taxas de variações serão denotadas por q , p e r [rad/s]. O ângulo de deslizamento lateral será dado por β [rad], as deflexões dos estabilizadores horizontais direito e esquerdo por $\delta_{HR}(t)$ e $\delta_{HL}(t)$ [rad], respectivamente; os parâmetros $\delta_{FR}(t)$ e $\delta_{FL}(t)$ [rad] corresponderão às deflexões dos flaps direito e esquerdo, respectivamente; $\delta_C(t)$ e $\delta_R(t)$ [rad] às deflexões do comando dianteiro e do leme, respectivamente.

A altitude do avião está relacionada com o seu ângulo de inclinação, o qual indica se este está direcionado para cima ou para baixo. Normalmente o avião ascende suavemente, de modo que a aerodinâmica de suas asas produz a elevação suficiente para vencer a gravidade. Não se tem nível de vôo quando $\theta = 0$ e sim quando $\theta = \alpha$. A Fig. 5.2 mostra o avião ascendendo pois, tem-se $\theta > \alpha$. Para análise das condições de vôo a quantidade de elevação aumenta quando α também aumenta.

A dinâmica lateral-longitudinal de uma aeronave pode ser estudada aplicando-se o modelo de espaço de estado, resultando em

$$\dot{x}(t) = Ax(t) + Bu(t), \quad (5.53)$$

onde $x(t) = x$ é o vetor de estado e, $u(t) = u$, o controle de entrada, denotados por

$$x = \begin{bmatrix} v \\ \alpha \\ q \\ \theta \\ \beta \\ p \\ r \\ \varphi \\ \psi \end{bmatrix} \quad \text{e} \quad u = \begin{bmatrix} \delta_{HR} \\ \delta_{HL} \\ \delta_{FR} \\ \delta_{FL} \\ \delta_C \\ \delta_R \end{bmatrix}. \quad (5.54)$$

Considerando-se os coeficientes matriciais A e B dados por

$$A = \begin{bmatrix} -0.016 & 8.4 & -0.9 & -9.6 & -1.5 & -0.27 & -0.086 & 0 & 0 \\ -0.003 & -1.2 & 1 & 0 & 0.08 & 0.062 & 0.009 & 0 & 0 \\ -0.0001 & 3.9 & -0.85 & 0 & 0.017 & 0.0038 & 0.04 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.003 & 0.15 & 0.02 & 0.97 & -0.56 & 0.13 & -0.91 & 0 & 0 \\ -0.00001 & 0.71 & 0.03 & 0.01 & -48 & -3.5 & 0.22 & 0 & 0 \\ 0.00001 & -0.94 & 0.06 & 0.005 & 9.2 & -0.028 & -0.51 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 0.12 & 0.12 & -0.38 & -3.8 & 0 & 0 \\ -0.16 & -0.16 & -0.27 & -0.27 & 0 & 0 \\ -9.5 & -9.5 & -2.5 & -2.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.019 & -0.019 & -0.001 & 0.001 & 0.42 & 0.053 \\ -2.9 & 2.9 & -3.1 & 3.1 & 0.73 & 0.92 \\ 3.1 & -3.1 & 0.78 & -0.78 & 0.61 & -0.45 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

o modelo é instável. Então o modelo é estabilizado com o uso da técnica de desenho do regulador quadrático, que determina uma matriz de ganho ótima. Para uma formulação detalhada veja-se [Rohrs, Bryson, Lyshevski]. As simulações numéricas para a obtenção da matriz de ganho foram realizadas usando-se o software MATLAB. Donde resulta o sistema de estado aumentado estável representado por

$$\dot{\mathcal{X}}_{\Sigma}(t) = \mathcal{A}\mathcal{X}_{\Sigma}(t) + f(t), \quad (5.55)$$

com a saída, $\mathcal{X}_{\Sigma}(t) = \mathcal{X}_{\Sigma}$ e, a entrada, $f(t) = f$, sendo vetores de ordem 12 dados por

$$\mathcal{X}_{\Sigma} = \begin{bmatrix} v \\ \alpha \\ q \\ \theta \\ \beta \\ p \\ r \\ \varphi \\ \psi \\ x_{\theta}^{ref} \\ x_{\phi}^{ref} \\ x_{\psi}^{ref} \end{bmatrix} \quad \text{e} \quad f = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \quad (5.56)$$

O coeficiente matricial

$$A = [\mathcal{A}_1 \quad \mathcal{A}_2],$$

do sistema tem como valores

$$\mathcal{A}_1 = \begin{bmatrix} -0.1204 & 7.8278 & -0.8666 & -8.9564 & -1.5962 & -0.2636 \\ -0.0541 & -1.6149 & 0.7321 & -1.4244 & -0.82660.1466 & \\ 0.2823 & -0.2155 & -12.8651 & -74.7071 & -35.7809 & 3.4039 \\ 0 & 0 & 1.0000 & 0 & 0 & 0 \\ 0.0040 & 0.0368 & 0.0047 & 0.8903 & -3.6061 & 0.3003 \\ 0.0503 & -0.2064 & -0.0610 & -0.5143 & -47.4985 & -3.4888 \\ -0.0176 & 0.1756 & 0.2366 & 2.1404 & 20.0887 & -2.2848 \\ 0 & 0 & 0 & 0 & 0 & 1.0000 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1.0000 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathcal{A}_2 = \begin{bmatrix} -0.0721 & 0.0418 & -0.0964 & 0.2968 & -0.0056 & 0.0686 \\ 0.1630 & 0.5441 & -0.2064 & 4.6170 & -0.7051 & 0.5457 \\ 6.1769 & 21.8555 & -6.4212 & 186.8032 & -29.4511 & 21.2980 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -0.7924 & 1.1833 & -2.4891 & -1.6853 & -1.2657 & 3.3541 \\ 8.3585 & -3.4970 & 28.1265 & -3.8232 & 27.0789 & -56.5742 \\ -11.5978 & -12.6074 & -34.7830 & 1.0057 & 10.0947 & 58.8791 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1.0000 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1.0000 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1.0000 & 0 & 0 & 0 \end{bmatrix}.$$

Com o objetivo de estudar o problema de controle do erro de rastreamento considera-se o sistema sujeito a condições iniciais nulas e deste modo, calcula-se a resposta forçada do mesmo.

O sistema de primeira ordem na formulação de espaço de estado dado pela equação (5.55) foi resolvido utilizando-se os métodos espectral e o método não-espectral operacional, cujas variáveis fornecem a dinâmica dos movimentos laterais e longitudinais do problema aerodinâmico. Os resultados obtidos através destas técnicas são apresentados nas Figs. 5.6, 5.7 e 5.8.

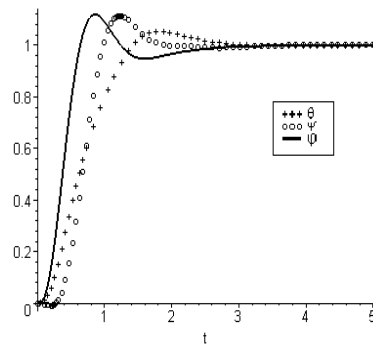


Figura 5.6 – Ângulos de Euler

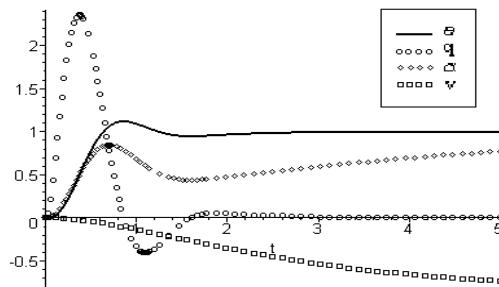


Figura 5.7 – Dinâmica Lateral

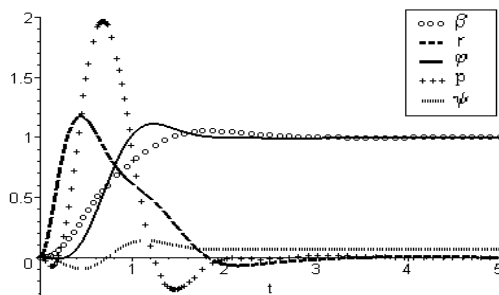


Figura 5.8 – Dinâmica Longitudinal

O erro de rastreamento é dado por $e(t) = r_e(t) - y(t)$ onde $r_e(t)$ é uma entrada de referência associada aos ângulos de Euler, denotada por

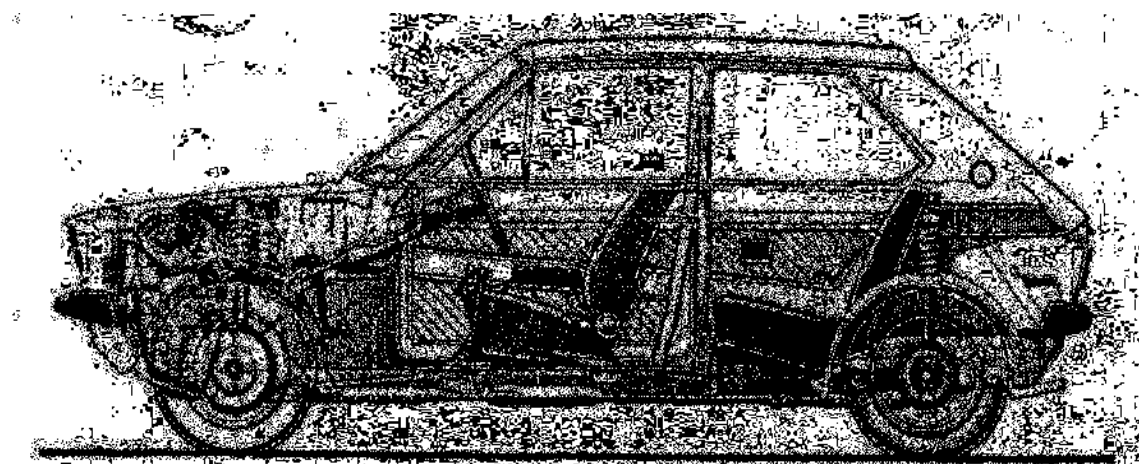
$$r_e(t) = \begin{bmatrix} r_\theta(t) \\ r_\phi(t) \\ r_\psi(t) \end{bmatrix}$$

e

$$y = Hx = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \alpha \\ q \\ \theta \\ \beta \\ p \\ r \\ \varphi \\ \psi \end{bmatrix}. \quad (5.57)$$

5.7 Um modelo de segunda ordem para a suspensão de um carro

Aborda-se nesta seção um modelo de suspensão de um veículo. Tais modelos, embora apresentados de forma simplificada, são bastante usados na teoria de controle de sistemas dinâmicos, sendo que uma das aplicações desse tipo de formulação é a análise dos efeitos das vibrações sobre os veículos em relação ao conforto dos passageiros e à capacidade dos mesmos manterem-se na estrada, isto é, a eficiência do sistema de suspensão do carro. Estes efeitos podem ser observados a partir da variação de diferentes parâmetros tais como os coeficientes da suspensão, as irregularidades da estrada e a posição do assento, entre outros.



Considere-se um modelo simplificado da metade de um carro, conforme mostra a Fig. 5.9. De acordo com [Lyshevski] considere-se o sistema sem movimento de rolamentos laterais.

Figura 5.9 – Modelo da suspensão da metade de um carro

Os deslocamentos y_1 e y_5 são as entradas; os parâmetros m_1 e m_2 representam as massas equivalentes das rodas dianteira e traseira, respectivamente; enquanto m_3 representa a massa da estrutura principal do automóvel e m_4 a massa equivalente dos assentos e dos passageiros. As propriedades de elasticidade e dissipação de energia (atrito viscoso) dos pneus são dadas por (k_{s1}, k_{s3}) e $(B_{\nu1}, B_{\nu3})$, respectivamente; para o sistema de suspensão do carro representa-se por (k_{s2}, k_{s4}) e $(B_{\nu2}, B_{\nu4})$, respectivamente. A rigidez dos assentos é expressa por k_{s5} .

Aplicando-se a 2ª Lei de Newton às rodas obtém-se

$$\begin{cases} m_1 \ddot{y}_2 &= k_{s1}(y_1 - y_2) + B_{\nu1}(\dot{y}_1 - \dot{y}_2) + k_{s2}(y_3 - y_2) + B_{\nu2}(\dot{y}_3 - \dot{y}_2) \\ m_2 \ddot{y}_6 &= k_{s3}(y_5 - y_6) + B_{\nu3}(\dot{y}_5 - \dot{y}_6) + k_{s4}(y_7 - y_6) + B_{\nu4}(\dot{y}_7 - \dot{y}_6) \end{cases} \quad (5.58)$$

Denotando-se por l_1 e l_2 as distâncias das extremidades esquerda e direita ao centro de massa do carro, respectivamente e, supondo ângulos pequenos, as relações geométricas relativas ao deslocamento do centro de massa da estrutura do carro e o ângulo de giro podem ser expressas por

$$y_{CM} = y_3 + \frac{l_1}{l_1 + l_2}(y_7 - y_3),$$

$$y_s = y_3 + \frac{l_2}{l_1 + l_2}(y_7 - y_3),$$

$$\omega = \frac{y_7 - y_3}{l_1 + l_2}.$$

Assumindo-se forças sobre a massa m_3 da estrutura do carro, resulta

$$m_3\ddot{y}_{CM} = k_{s2}(y_2 - y_3) + B_{\nu2}(\dot{y}_2 - \dot{y}_3) + k_{s4}(y_6 - y_7) + B_{\nu4}(\dot{y}_6 - \dot{y}_7) + k_{s5}(y_4 - y_s). \quad (5.59)$$

Utilizando-se a lei rotacional de Newton tem-se a equação

$$J\ddot{\omega} = -l_1k_{s2}(y_2 - y_3) - l_1B_{\nu2}(\dot{y}_2 - \dot{y}_3) + l_2k_{s4}(y_6 - y_7) + l_2B_{\nu4}(\dot{y}_6 - \dot{y}_7) - (l_1 - l_2)k_{s5}(y_4 - y_s). \quad (5.60)$$

Finalmente, assumindo forças sobre a massa dos passageiros e dos assentos, obtém-se a expressão

$$m_4\ddot{y}_4 = k_{s5}(y_s - y_4). \quad (5.61)$$

As equações diferenciais ordinárias de segunda ordem acopladas dadas por (5.58), (5.59), (5.60) e (5.61) definem o modelo matemático para o sistema de parâmetros concentrados da suspensão da metade de um carro, o qual é denotado matricialmente por

$$A_2 \frac{d^2 y}{dt^2}(t) + A_1 \frac{dy}{dt}(t) + A_0 y(t) = F(t) \quad (5.62)$$

onde

$$F(t) = B_0 u(t) + B_1 \frac{du}{dt}(t). \quad (5.63)$$

Os vetores correspondentes a entrada $u(t)$ e saída $y(t)$ do sistema são de segunda e quinta ordem, respectivamente, denotados por

$$y(t) = \begin{bmatrix} y_2 \\ y_6 \\ y_{CM} \\ \omega \\ y_4 \end{bmatrix} \quad \text{e} \quad u(t) = \begin{bmatrix} y_1 \\ y_5 \end{bmatrix}.$$

Os coeficientes matriciais A_2 , A_1 , A_0 , B_0 e B_1 são da forma

$$A_2 = \begin{bmatrix} m_1 & 0 & 0 & 0 & 0 \\ 0 & m_2 & 0 & 0 & 0 \\ 0 & 0 & m_3 & 0 & 0 \\ 0 & 0 & 0 & J & 0 \\ 0 & 0 & 0 & 0 & m_4 \end{bmatrix},$$

$$A_1 = \begin{bmatrix} B_{\nu1} + B_{\nu2} & 0 & -B_{\nu2} & B_{\nu2}l_1 & 0 \\ 0 & B_{\nu3} + B_{\nu4} & -B_{\nu4} & -B_{\nu4}l_2 & 0 \\ -B_{\nu2} & -B_{\nu4} & B_{\nu2} + B_{\nu4} & B_{\nu4}l_2 - B_{\nu2}l_1 & 0 \\ B_{\nu2}l_1 & -B_{\nu4}l_2 & B_{\nu4}l_2 - B_{\nu2}l_1 & B_{\nu2}l_1^2 + B_{\nu4}l_2^2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_0 = \begin{bmatrix} k_{s1} + k_{s2} & 0 & -k_{s2} & k_{s2}l_1 & 0 \\ 0 & k_{s3} + k_{s4} & -k_{s4} & -k_{s4}l_2 & 0 \\ -k_{s2} & -k_{s4} & k_{s2} + k_{s4} + k_{s5} & k_{s4}l_2 - k_{s2}l_1 + k_{s5}(l_2 - l_1) & -k_{s5} \\ k_{s2}l_1 & -k_{s4}l_2 & k_{s4}l_2 - k_{s2}l_1 + k_{s5}(l_2 - l_1) & k_{s2}l_1^2 + k_{s4}l_2^2 - k_{s5}(l_2 - l_1)^2 & -k_{s5}(l_2 - l_1) \\ 0 & 0 & -k_{s5} & -k_{s5}(l_2 - l_1) & k_{s5} \end{bmatrix},$$

$$B_0 = \begin{bmatrix} k_{s1} & 0 \\ 0 & k_{s3} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \text{e} \quad B_1 = \begin{bmatrix} B_{\nu 1} & 0 \\ 0 & B_{\nu 3} \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Como pode ser observado, neste problema os coeficientes matriciais A_j são simétricos.

Supondo-se que a entrada no sistema se dá apenas na posição dos pneus de acordo com as irregularidades da estrada, considere-se, para efeitos de ilustração, entradas do tipo oscilatórias nas duas rodas dadas por

$$y_1 = y_5 = \begin{cases} 0, & 0 \leq t < a \\ A - A \cos(2\pi\omega t), & a \leq t < b \\ 0, & t \geq b \end{cases}$$

com amplitude máxima de A . Esta entrada é mostrada na Fig. 5.10 com $A=0.075$, $a=0.25$, $b=0.5$, $\omega = 4$.

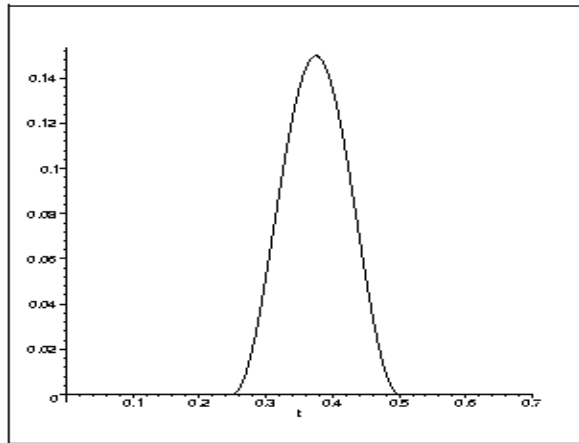


Figura 5.10 – Entrada tipo oscilatória nas rodas

5.8 A equação $Mu'' + Cu' + Ku = F(t)$

No estudo do sistema matricial de segunda ordem

$$Mu'' + Cu' + Ku = F(t) \quad (5.64)$$

com coeficientes M , C , K matrizes constantes de ordem $n \times n$, u , F vetores $n \times 1$ e M não-singular, se fará uso dos três princípios das equações lineares, citados no início deste capítulo.

A *solução matricial fundamental* ou *resposta impulso* ou *solução dinâmica* da equação $Mu'' + Cu' + Ku = F(t)$ é definida como sendo a matriz $h(t)$ de ordem $n \times n$ que satisfaz o problema de valor inicial homogêneo

$$\begin{aligned} Mh'' + Ch' + Kh &= 0, \\ h(0) &= 0, \quad Mh'(0) = I, \end{aligned} \quad (5.65)$$

onde I denota a matriz identidade $n \times n$ e 0 a matriz nula.

Do Princípio da Decomposição, tem-se que a solução da equação diferencial não-homogênea

$$Mu'' + Bu' + Ku = F(t)$$

pode ser escrita na forma

$$u(t) = u_h(t) + u_p(t)$$

onde u_h , é solução da equação homogênea

$$Mu'' + Cu' + Ku = 0 \quad (5.66)$$

e u_p , uma solução particular da equação não-homogênea.

Para qualquer vetor constante c de ordem $n \times 1$, tem-se que $u(t) = h(t)c$ é solução homogênea. Derivando (5.66) segue que $h'(t)$ também é solução. Assim, pelo segundo princípio

$$u(t) = h(t)a + h'(t)b$$

é solução homogênea. As matrizes $h(t)$, $h'(t)$ formam uma base matricial, denominada *base dinâmica*; as constantes a , b podem ser determinadas de maneira única em termos de valores iniciais arbitrários $u(0)$ e $u'(0)$:

$$\begin{aligned} u(0) &= h(0)a + h'(0)b = M^{-1}b \rightarrow b = Mu(0) \\ u'(0) &= h'(0)a + h''(0)b = M^{-1}a - M^{-1}Ch'(0)b \rightarrow a = Mu'(0) + Cu(0). \end{aligned}$$

Decorre daí que

$$u(t) = [h'(t)M + h(t)C]u_o + h(t)Mu'_o \quad (5.67)$$

satisfaz o problema de valor inicial homogêneo

$$\begin{aligned} Mu'' + Cu' + Ku &= 0, \\ u(0) &= u_o, \quad u'(0) = u'_o. \end{aligned} \quad (5.68)$$

As matrizes

$$h_o(t) = h'(t)M + h(t)C, \quad (5.69)$$

$$h_1(t) = h(t)M \quad (5.70)$$

formam outra base para as soluções homogêneas. Ela é dita *normalizada*, devido que

$$h_o(0) = I, \quad h'_o(0) = 0, \quad (5.71)$$

$$h_1(0) = 0, \quad h'_1(0) = I. \quad (5.72)$$

Uma solução particular não-homogênea é fornecida pela *resposta forçada*. Ela possui condições iniciais nulas em $t = 0$ e é dada pela integral

$$u(t) = \int_0^t h(t-\tau)F(\tau)d\tau.$$

De fato, pela regra de Leibniz e os valores iniciais de $h(t)$, vem

$$\begin{aligned} u'(t) &= h(t-\tau)F(\tau)|_{\tau=t} + \int_0^t h'(t-\tau)F(\tau)d\tau = \int_0^t h'(t-\tau)F(\tau)d\tau \\ u''(t) &= h'(t-\tau)F(\tau)|_{\tau=t} + \int_0^t h''(t-\tau)F(\tau)d\tau = h'(0)F(t) + \int_0^t h''(t-\tau)F(\tau)d\tau \end{aligned}$$

Assim,

$$Mu'' + Cu' + Ku = Mh'(0)F(t) + \int_0^t [Mh''(t-\tau) + Ch'(t-\tau) + Kh(t-\tau)]F(\tau)d\tau = F(t)$$

Dessas relações é claro que $u(0) = 0$, $u'(0) = 0$.

Assim, a solução do problema de valor inicial

$$\begin{aligned} Mu'' + Cu' + Ku &= F(t) \\ u(0), \quad u'(0) &\text{ dados} \end{aligned} \quad (5.73)$$

é dada pela fórmula de variação de parâmetros

$$u(t) = [h'(t)M + h(t)C]u(0) + h(t)Mu'(0) + \int_0^t h(t-\tau)F(\tau)d\tau. \quad (5.74)$$

5.8.1 O método de Cauchy

A solução matricial fundamental $h(t)$ pode, em princípio, ser descrita por uma série de Taylor

$$h(t) = \sum_{j=0}^{\infty} h_j \frac{t^j}{j!}, \quad (5.75)$$

onde $h_j = h^{(j)}(0)$. Substituindo em (5.66), e utilizando (5.117), obtém-se a equação recursiva

$$\begin{aligned} Mh_{j+2} + Ch_{j+1} + Kh_j &= h_{j+2}M + h_{j+1}C + h_jK = 0, \\ Mh_1 &= I, \quad h_0 = 0. \end{aligned} \quad (5.76)$$

Na resolução de (5.76), ser são considerados os casos:

1. Não-conservativo (clássico e não-clássico) $C \neq 0$
2. Conservativo, $C=0$.

Infelizmente, devido ao fato de que, em geral, os coeficientes matriciais M , C , e K não necessariamente comutam, a obtenção analítica de h_k não é tarefa simples. Porém, o caso conservativo, ou não-amortecido $C = 0$, pode ser facilmente resolvido pelo método de Cauchy.

Quando $\mathbf{C} = 0$, a equação matricial torna-se

$$Mu'' + Ku = F(t) \quad (5.77)$$

e é dita não-amortecida ou conservativa. Neste caso, o problema de valor inicial

$$\begin{aligned} Mh_{j+2} + Kh_j &= h_{j+2}M + h_jK = 0, \\ Mh_1 &= I, \quad h_0 = 0, \end{aligned} \quad (5.78)$$

pode ser resolvido por indução. Tem-se que

$$h_{2k} = 0, \quad h_{2k+1} = (-1)^k (M^{-1}K)^k M^{-1}. \quad (5.79)$$

Desta maneira, a solução fundamental é dada por

$$h(t) = \sum_{k=0}^{\infty} (-1)^k (M^{-1}K)^k M^{-1} \frac{t^{2k+1}}{(2k+1)!}.$$

e em analogia ao caso escalar tem-se que

$$h(t) = \frac{\text{sen}(\sqrt{M^{-1}K}t)}{\sqrt{M^{-1}K}M^{-1}}, \quad (5.80)$$

onde o símbolo funcional para $h(t)$ tem como sentido sendo uma série de potências, bem definida para qualquer par de matrizes M, K , com M não-singular.

No caso geral, multiplicando por M^{-1} a equação

$$Mu'' + Cu' + Ku = 0$$

vem uma equação do tipo

$$u'' + Bu' + Au = 0, \quad B = M^{-1}C, \quad A = M^{-1}K$$

Se os coeficientes B,C comutarem (BC=CB), então, a solução fundamental $h(t)$ desta última equação é dada por

$$h(t) = e^{-Bt/2} \frac{\text{senh}(\sqrt{B^2 - 4A}t/2)}{\sqrt{B^2 - 4A}/2} \quad (5.81)$$

onde

$$\frac{\text{senh}(\sqrt{B^2 - 4A}t/2)}{\sqrt{B^2 - 4A}/2} = \sum_{k=0}^{\infty} (-1)^k ((B^2 - A)/4)^k \frac{t^{2k+1}}{(2k+1)!}$$

Quando os coeficientes não comutarem, o que ocorre na maioria das vezes, o processo iterativo para obter os coeficientes h_k torna-se dificultoso e sua identificação resulta praticamente inviável.

5.8.2 O método espectral

O método espectral para a resolução de sistemas lineares de segunda ordem é muito utilizado com sistemas conservativos e com sistemas não-conservativos em que a matriz de atrito C possui uma estrutura particular. Em geral, funciona bem com sistemas que possuem autovalores distintos.

5.8.3 Caso homogêneo

Considere-se a equação homogênea

$$Mu'' + Cu' + Ku = 0.$$

Por substituição direta na equação, a procura de soluções do tipo exponencial

$$u = e^{\lambda t}v, \quad v \neq 0,$$

segue a equação algébrica ou *problema de autovalor*

$$[\lambda^2 M + \lambda C + K]v = 0. \quad (5.82)$$

Esta equação possui solução v não-nula se o determinante do sistema é nulo, ou seja, λ deve ser uma raiz do polinômio característico

$$P(\lambda) = \det[\lambda^2 M + \lambda C + K]. \quad (5.83)$$

Nesta situação, λ é dito *autovalor* e v um *autovetor* associado. Se os coeficientes M, C, K são $n \times n$, então, tem-se $2n$ autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_{2n}\}$ e os $2n$ autovetores correspondentes $\{v_1, v_2, \dots, v_{2n}\}$. Pelo princípio da superposição linear e (5.21), vem

$$u(t) = c_1 e^{\lambda_1 t} v_1 + c_2 e^{\lambda_2 t} v_2 + \dots + c_{2n} e^{\lambda_{2n} t} v_{2n} = V e^{\Lambda t} c$$

é solução homogênea. Aqui c é o vetor cujas componentes são as constantes c_k , V é a matriz $n \times 2n$ cujas colunas são os autovetores v_k , chamada de *matriz modal* e a matriz diagonal

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \dots & \lambda_{2n} \end{bmatrix}, \quad (5.84)$$

formada com os autovalores do sistema (5.82), vem a ser a *matriz espectral*.

Para determinar as constantes c_k , utilizam-se os valores iniciais $u_o = u(0)$, $u'(0) = u'_o$ da solução. Assim,

$$\begin{aligned} u(0) &= c_1 v_1 + c_2 v_2 + \dots + c_{2n} v_{2n} = Vc \\ u'(0) &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \dots + c_{2n} \lambda_{2n} v_{2n} = V\Lambda c \end{aligned}$$

Tem-se o sistema algébrico linear

$$\begin{bmatrix} V \\ V\Lambda \end{bmatrix} c = \begin{bmatrix} u_o \\ u'_o \end{bmatrix} \quad (5.85)$$

Este sistema possui solução única c para qualquer u_o, u'_o dados, unicamente se a matriz do sistema for não singular. Isto equivale a que seu determinante seja não nulo, ou, também, que suas $2n$ colunas

$$\begin{bmatrix} v_1 & v_2 & \dots & v_{2n} \\ \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_{2n} v_{2n} \end{bmatrix} \quad (5.86)$$

sejam vetores linearmente independentes.

Sob a hipótese de que as $2n$ colunas dadas em (5.86) são linearmente independentes, tem-se que a solução da equação homogênea

$$Mu'' + Cu' + CKu = 0$$

é dada por

$$u(t) = Ve^{\Lambda t}c,$$

com V a matriz modal (??), Λ a matriz espectral (??) e c solução de (5.85). *Esta hipótese é sempre válida no caso em que todos os autovalores são distintos.*

Modos Normais

No caso conservativo

$$M\ddot{u} + Ku = F(t)$$

$C = 0$, suponha-se que M e K são matrizes simétricas reais, com M positiva definida. O problema de autovalor

$$(\lambda^2 M + K)v = 0 \tag{5.87}$$

possui autovalores que são reais ou puramente imaginários ($\lambda = i\omega$). Pois, sendo M , K matrizes reais segue que $\lambda^2 = -\frac{v^T K v}{v^T M v}$ é um número real. Dai que o autovalor λ será real quando o quociente for negativo ou puramente imaginário quando é positivo. Como M é positiva definida, segue que todos os autovalores serão puramente imaginários não nulos quando K for positiva definida e será incluído o zero como autovalor quando K for matriz positiva semi-definida².

Assuma-se que M e K são matrizes simétricas positivas definidas.

Com esta hipótese todos os autovalores são puramente imaginários, ou seja, existem soluções oscilatórias do tipo $u = e^{i\omega t}v$ ou, $u = \cos(\omega t)a + \sin(\omega t)b$, e origina-se o problema *generalizado* de autovalor

$$(-\omega^2 M + K)v = 0$$

ou

$$Kv = \omega^2 Mv, \quad v \neq 0. \tag{5.88}$$

As soluções deste problema são denominadas *modos*. Observe-se que v_k é o correspondente autovetor para os autovalores $i\omega_k$ e $-i\omega_k$.

Os autovetores possuem uma conveniente propriedade de ortogonalidade. Suponha-se que v, u são soluções correspondentes a ω e γ , respectivamente, ou seja

$$Kv = \omega^2 Mv, \quad v \neq 0$$

$$Ku = \gamma^2 Mu, \quad u \neq 0$$

Da simetria de K e M , vem

²A matriz K satisfaz $v^T K v \leq 0$ para qualquer vetor v .

$$0 = u^t K v - v^t K u = \omega^2 u^t M v - \gamma^2 v^t M u = [\omega^2 - \gamma^2] v^t M u.$$

Portanto, para $\omega \neq \gamma$,

$$u^t M v = 0$$

$$u^t K v = 0.$$

Como M, K são matrizes positivas definidas, tem-se $v^t M v > 0$, $v^t K v > 0$.

Os autovetores podem ser escolhidos de modo que $v_k^t M v_k = 1$, $k = 1 : n$. Para tanto, as soluções v^* obtidas resolvendo o problema de autovalor $(K - \omega^2 M)v = 0$ são normalizados com respeito da matriz M, isto é

$$v_k = v_k^* / \sqrt{(v_k^*)^t M v_k^*}, \quad k = 1 : 3.$$

Resumindo, são obtidas soluções

$$(-\omega_k^2 M + K)v_k = 0, \quad k = 1 : n$$

do problema de autovalor (5.88) de modo que

$$v_k^t M v_j = \begin{cases} 0, & k \neq j \\ 1, & k = j \end{cases}, \quad v_k^t K v_j = \begin{cases} 0, & k \neq j \\ \omega_k^2, & k = j. \end{cases} \quad (5.89)$$

A matriz V cujas colunas são os autovetores v_k é referida como sendo matriz modal. A matriz diagonal Ω cujos elementos diagonais são as frequências ω_k é referida como matriz espectral. Matricialmente, tem-se as relações

$$V^t M V = I, \quad V^t K V = \Omega^2. \quad (5.90)$$

Os autovetores v_k são referidos como sendo *modos normais* e os ω_k como *frequências naturais* do sistema conservativo M, K.

Para obter a solução do sistema homogêneo $M\ddot{u} + Ku = 0$ que satisfaz as condições iniciais $u(0) = u_o$, $\dot{u}(0) = \dot{u}_o$, considera-se a solução geral obtida do princípio da superposição linear com as soluções do tipo exponencial, isto é

$$u(t) = \sum_{k=1}^n [c_{-k} e^{-i\omega_k t} + c_k e^{i\omega_k t}] v_k \quad (5.91)$$

pois, v_k é o mesmo autovetor para os autovalores $-i\omega_k, i\omega_k$.

Para determinar as constantes c_k , utilizam-se os valores iniciais $u_o = u(0)$, $\dot{u}'(0) = \dot{u}'_o$ da solução e a propriedade de ortogonalidade dos modos.

OBSERVAÇÃO

No caso em que $\lambda = 0$ é autovalor, o autovetor correspondente é chamado de *modo-rígido*. O sistema pode movimentar-se como corpo rígido sem deformação nos elementos elásticos.

REPRESENTAÇÃO MODAL DE $h(t)$

O processo anterior para obtenção da solução do problema de valor inicial

$$M\ddot{u} + Ku = 0 \quad (5.92)$$

$$u(0) = u_o, \quad \dot{u}(0) = \dot{u}_o \quad (5.93)$$

pode ser simplificado com o uso da resposta impulso. O cálculo modal de $h(t)$ é como segue.

Da primeira relação em (5.90), vem

$$M^{-1} = VV^t,$$

Assim, substituindo em (5.80), decorre

$$h(t) = V \frac{\text{sen}\Omega t}{\Omega} V^t = \sum_{k=1}^n \frac{\text{sen}(\omega_k t)}{\omega_k} v_k v_k^t \quad (5.94)$$

Decorre que a solução do problema de valor inicial

$$M\ddot{u} + Ku = 0 \quad (5.95)$$

$$(5.96)$$

$$u(0) = u_o, \quad \dot{u}(0) = \dot{u}_o \quad (5.97)$$

é dada por

$$u(t) = \sum_{k=1}^n \cos(\omega_k t) v_k v_k^t u(0) + \sum_{k=1}^n \frac{\text{sen}(\omega_k t)}{\omega_k} v_k v_k^t \dot{u}(0). \quad (5.98)$$

Exemplo 5.63

Considere-se um sistema torsional com três discos que possuem momento de inércia de massa $I_1 = 2 \times 10^3 \text{kg.m}^2$, $I_2 = 3 \times 10^3 \text{kg.m}^2$, e $I_3 = 4 \times 10^3 \text{kg.m}^2$. Os coeficientes de rigidez dos eixos conectando esses discos são $k_1 = 12 \times 10^5 \text{N.m}$, $k_2 = 24 \times 10^5 \text{N.m}$, e $k_3 = 36 \times 10^5 \text{N.m}$.

1. Obter a matriz modal e a matriz espectral do sistema.
2. Determinar a resposta livre como resultado das condições iniciais

$$\theta(0) = \begin{bmatrix} 0,1 \\ 0,05 \\ 0,01 \end{bmatrix}, \quad \dot{\theta}(0) = \begin{bmatrix} 10 \\ 15 \\ 20 \end{bmatrix}.$$

As equações do movimento são

$$M\ddot{\theta} + K\theta = 0$$

onde

$$M = \begin{bmatrix} I_1 & 0 & 0 \\ 0 & I_2 & 0 \\ 0 & 0 & I_3 \end{bmatrix} = 2 \times 10^3 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1,5 & 0 \\ 0 & 0 & 2 \end{bmatrix} \text{ kg.m}^2$$

$$K = \begin{bmatrix} k_1 & -k_1 & 0 \\ -k_1 & k_1 + k_2 & -k_2 \\ 0 & -k_2 & k_2 + k_3 \end{bmatrix} = 12 \times 10^5 \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & -2 \\ 0 & -2 & 5 \end{bmatrix} \text{ N.m}$$

$$\theta = [\theta_1 \ \theta_2 \ \theta_3]^t$$

Aqui $\theta_1, \theta_2, \theta_3$ são as oscilações torsionais dos discos.

A procura de soluções oscilatórias $u = \text{sen}(\omega t + \phi)v$ do sistema, conduz ao problema de autovalor $(-\omega^2 M + K)v = 0$. Substituindo valores e simplificando, decorre o sistema

$$\left(-\beta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1,5 & 0 \\ 0 & 0 & 2 \end{bmatrix} + \begin{bmatrix} 1 & -1 & 0 \\ -1 & 3 & -2 \\ 0 & -2 & 5 \end{bmatrix} \right) v = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \beta = \frac{\omega^2}{6 \times 10^2}$$

ou

$$\begin{bmatrix} 1 - \beta & -1 & 0 \\ -1 & 3 - 1,5\beta & -2 \\ 0 & -2 & 5 - 2\beta \end{bmatrix} \begin{bmatrix} v_{11} \\ v_{21} \\ v_{31} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \quad (5.99)$$

Este sistema possui solução v não nula para um certo valor de β somente se o determinante do sistema anula-se. Decorre a equação característica

$$\beta^3 - 5,5\beta^2 + 7,5\beta - 2 = 0,$$

a qual possui as raízes

$$\beta_1 = 0,3516 \quad \beta_2 = 1,606 \quad \beta_3 = 3,542.$$

Como $\omega^2 = 600\beta$, as frequências naturais associadas com as raízes $\beta_1, \beta_2, \beta_3$ são dadas, respectivamente, por

$$\omega_1 = 14,52 \text{ rad/s} \quad \omega_2 = 31,05 \text{ rad/s} \quad \omega_3 = 46,1 \text{ rad/s}.$$

Os modos são obtidos resolvendo por eliminação (5.99), para cada valor da raiz $\beta_k, k = 1 : 3$. Escolhe-se os autovetores

$$v_1^* = \begin{bmatrix} 1 \\ 0,649 \\ 0,302 \end{bmatrix}, \quad v_2^* = \begin{bmatrix} 1 \\ -0,607 \\ 0,679 \end{bmatrix}, \quad v_3^* = \begin{bmatrix} 1 \\ -2,54 \\ 2,438 \end{bmatrix}.$$

Para obter a relação $V^t M V = I$, os vetores acima são normalizados

$$v_k = v_k^* / \sqrt{(v_k^*)^t M v_k^*}, \quad k = 1 : 3.$$

Assim,

$$V = \begin{bmatrix} .017 & .014 & .0048 \\ .011 & -.0085 & -.012 \\ .0051 & -.0095 & .012 \end{bmatrix}$$

é a matriz modal e

$$\Omega = \begin{bmatrix} 14,52 & 0 & 0 \\ 0 & 31,05 & 0 \\ 0 & 0 & 46,1 \end{bmatrix}$$

é a matriz espectral do sistema de discos dado.

A forma dos modos é como segue. Consiste em desenhar um autovetor v como função de variável discreta, isto é, unir os pontos $(1, v_1), (2, v_2), \dots, (n, v_n)$ onde v_k é a k -ésima componente do vetor v .

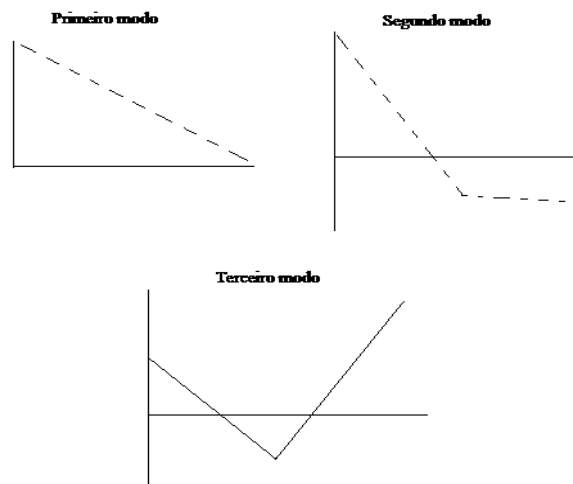


Figura 5.11 – Modos dos sistemas de discos

A resposta livre é dada por (5.91) com $n=3$

$$u(t) = \sum_{k=1}^3 (c_{-k} e^{-i\omega_k t} + c_k e^{i\omega_k t}) v_k$$

segue que

$$u(0) = \sum_{k=1}^3 (c_{-k} + c_k) v_k$$

$$u'(0) = \sum_{k=1}^3 i\omega_k (c_k - c_{-k}) v_k$$

Utilizando a propriedade de modos normais (5.89) com o primeiro modo, isto é, $v_1^t M v_1 = 1$, $v_1^t M v_2 = 0$, $v_1^t M v_3 = 0$, vem

$$c_1 + c_{-1} = v_1^t M \theta(0) = 5.254$$

$$c_1 - c_{-1} = \frac{v_1^t M \dot{\theta}(0)}{i\omega_1} = -85.61i$$

ou,

$$\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_{-1} \end{bmatrix} = \begin{bmatrix} 5.254 \\ -85.61i \end{bmatrix}$$

Resolvendo este sistema 2×2 , vem

$$\begin{bmatrix} c_1 \\ c_{-1} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 5.254 \\ -85.61i \end{bmatrix} = \begin{bmatrix} 2.627 - 42.81i \\ 2.627 + 42.81i \end{bmatrix}$$

Similarmente com os modos v_2 e v_3 ,

$$\begin{bmatrix} c_2 \\ c_{-2} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} -0.36 \\ 27.78i \end{bmatrix} = \begin{bmatrix} -0.18 + 13.89i \\ -0.18 - 13.89i \end{bmatrix}$$

$$\begin{bmatrix} c_3 \\ c_{-3} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1.1 \\ -11.19i \end{bmatrix} = \begin{bmatrix} 0.55 + 5.595i \\ 0.55 - 5.595i \end{bmatrix}$$

Assim,

$$u(t) = \begin{bmatrix} 0.0893\cos(14.52t) + 1.46\sen(14.52t) - 0.00504\cos(31.05 * t) - 0.389\sen(31.05t) + 0.00528\cos(46.1t) + 0.0538\sen(46.1t) \\ 0.0578\cos(14.52t) + 0.942\sen(14.52t) + 0.00306\cos(31.05t) + 0.236\sen(31.05t) - 0.0132\cos(46.1t) - 0.134\sen(46.1t) \\ 0.0268\cos(14.52t) + 0.437\sen(14.52t) + 0.00342\cos(31.05t) + 0.264\sen(31.05t) + 0.0132\cos(46.1t) + 0.134\sen(46.1t) \end{bmatrix}$$

Os gráficos das componentes da resposta livre, são dados a seguir.

Figura 5.12 – Componentes da resposta livre

O caso simétrico não-conservativo

No caso em que os coeficientes M , C , K são matrizes simétricas reais, M , K não-singulares, e todos os autovalores $\lambda_1, \lambda_2, \dots, \lambda_{2n}$ são distintos, pode ser obtida a fórmula espectral para $h(t)$

$$h(t) = \sum_{k=0}^{2n} \frac{\lambda_k}{s_k} e^{\lambda_k t} v_k v_k^t \quad (5.100)$$

onde

$$s_k = \lambda_k^2 v_k^t M v_k - v_k^t C v_k \quad (5.101)$$

para cada autovetor v_k correspondente ao autovalor λ_k . Para uma discussão do método espectral e extensão ao caso não-simétrico veja-se [Copetti,R.].

COMENTÁRIO

O estudo de sistemas com coeficientes M , C , K que gerem autovalores puramente imaginários é um tema de continuada pesquisa. Eles aparecem frequentemente com sistemas de natureza giroscópica ou que envolvam rotação. Por outro lado, a verificação de que os autovalores de um sistema M, C, K são todos distintos nem sempre é uma tarefa simples. O caso conservativo, $C = 0$, e com M , K simétricas positivas definidas é o mais conhecido. Os autovalores são todos puramente imaginários e distintos. Resulta que as matrizes M , K podem ser simultaneamente diagonalizadas por uma mesma matriz V , segundo a discussão anterior sobre modos normais, isto é

$$V^t M V = I, \quad V^t K V = \Omega^2.$$

O uso de modos normais é muito utilizado ao supor que a matriz C satisfaz a condição de Rayleigh

$$C = \alpha M + \beta K,$$

para certas constantes α , β . Pois, fazendo a mudança de variável

$$u = V q$$

e pré-multiplicando a esquerda por V^t a equação

$$M\ddot{u} + C\dot{u} + Ku = F(t)$$

decorre

$$V^t M V \ddot{q} + V^t C V \dot{q} + V^t K V q = V^t F(t) = f(t)$$

$$\ddot{q} + (\alpha I + \beta \Omega) \dot{q} + \Omega^2 q = f(t)$$

Esta última equação está na forma desacoplada, isto é, equivale a resolver para cada componente q_k do vetor q , uma equação escalar

$$\ddot{q}_k + (\alpha + \beta\omega_k)\dot{q}_k + \omega_k^2 q_k = f_k(t), \quad k = 1 : n.$$

5.8.4 O caso não-homogêneo

A solução da equação não-homogênea

$$Mu'' + Cu' + Ku = F(t) \quad (5.102)$$

pode ser obtida com o uso do método de variação de parâmetros. Para isto, procura-se uma resposta da forma

$$u(t) = c_1(t)e^{\lambda_1 t}v_1 + c_2(t)e^{\lambda_2 t}v_2 + \dots + c_{Nn}(t)e^{\lambda_{Nn} t}v_{Nn}. \quad (5.103)$$

onde os λ_k , v_k são os autovalores e autovetores do sistema e os coeficientes são funções a serem determinadas. O cálculo das derivadas de $u(t)$ é simplificado com a condição de Lagrange

$$\sum_{j=1}^{2n} \dot{c}_j(t)\lambda_j^{k-1}e^{\lambda_j t}v_j = 0, \quad (5.104)$$

de modo que,

$$u'(t) = \sum_{j=1}^{2n} c_j(t)\lambda_j^k e^{\lambda_j t}v_j, \quad (5.105)$$

Substituindo-se a expressão para $u(t)$ e suas derivada (5.105) na equação (5.102), resulta que

$$M \sum_{j=1}^{2n} \dot{c}_j(t)\lambda_j e^{\lambda_j t}v_j = F(t). \quad (5.106)$$

Assim, as duas equações descritas por (5.104) e (5.106) formam o sistema linear

$$\begin{aligned} \sum_{j=1}^{2n} \dot{c}_j(t)e^{\lambda_j t}v_j &= 0 \\ \sum_{j=1}^{2n} \dot{c}_j(t)\lambda_j e^{\lambda_j t}Mv_j &= F(t) \end{aligned}$$

em $\dot{c}_j(t)$, que pode ser escrito na forma matricial como

$$\begin{bmatrix} v_1 & v_2 & \dots & v_{2n} \\ \lambda_1 v_1 & \lambda_2 v_2 & \dots & \lambda_{2n} v_{2n} \end{bmatrix} \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & e^{\lambda_{2n} t} \end{bmatrix} \begin{bmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \\ \vdots \\ \dot{c}_{2n}(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ M^{-1}F(t) \end{bmatrix}. \quad (5.107)$$

O sistema pode ser representado de forma compacta por

$$\mathbf{V}e^{Dt}\dot{\mathbf{c}}(t) = \mathcal{F}(t), \quad (5.108)$$

onde denota-se

$$\mathbf{V} = \begin{bmatrix} \mathbf{V} \\ \mathbf{VD} \end{bmatrix}, \quad e^{Dt} = \begin{bmatrix} e^{\lambda_1 t} & 0 & \dots & 0 \\ 0 & e^{\lambda_2 t} & \dots & 0 \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & e^{\lambda_{2n} t} \end{bmatrix}, \quad (5.109)$$

$$\dot{\mathbf{c}}(t) = \begin{bmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \\ \vdots \\ \dot{c}_{2n}(t) \end{bmatrix} \quad \text{e} \quad \mathcal{F}(t) = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ M^{-1}F(t) \end{bmatrix}. \quad (5.110)$$

Resolvendo-se (5.108) em termos das componentes do vetor $\dot{\mathbf{c}}(t)$ resulta

$$\dot{\mathbf{c}}(t) = (\mathbf{V}e^{Dt})^{-1} \mathcal{F}(t) = \mathbf{V}^{-1}e^{-Dt} \mathcal{F}(t), \quad (5.111)$$

sendo que a solução desta equação diferencial matricial de primeira ordem, para $t = 0$, pode ser integrada componente a componente. As condições iniciais para as $c_j(t)$ são determinadas a partir das condições iniciais para o sistema, isto é

$$\mathbf{c}(0) = \mathbf{V}^{-1} \begin{bmatrix} u(0) \\ \dot{u}(0) \end{bmatrix}. \quad (5.112)$$

5.8.5 Métodos não-espectrais

Este métodos não requerem o uso de autovetores e são apresentados diretamente com a equação de segunda ordem.

O método operacional

Outra maneira de obter a fórmula de variação de parâmetros (5.74) e obter outras propriedades da solução fundamental $\mathbf{h}(t)$, consiste na utilização do método operacional da transformada de Laplace. Considere-se a equação diferencial

$$\begin{aligned} Mu'' + Cu' + Ku &= F(t), \\ u(0) &= u_0, \quad u'(0) = u'_0 \end{aligned}$$

onde M, C, K são matrizes arbitrárias $n \times n$. Aplicando a transformada de Laplace a esta equação, obtém-se

$$Ms^2U(s) - M\dot{u}(0) - Msu(0) + C[sU(s) - u(0)] + CU(s) = F(s),$$

onde $U(s), F(s)$ são as transformadas de Laplace de $u(t)$ e $F(t)$, respectivamente. Através de simplificações, decorre a equação operacional

$$\Delta(s)U(s) = (sM + C)u(0) + M\dot{u}(0) + f(s), \quad (5.113)$$

onde

$$\Delta(s) = s^2M + sC + K.$$

De (5.113), segue que a transformada de Laplace $H(s)$ da solução fundamental $h(t)$ satisfaz a equação operacional

$$\Delta(s)H(s) = I, \quad (5.114)$$

ou seja, a *matriz de transferência* $H(s)$ é a inversa do polinômio matricial $\Delta(Is)$.

$$H(s) = \Delta(s)^{-1} = [s^2M + sC + c]^{-1}. \quad (5.115)$$

Assim, pré-multiplicando (5.113) por

$$\Delta^{-1}(s) = H(s),$$

decorre

$$U(s) = H(s)M\dot{u}(0) + (sH(s)M + H(s)C)u(0) + H(s)f(s).$$

e utilizando a transformada inversa de Laplace, obtém-se uma fórmula para as soluções de (5.113), em termos da solução fundamental $h(t)$,

$$u(t) = h(t)M\dot{u}(0) + (\dot{h}(t)M + h(t)C)u(0) + \int_0^t h(t-\tau)f(\tau)d\tau. \quad (5.116)$$

A equação (5.116) mostra que, para conhecer a *solução* do sistema (5.113), é suficiente determinar a solução fundamental ou resposta impulso associada ao mesmo.

A seguinte propriedade de comutatividade é válida para $h(t)$. Ela segue de tomar a transformada inversa de Laplace em (5.114), pois $H(s)$ é inversa à direita e à esquerda de $\Delta(s)$. Assim,

$$\boxed{Mh'' + Ch' + Kh = h''M + h'C + hK = 0}. \quad (5.117)$$

Exemplo 5.64

Determinar a solução do sistema homogêneo

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

sujeito às condições iniciais

$$u(0) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \dot{u}(0) = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}.$$

SOLUÇÃO

A equação operacional, que decorre de aplicar a transformada de Laplace,

$$(s^2M + sC + K)U(s) = (sM + C)u(0) + Mu'(0),$$

é dada por

$$\begin{bmatrix} s^2 + s + 1 & s & -s^2 + 1 \\ s & s^2 + s + 1 & 0 \\ s^2 - 1 & 0 & s^2 + s + 1 \end{bmatrix} U(s) = \begin{bmatrix} -2s + 5 \\ 2s + 5 \\ 4s + 7 \end{bmatrix}.$$

Invertendo a matriz dos coeficientes, para obter $U(s)$, vem

$$U(s) = \begin{bmatrix} U_1(s) \\ U_2(s) \\ U_3(s) \end{bmatrix} = \frac{\begin{bmatrix} (2s^3 + 8s^2 - 6s - 2)(s^2 + s + 1) \\ 4s^5 + 12s^4 + 2s^3 + 8s^2 + 14s + 5 \\ 2(s + 1)(3s^4 + 4s^3 + 7s^2 + 2s + 6) \end{bmatrix}}{P(s)}$$

onde $P(s)$ é o polinômio característico

$$P(s) = \det(s^2M + sC + K) = s^4 + 6s^3 + 19s^2 + 37s + 26.$$

Decorre

$$u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{bmatrix} = \begin{bmatrix} \mathcal{L}^{-1}(U_1(s)) \\ \mathcal{L}^{-1}(U_2(s)) \\ \mathcal{L}^{-1}(U_3(s)) \end{bmatrix},$$

onde

$$\begin{aligned} u_1(t) &= \frac{15t - 9}{9}e^{-t} + 2e^{t/2} \left[\cos \frac{\sqrt{3}}{2}t - \frac{2}{3\sqrt{3}}\text{sen} \frac{\sqrt{3}}{2}t \right] \\ u_2(t) &= \frac{5t - 14}{6}e^{-t} - e^{t/2} \left[\frac{5}{12}\cos \frac{\sqrt{3}}{2}t - \frac{7}{12\sqrt{3}}\text{sen} \frac{\sqrt{3}}{2}t \right] + e^{-t/2} \left[\frac{19}{4}\cos \frac{\sqrt{3}}{2}t + \frac{9}{4\sqrt{3}}\text{sen} \frac{\sqrt{3}}{2}t \right] \\ u_3(t) &= \frac{10}{3}e^{-t} + e^{t/2} \left[\frac{2}{3}\cos \frac{\sqrt{3}}{2}t + \text{sen} \frac{\sqrt{3}}{2}t \right] - e^{-t/2} \left[\cos \frac{\sqrt{3}}{2}t - \frac{4}{\sqrt{3}}\text{sen} \frac{\sqrt{3}}{2}t \right] \end{aligned}$$

Uma fórmula analítica para a solução $h(t)$

Tem sido possível estabelecer uma fórmula analítica para a solução fundamental $h(t)$ e a matriz de transferência $H(s)$ [Claeyssen]. Esta fórmula envolve três equações características dos tipos algébrica, diferencial e em diferenças. Mais precisamente,

$$\boxed{h(t) = \sum_{j=1}^{2n} \sum_{i=0}^{j-1} b_i d^{j-i-1}(t) h_{2n-j}}, \quad (5.118)$$

onde $h_k = h^k(0)$ satisfaz a equação matricial em diferenças

$$Mh_{k+2} + Ch_{k+1} + Ch_k = 0, \quad (5.119)$$

com valores iniciais

$$h_0 = 0, \quad Mh_1 = I.$$

O polinômio de grau $2n$, denotado por

$$P(s) = \det [Ms^2 + Cs + C] = \sum_{k=0}^{2n} b_k s^k, \quad (5.120)$$

corresponde ao polinômio característico associado ao sistema, e $d(t)$ representa a solução do problema de valor inicial expresso por

$$b_{2n}d^{(2n)}(t) + b_{2n-1}d^{(2n-1)}(t) + \cdots + b_1d'(t) + b_0d(t) = 0, \quad (5.121)$$

com os valores iniciais dados por

$$d(0) = 0, \quad \dot{d}(0) = 0, \quad \dots, \quad d^{(2n-2)}(0) = 0, \quad b_{2n}d^{(2n-1)}(0) = 1. \quad (5.122)$$

Aplicando-se a transformada de Laplace em (5.118), obtém-se que a matriz de transferência é dada pela expressão

$$H(s) = [s^2M + sC + K]^{-1} = \sum_{j=1}^{2n} \sum_{i=0}^{j-1} b_i \frac{s^{j-i-1}}{P(s)} h_{2n-j} = \frac{Q(s)}{P(s)}, \quad (5.123)$$

onde

$$Q(s) = \sum_{j=1}^{2n} \sum_{i=0}^{j-1} b_i s^{j-i-1} h_{2n-j}. \quad (5.124)$$

e $P(s)$ definido por (4.106)

Fórmula para o caso de raízes simples

Quando todas as raízes s_k de $P(s)$ forem distintas, a fórmula para $h(t)$ pode ser simplificada. Tem-se que $d(t)$ dada por

$$d(t) = \sum_{k=0}^{2n} \frac{e^{s_k t}}{P'(s_k)}$$

e, com a introdução dos polinômios

$$q_j(s) = \sum_{i=0}^{j-1} b_i s^{j-1-i}, \quad j = 1, 2, \dots, 2n, \quad (5.125)$$

decorre que

$$h(t) = \sum_{j=1}^{2n} \sum_{i=0}^{j-1} b_i \sum_{k=0}^{2n} \frac{s_k^{j-1-i} e^{s_k t} h_{2n-j}}{P'(s_k)}.$$

Deste modo,

$$h(t) = \sum_{k=0}^{2n} E_k e^{s_k t}, \quad H(s) = \sum_{k=0}^n \frac{E_k}{s - s_k}, \quad (5.126)$$

onde

$$E_k = \frac{1}{P'(s_k)} \sum_{j=1}^{2n} q_j(s_k) h_{2n-j} \quad (5.127)$$

As fórmulas para $h(t)$ e $H(s)$ são obtidas facilmente, na implementação de sistemas de pequeno porte, com o uso de softwares simbólicos.

Exemplo 5.65

Determinar a solução do problema de valor inicial

$$\begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} \ddot{u}_1(t) \\ \ddot{u}_2(t) \end{bmatrix} + \begin{bmatrix} 3 & -4 \\ -4 & 8 \end{bmatrix} \begin{bmatrix} \dot{u}_1(t) \\ \dot{u}_2(t) \end{bmatrix} + \begin{bmatrix} 5 & -3 \\ -3 & 7 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = \begin{bmatrix} 4 \cos 2t \\ 3t \end{bmatrix}$$

$$u(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \dot{u}(0) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Solução

Tem-se o polinômio característico

$$P(s) = \det(s^2 M + sC + K) = s^4 + 6s^3 + 19s^2 + 37s + 26.$$

A solução da equação diferencial característica

$$d^{(4)}(t) + 6d^{(3)}(t) + 19\ddot{d}(t) + 37\dot{d}(t) + 26d(t) = 0,$$

com os valores iniciais

$$d(0) = \dot{d}(0) = \ddot{d}(0) = 0, \quad d^{(3)}(0) = 1,$$

é dada, de modo aproximado, por

$$d(t) = -0.072239 e^{(-2.7255 t)} + 0.10396 e^{(-1.2665 t)} - 0.031716 e^{(-1.0040 t)} \cos(2.5543 t) - 0.038004 e^{(-1.0040 t)} \operatorname{sen}(2.5543 t).$$

Por recursão, obtém-se da equação característica em diferenças

$$Mh_{k+2} + Ch_{k+1} + Kh_k = 0, \quad h_0 = 0, \quad Mh_1 = I, \quad k = 0 : 1,$$

que

$$\begin{aligned} h_1 &= M^{-1} \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \\ h_2 &= -M^{-1}Ch_1 = \begin{bmatrix} -4 & -2 \\ -2 & -3 \end{bmatrix}, \\ h_3 &= -M^{-1}Ch_2 - M^{-1}Kh_1 = \begin{bmatrix} -7 & -4 \\ -4 & 4 \end{bmatrix}. \end{aligned}$$

Substituindo os valores obtidos acima em

$$h(t) = \sum_{j=1}^4 \sum_{i=0}^{j-1} b_i h^{(j-1-i)}(t) h_{4-j}$$

e, com o cálculo da integral

$$u(t) = \int_0^t h(t-\tau) f(\tau) d\tau,$$

obtem-se as componentes

$$u_1(t) = 0.982 \cos 2t + 1.13 \operatorname{sen} 2t + 0.345t - 0.0307 + 0.0182 e^{-2.73t} - 0.0979 e^{-1.27t} - 0.872 e^{-t} \cos 2.55t - 1.39 e^{-t} \operatorname{sen} 2.55t$$

$$u_2(t) = 0.528 \cos 2t + 0.536 \operatorname{sen} 2t + 0.577t - 0.475 - 0.156 e^{-2.73t} + 0.590 e^{-1.27t} - 0.486 e^{-t} \cos 2.55t - 0.712 e^{-t} \operatorname{sen} 2.55t$$

Observe-se que a solução $u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$ é a superposição da solução permanente $u_p(t)$, a qual corresponde ao termo forçante dado, e a solução transiente $u_h(t)$ que é uma solução do sistema homogêneo, como mostra a figura (5.13).

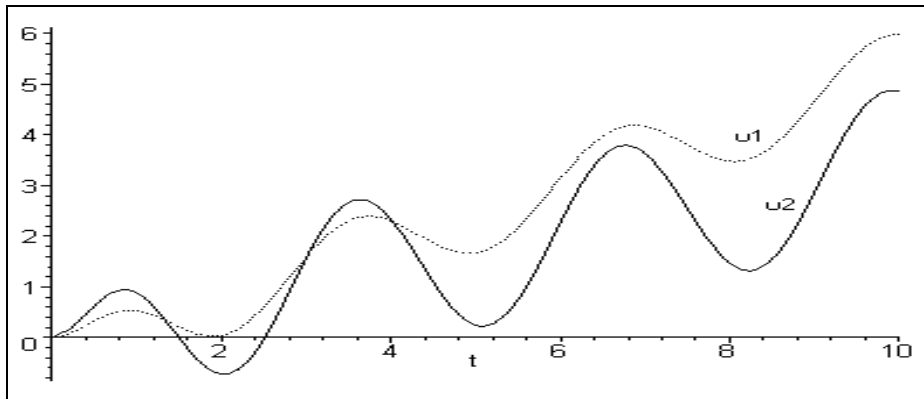


Figura 5.13 – Respostas $u_1(t)$ e $u_2(t)$

5.8.6 A formulação de estado

Na literatura clássica [Kailath], [Ginsberg], o estudo da equação de segunda ordem (5.64) consiste em reduzi-la a uma equação de primeira ordem com a transformação de Hamilton, ou formulação no espaço de estado. Fazendo $z_1 = u$, $z_2 = u'$, tem-se que $z_1' = u' = z_2$, $Mz_2' = Mu'' = -Cu - Cu' + F(t) = -Cz_1 - Cz_2 + F(t)$. Assim, o sistema

$$\begin{aligned} z_1' &= z_2, \\ Mz_2' &= -Cz_1 - Cz_2 + F(t) \end{aligned}$$

pode ser escrito como

$$M \frac{dz}{dt} + Cz = f(t) \quad (5.128)$$

onde

$$A = \begin{bmatrix} I & 0 \\ 0 & M \end{bmatrix}, \quad B = \begin{bmatrix} 0 & -I \\ C & C \end{bmatrix}, \quad (5.129)$$

$$z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}, \quad \frac{dz}{dt} = z' = \begin{bmatrix} dz_1/dt \\ dz_2/dt \end{bmatrix}, \quad f = \begin{bmatrix} 0 \\ F(t) \end{bmatrix}$$

Decorre daí que

$$\boxed{z(t) = h(t)u(0) + \int_0^t h(t-\tau)F(\tau)d\tau, \quad h(t) = e^{-At} \mathcal{A}^{-1}} \quad (5.130)$$

onde a matriz de ordem $2n \times 2n$

$$\mathcal{A} = -A^{-1}B = \begin{bmatrix} 0 & I \\ -A^{-1}C & -A^{-1}C \end{bmatrix} \quad (5.131)$$

é referida como sendo a matriz companheira de A , C e C . Aqui

$$Ah'(t) + Bh(t) = h'(t)A + h(t)B = 0, \quad (5.132)$$

$$h(0) = 0, \quad Ah'(0) = I. \quad (5.133)$$

Qualquer informação relativa a solução $u(t)$ da equação de segunda ordem (5.64) deve ser obtida de (5.130) através das n primeiras componentes de z . Por outro lado, tendo em conta a equação (5.74), várias propriedades para as equações de segunda ordem podem ser induzidas convenientemente a partir da formulação de estado. Por exemplo, pode ser estabelecida a seguinte relação entre as soluções fundamentais $h(t)$ e $h(t)$

$$h(t) = \begin{bmatrix} h_o(t) & h_1(t) \\ h'_o(t) & h'_1(t) \end{bmatrix} = \begin{bmatrix} h'(t)A + h(t)C & h(t)A \\ h''(t)A + h'(t)C & h'(t)A \end{bmatrix}$$

Também, utilizando a relação

$$d^k e^{tA} / dt^k = \mathcal{A}^k e^{tA},$$

segue que as potências da matriz companheira são dadas por

$$\mathcal{A} = \begin{bmatrix} h_{o,k} & h_{1,k} \\ h_{o,k+1} & h_{1,k+1} \end{bmatrix} = \begin{bmatrix} h_{k+1}A + h_kC & h_k \\ h_{k+2}A + h_{k+1}C & h_{k+1} \end{bmatrix} \quad (5.134)$$

Aqui $h_k = h^{(k)}(0)$ é a solução do problema discreto de valor inicial

$$\begin{aligned} Ah_{k+2} + Ch_{k+1} + Ch_k &= 0 \\ h_o &= 0, Ah_1 = I \end{aligned}$$

e $h_{o,k} = h_o^{(k)}(0)$.

Existem outras mudanças de variável para obter equações equivalentes de primeira ordem, porém não acrescentam nada de novo. Ainda que este caminho da redução possa simplificar as demonstrações de resultados obtidos por outros meios ou ser útil para checar resultados esperados, ele apresenta uma visão estreita sobre a dinâmica das equações de segunda ordem e dificuldades na sua manipulação. Observe-se que a variável dependente z possui agora o dobro de componentes da variável original u e que a solução fundamental $h(t)$ é de ordem $2n \times 2n$.

5.9 Comportamento assintótico

Como foi visto com sistemas de 1ª ordem, o sistema de 2ª ordem é dito assintoticamente estável quando as soluções do sistema homogêneo tendem para zero quando t tende para infinito.

As soluções do sistema $Mu'' + Cu' + Cu = 0$ são da forma $u(t) = h(t)Mu'(0) + [h'(t)M + h(t)C]u(0)$. Seu comportamento no infinito é o da solução fundamental $h(t)$ e de sua derivada $h'(t)$. Utilizando as fórmulas (5.118) e (??), decorre que este comportamento é determinado pelo comportamento de $d(t)$. Assim, $h(t)$ é assintoticamente estável se, e somente se, todas as raízes de $P(s)$, isto é, os autovalores do problema de autovalor $[\lambda^2 M + \lambda C + C]v = 0$, possuem parte real negativa.

5.10 Filtragem

As características de filtro da equação

$$Mu'' + Cu' + Ku = F(t, x)$$

são observadas com entradas

$$F(t, x) = e^{i\omega t}v.$$

A resposta particular $u = e^{i\omega t}z$ pode ter como resultado que vetor z de menor (ou maior) amplitude ("norma") do que da entrada v . Substituindo u e $F(t, x)$ na equação tem-se

$$[(i\omega)^2 M + i\omega C + K]z = v \quad (5.135)$$

A equação (5.135) possui solução bem definida

$$z = [(i\omega)^2 M + i\omega C + K]^{-1}v = H(i\omega)v, \quad (5.136)$$

com

$$H(i\omega) = \frac{Q(i\omega)}{P(i\omega)}v$$

onde $Q(s) = \text{adj}[s^2 M + sC + C]^3$ e $P(s) = \det[s^2 M + sC + C]$ é o polinômio característico, desde que $s = i\omega$ não seja uma raiz de $P(s)$, isto é, um autovalor do sistema M,C,K. Se $\lambda = i\omega_s$ é um autovalor do sistema e ω está perto de ω_s então $P(\omega)$ será muito pequeno e, em consequência, w será um vetor de grande amplitude. (*fenômeno da ressonância*).

³Aqui $\text{adj}(A)$ indica transposta da matriz dos cofatores de A , isto é, a matriz adjunta ou, também dita matriz adjugada

CAPÍTULO 6

O Método Dos Mínimos Quadrados

Um sistema arbitrário de equações lineares $\mathbf{Ax} = \mathbf{b}$ pode não, ter uma solução, isto é, o dado \mathbf{b} é inconsistente. Ou seja, não é possível realizar a retrosubstituição como resultado da eliminação Gaussiana. No entanto, em certas situações é necessário resolver o sistema de alguma maneira, por exemplo, quando o número de equações sobrepuja o número de incógnitas. Na prática, isto ocorre na observação ou aquisição de dados. Há mais dados disponíveis do que incógnitas. Uma vez que, como resultado de medições experimentais deve-se prever algum tipo de solução, faz-se necessário *ampliar* ou *reinterpretar* o conceito de solução.

Suponha-se, por exemplo, que haja uma relação linear entre uma variável dependente y e uma variável independente x , ou seja,

$$y = \alpha x + \beta \quad (6.1)$$

e que sejam realizadas medições em y para vários valores de x , originando a seguinte tabela

x	0	3	5	8	10
y	2	5	6	9	11

Graficando estes pontos, observa-se na fig. 6.1 que não estão situados sobre uma linha reta.

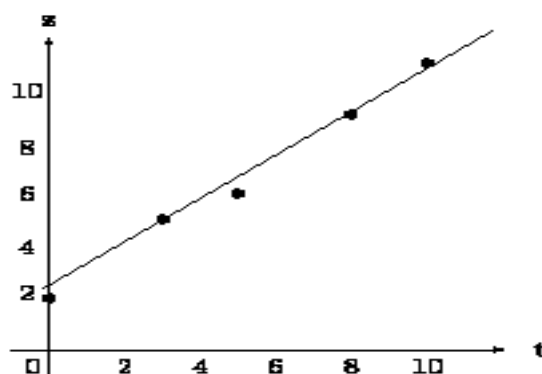


Figura 6.1 – Método dos Mínimos Quadrados

Este fato é devido a erros nas medições ou aos parâmetros α ou β não serem constantes. Ignore-se a segunda possibilidade. Considere-se que as medições estão sujeitas a erros.

Se y_i denota o valor medido para y quando $x = x_i$, então a relação entre y_i e x_i pode ser escrita como

$$y_i = \alpha x_i + \beta + e_i, \quad (6.2)$$

onde e_i denota o erro que provém da medição. Agora, o problema consiste em deduzir “os melhores valores possíveis” de α e β ou, em outras palavras, encaixar a “melhor” linha reta entre os dados graficados na figura 6.1. Este procedimento foi proposto por Legendre, que o denominou de *método dos mínimos quadrados*. Uma fundamentação do método, em conexão com a teoria da probabilidade, foi feita por Gauss.

O critério, para determinar os parâmetros α e β , é o de minimizar a soma dos quadrados dos erros devido a m medições:

$$E = \sum_{i=1}^m e_i^2 = \sum_{i=1}^m (y_i - \alpha x_i - \beta)^2, \quad (6.3)$$

na qual m denota o número total de observações. Daí o termo de “mínimos quadrados”.

Para minimizar E , calcula-se as derivadas parciais em relação a α e β , obtendo-se as condições

$$\frac{\partial E}{\partial \alpha} = E_\alpha = -2 \sum_{i=1}^m (y_i - \alpha x_i - \beta) x_i = 0 \quad (6.4)$$

e

$$\frac{\partial E}{\partial \beta} = E_\beta = -2 \sum_{i=1}^m (y_i - \alpha x_i - \beta) = 0, \quad (6.5)$$

as quais dão origem ao denominado “sistema de equações normais”. Observe-se que, para α , β que satisfazem essas equações, $E_{\alpha\beta} > 0$ e $E_{\alpha\alpha}E_{\beta\beta} - (E_{\alpha\beta})^2 > 0$ então, E atinge um valor mínimo. Visando discutir estas equações, são introduzidos o vetor dos dados y_i , a matriz de dados x_i e o vetor incógnita com os coeficientes desconhecidos α e β representados por

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}, \quad \mathbf{A} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_m & 1 \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad (6.6)$$

respectivamente. Se \mathbf{e} denota o vetor erro com elementos

$$e_i = y_i - (\alpha x_i + \beta), \quad (6.7)$$

então

$$\mathbf{e} = \mathbf{y} - \mathbf{A}\mathbf{x} \quad (6.8)$$

e as condições $E_\alpha = E_\beta = 0$ podem ser escritas como

$$0 = \sum_{i=1}^m e_i x_i \quad \text{e} \quad 0 = \sum_{i=1}^m e_i, \quad (6.9)$$

respectivamente. Na forma matricial, tem-se

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_m \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

isto é,

$$\mathbf{A}^t \mathbf{e} = \mathbf{0}, \quad (6.10)$$

ou, de modo equivalente,

$$\mathbf{A}^t(\mathbf{y} - \mathbf{Ax}) = \mathbf{0}, \quad (6.11)$$

ou

$$\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{y} \quad (6.12)$$

que é denominada *equação normal* associada ao sistema $\mathbf{Ax} = \mathbf{b}$.

Substituindo valores, decorre

$$\begin{bmatrix} \sum_{i=1}^m x_i^2 & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & m \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \sum_{i=1}^m x_i y_i \\ \sum_{i=1}^m y_i \end{bmatrix},$$

cuja solução é

$$\alpha = \bar{y} - \beta \bar{x}, \quad \beta = \frac{\sigma_{xy}}{\sigma_x^2} \quad (6.13)$$

onde

$$\bar{y} = \frac{\sum_{i=1}^m y_i}{m}, \quad \bar{x} = \frac{\sum_{i=1}^m x_i}{m} \quad (6.14)$$

são os valores médios de x e y , respectivamente,

$$\sigma_x^2 = \sum_{i=1}^m \frac{x_i^2}{m} - \bar{x}^2, \quad (6.15)$$

é o desvio padrão de x_1, x_2, \dots, x_m , e

$$\sigma_{xy} = \sum_{i=1}^m \frac{x_i y_i}{m} - \bar{x} \bar{y}, \quad \bar{x} \bar{y} = \sum_{i=1}^m \frac{x_i y_i}{m} \quad (6.16)$$

é a covariância entre os valores x_1, x_2, \dots, x_m e y_1, y_2, \dots, y_m . Para estes valores de α e β , a reta $y = \alpha x + \beta$ é denominada *linha de regressão* de y sobre x .

No exemplo numérico, dado pela tabela anterior, a equação normal é dada pelo sistema

$$\begin{bmatrix} 5 & 25 \\ 26 & 198 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} 33 \\ 227 \end{bmatrix},$$

Resolvendo este sistema, obtém-se $\alpha = 2.01$ e $\beta = 0.88$. A linha de regressão deste exemplo é a do gráfico anterior.

Exemplo 6.66

Existe evidência teórica para supor que os dados na tabela

x	1.0	1.5	2.0	2.5	3.0	3.5
z	3.00	2.34	1.82	1.42	1.10	0.86

devem estar sobre uma exponencial decrescente $z = Me^{-kx}$. Determinar os valores M e k utilizando o método dos mínimos quadrados.

Solução

Tem-se seis dados para determinar dois parâmetros. Tomando o logaritmo, segue $\ln z = \ln M - kx$. Fazendo, $y = \ln z$, $\beta = \ln M$ e $\alpha = -k$, obtém-se a relação linear $y = \alpha x + \beta$. Em termos de x e y , os dados da tabela são

x	1.0	1.5	2.0	2.5	3.0	3.5
$y = \ln z$	1.0986	0.8502	0.5988	0.3507	0.0953	-0.1508

Tem-se o sistema sobredeterminado

$$\begin{aligned}\alpha 1.0 + \beta &= 1.0986 \\ \alpha 1.5 + \beta &= 0.8502 \\ \alpha 2.0 + \beta &= 0.5988 \\ \alpha 2.5 + \beta &= 0.3507 \\ \alpha 3.0 + \beta &= 0.0953 \\ \alpha 3.5 + \beta &= -0.1508\end{aligned}$$

Aplicando o método dos mínimos quadrados, isto é, formando a equação normal e resolvendo, decorre que a linha de regressão de $\ln z$ sobre x é calculada determinando (com 4 casas decimais) os valores médios $\bar{x} = 2.2500$, $\bar{y} = 0.4738$, o desvio padrão $\sigma_x^2 = 0.7292$ e a covariância $\sigma_{xy} = -0.3650$.

Assim,

$$\alpha = -k = -0.5005, \quad \beta = \ln M = 1.5999$$

Portanto,

$$z = 4.95e^{-0.5x}$$

6.1 A Equação Matricial Normal

A discussão anterior pode ser generalizada rapidamente para o caso em que y depende linearmente de n parâmetros $\alpha_1, \alpha_2, \dots, \alpha_n$, ou seja,

$$y = \phi_1(x)\alpha_1 + \phi_2(x)\alpha_2 + \dots + \phi_n(x)\alpha_n \quad (6.17)$$

onde $\phi_1, \phi_2, \dots, \phi_n$ são funções simples. Por exemplo,

$$\phi_1(x) = x, \quad \phi_2(x) = x^2, \quad \phi_3(x) = x^2, \quad \dots \quad \phi_{n-1}(x) = x^{n-1}, \quad \phi_n(x) = 1.$$

Para cada um dos valores x_1, x_2, \dots, x_n , tem-se as medições

$$y_i = \alpha_1 a_{i1} + \alpha_2 a_{i2} + \dots + \alpha_n a_{in}, \quad i = 1 : m,$$

onde $a_{ij} = \phi_j(x_i)$. Matricialmente,

$$\mathbf{Ax} = \mathbf{b}, \quad (6.18)$$

com

$$\mathbf{A} = [\phi_j(x_i)] = \begin{bmatrix} \phi_1(x_1) & \phi_2(x_1) & \cdots & \phi_n(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & \cdots & \phi_n(x_2) \\ \vdots & \vdots & & \vdots \\ \phi_1(x_m) & \phi_2(x_m) & \cdots & \phi_n(x_m) \end{bmatrix}$$

e

$$\mathbf{x} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Se o número de observações for maior do que o número de variáveis ($m > n$), o sistema linear

$$\mathbf{b} = \mathbf{Ax}, \quad (6.19)$$

em geral, será inconsistente. Em consequência, os valores $\alpha_1, \alpha_2, \dots, \alpha_n$ serão indeterminados. Suponha-se um erro e_i em cada medição, isto é

$$y_i = \alpha_1 a_{i1} + \alpha_2 a_{i2} + \cdots + \alpha_n a_{in} + e_i, \quad i = 1, 2, \dots, m. \quad (6.20)$$

Então, os valores $\alpha_1, \alpha_2, \dots, \alpha_n$ podem ser selecionados de modo a minimizar a soma dos quadrados dos erros

$$e_i = b_i - \sum_{k=1}^n \alpha_k a_{ik}, \quad (6.21)$$

ou seja,

$$E = \sum_{i=1}^m (b_i - \alpha_1 a_{i1} - \alpha_2 a_{i2} - \cdots - \alpha_n a_{in})^2. \quad (6.22)$$

Escrevendo

$$\mathbf{e} = \mathbf{b} - \mathbf{Ax}, \quad (6.23)$$

decorre

$$E = (\mathbf{b} - \mathbf{Ax})^t (\mathbf{b} - \mathbf{Ax}) = \|\mathbf{b} - \mathbf{Ax}\|^2. \quad (6.24)$$

As condições necessárias para a minimização do erro quadrático total E

$$0 = E_{\alpha_j} = -2 \sum_{i=1}^m (b_i - \alpha_1 a_{i1} - \cdots - \alpha_n a_{in}) a_{ij}; \quad j = 1, 2, \dots, n, \quad (6.25)$$

originam a *equação normal*

$$\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}. \quad (6.26)$$

No caso em que $\phi_i(x) = x^i$, $i = 1 : n-1$ e $\phi_n(x) = 1$, as equações normais definem os coeficientes de uma *regressão polinomial*.

Surge a pergunta: como é que a solução por mínimos quadrados estende o conceito de solução clássica? Em verdade, tem-se que

Qualquer solução de um sistema consistente $\mathbf{Ax} = \mathbf{b}$ é uma solução por mínimos quadrados, isto é, $\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}$.

De fato, se o sistema $\mathbf{Ax} = \mathbf{b}$ possui solução \mathbf{x} , então multiplicando-o por \mathbf{A}^t decorre que \mathbf{x} satisfaz a equação normal $\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}$ e, portanto, é uma solução por mínimos quadrados. Deve ser salientado que o recíproco nem sempre é válido.

6.1.1 Resolução da Equação Normal

Dado o sistema $\mathbf{Ax} = \mathbf{b}$ com \mathbf{A} de ordem $m \times n$, considere-se a equação normal

$$\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}. \quad (6.27)$$

Com o auxílio da decomposição em valores singulares de \mathbf{A} pode ser estabelecido que a equação normal é sempre consistente. Mais precisamente,

A equação normal $\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}$ sempre possui solução, a qual não é necessariamente única.

Suponha-se que a matriz \mathbf{A} de ordem $m \times n$ possui posto r . Então, substituindo na equação normal

$$\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b}, \quad (6.28)$$

a decomposição em valores singulares da matriz \mathbf{A} ,

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t, \quad (6.29)$$

onde \mathbf{U} é de ordem $m \times m$, $\mathbf{\Sigma}$ de ordem $m \times n$ e \mathbf{V} de orden $n \times n$, decorre

$$\mathbf{V}\mathbf{\Sigma}^t \mathbf{U}^t \mathbf{U}\mathbf{\Sigma}\mathbf{V}^t \mathbf{x} = \mathbf{V}\mathbf{\Sigma}^t \mathbf{U}^t \mathbf{b}. \quad (6.30)$$

Utilizando a ortogonalidade das matrizes \mathbf{U} , \mathbf{V} e introduzindo a mudança de variáveis

$$\mathbf{z} = \mathbf{V}^t \mathbf{x}, \quad \mathbf{d} = \mathbf{U}^t \mathbf{b} \quad (6.31)$$

com \mathbf{z} de ordem $n \times 1$ e \mathbf{d} de ordem $m \times 1$, obtém-se o sistema

$$\mathbf{\Sigma}^t \mathbf{\Sigma} \mathbf{z} = \mathbf{\Sigma}^t \mathbf{d}. \quad (6.32)$$

Aqui

$$\mathbf{\Sigma}^t \mathbf{\Sigma} = \begin{bmatrix} 1/\sigma_1^2 & 0 & \cdots & 0 & 0 \\ 0 & 1/\sigma_2^2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 1/\sigma_r^2 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

é uma matriz $n \times n$ e

$$\Sigma^t \mathbf{d} = \begin{bmatrix} d_1/\sigma_1 \\ d_2/\sigma_2 \\ \vdots \\ d_r/\sigma_r \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

é um vetor $n \times 1$. Resolvendo o sistema, tem-se a solução geral

$$z_1 = \frac{d_1}{\sigma_1}, \quad \dots, \quad z_r = \frac{d_r}{\sigma_r}, \quad (6.33)$$

sendo

$$z_{r+1}, \quad \dots, \quad z_n \quad \text{arbitrários.}$$

Assim,

$$\mathbf{x} = \mathbf{V}\mathbf{z} = \sum_{k=1}^r \frac{d_k}{\sigma_k} \mathbf{q}_k + \sum_{k=r+1}^n c_k \mathbf{q}_k, \quad (6.34)$$

com c_k uma constante arbitrária. Ou seja, a equação normal é sempre consistente.

Por outro lado, a solução por mínimos quadrados não é necessariamente única, porém, é possível escolher, com um certo critério, uma destas soluções. Para isto, observe-se que, devido à ortogonalidade dos vetores \mathbf{q}_k , tem-se

$$\|\mathbf{x}\|^2 = \sum_{k=1}^r \left(\frac{d_k}{\sigma_k} \right)^2 + \sum_{k=r+1}^n c_k^2. \quad (6.35)$$

A solução por mínimos quadrados

$$\mathbf{x}^+ = \sum_{k=1}^r \frac{d_k}{\sigma_k} \mathbf{q}_k \quad (6.36)$$

é a que possui a menor norma euclidiana. Esta solução pode ser escrita na forma

$$\mathbf{x}^+ = [\mathbf{q}_1 \quad \mathbf{q}_2 \quad \dots \quad \mathbf{q}_r] \begin{bmatrix} d_1/\sigma_1 \\ d_2/\sigma_2 \\ \vdots \\ d_r/\sigma_r \end{bmatrix}$$

ou, simplesmente,

$$\mathbf{A} = \begin{bmatrix} -3 & 1 \\ -2 & 1 \\ -1 & 1 \\ 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \end{bmatrix}.$$

Tem-se que

$$\mathbf{A}^t \mathbf{A} = \begin{bmatrix} 28 & 0 \\ 0 & 7 \end{bmatrix}$$

possui os autovalores 28 e 7, com correspondentes autovetores ortonormais $\mathbf{q}_1 = [1 \ 0]^t$ e $\mathbf{q}_2 = [0 \ 1]^t$. Similarmente, a matriz $\mathbf{A} \mathbf{A}^t$ de ordem 7×7 possui os autovalores 28,7 e 0 (repetido 5 vezes) e a matriz modal obtida é

$$\mathbf{U} = \begin{bmatrix} \frac{-3}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{15}{\sqrt{420}} & 0 & 0 & 0 & 0 \\ \frac{-2}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{-10}{\sqrt{420}} & \frac{10}{\sqrt{210}} & 0 & 0 & 0 \\ \frac{-1}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{-7}{\sqrt{420}} & \frac{-8}{\sqrt{210}} & \frac{2}{\sqrt{10}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{7}} & \frac{-4}{\sqrt{420}} & \frac{-5}{\sqrt{210}} & \frac{-2}{\sqrt{10}} & \frac{3}{\sqrt{30}} & 0 \\ \frac{1}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{-1}{\sqrt{420}} & \frac{-2}{\sqrt{210}} & \frac{-1}{\sqrt{10}} & \frac{-4}{\sqrt{30}} & \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{2}{\sqrt{420}} & \frac{1}{\sqrt{210}} & 0 & \frac{-1}{\sqrt{30}} & \frac{-2}{\sqrt{6}} \\ \frac{3}{\sqrt{28}} & \frac{1}{\sqrt{7}} & \frac{4}{\sqrt{420}} & \frac{4}{\sqrt{210}} & \frac{1}{\sqrt{10}} & \frac{2}{\sqrt{30}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

e, calculando a matriz dos valores singulares, tem-se

$$\mathbf{A}^+ = \mathbf{V} \begin{bmatrix} \frac{1}{\sqrt{28}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{7}} & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{U}^t.$$

Assim,

$$\mathbf{x}^+ = \mathbf{A}^+ \mathbf{b}$$

6.1.2 Inversa Generalizada de uma Matriz

Na álgebra das matrizes, o método dos mínimos quadrados permite estender o conceito de inversa. Deste modo, todas as matrizes singulares ou inversíveis, quadradas ou retangulares possuirão uma inversa *generalizada*. Este conceito será introduzido gradualmente e, uma vez que esteja bem estabelecido, poder-se-á formular a noção de inversa generalizada de uma matriz de maneira algébrica; isto é, somente em base de definições e propriedades. Daí que, na literatura, a inversa generalizada é referida frequentemente como a inversa do Moore-Penrose .

Nesta secção a inversa generalizada será introduzida de maneira simples e concreta. Para cada dado \mathbf{b} , o sistema $\mathbf{A} \mathbf{x} = \mathbf{b}$ possui uma solução por mínimos quadrados de mínima norma \mathbf{x}^+ . Escrevendo

$$\mathbf{x}^+ = \mathbf{A}^+ \mathbf{b} \tag{6.38}$$

decorre que

$$\mathbf{A}^+ = \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^t, \quad (6.39)$$

onde

$$\mathbf{\Sigma}^+ = \begin{bmatrix} 1/\sigma_1 & & & & & \\ & \ddots & & & & \\ & & 1/\sigma_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{bmatrix}.$$

Na literatura, a matriz \mathbf{A}^+ definida acima é chamada de **pseudo-inversa** de \mathbf{A} ou *inversa generalizada* de Moore-Penrose de \mathbf{A} . A pseudo-inversa possui as seguintes propriedades algébricas:

$$\begin{aligned} \mathbf{A}\mathbf{A}^+\mathbf{A} &= \mathbf{A}, & \mathbf{A}^+\mathbf{A}\mathbf{A}^+ &= \mathbf{A}^+ \\ \mathbf{A}\mathbf{A}^+ &= (\mathbf{A}\mathbf{A}^+)^t, & \mathbf{A}^+\mathbf{A} &= (\mathbf{A}^+\mathbf{A})^t \end{aligned} \quad (6.40)$$

Resumindo:

Com a introdução da pseudo-inversa de uma matriz \mathbf{A} de ordem $m \times n$, tem-se que o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ possui, para \mathbf{b} arbitrário, uma única solução por mínimos quadrados de mínima norma, mais especificamente,

$$\mathbf{x} = \mathbf{A}^+\mathbf{b} = \begin{cases} \mathbf{A}^{-1}\mathbf{b}, & \text{se } \det(\mathbf{A}) \neq 0; \\ (\mathbf{A}^t\mathbf{A})^{-1}\mathbf{A}^t\mathbf{b}, & \text{se posto de } \mathbf{A} = n; \\ \mathbf{V}\mathbf{\Sigma}^+\mathbf{U}^t\mathbf{b}, & \text{se posto de } \mathbf{A} < n. \end{cases} \quad (6.41)$$

6.2 A Fatorização QR

A resolução da equação normal pode ser, em princípio, realizada através de qualquer método para matrizes simétricas. Entretanto, o sistema $\mathbf{A}\mathbf{x} = \mathbf{b}$ pode ser facilmente resolvido, se a matriz \mathbf{A} for ortogonal. Pois, pela ortogonalidade, a equação normal

$$\mathbf{A}^t\mathbf{A}\mathbf{x} = \mathbf{A}^t\mathbf{b} \quad (6.42)$$

é simplificada e obtém-se

$$\mathbf{x} = \mathbf{A}^t\mathbf{b}. \quad (6.43)$$

Em geral, as matrizes não-singulares ou as matrizes retangulares cujo posto é igual ao seu número de colunas, não são ortogonais, contudo é possível “ortogonalizá-las”, através do processo de Gram-Schmidt descrito a seguir. Este processo, de ortonormalização das colunas de uma matriz \mathbf{A} , induz uma importante fatorização. Por outro lado, do ponto de vista numérico, deve ser observado que, como a inversa de uma matriz ortogonal é sua transposta, este tipo de matrizes é bem condicionado.

Os vetores gerados pelo processo de Gram-Schmidt (??) podem ser escritos

$$\mathbf{w}_1 = \|\mathbf{w}_1\|\mathbf{q}_1, \quad (6.44)$$

e

$$\mathbf{w}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} (\mathbf{q}_i^t \mathbf{a}_j) \mathbf{q}_i, \quad j = 2 : n \mathbf{q}_i \quad (6.45)$$

onde os \mathbf{q}_i são ortonormais. Definindo

$$\mathbf{w}_i = r_{ii} \mathbf{q}_i, \quad r_{ii} = \|\mathbf{w}_i\|. \quad (6.46)$$

e

$$r_{ij} = \mathbf{q}_i^t \mathbf{a}_j, \quad \text{para } i < j \quad (6.47)$$

obté-m-se

$$\mathbf{w}_j = \mathbf{a}_j - \sum_{i=1}^{j-1} r_{ij} \mathbf{q}_i. \quad (6.48)$$

Assim,

$$\begin{aligned} \mathbf{a}_j &= r_{jj} \mathbf{q}_j + \sum_{i=1}^{j-1} r_{ij} \mathbf{q}_i \\ &= r_{1j} \mathbf{q}_1 + r_{2j} \mathbf{q}_2 + \cdots + r_{jj} \mathbf{q}_j \end{aligned} \quad (6.49)$$

logo

$$\mathbf{a}_j = [\mathbf{q}_1 \quad \mathbf{q}_2 \cdots \mathbf{q}_n] \begin{bmatrix} r_{1j} \\ r_{2j} \\ \vdots \\ r_{jj} \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \quad (6.50)$$

Segue-se que

$$\mathbf{A} = [\mathbf{a}_1 \quad \mathbf{a}_2 \cdots \mathbf{a}_n] = [\mathbf{q}_1 \quad \mathbf{q}_2 \cdots \mathbf{q}_n] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & \cdots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{nn} \end{bmatrix}$$

ou, simplesmente,

$$\mathbf{A} = \mathbf{QR}. \quad (6.51)$$

Aqui \mathbf{Q} é uma matriz de ordem $m \times n$, cujas colunas são ortonormais, e \mathbf{R} é uma matriz triangular superior. Tem-se o seguinte resultado

Qualquer matriz real \mathbf{A} de ordem $m \times n$, com colunas linearmente independentes, pode ser fatorizada em $\mathbf{A} = \mathbf{QR}$. As colunas da matriz \mathbf{Q} , de ordem $m \times n$, são ortonormais e \mathbf{R} é uma matriz triangular superior inversível de ordem $n \times n$. Se a matriz original é quadrada,

então \mathbf{Q} é uma matriz quadrada e, por tanto, uma matriz ortogonal.

Exemplo 6.68

Construir uma fatorização \mathbf{QR} para a matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} .$$

Solução

Tem-se que

$$\mathbf{v}_1 = \mathbf{a}_1 ,$$

$$r_{11} = \|\mathbf{v}_1\| = \sqrt{2} ,$$

$$\mathbf{q}_1 = \frac{\mathbf{v}_1}{r_{11}} = \left[\frac{1}{\sqrt{2}} \quad \frac{1}{\sqrt{2}} \quad 0 \right]^t ,$$

$$r_{12} = \mathbf{a}_2^t \mathbf{q}_1 = \frac{1}{\sqrt{2}} ,$$

e

$$r_{13} = \mathbf{a}_3^t \mathbf{q}_1 = \frac{1}{\sqrt{2}} .$$

Assim,

$$\mathbf{a}_1 = r_{11} \mathbf{q}_1 = [1 \quad 1 \quad 0]^t ,$$

$$\mathbf{v}_2 = \mathbf{a}_2 - r_{12} \mathbf{q}_1 = [1/2 \quad -1/2 \quad 1]^t ,$$

$$r_{22} = \|\mathbf{v}_2\| = \sqrt{2/3} ,$$

$$\mathbf{q}_2 = \frac{\mathbf{v}_2}{r_{22}} = \sqrt{2/3} [1/2 \quad -1/2 \quad 1]^t ,$$

$$\mathbf{a}_2 = r_{12} \mathbf{q}_1 + r_{22} \mathbf{q}_2 = \sqrt{1/2} \mathbf{q}_1 + \sqrt{3/2} \mathbf{q}_2 ,$$

$$r_{23} = \mathbf{a}_3^t \mathbf{q}_2 = \sqrt{2/3} (1/2) = \sqrt{1/6} ,$$

$$\mathbf{v}_3 = \mathbf{a}_3 - r_{13} \mathbf{q}_1 - r_{23} \mathbf{q}_2 = [-2/3 \quad 2/3 \quad 2/3]^t ,$$

$$r_{33} = \|\mathbf{v}_3\| = \sqrt{4/3} ,$$

$$\mathbf{q}_3 = \frac{\mathbf{v}_3}{r_{33}} = \sqrt{3/4} [-2/3 \quad 2/3 \quad 2/3]^t ,$$

e

$$\begin{aligned} \mathbf{a}_3 &= r_{13} \mathbf{q}_1 + r_{23} \mathbf{q}_2 + r_{33} \mathbf{q}_3 \\ &= \sqrt{1/2} \mathbf{q}_1 + \sqrt{1/6} \mathbf{q}_2 + \sqrt{4/3} \mathbf{q}_3 . \end{aligned}$$

Portanto,

$$[\mathbf{a}_1 \ \mathbf{a}_2 \ \mathbf{a}_3] = [\mathbf{q}_1 \ \mathbf{q}_2 \ \mathbf{q}_3] \begin{bmatrix} \sqrt{2} & \sqrt{1/2} & \sqrt{1/2} \\ 0 & \sqrt{3/2} & \sqrt{1/6} \\ 0 & 0 & \sqrt{4/3} \end{bmatrix}.$$

Exemplo 6.69

Obter a fatorização **QR** da matriz

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix}.$$

Solução

Pelo processo de Gram-Schmidt, tem-se

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 1/2 \\ -1/2 \\ -1/2 \\ 1/2 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} -1 \\ 1 \\ -1 \\ 1 \end{bmatrix}.$$

Então,

$$\mathbf{q}_1 = \frac{1}{2}\mathbf{w}_1, \quad \mathbf{q}_2 = \mathbf{w}_2, \quad \mathbf{q}_3 = \frac{1}{2}\mathbf{w}_3$$

$$r_{11} = \|\mathbf{w}_1\| = 2, \quad r_{12} = \mathbf{a}_2^t \mathbf{q}_1 = 1, \quad r_{13} = \mathbf{a}_3^t \mathbf{q}_1 = -1$$

$$r_{22} = \|\mathbf{w}_2\| = 1, \quad r_{23} = \mathbf{a}_3^t \mathbf{q}_2 = 1$$

e

$$r_{33} = \|\mathbf{w}_3\| = 2.$$

Assim,

$$\mathbf{A} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

A fatorização **QR** de uma matriz quadrada **A** pode ser utilizada na resolução de um sistema consistente $\mathbf{Ax} = \mathbf{b}$. Pois, substituindo $\mathbf{A} = \mathbf{QR}$ na equação, vem

$$\mathbf{QRx} = \mathbf{b},$$

ou, por ser **Q** ortogonal

$$\mathbf{Rx} = \mathbf{Q}^t \mathbf{b}.$$

Sendo **R** triangular superior, pode-se resolver por retrosubstituição.

Exemplo 6.70

Utilizar a fatorização **QR** para resolver

$$\begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ -2 \end{bmatrix}.$$

Solução

Do exemplo anterior, pode-se escrever

$$\mathbf{QRx} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ 1 \\ 3 \\ -2 \end{bmatrix}$$

Multiplicando este sistema por \mathbf{Q}^t , obtém-se

$$\mathbf{Q}^t \mathbf{QRx} = \mathbf{Rx} = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 3 \\ -2 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \\ -2 \end{bmatrix}.$$

Por retrosubstituição, $x_3 = -1$, $x_2 = -2$, $x_1 = 1$. Assim,

$$\mathbf{x} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}.$$

O uso da fatorização **QR** na equação normal

$$\mathbf{A}^t \mathbf{Ax} = \mathbf{A}^t \mathbf{b},$$

também, fornece uma solução imediata. Pois,

$$\mathbf{R}^t \mathbf{Rx} = \mathbf{R}^t \mathbf{Q}^t \mathbf{b}$$

ou, para \mathbf{A} $m \times n$ com posto n

$$\mathbf{x} = \mathbf{R}^{-1} \mathbf{Q}^t \mathbf{b}.$$

6.2.1 Matrizes de Householder

Do ponto de vista numérico, o processo de Gram-Schmidt deve ser modificado. Os vetores calculados \mathbf{q}_k poderão ser não ortogonais, devido aos erros de arredondamento e às subtrações. Para evitar este problema, o cálculo dos \mathbf{w}_k é feito gradualmente, sem a subtração sucessiva e simultânea com todos os múltiplos dos \mathbf{w}_j para $j < k$ como no Gram-Schmidt.

A obtenção da fatorização **QR**, ainda com a modificação do processo de Gram-Schmidt, pode ser inexata. Porém, esta fatorização pode ser obtida com o auxílio de sucessivas matrizes de Householder. Uma *matriz de Householder*, ou um refletor elementar, é uma matriz da forma

$$\mathbf{H} = \mathbf{I} - 2 \frac{\mathbf{v}\mathbf{v}^t}{\|\mathbf{v}\|^2}. \quad (6.52)$$

Usualmente, o vetor \mathbf{v} é normalizado e \mathbf{H} reduz-se a $\mathbf{I} - 2\mathbf{u}\mathbf{u}^t$. Em qualquer caso, \mathbf{H} é uma matriz simétrica e ortogonal

$$\mathbf{H}^t = \mathbf{I} - 2(\mathbf{u}\mathbf{u}^t)^t = \mathbf{I} - 2\mathbf{u}^t\mathbf{u} = \mathbf{H} \quad (6.53)$$

e

$$\mathbf{H}^t\mathbf{H} = (\mathbf{I} - 2\mathbf{u}\mathbf{u}^t)^t(\mathbf{I} - 2\mathbf{u}\mathbf{u}^t) = \mathbf{I} - 4\mathbf{u}\mathbf{u}^t + 4\mathbf{u}\mathbf{u}^t\mathbf{u}\mathbf{u}^t = \mathbf{I}. \quad (6.54)$$

Assim,

$$\mathbf{H} = \mathbf{H}^t = \mathbf{H}^{-1}. \quad (6.55)$$

As matrizes de Householder possuem a seguinte propriedade, a qual é muito importante no pré-processamento matricial em métodos numéricos:

Dado o vetor \mathbf{a} , então a matriz de Householder $\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^t/\|\mathbf{v}\|^2$, com $\mathbf{v} = \mathbf{a} - \mathbf{r}$, transforma \mathbf{a} em \mathbf{r} .

De fato,

$$\mathbf{H}\mathbf{a} = \mathbf{a} - 2\frac{(\mathbf{a} - \mathbf{r})(\mathbf{a} - \mathbf{r})^t\mathbf{a}}{(\mathbf{a} - \mathbf{r})^t(\mathbf{a} - \mathbf{r})}. \quad (6.56)$$

Como \mathbf{H} deve ser ortogonal, $\|\mathbf{H}\mathbf{a}\| = \|\mathbf{a}\| = \|\mathbf{r}\|$, tem-se $\mathbf{a}^t\mathbf{a} = \mathbf{r}^t\mathbf{r}$. Assim,

$$\begin{aligned} \frac{2(\mathbf{a} - \mathbf{r})^t\mathbf{a}}{(\mathbf{a} - \mathbf{r})^t(\mathbf{a} - \mathbf{r})} &= \frac{2(\mathbf{a}^t\mathbf{a} - \mathbf{r}^t\mathbf{a})}{(\mathbf{a}^t\mathbf{a} - 2\mathbf{a}^t\mathbf{r} + \mathbf{r}^t\mathbf{r})} \\ &= \frac{2(\mathbf{a}^t\mathbf{a} - \mathbf{r}^t\mathbf{a})}{2(\mathbf{a}^t\mathbf{a} - \mathbf{r}^t\mathbf{a})} = 1. \end{aligned}$$

Decorre que

$$\mathbf{H}\mathbf{a} = \mathbf{a} - (\mathbf{a} - \mathbf{r}) = \mathbf{r}. \quad (6.57)$$

Exemplo 6.71

Utilizando uma matriz de Householder, triangularizar a matriz

$$\mathbf{A} = \begin{bmatrix} 3 & 4 \\ 4 & 0 \end{bmatrix}.$$

Solução

Transforma-se o vetor

$$\mathbf{a} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

no vetor

$$\mathbf{r} = \begin{bmatrix} \|\mathbf{a}\| \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

A matriz de Householder, com

$$\mathbf{v} = \mathbf{a} - \mathbf{r} = \begin{bmatrix} -2 \\ 0 \end{bmatrix},$$

é

$$\mathbf{H} = \mathbf{I} - 2\mathbf{v}\mathbf{v}^t / \|\mathbf{v}\|^2 = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix}.$$

A matriz \mathbf{H} é simétrica e ortogonal, entretanto $\mathbf{H}\mathbf{A}$ é triangular:

$$\mathbf{H}\mathbf{A} = \begin{bmatrix} 0.6 & 0.8 \\ 0.8 & -0.6 \end{bmatrix} \begin{bmatrix} 3 & 4 \\ 4 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 2.4 \\ 0 & 3.2 \end{bmatrix}.$$

No exemplo anterior, deve ser observado que $\mathbf{A} = \mathbf{Q}\mathbf{R}$ com $\mathbf{Q} = \mathbf{H}^t = \mathbf{H}$. Em geral, a fatorização $\mathbf{Q}\mathbf{R}$ pode ser obtida com auxílio de sucessivas matrizes de Householder. A primeira delas, \mathbf{H}_1 , transforma a primeira coluna de \mathbf{A} na primeira coluna de \mathbf{R} . A segunda, \mathbf{H}_2 , transforma a segunda coluna de $\mathbf{H}_1\mathbf{A}$ na segunda coluna de \mathbf{R} e, assim, sucessivamente. O produto dessas matrizes é a matriz ortogonal \mathbf{Q} .

Métodos Iterativos e Equações em Diferenças

7.1 Introdução

O método da eliminação Gaussiana (ou qualquer uma de suas variantes) aplicado no sistema não-singular $\mathbf{A}\mathbf{u} = \mathbf{b}$, é denominado *método direto*. Pois, com os dados \mathbf{A} e \mathbf{b} é realizado um número finito de operações alagéricas para obter a solução.

Se \mathbf{A} é *esparsa*, no sentido que um alto percentual de seus elementos são zeros, a eliminação Gaussiana, usualmente, não é beneficiada pela presença desses zeros. Os elementos zero, num passo da eliminação, não necessariamente permanecem os mesmos nos passos seguintes (“*fill-in*”). Ainda que, hoje em dia, existam técnicas de armazenamento para diminuir este “*fill-in*”, há outros métodos, chamados de *métodos iterativos* que visam aproximar a solução da equação $\mathbf{A}\mathbf{u} = \mathbf{b}$, quando \mathbf{A} é grande e esparsa. Em muitos casos, estes métodos preservam os zeros e o padrão dos elementos não nulos.

A descrição de um método iterativo é a seguinte. Dada uma aproximação inicial \mathbf{u}_0 da solução e, utilizando-a \mathbf{u}_0 como dado de “entrada”, obtém-se uma nova aproximação \mathbf{u}_1 por algum algoritmo que utilize \mathbf{A} e \mathbf{b} denominado **passo de refinamento**. O processo é repetido com \mathbf{u}_1 , usando o mesmo algoritmo, para produzir \mathbf{u}_2 . Em geral, \mathbf{u}_{k+1} é um refinamento de \mathbf{u}_k . Espera-se que a seqüência de vetores $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_k, \dots$ convirja para a solução exata \mathbf{u} . A construção de um método iterativo segue o seguinte principio. A matriz \mathbf{A} é substituída por uma matriz mais simples e não singular \mathbf{B} de modo que produz o sistema equivalente

$$\mathbf{B}\mathbf{u} = (\mathbf{B} - \mathbf{A})\mathbf{u} + \mathbf{b} . \quad (7.1)$$

Multiplicando ambos membros deste sistema pela inversa de \mathbf{B} , decorre

$$\mathbf{u} = \mathbf{M}\mathbf{u} + \mathbf{d} , \quad \mathbf{M} = \mathbf{B}^{-1}(\mathbf{B} - \mathbf{A}) , \quad \mathbf{d} = \mathbf{B}^{-1}\mathbf{b} . \quad (7.2)$$

O membro da direita é considerado como um algoritmo de refinamento e define-se o método iterativo

$$\mathbf{u}_{k+1} = \mathbf{M}\mathbf{u}_k + \mathbf{d} , \quad (7.3)$$

que gera a seqüência de aproximações

$$\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_k, \dots . \quad (7.4)$$

7.2 Métodos de Jacobi e de Gauss-Seidel

O método mais antigo é o de Jacobi. A matriz \mathbf{B} é a parte diagonal da matriz \mathbf{A} . Se todos os elementos a_{ii} são não nulos, então o cálculo da inversa dessa matriz diagonal

é extremamente simples. A matriz \mathbf{B} para o método de Jacobi é denotada por $\mathbf{D} = \text{diag}[a_{11}, a_{12}, \dots, a_{nn}]$. A iteração

$$\mathbf{D}\mathbf{u}_{k+1} = (\mathbf{D} - \mathbf{A})\mathbf{u}_k + \mathbf{b}, \quad (7.5)$$

escreve-se em componentes

$$\begin{aligned} a_{11}(u_1)_{k+1} &= (-a_{12}u_2 - a_{13}u_3 - \dots - a_{1n}u_n)_k + b_1 \\ &\vdots \\ a_{nn}(u_n)_{k+1} &= (-a_{n1}u_1 - a_{n2}u_2 - \dots - a_{n,n-1}u_{n-1})_k + b_n. \end{aligned}$$

Assim,

$$(u_i)_{k+1} = -\frac{1}{a_{ij}} \left[\sum_{j \neq i} a_{ij}(u_j)_k + b_i \right], \quad i = 1, 2, \dots, n.$$

Quando \mathbf{A} é esparsa, na expressão acima, os termos à direita são em maioria zeros e o passo de \mathbf{u}_k para u_{k+1} , torna-se fácil. A questão importante, é saber quando e com que rapidez a iteração converge.

Exemplo 7.72

Para o sistema

$$\begin{bmatrix} 3 & 2 & 0 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \\ 3 \end{bmatrix},$$

as primeiras aproximações da solução exata, começando com $x^0 = 0$, são

$$\begin{bmatrix} 5/3 \\ 3/2 \\ 3/2 \end{bmatrix}, \quad \begin{bmatrix} 2/3 \\ 3/4 \\ 2/3 \end{bmatrix}, \quad \begin{bmatrix} 7/6 \\ 7/6 \\ 7/6 \end{bmatrix}, \quad \begin{bmatrix} 8/9 \\ 11/12 \\ 11/12 \end{bmatrix}$$

e observa-se uma rápida convergência para a solução exata

$$\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

A matriz de iteração é

$$\mathbf{M} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A} = \begin{bmatrix} 0 & 2/3 & 0 \\ 0 & 0 & 1/2 \\ 1/2 & 0 & 0 \end{bmatrix}.$$

Exemplo 7.73

Para o sistema

$$\begin{bmatrix} 3 & 2 & 1 \\ 0 & 2 & 3 \\ 2 & 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ 4 \end{bmatrix}$$

obtem-se as aproximações sucessivas

$$\begin{bmatrix} 2 \\ 5/2 \\ 2 \end{bmatrix}, \quad \begin{bmatrix} -1/3 \\ -1/2 \\ 0 \end{bmatrix}, \quad \begin{bmatrix} 7/3 \\ 5/2 \\ 7/3 \end{bmatrix}, \quad \dots$$

e não existe convergência para a solução exata $\mathbf{u} = [1 \ 1 \ 1]^t$.

Para discutir a convergência do método, é conveniente desenvolver conceitos básicos de equações em diferenças. Isto será feito na seguinte seção.

O método de Jacobi é sempre convergente, quando \mathbf{A} é *diagonal dominante*, isto é, com propriedade seguinte

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, 2, \dots, n. \quad (7.6)$$

Para matrizes \mathbf{A} simétricas e positivas definidas cujos autovalores são menores que 2, o método de Jacobi também é convergente.

7.2.1 Método de Gauss-Seidel

Para uma matriz \mathbf{A} de grande porte, a iteração de Jacobi requer que sejam mantidas todas as componentes de \mathbf{u}_k na memória, até que o cálculo de \mathbf{u}_{k+1} seja realizado. Uma variação deste método, consiste em utilizar as componentes de u_{k+1} logo que forem calculadas. Assim, u_{k+1} coloca-se na memória utilizada por \mathbf{u}_k , conforme cada componente vai sendo calculada. Por exemplo, para o sistema

$$\begin{bmatrix} 10^6 & 1 & 1 \\ 1 & 10^6 & 1 \\ 1 & 1 & 10^6 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \quad (7.7)$$

escreve-se a iteração de Jacobi

$$\begin{aligned} 10^6(u_1)_{k+1} &= b_1 - (u_2)_k - (u_3)_k \\ 10^6(u_2)_{k+1} &= b_2 - (u_1)_k - (u_3)_k \\ 10^6(u_3)_{k+1} &= b_3 - (u_1)_k - (u_2)_k. \end{aligned} \quad (7.8)$$

Com esta variação, a aproximação da primeira componente é determinada pela primeira equação. Este valor pode ser utilizado na segunda equação, para obter uma aproximação da segunda componente. As duas aproximações obtidas são utilizadas para obter a aproximação da terceira componente da terceira equação. Assim,

$$\begin{aligned} 10^6(u_1)_{k+1} &= b_1 - (u_2)_k - (u_3)_k \\ 10^6(u_2)_{k+1} &= b_2 - (u_1)_{k+1} - (u_3)_k \\ 10^6(u_3)_{k+1} &= b_3 - (u_1)_{k+1} - (u_2)_{k+1} \end{aligned} \quad (7.9)$$

Este método, atribuído a **Gauss-Seidel**, é escrito na forma matricial

$$\mathbf{D}\mathbf{u}_{k+1} = \mathbf{b} - \mathbf{L}\mathbf{u}_{k+1} - \mathbf{U}\mathbf{u}_k, \quad (7.10)$$

onde

$$\mathbf{D} = \text{diag}[a_{11}, a_{22}, \dots, a_{nn}] ,$$

$$\mathbf{L} = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ a_{21} & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{n,n-1} & 0 \end{bmatrix} ,$$

$$\mathbf{U} = \begin{bmatrix} 0 & a_{12} & \cdots & \cdots & a_{1n} \\ 0 & 0 & a_{23} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & 0 \end{bmatrix} .$$

Definindo

$$\mathbf{M} = (\mathbf{D} + \mathbf{L})^{-1}\mathbf{U} , \quad \mathbf{d} = (\mathbf{D} + \mathbf{L})^{-1}\mathbf{b} . \quad (7.11)$$

decorre o esquema iterativo de Gauss-Seidel

$$\mathbf{u}_{k+1} = \mathbf{M}\mathbf{u}_k + \mathbf{d} , \quad (7.12)$$

O método de Gauss-Seidel converge para os seguintes tipos de matrizes

1. \mathbf{A} é diagonal dominante.
2. \mathbf{A} é simétrica positiva definida.

7.3 Equações em Diferenças

Uma equação em diferenças é uma equação que relaciona o valor

$$\mathbf{u}_k = \begin{bmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_i(k) \end{bmatrix} \quad (7.13)$$

com outros valores (usualmente próximos) de uma seqüência de vetores $\mathbf{u}_0, \mathbf{u}_1, \mathbf{u}_2, \dots$. Por uma *equação linear em diferenças de primeira ordem com coeficientes constantes*, entende-se uma equação da forma

$$\mathbf{D}\mathbf{u}_{k+1} = \mathbf{B}\mathbf{u}_k + \mathbf{g}_k , \quad (7.14)$$

onde os coeficientes \mathbf{B} e \mathbf{D} são matrizes quadradas de ordem n e \mathbf{g}_k é um vetor $n \times 1$. A equação é considerada como *regular*, quando \mathbf{D} é não-singular; caso contrário como *singular*. Um equação regular é escrita como

$$\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k + \mathbf{f}_k \quad (7.15)$$

Este tipo de equações surge naturalmente na solução numérica de equações diferenciais ou em modelos, onde a variável independente é discreta (inteira) ou quando convém, matematicamente, considerá-la como tal. Parte da especificação, de uma equação em diferenças, é a coleção dos inteiros k , para os quais a relação 7.14 é válida. Esta coleção pode ser finita ou infinita. Por exemplo, $k = 0, 1, 2, \dots, N$ ou $k = p, p+1, \dots, p+N$ ou $k = 0, 1, 2, 3, \dots$.

7.3.1 Equações Homogêneas

No método de Jacobi ou de Gauss-Seidel, o erro

$$\mathbf{e}_k = \mathbf{u}_k - \mathbf{u} \quad (7.16)$$

satisfaz uma equação do tipo

$$\mathbf{e}_{k+1} = \mathbf{M}\mathbf{e}_k. \quad (7.17)$$

Quando o termo \mathbf{f}_k está ausente na equação 7.15, isto é,

$$\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k, \quad (7.18)$$

a equação em diferenças é dita *homogênea*.

A solução pode ser obtida por simples recursão, uma vez que um valor inicial \mathbf{u}_0 seja especificado. Por repetidas substituições,

$$\begin{aligned} \mathbf{u}_1 &= \mathbf{A}\mathbf{u}_0 \\ \mathbf{u}_2 &= \mathbf{A}\mathbf{u}_1 = \mathbf{A}\mathbf{A}\mathbf{u}_0 = \mathbf{A}^2\mathbf{u}_0 \\ \mathbf{u}_3 &= \mathbf{A}\mathbf{A}^2\mathbf{u}_0 = \mathbf{A}^3\mathbf{u}_0 \end{aligned} \quad (7.19)$$

e, de modo genérico,

$$\mathbf{u}_k = \mathbf{A}^k\mathbf{u}_0, \quad k = 0, 1, 2, \dots \quad (7.20)$$

Deste modo, a convergência dos métodos de Jacobi e de Gauss Seidel está diretamente relacionada com a convergência do termo $\mathbf{M}^k\mathbf{e}_0$ para zero. Isto pode ser feito através do seguinte método.

7.3.2 O Método Espectral: Soluções Exponenciais

A equação $\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k$ possui soluções da forma

$$\mathbf{u}_k = \lambda^k\mathbf{v}, \quad \mathbf{v} \neq 0 \quad (7.21)$$

quando \mathbf{v} é uma autovetor de \mathbf{A} . Pois, substituindo e simplificando

$$\lambda^{k+1}\mathbf{v} = \mathbf{A}(\lambda^k\mathbf{v}) = \lambda^k\mathbf{A}\mathbf{v} \quad (7.22)$$

decorre que $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

Se \mathbf{A} é uma matriz simétrica (ou é uma matriz não defeituosa) de ordem n , então tem-se n soluções exponenciais

$$(\lambda_1)^k\mathbf{v}_1, \quad (\lambda_2)^k\mathbf{v}_2, \quad \dots, \quad (\lambda_n)^k\mathbf{v}_n,$$

correspondentes a cada um dos n autovetores, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$, associados aos n autovalores $\lambda_1, \lambda_2, \dots, \lambda_n$ da matriz \mathbf{A} , os quais são linearmente independentes (ortogonais no caso \mathbf{A} simétrica). Nesta situação, a *superposição ou combinação linear* destas soluções exponenciais

$$\mathbf{u}_k = c_1(\lambda_1)^k\mathbf{v}_1 + c_2(\lambda_2)^k\mathbf{v}_2 + \dots + c_n(\lambda_n)^k\mathbf{v}_n, \quad (7.23)$$

é a *solução geral* da equação linear em diferenças

$$\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k. \quad (7.24)$$

De fato, esta combinação é sempre solução da equação, pois

$$\begin{aligned}
\mathbf{u}_{k+1} &= c_1(\lambda_1)^{k+1}\mathbf{v}_1 + \cdots + c_n(\lambda_n)^{k+1}\mathbf{v}_n \\
&= c_1(\lambda_1)^k\mathbf{A}\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{A}\mathbf{v}_n \\
&= \mathbf{A}[c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n] = \mathbf{A}\mathbf{u}_k .
\end{aligned} \tag{7.25}$$

De outro lado, como os n autovetores \mathbf{v}_j são linearmente independentes, para qualquer vetor $n \times 1$, considerado como dado inicial \mathbf{u}_0 , é possível determinar coeficientes c_j (ou $c_j = (\mathbf{v}_j)^t\mathbf{u}_0/(\mathbf{v}_j)^t\mathbf{v}_j$, quando os \mathbf{v}_j são ortogonais), tais que

$$\mathbf{u}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n . \tag{7.26}$$

Deste modo a equação $\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k$, com \mathbf{u}_0 dado, tem como solução

$$\begin{aligned}
\mathbf{u}_k = \mathbf{A}^k\mathbf{u}_0 &= c_1\mathbf{A}^k\mathbf{v}_1 + \cdots + c_n\mathbf{A}^k\mathbf{v}_n \\
&= c_1(\lambda_1)^k\mathbf{v}_1 + \cdots + c_n(\lambda_n)^k\mathbf{v}_n ,
\end{aligned} \tag{7.27}$$

que é uma combinação linear das soluções exponenciais.

Conclui-se que *as soluções da equação $\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k$ convergem para vetor $\mathbf{0}$ se todos os autovalores da matriz \mathbf{A} possui parte real menor que a unidade*. A taxa de convergência estará determinada pelo *raio espectral* da matriz \mathbf{A} , isto é, pelo autovalor com maior valor absoluto.

Exemplo 7.74

Resolver $\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k$ considerando a matriz

$$\mathbf{A} = \begin{bmatrix} 7 & -2 & 1 \\ -2 & 10 & -2 \\ 1 & -2 & 7 \end{bmatrix} .$$

Solução

Os autovalores de \mathbf{A} são 6 (duplo) e 12. Para $\lambda = 6$, tem-se os autovetores

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \quad \text{e} \quad \mathbf{v}_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} ;$$

Para $\lambda = 12$, o autovetor associado

$$\mathbf{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}$$

é ortogonal a \mathbf{v}_1 e \mathbf{v}_2 .

Escrevendo o vetor \mathbf{u}_0 na forma

$$\mathbf{u}_0 = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + c_3\mathbf{v}_3 ,$$

onde \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 são os autovetores de \mathbf{A} , decorre

$$\mathbf{u}_k = c_1 6^k \mathbf{v}_1 + c_2 6^k \mathbf{v}_2 + c_3 12^k \mathbf{v}_3 .$$

Exemplo 7.75

Verificar que o método de Jacobi converge para um sistema com $\mathbf{Ax} = \mathbf{b}$, com

$$\mathbf{A} = \begin{bmatrix} 3 & 2 & 2 \\ 0 & 2 & 1 \\ 1 & 0 & 2 \end{bmatrix},$$

Solução

Tem-se a matriz de iteração

$$\mathbf{M} = \mathbf{I} - \mathbf{D}^{-1}\mathbf{A} = \begin{bmatrix} 0 & 2/3 & 2/3 \\ 0 & 0 & 1/2 \\ 1/2 & 0 & 0 \end{bmatrix},$$

cujo raio espectral é cerca de 0.75. Por tanto, há convergência.

7.3.3 Desacoplamento

O método espectral, descrito na secção anterior em termos das soluções exponenciais, pode ser analisado matricialmente como segue.

A matriz modal

$$\mathbf{V} = [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] \quad (7.28)$$

juntamente com a matriz espectral

$$\mathbf{D} = \text{diag}[\lambda_1 \ \cdots \ \lambda_n] \quad (7.29)$$

ambas correspondentes à matriz \mathbf{A} , tornam possível a fatorização

$$\mathbf{A} = \mathbf{VDV}^{-1}, \quad (7.30)$$

No caso, \mathbf{A} simétrica, tem-se que $V^{-1} = V^t$. Segue,

$$\mathbf{A}^2 = (\mathbf{VDV}^{-1})(\mathbf{VDV}^{-1}) = \mathbf{VD}^2\mathbf{V}^{-1} \quad (7.31)$$

e

$$\mathbf{A}^3 = \mathbf{VD}^3\mathbf{V}^{-1} \quad (7.32)$$

e, genericamente

$$\mathbf{A}^k = \mathbf{VD}^k\mathbf{V}^{-1}. \quad (7.33)$$

Sustituindo esta expressão para \mathbf{A}^k em 7.20, resulta

$$\mathbf{u}_k = (\mathbf{VDV}^{-1})^k \mathbf{u}_0 = \mathbf{VD}^k\mathbf{V}^{-1}\mathbf{u}_0, \quad k = 1, 2, \dots \quad (7.34)$$

Esta fórmula, está relacionada com a combinação linear de soluções exponenciais. De fato, se

$$\mathbf{c} = \mathbf{V}^{-1} \mathbf{u}_0 = [c_1 \ \cdots \ c_n]^t, \quad (7.35)$$

então

$$\mathbf{Vc} = \mathbf{u}_0 \quad (7.36)$$

e

$$\begin{aligned}
\mathbf{u}_k &= \mathbf{V} \mathbf{D}^k \mathbf{V}^{-1} \mathbf{u}_0 \\
&= [\mathbf{v}_1 \ \cdots \ \mathbf{v}_n] \begin{bmatrix} (\lambda_1)^k & & & \\ & (\lambda_2)^k & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \\
&= c_1 (\lambda_1)^k \mathbf{v}_1 + \cdots + c_n (\lambda_n)^k \mathbf{v}_n .
\end{aligned} \tag{7.37}$$

Em particular, os coeficientes c_j para o valor inicial \mathbf{u}_0 são tais que

$$\mathbf{u}_0 = c_1 (\lambda_1)^0 \mathbf{v}_1 + \cdots + c_n (\lambda_n)^0 \mathbf{v}_n , \tag{7.38}$$

ou, equivalentemente, $\mathbf{V} \mathbf{c} = \mathbf{u}_0$, ou seja, $\mathbf{c} = \mathbf{V}^{-1} \mathbf{u}_0$.

Uma outra maneira de realizar essa operação, que produz o mesmo resultado, é substituir $\mathbf{A} = \mathbf{V} \mathbf{D} \mathbf{V}^{-1}$ diretamente na equação em diferenças $\mathbf{u}_{k+1} = \mathbf{A} \mathbf{u}_k$ e, após, multiplicar pela inversa de \mathbf{V} à esquerda de ambos os membros, obtendo-se

$$\mathbf{V}^{-1} \mathbf{u}_{k+1} = \mathbf{D} \mathbf{V}^{-1} \mathbf{u}_k . \tag{7.39}$$

Através da mudança de variável

$$\mathbf{z}_k = \mathbf{V}^{-1} \mathbf{u}_k , \quad \mathbf{u}_k = \mathbf{V} \mathbf{z}_k , \tag{7.40}$$

decorre

$$\mathbf{z}_{k+1} = \mathbf{D} \mathbf{z}_k , \tag{7.41}$$

cuja solução é

$$\mathbf{z}_k = \mathbf{D}^k \mathbf{z}_0 = \begin{bmatrix} (\lambda_1)^k & & & \\ & (\lambda_2)^k & & \\ & & \ddots & \\ & & & (\lambda_n)^k \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} c_1 (\lambda_1)^k \\ c_2 (\lambda_2)^k \\ \vdots \\ c_n (\lambda_n)^k \end{bmatrix} . \tag{7.42}$$

Aqui c_j denota a j -ésima componente de \mathbf{z}_0 .

Observe-se que a solução \mathbf{z}_k dada por 7.42, pode ser obtida diretamente do sistema 7.41, o qual está na forma **desacoplada**

$$\begin{aligned}
(z_{k+1})_1 &= (\lambda_1)(z_k)_1 \\
(z_{k+1})_2 &= (\lambda_2)(z_k)_2 \\
&\vdots \\
(z_{k+1})_n &= (\lambda_n)(z_k)_n .
\end{aligned} \tag{7.43}$$

Deste modo a determinação de cada componente de z_k não estará vinculada às demais.

7.4 Soluções Transientes e Estacionárias

Considere-se a equação não-homogênea 7.15, em que o termo f_k independe de k , isto é,

$$\mathbf{u}_{k+1} = \mathbf{A}\mathbf{u}_k + \mathbf{f} . \quad (7.44)$$

A solução geral escreve-se

$$u_k = u_{h,k} + u_{p,k} \quad (7.45)$$

onde $u_{h,k}$ é a solução geral homogênea e $u_{p,k}$ é alguma solução particular não-homogênea. Supondo que a matriz \mathbf{A} é de ordem $n \times n$, não defeituosa e com autovalores e autovetores λ_j e \mathbf{v}_j , respectivamente, tem-se

$$u_{h,k} = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k + \cdots + c_n \lambda_n^k \mathbf{v}_n \quad (7.46)$$

Como o termo não homogêneo é um vetor constante, procura-se determinar uma *solução estacionária*

$$\mathbf{u}_{p,k} = \mathbf{c} \quad (7.47)$$

que não depende de k . Assim,

$$\mathbf{c} = \mathbf{A}\mathbf{c} + \mathbf{f} , \quad (7.48)$$

ou seja,

$$(\mathbf{I} - \mathbf{A})\mathbf{c} = \mathbf{f} . \quad (7.49)$$

Se $(\mathbf{I} - \mathbf{A})$ é uma matriz não singular, isto é, $\lambda = 1$ não é autovalor de \mathbf{A} , então

$$\mathbf{c} = (\mathbf{I} - \mathbf{A})^{-1} \mathbf{f} , \quad (7.50)$$

é a única solução estacionária de ??.

Se a matriz \mathbf{A} possui raio espectral menor que 1, decorre que a solução geral

$$u_k = u_{h,k} + u_{p,k} = c_1 \lambda_1^k \mathbf{v}_1 + c_2 \lambda_2^k + \cdots + c_n \lambda_n^k \mathbf{v}_n + (\mathbf{I} - \mathbf{A})^{-1} \mathbf{f} \quad (7.51)$$

converge para a solução estacionária quando $k \rightarrow \infty$. Nesta situação, a solução homogênea é dita *transiente* e a solução estacionária é dita, também, de *solução permanente*.

7.4.1 Método Operacional

A resolução de equações em diferenças pode ser realizada com o uso da transformada z . Para \mathbf{u}_k , define-se

$$U(z) = \sum_{k=0}^{\infty} \frac{\mathbf{u}_k}{z^k} \quad (7.52)$$

como sendo a *transformada* z da função de variável discreta $\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_k, \dots$. A propriedade de translação

$$z U(z) - z \mathbf{u}_0 = \sum_{k=0}^{\infty} \frac{\mathbf{u}_{k+1}}{z^k} \quad (7.53)$$

permite transformar uma equação em diferenças

$$\mathbf{u}_{k+1} = \mathbf{A} \mathbf{u}_k + \mathbf{f}_k, \quad (7.54)$$

com valor inicial \mathbf{u}_o , em uma equação algébrica para $U(z)$, isto é,

$$z U(z) - z \mathbf{u}_o = \mathbf{A} U(z) + F(z) \quad (7.55)$$

ou

$$U(z) = [z \mathbf{I} - \mathbf{A}]^{-1} z \mathbf{u}_o + [z \mathbf{I} - \mathbf{A}]^{-1} F(z) \quad (7.56)$$

A solução \mathbf{u}_k vem a ser a *transformada inversa* de $U(z)$.

Para uma equação em diferenças de segunda ordem,

$$\mathbf{A} \mathbf{u}_{k+2} + \mathbf{B} \mathbf{u}_{k+1} + \mathbf{C} \mathbf{u}_k = \mathbf{f}_k \quad (7.57)$$

com valores iniciais \mathbf{u}_o e \mathbf{u}_1 , o método operacional é o mesmo, tendo em conta a propriedade de translação

$$z^2 U(z) - z^2 [\mathbf{u}_o + \mathbf{u}_1 z^{-1}] = \sum_{k=0}^{\infty} \frac{\mathbf{u}_{k+2}}{z^k}. \quad (7.58)$$

No MATLAB, os comandos para a transformada z e sua inversa são $ztrans(u)$ e $invztrans(U)$, respectivamente. Por exemplo,

$$u = 2^n/7 - (-5)^n/7;$$

$$U = trans(u) = \frac{z}{(z-2)(z+5)};$$

$$u = invztrans(U) = 2^n/7 - (-5)^n/7.$$

No caso de vetores e matrizes, a transformada z e sua inversa devem ser aplicadas a cada componente.

No MAPLE, os comandos são os mesmos. Somente muda a maneira de entrar a função

$$ztrans(u(n), n, z); \quad invztrans(U(z), z, n)$$

e mantém a notação para as variáveis independentes original e da transformada. No caso de vetores e matrizes, a transformada z e sua inversa devem ser aplicadas a cada componente. No Maple, é utilizando o comando adicional *map*.

7.5 Método da Potência

Diversos métodos iterativos para o cálculo dos autovalores de uma matriz têm sido desenvolvidos. O mais simples, e o menos eficiente, é o chamado *método da potência*. Se \mathbf{A} é uma matriz não-defeituosa com um autovalor dominante, digamos λ_1 , então da expansão espectral

$$\mathbf{z}_o = c_1 \lambda_1 \mathbf{v}_1 + c_2 \lambda_2 \mathbf{v}_2 + \cdots + c_n \lambda_n \mathbf{v}_n, \quad (7.59)$$

obtem-se que

$$\mathbf{z}_k = \mathbf{A} \mathbf{z}_{k-1} = \mathbf{A}^k \mathbf{z}_0 = \sum_{j=1}^n c_j \lambda_j^k \mathbf{v}_j . \quad (7.60)$$

Portanto,

$$\frac{\mathbf{z}_k}{\lambda_1^k} = c_1 \mathbf{v}_1 + \sum_{j=2}^n c_j \frac{\lambda_j^k}{\lambda_1^k} \mathbf{v}_j \quad (7.61)$$

e, no limite, quando k tende para infinito, a seqüência converge para um autovetor associado ao maior autovalor. De outro lado, o quociente de i -ésimas componentes arbitrárias, entre dois iterados consecutivos

$$\frac{z_{k+1, i}}{z_{k, i}} , \quad (7.62)$$

tende para o maior autovalor λ_1 .

Outros métodos iterativos consistem em considerar uma seqüência de matrizes similares $\mathbf{A}_{k+1} = \mathbf{Q}_k^{-1} \mathbf{A}_k \mathbf{Q}_k$, tais que possuam os mesmos autovalores da matriz \mathbf{A} e que converjam para uma matriz que exiba os autovalores de \mathbf{A} (por exemplo, uma matriz diagonal). Em particular, tem-se o método QR, no qual a seqüência é gerada por

$$\mathbf{A}_k = \mathbf{Q}_k \mathbf{R}_k, \quad \mathbf{R}_k \mathbf{Q}_k = \mathbf{A}_{k+1}, \quad k = 0, 1, 2, \dots , \quad (7.63)$$

com \mathbf{Q}_k ortogonal e \mathbf{R}_k triangular superior. A seqüência converge para uma matriz triangular superior. Se \mathbf{A} for uma matriz simétrica, ela poderá ser pré-processada (reduzida para uma matriz tridiagonal) e o método será aplicado a esta última matriz. este método é utilizado no MATALB.

CAPÍTULO 8

Métodos Variacionais

Uma maneira diferente de resolver o sistema $\mathbf{Ax} = \mathbf{b}$, consiste em considerá-lo como proveniente de um processo de otimização. Por exemplo, no caso escalar, considere-se a parábola

$$q(x) = \frac{1}{2} a x^2 - b x. \quad (8.1)$$

Esta função possui um valor crítico quando

$$q'(x) = a x - b = 0. \quad (8.2)$$

Ou seja,

$$x = a^{-1} b \quad (8.3)$$

Sendo que $q''(x) = a$, decorre que a parábola possui um mínimo global para esse valor de x quando $a > 0$ e um máximo global se $a < 0$. Então, *resolver $ax = b$ equivale a minimizar (ou maximizar) a função quadrática $q(x) = \frac{1}{2} a x^2 - b x$.*

No caso matricial, considere-se a função

$$\mathbf{q}(\mathbf{x}) = \frac{1}{2} \mathbf{x}^t \mathbf{A} \mathbf{x} - \mathbf{x}^t \mathbf{b}. \quad (8.4)$$

Suponha-se que a matriz \mathbf{A} é simétrica e positiva definida. No lugar de utilizar as condições do cálculo, pode ser estabelecido que se \mathbf{x}_m for a solução de $\mathbf{Ax} = \mathbf{b}$, então, para um outro vetor \mathbf{y} , tem-se

$$\begin{aligned} \mathbf{q}(\mathbf{y}) - \mathbf{q}(\mathbf{x}_m) &= \frac{1}{2} \mathbf{y}^t \mathbf{A} \mathbf{y} - \mathbf{y}^t \mathbf{b} - \frac{1}{2} \mathbf{x}_m^t \mathbf{A} \mathbf{x}_m + \mathbf{x}_m^t \mathbf{b} \\ &= \frac{1}{2} (\mathbf{y} - \mathbf{x}_m)^t \mathbf{A} (\mathbf{y} - \mathbf{x}_m). \end{aligned} \quad (8.5)$$

Como \mathbf{A} é positiva definida, a última expressão nunca pode ser negativa, e é zero se $\mathbf{y} - \mathbf{x}_m = \mathbf{0}$. Logo, $\mathbf{q}(\mathbf{x})$ assume um valor mínimo para $\mathbf{x}_m = \mathbf{A}^{-1} \mathbf{b}$. Neste ponto

$$\begin{aligned} q_{min} &= \frac{1}{2} \mathbf{A}^{-1} \mathbf{b}^t \mathbf{A} \mathbf{A}^{-1} \mathbf{b} - \mathbf{A}^{-1} \mathbf{b}^t \mathbf{b} \\ &= -\frac{1}{2} \mathbf{A}^{-1} \mathbf{b}^t \mathbf{b}. \end{aligned} \quad (8.6)$$

O recíproco também é válido. Se \mathbf{x}_m minimiza $q(\mathbf{x})$, então $\mathbf{Ax}_m = \mathbf{b}$.

Resumindo,

Se \mathbf{A} é simétrica e positiva definida, então

$q(\mathbf{x})$ atinge um valor mínimo no ponto onde $\mathbf{Ax} = \mathbf{b}$.

O valor mínimo é $\mathbf{q}(\mathbf{A}^{-1} \mathbf{b}) = -\frac{1}{2} \mathbf{b}^t \mathbf{A}^{-1} \mathbf{b}$.

8.0.1 Método de Rayleigh-Ritz

Considere-se o problema de temperatura estacionária

$$-\nabla^2 T(x, y) = f(x, y) \quad (8.7)$$

numa placa que ocupa a região plana Ω , e que assume o valor de contorno

$$T(x, y) = 0 \quad (8.8)$$

para pontos no bordo Γ da placa.

O método de Rayleigh-Ritz consiste de três passos.

1. A equação dada é escrita na forma compacta

$$Ku = F \quad (8.9)$$

onde $K = -\nabla^2$, $u = T(x, y)$ e $F = f(x, y)$. Considera-se a forma quadrática funcional

$$q(\mathbf{u}) = \frac{1}{2} u^t K u - u^t F \quad (8.10)$$

onde é definido o produto escalar funcional

$$u^t v = \int_{\Omega} u(x, y) v(x, y) dx dy. \quad (8.11)$$

Assim

$$q = \frac{1}{2} \int_{\Omega} u(x, y) \nabla^2 u(x, y) dx dy - \int_{\Omega} u(x, y) f(x, y) dx dy \quad (8.12)$$

Utilizando integração por partes, tem-se

$$\int_{\Omega} u(x, y) \frac{\partial u}{\partial x} dx dy = u(x, y) \frac{\partial u}{\partial x} \Big|_{\Gamma} - \int_{\Omega} \frac{\partial u}{\partial x} \frac{\partial u}{\partial x} dx dy \quad (8.13)$$

Com a condição de contorno $u(x, y) = 0$ para pontos em Γ , resulta

$$\int_{\Omega} u(x, y) \frac{\partial u}{\partial x} dx dy = - \int_{\Omega} \left(\frac{\partial u}{\partial x} \right)^2 dx dy \quad (8.14)$$

Similarmente, com a integração com a derivada parcial com relação a y . Observe-se que essas integrações podem ser simplificadas com o auxílio da identidade de Green

$$\int_{\Omega} u(x, y) \nabla^2 v(x, y) dx dy = \int_{\Omega} \nabla v(x, y) \cdot \nabla u(x, y) dx dy - \int_{\Gamma} u(x, y) \frac{\partial v(x, y)}{\partial n} dS$$

onde $\frac{\partial v}{\partial n} = \nabla v \cdot \mathbf{n}$ é a derivada na direção normal exterior ao contorno Γ . Assim,

$$q = \int_{\Omega} \left[\left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2 \right] dx dy - \int_{\Omega} f(x, y) u(x, y) dx dy \quad (8.15)$$

Esta integração apresenta duas características. Na equação original aparece derivada parcial segunda, entretanto, após a integração, somente aparece a primeira derivada. Durante a integração é incorporada a condição de contorno.

2. É introduzida uma aproximação funcional para a temperatura

$$u = T(x, y) = u_1 \phi_1(x, y) + u_2 \phi_2(x, y) + \cdots + u_N \phi_N(x, y) \quad (8.16)$$

onde as funções $\phi_i(x, y)$ satisfazem as condições de contorno do problema.

Substituindo u , $\frac{\partial u}{\partial x}$ e $\frac{\partial u}{\partial y}$ na forma q , decorre

$$q = \frac{1}{2} \mathbf{U}^t \mathbf{K} \mathbf{U} - \mathbf{U}^t \mathbf{F} \quad (8.17)$$

onde

$$\mathbf{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_N \end{bmatrix}, \quad \mathbf{K} = [k_{ij}] \quad (8.18)$$

onde

$$f_k = \int_{\Omega} f(x, y) \phi_k(x, y) dx dy, \quad k_{ij} = \int_{\Omega} \left[\frac{\partial \phi_i}{\partial x} \frac{\partial \phi_j}{\partial x} + \frac{\partial \phi_i}{\partial y} \frac{\partial \phi_j}{\partial y} \right] dx dy \quad (8.19)$$

3. O sistema

$$\mathbf{K} \mathbf{U} = \mathbf{F} \quad (8.20)$$

é resolvido para determinar os valores u_k nos quais q atinge valor mínimo.

Deve ser salientado que o método de Rayleigh-Ritz é válido para *problemas simétricos e positivos definidos*, isto é o operador \mathbf{K} verifica

$$\mathbf{u}^t \mathbf{K} \mathbf{v} = \mathbf{v}^t \mathbf{K} \mathbf{u}, \quad \mathbf{u}^t \mathbf{K} \mathbf{u} > c \mathbf{u}^t \mathbf{u} \quad (8.21)$$

onde c é uma constante positiva. A matriz \mathbf{K} que resulta do processo de minimização da forma quadrática q preserva essas propriedades, isto é, é simétrica e positiva definida.