

Capítulo 23

Métodos numéricos para autovalores

Consideramos o problema linear de autovalor na forma $(\mathbf{K} - \lambda\mathbf{M})\phi = \mathbf{0}$, com matrizes \mathbf{K} e \mathbf{M} simétricas de ordem N . A determinação da solução deste problema pode ser feita usando um dos vários métodos disponíveis em álgebra, como a família de métodos de Jacobi, ou ainda o método QR se uma das matrizes for não simétrica (como ocorre, por exemplo, em problemas com efeito giroscópico como em rotores e discos girantes). Estes métodos são iterativos e produzem todos os N autopares do problema. São métodos adequados apenas para lidar com problemas pequenos, algo da ordem de, no máximo, algumas centenas ou alguns milhares de graus de liberdade. A partir de um certo tamanho dois efeitos indesejáveis ocorrem. Por um lado, o tempo computacional cresce insuportavelmente, e por outro lado, o número de operações e iterações necessárias cresce de tal forma que os erros de truncamento inutilizam os resultados ou impedem a convergência. Estes métodos não serão vistos aqui, pois podem ser revisados em livros clássicos de álgebra. Também os algoritmos encontram-se programados na maioria das bibliotecas de rotinas comerciais, em Fortran, como a biblioteca IMSL[®] [52] e outras, e como função em programas como Matlab e Mathematica.

Neste capítulo nos dedicamos aos métodos adequados ao trabalho com sistemas de grande porte, como se torna necessário quando as matrizes são geradas pelo método de elementos finitos para modelamento detalhado de peças e sistemas de geometrias complexas. Os métodos que se prestam a isto buscam a determinação de um número pequeno de autopares, m , com $m \ll N$. Essa é exatamente a situação desejada no método de sobreposição modal, por exemplo (Capítulo 19). Este número pequeno deve sempre permanecer dentro do limite de uso dos métodos “algébricos” referidos no parágrafo acima, métodos de Jacobi ou QR para problemas de autovalores reais ou complexos, respectivamente.

23.1 Propriedades das matrizes \mathbf{K} e \mathbf{M} e dos autoproblemas

A energia de deformação e a energia cinética de um corpo que ocupa o volume V são dadas por

$$U = \frac{1}{2} \int_V \boldsymbol{\varepsilon} \cdot \boldsymbol{\sigma} dV, \quad \text{e} \quad T = \frac{1}{2} \int_V \rho \dot{\mathbf{u}} \cdot \dot{\mathbf{u}} dV. \quad (23.1)$$

e se o material for elástico linear a energia de deformação é $U = \frac{1}{2} \int_V \boldsymbol{\varepsilon} \cdot \mathbf{C} \boldsymbol{\varepsilon} dV$. Observe que, $dT = \frac{1}{2} \rho \dot{\mathbf{u}} \cdot \dot{\mathbf{u}} dV$ é uma quantidade sempre positiva, o que implica que T também o é. Fazendo a discretização por elementos finitos tem-se as aproximações das energias em suas formas semi-discretizadas:

$$\begin{aligned} U &= \frac{1}{2} \mathbf{U}^T(t) \underbrace{\left[\int_V \mathbf{B}(\mathbf{x})^T \mathbf{C} \mathbf{B}(\mathbf{x}) dV \right]}_{\mathbf{K}} \mathbf{U}(t), \\ T &= \frac{1}{2} \dot{\mathbf{U}}^T(t) \underbrace{\left[\int_V \rho \mathbf{N}(\mathbf{x})^T \mathbf{N}(\mathbf{x}) dV \right]}_{\mathbf{M}} \dot{\mathbf{U}}(t). \end{aligned} \quad (23.2)$$

Os termos entre colchetes podem ser reconhecidos como as matrizes de rigidez e massa \mathbf{K} e \mathbf{M} , de tal forma que as energias ficam

$$U = \frac{1}{2} \mathbf{U}^T(t) \mathbf{K} \mathbf{U}(t), \quad \text{e} \quad T = \frac{1}{2} \dot{\mathbf{U}}^T(t) \mathbf{M} \dot{\mathbf{U}}(t). \quad (23.3)$$

Consideremos o conceito de **função positiva definida**. Uma função de várias variáveis é dita ser positiva definida se ela é sempre não negativa, e é zero **se e apenas se** todas as variáveis são zero. A energia cinética em (23.3) é uma função das velocidades nodais $U_j(t)$, e é totalmente definida pelos coeficientes da matriz massa, M_{ij} . Então a positividade de T é ditada pelas características de \mathbf{M} . Como, fisicamente, sabe-se que a energia cinética é uma grandeza positiva definida, a matriz \mathbf{M} (consistente) deve ser *positiva definida*. Isto significa que \mathbf{M} deve ser tal que, dado um vetor \mathbf{V} arbitrário de ordem N ,

$$\begin{aligned} \mathbf{V}^T \mathbf{M} \mathbf{V} &\geq 0 & \text{para} & \quad \forall \mathbf{V} \neq \mathbf{0}, \text{ e} \\ \mathbf{V}^T \mathbf{M} \mathbf{V} &= 0 & \iff & \quad \mathbf{V} = \mathbf{0}. \end{aligned} \quad (23.4)$$

Esta é a definição de uma matriz positiva definida. Em (23.3), \mathbf{M} é tal que, qualquer que sejam as velocidades nodais, T será sempre não negativo.

A importância de se saber se uma matriz é ou não positiva definida surge no momento de se construir ou escolher algoritmos de solução aos problemas onde ela aparece. Existem algoritmos que tiram partido do fato, e só admitem, matrizes simétricas ou positiva-definida.

Quanto à **matriz de rigidez** a análise tem dois casos distintos. Observe que um deslocamento de corpo rígido num corpo, como uma translação, não envolve deformações. Neste caso o vetor de deslocamentos nodais $\mathbf{U}(t)$ é não nulo, porém a deformação ε e consequentemente a energia de deformação em (23.1) é nula. Neste caso a matriz de rigidez em (23.3) não é positiva definida, uma vez que não satisfaz as condições de (23.4). Como U neste caso não pode ser negativo, \mathbf{K} é dita matriz **positiva semidefinida**.

Por outro lado, se condições de contorno tiverem sido aplicadas no sistema de forma a impossibilitar movimentos de corpo rígido, a deformação ε será não nula se o campo $\mathbf{u}(t)$ for não trivial. Então $U \geq 0$ e \mathbf{K} é positiva definida. Maiores detalhes sobre a positividade da rigidez são vistos na seção 377.

A seção 19.2.3, página 509, mostra que, se \mathbf{M} e \mathbf{K} são simétricos, dois autovetores quaisquer ϕ^i e ϕ^j do autoproblema

$$[\mathbf{K} - \omega^2 \mathbf{M}] \phi = \mathbf{0} \quad (23.5)$$

são **M** ortogonais, isto é, $\phi^i \mathbf{M} \phi^j = \delta_{ij}$, se $\omega_i \neq \omega_j$.

23.2 Método da potência ou de Stodola

Este é um método iterativo para aproximar a solução do **primeiro autopar** do problema de autovalor (23.5), quando \mathbf{K} e \mathbf{M} são simétricos e \mathbf{K} é não-singular, isto é, o corpo está vinculado de forma a evitar deslocamentos e rotações de corpo rígido. Primeiro autopar em problema significa o autopar composto pelo menor autovalor e o correspondente autovetor.

Considere por um momento que conhecêssemos, de alguma maneira, o valor do primeiro autovalor, ω_1 , e buscássemos uma aproximação para o correspondente autovetor ϕ^1 . Observe que num sistema de grande porte não temos nenhuma esperança de obter o primeiro autopar verdadeiro, exato, de forma que o desafio consiste em “apenas” obter valores aproximados. Note que isto representa um novo nível de aproximação. Tínhamos inicialmente o sistema físico, em seu comportamento próprio. Aproximamos este comportamento com as equações de equilíbrio e as demais relações de elasticidade. Como não podíamos resolver precisamente estas equações diferenciais, as aproximamos por elementos finitos no problema semi-discretizado representado pelas eqs. (23.5). Também estas equações não podem ser resolvidas precisamente e buscamos apenas uma aproximação de seu resultado.

anterior de y_j^1 em termos de y_j^o em (23.10) temos

$$y_j^2 = \left(\frac{\omega_1}{\omega_j}\right)^4 y_j^o. \quad (23.13)$$

Isto resulta para a segunda iteração,

$$\mathbf{v}_2 = \Phi \mathbf{y}^2. \quad (23.14)$$

Para as próximas iterações o padrão se repete:

$$\begin{aligned} \mathbf{v}_o &= \Phi \mathbf{y}^o \\ \mathbf{v}_1 &= \Phi \mathbf{y}^1 \quad \text{onde } y_j^1 = \left(\frac{\omega_1}{\omega_j}\right)^2 y_j^o, \\ \mathbf{v}_2 &= \Phi \mathbf{y}^2 \quad \text{onde } y_j^2 = \left(\frac{\omega_1}{\omega_j}\right)^4 y_j^o, \\ \mathbf{v}_3 &= \Phi \mathbf{y}^3 \quad \text{onde } y_j^3 = \left(\frac{\omega_1}{\omega_j}\right)^6 y_j^o, \\ &\vdots \\ \mathbf{v}_s &= \Phi \mathbf{y}^s \quad \text{onde } y_j^s = \left(\frac{\omega_1}{\omega_j}\right)^{2s} y_j^o. \end{aligned} \quad (23.15)$$

Numa iteração s qualquer, (23.15) pode ser expandido na forma

$$\mathbf{v}_s = \Phi \mathbf{y}^s = \phi^1 y_1^o + \underbrace{\phi^2 \left(\frac{\omega_1}{\omega_2}\right)^{2s}}_{\leq 1} y_2^o + \underbrace{\phi^3 \left(\frac{\omega_1}{\omega_3}\right)^{2s}}_{\leq 1} y_3^o + \dots + \underbrace{\phi^N \left(\frac{\omega_1}{\omega_N}\right)^{2s}}_{\leq 1} y_N^o. \quad (23.16)$$

Lembramos que nesta notação estamos considerando os autovalores ordenados de forma crescente, isto é, $\omega_1 < \omega_2 < \dots < \omega_j < \dots < \omega_N$. Então $0 < \omega_1/\omega_j \leq 1$. Ao mesmo tempo, $\lim_{s \rightarrow \infty} \alpha^{2s} = 0$ se α for um número real no intervalo aberto $(0; 1)$. Como consequência, os termos entre parêntesis são tais que

$$\lim_{s \rightarrow \infty} \left(\frac{\omega_1}{\omega_N}\right)^{2s} = 0 \quad \text{para } N \neq 1. \quad (23.17)$$

De (23.16) tem-se que, conforme o número de iterações s cresce, todos os termos à direita da igualdade tendem a zero com exceção do primeiro, de forma que

$$\boxed{\lim_{s \rightarrow \infty} \mathbf{v}_s = y_1^o \phi^1} \quad (23.18)$$

se os vetores \mathbf{v}_s são gerados como em (23.7) e o autovalor ω_1 for conhecido. Observe que se ϕ^1 é um autovetor de (23.5), $y_1^o \phi^1$ também o é. Uma constante afeta apenas a norma do vetor. Assim, o fato de se ignorar os valores y 's utilizados na dedução, não afeta a capacidade do procedimento em obter o autovetor.

Nesse ponto deve-se retornar ao fato de que, de fato, não se conhece previamente o primeiro autovalor ω_1 . Porém, considere-se uma iteração arbitrária s :

$$\frac{1}{\omega_1^2} \mathbf{v}_{s+1} = \underbrace{\mathbf{K}^{-1} \mathbf{M} \mathbf{v}_s}_{\bar{\mathbf{v}}_{s+1}}, \quad (23.19)$$

se ω_1 for conhecido. Normalmente, toma-se \mathbf{v}_s e calcula-se o lado direito, obtendo-se um vetor $\bar{\mathbf{v}}_{s+1}$. Nesta expressão \mathbf{v}_{s+1} é desconhecido e busca-se o valor de ω_1^2 . Uma das maneiras existentes de obter aproximações de ω_1 consiste em aproximar (23.19) usando o vetor da iteração anterior, \mathbf{v}_s que é conhecido, em lugar de \mathbf{v}_{s+1} , isto é,

$$\mathbf{v}_s = \alpha_{s+1}^2 \bar{\mathbf{v}}_{s+1}. \quad (23.20)$$

A princípio, se poderia estimar α_s^2 como o quociente entre um termo do vetor \mathbf{v}_s pelo termo correspondente de $\bar{\mathbf{v}}_{s+1}$, isto é,

$$\alpha_{s+1}^2 = \frac{\mathbf{v}_{s,1}}{\mathbf{v}_{s+1,1}} = \frac{\mathbf{v}_{s,2}}{\mathbf{v}_{s+1,2}} = \dots = \frac{\mathbf{v}_{s,j}}{\mathbf{v}_{s+1,j}} \quad (23.21)$$

se tivéssemos os vetores exatos. Como são aproximados, cada quociente tem um valor distinto. O melhor é tomar uma média. Fisicamente é mais adequado usar um produto escalar com a matriz massa como peso, em lugar da matriz identidade que resultaria num produto escalar Euclidiano. Então, pré-multiplicamos (23.20) por $\bar{\mathbf{v}}_{s+1}^T \mathbf{M}$:

$$\alpha_{s+1}^2 = \frac{\bar{\mathbf{v}}_{s+1}^T \mathbf{M} \mathbf{v}_s}{\bar{\mathbf{v}}_{s+1}^T \mathbf{M} \bar{\mathbf{v}}_{s+1}}. \quad (23.22)$$

Observe que se os vetores fossem a solução do problema, se teria $\mathbf{v}_s = \phi^1$ e $\bar{\mathbf{v}}_{s+1} = \mathbf{K}^{-1} \mathbf{M} \phi^1$. Então

$$\alpha_{s+1}^2 = \frac{\phi^{1T} \mathbf{M} \mathbf{K}^{-1} \mathbf{M} \phi^1}{\phi^{1T} \mathbf{M} \mathbf{K}^{-1} \mathbf{M} \mathbf{K}^{-1} \mathbf{M} \phi^1}.$$

Mas, considerando que $\mathbf{K}^{-1} \mathbf{M} \phi^1 = (1/\omega_1^2) \phi^1$, a expressão acima reduz-se a

$$\alpha_{s+1}^2 = \frac{\omega_1^{-2}}{\omega_1^{-4}} \quad \Rightarrow \quad \boxed{\alpha_{s+1}^2 = \omega_1^2} \quad (23.23)$$

Então, como os vetores \mathbf{v}_s convergem para ϕ^1 , a tendência é que α_s convirja para ω_1 .

A **síntese do algoritmo** é a mostrada na Tabela 23.1

Tabela 23.1: Fluxograma do método da potência ou de Stodola.

Etapa 1 - Inicializar $\mathbf{v}_o = \{1, 1, 1, \dots, 1\}^T$, $s = -1$;
Etapa 2 - Fatorar \mathbf{K} .
Etapa 3 - Início das iterações: $s = s + 1$, $\bar{\mathbf{v}}_{s+1} = \mathbf{K}^{-1} \mathbf{M} \mathbf{v}_s$, eq. (23.19)
Etapa 4 - $\check{\mathbf{v}}_{s+1} = \mathbf{M} \bar{\mathbf{v}}_{s+1}$,
Etapa 5 - $\alpha_{s+1}^2 = \frac{\check{\mathbf{v}}_{s+1}^T \mathbf{v}_s}{\bar{\mathbf{v}}_{s+1}^T \check{\mathbf{v}}_{s+1}}$, eq. (23.22)
Etapa 6 - $\mathbf{v}_{s+1} = \frac{\bar{\mathbf{v}}_{s+1}}{(\bar{\mathbf{v}}_{s+1}^T \mathbf{M} \bar{\mathbf{v}}_{s+1})^{1/2}} = \frac{\bar{\mathbf{v}}_{s+1}}{(\bar{\mathbf{v}}_{s+1}^T \check{\mathbf{v}}_{s+1})^{1/2}}$,
Etapa 7 - Testar convergência - Se $\frac{ \alpha_s - \alpha_{s+1} }{\alpha_{s+1}} > tol$, voltar à Etapa 3 , do contrário, $\phi^1 = \mathbf{v}_{s+1}$ e $\omega_1 = \alpha_{s+1}$.

Com esta sequência de operações, cada iteração envolve dois produtos vetor \times vetor e dois produtos matriz \times vetor. Previamente ao início das iterações a matriz de rigidez é triangularizada e armazenada. A cada iteração a operação na Etapa 3 consiste apenas na substituição retroativa para a determinação de $\bar{\mathbf{v}}_{s+1}$.

Dois aspectos afetam a **convergência do método**. Observe (23.16). Se, por exemplo, ω_2 for muito próximo de ω_1 , o valor $(\omega_1/\omega_2)^{2s}$ tende a zero muito lentamente, consequentemente retardando a convergência de \mathbf{v}_s , isto é, requerendo um número maior de iterações. Então a taxa de convergência do método depende das características de \mathbf{K} e \mathbf{M} que se refletem nos autovalores. Um valor usual para tol é 10^{-6} .

O segundo aspecto sobre a convergência pode ser identificando observando a eq. (23.8) que mostra a expansão do vetor inicial \mathbf{v}_o , em termos da base de autovetores. Observe o que ocorre se \mathbf{v}_o for escolhido de forma a não ter uma componente na direção ϕ^1 , isto é, se y_1^o for nulo. Então \mathbf{v}_o é ortogonal a ϕ^1 , o que torna impossível a convergência ao primeiro modo. Felizmente, dadas as matrizes do sistema, é praticamente impossível que \mathbf{v}_o seja gerado excluindo ϕ^1 .

Definição de \mathbf{v}_o : a princípio qualquer vetor não nulo poderia ser usado, como por exemplo $\mathbf{v}_o = \{1; 0; \dots 0\}^T$. Entretanto a prática mostra que $\mathbf{v}_o = \{1; 1; 1; \dots; 1\}^T$ é melhor por ativar todos os termos da matriz \mathbf{M} .

Exemplo 23.1

Considere o autop problema dos Exemplos dos Capítulos 18 e 19 nas páginas 496 e 512, definido por

$$\left\{ \left[\begin{array}{ccc} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{array} \right] - \alpha_j^2 \left[\begin{array}{ccc} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{array} \right] \right\} \left[\begin{array}{c} \phi_1^j \\ \phi_2^j \\ \phi_3^j \end{array} \right] = \mathbf{0}. \quad (23.24)$$

Use o método de Stodola para estimar o primeiro autotar. (A solução analítica é $\alpha_1^2 = (11 - 6\sqrt{3})/13 = 0,04674578112 \text{ m}^{-2}\text{s}^{-4}$ e $\phi^1 = [0,170518; 0,295345; 0,341035]^T$.)

Solução

Arbitra-se o vetor inicial $\mathbf{v}_o = \{1; 1; 1\}^T$ e seguem-se as etapas listadas em (23.1). A primeira estimativa \mathbf{v}_1 é obtida resolvendo

$$\mathbf{K} \bar{\mathbf{v}}_1 = \mathbf{M} \mathbf{v}_o = \begin{Bmatrix} 5 \\ 6 \\ 3 \end{Bmatrix} \quad \longrightarrow \quad \bar{\mathbf{v}}_1 = \begin{Bmatrix} 14 \\ 23 \\ 26 \end{Bmatrix}.$$

$$\check{\mathbf{v}}_1 = \mathbf{M} \bar{\mathbf{v}}_1 = \begin{Bmatrix} 79 \\ 132 \\ 75 \end{Bmatrix}.$$

$$\alpha_1^2 = \frac{\check{\mathbf{v}}_1^T \mathbf{v}_o}{\check{\mathbf{v}}_1^T \check{\mathbf{v}}_1} = \frac{63}{6,092} \quad \longrightarrow \quad \alpha_1 = 0,0464468.$$

$$\mathbf{v}_1 = \frac{\bar{\mathbf{v}}_1}{\check{\mathbf{v}}_1^T \check{\mathbf{v}}_1} = \begin{Bmatrix} 0,17937 \\ 0,29468 \\ 0,333114 \end{Bmatrix},$$

com o que se conclui a primeira iteração com as aproximações α_1 e \mathbf{v}_1 para o primeiro autotar. As próximas iterações produzem $\alpha_2^2 = 0,0467475$, $\alpha_3^2 = 0,0467458$ e $\mathbf{v}_3 = \{0,170593; 0,295345; 0,340961\}$. Comparando com a solução analítica isso resulta numa convergência em torno de 10^{-3} na frequência e 10^{-6} no primeiro modo.

23.3 Método de iteração progressiva

Este método pode ser visto como o mesmo método de Stodola, apenas alterado de forma a aproximar o último autotar em vez do primeiro, isto é, as iterações convergem para o maior autotar do sistema, ω_N^2 e ϕ_N .

Reescreve-se o problema (23.5) como

$$\mathbf{M} \phi = \frac{1}{\omega^2} \mathbf{K} \phi. \quad (23.25)$$

e define-se $\beta^2 = 1/\omega^2$. Aplicando o método de Stodola, tem-se uma aproximação do menor autovalor do problema, β_1^2 , que corresponderá então ao maior valor do sistema, ω_N^2 . Se \mathbf{M} puder ser invertida, (e as matrizes consistentes do MEF sempre são positivas definidas) o algoritmo do método fica o mesmo que o de Stodola no Fluxograma da Tabela 23.1, trocando \mathbf{K} por \mathbf{M} e vice-versa.

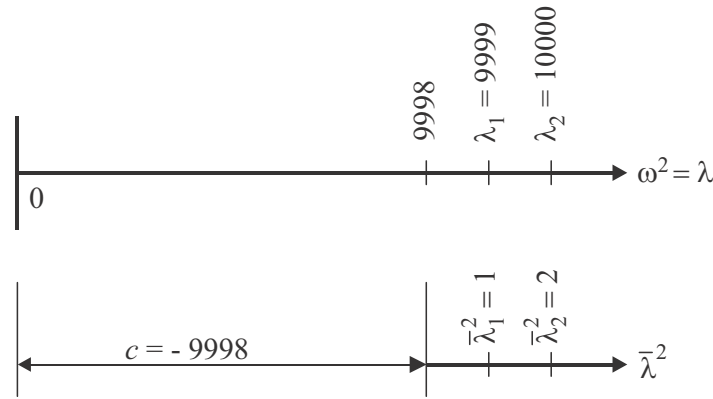


Figura 23.1: Deslocamento de autovalores para alterar taxa de convergência.

“shift” nos autovalores no valor de $c = -0,5$ e obtenha os novos valores $\bar{\lambda}$.

Solução

O novo autoproblema é obtido gerando uma nova matriz de rigidez $\bar{\mathbf{K}}$ a partir de (23.26), o que resulta em

$$\left\{ \left[\begin{array}{ccc} 0 & -1,5 & 0 \\ -1,5 & 0 & -1,5 \\ 0 & -1,5 & 0 \end{array} \right] - \bar{\lambda}_j \left[\begin{array}{ccc} 4 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 1 & 2 \end{array} \right] \right\} \begin{bmatrix} \phi_1^j \\ \phi_2^j \\ \phi_3^j \end{bmatrix} = \mathbf{0}.$$

O polinômio característico é

$$\rho(\bar{\lambda}) = [-32\bar{\lambda}^2 + 6(1,5 + \bar{\lambda})^2] = 0,$$

e suas raízes são $\bar{\lambda}_1 = (11 - 6\sqrt{3})/13 - 0,5 = -0,453234$, $\bar{\lambda}_2 = 0$, e $\bar{\lambda}_3 = (11 + 6\sqrt{3})/13 - 0,5 = 1,14556$. A matriz $\bar{\mathbf{K}}$ é então singular. Caso a matriz $\bar{\mathbf{K}}$ fosse originalmente singular um shift positivo a tornaria positiva-definida. Note que os autovetores do sistema acima são os mesmos obtidos no Exemplo da página 496 uma vez que eles não são afetados pelo shift.

23.4 Processo de ortogonalização de Gram-Schmidt

Como visto no Capítulo 19, métodos de solução do problema dinâmico, como o de sobreposição modal, além dos métodos de análise harmônica e análise espectral, dentre outros, dependem da prévia determinação da matriz modal, com todos os autovetores ortonormalizados. Entretanto, no caso de haverem pares de autovalores idênticos, os correspondentes autovetores não são naturalmente ortogonais entre si, de forma que torna-se necessária a realização de um processo de ortonormalização do conjunto.

Considere-se um espaço euclidiano N -dimensional V , isto é, V é o conjunto de todos os vetores de N componentes reais. Por exemplo, no espaço tridimensional, $N = 3$, cada vetor \mathbf{v} é definido por três componentes $\mathbf{v} = \{v_x; v_y; v_z\}^T$ em relação a base \hat{i}_x , \hat{i}_y e \hat{i}_z . Considere-se, inicialmente, dois vetores \mathbf{v}_1 e \mathbf{v}_2 , arbitrários, não ortogonais entre si, como na Figura 23.2.

Deseja-se usar o conjunto \mathbf{v}_1 e \mathbf{v}_2 de vetores para construir um outro conjunto, \mathbf{e}_1 e \mathbf{e}_2 que seja um **conjunto ortonormal de vetores**. (O termo “ortonormal” significa que os vetores são ortogonais entre si, e também que cada um deles é normalizado, isto é, tem norma unitária.) Um processo que se pode utilizar é o seguinte. Define-se o primeiro vetor \mathbf{e}_1 como um vetor unitário paralelo a \mathbf{v}_1 , isto é, $\mathbf{e}_1 = \mathbf{v}_1 / (\mathbf{v}_1 \cdot \mathbf{v}_1)^{1/2}$. Observando a figura nota-se que o módulo da componente de \mathbf{v}_2 na direção \mathbf{e}_1 é $\mathbf{v}_2 \cdot \mathbf{e}_1$ que forma um vetor $(\mathbf{v}_2 \cdot \mathbf{e}_1) \mathbf{e}_1$. Então

$$\mathbf{v}_2 = (\mathbf{v}_2 \cdot \mathbf{e}_1) \mathbf{e}_1 + \bar{\mathbf{e}}_2.$$

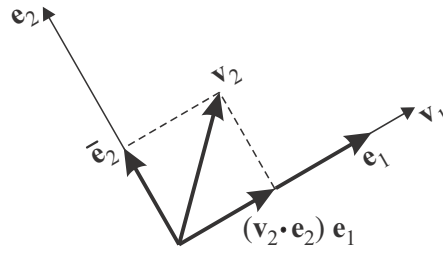


Figura 23.2: Ilustração de um processo de ortogonalização.

Logo, $\bar{\mathbf{e}}_2$ pode ser calculado através de

$$\bar{\mathbf{e}}_2 = \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{e}_1) \mathbf{e}_1. \quad (23.29)$$

Fazendo o produto interno de $\mathbf{e}_1 \cdot \bar{\mathbf{e}}_2$ é possível confirmar que estes vetores são ortogonais: $\mathbf{e}_1 \cdot \bar{\mathbf{e}}_2 = \mathbf{e}_1 \cdot \mathbf{v}_2 - (\mathbf{v}_2 \cdot \mathbf{e}_1) \mathbf{e}_1 \cdot \mathbf{e}_1 = 0$ uma vez que $\mathbf{e}_1 \cdot \mathbf{e}_1 = 1$. Finalmente $\bar{\mathbf{e}}_2$ pode ser normalizado gerando

$$\mathbf{e}_2 = \frac{\bar{\mathbf{e}}_2}{(\bar{\mathbf{e}}_2 \cdot \bar{\mathbf{e}}_2)^{1/2}}. \quad (23.30)$$

Dado um terceiro vetor no conjunto, \mathbf{v}_3 , não paralelo a \mathbf{v}_1 ou a \mathbf{v}_2 , pode-se gerar um terceiro vetor unitário \mathbf{e}_3 ortogonal a \mathbf{e}_1 e \mathbf{e}_2 por

$$\bar{\mathbf{e}}_3 = \mathbf{v}_3 - (\mathbf{v}_3 \cdot \mathbf{e}_1) \mathbf{e}_1 - (\mathbf{v}_3 \cdot \mathbf{e}_2) \mathbf{e}_2, \quad (23.31)$$

que seria em seguida normalizado. **De forma geral**, dados N vetores, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$, linearmente independentes, pode-se gerar um conjunto ortonormal de N vetores $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_N$, tal que $\mathbf{e}_i \cdot \mathbf{e}_j = \delta_{ij}$, da seguinte forma:

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{(\mathbf{v}_1 \cdot \mathbf{v}_1)^{1/2}}, \quad \bar{\mathbf{e}}_j = \mathbf{v}_j - \sum_{k=1}^{j-1} (\mathbf{v}_j \cdot \mathbf{e}_k) \mathbf{e}_k, \quad \mathbf{e}_j = \frac{\bar{\mathbf{e}}_j}{(\bar{\mathbf{e}}_j \cdot \bar{\mathbf{e}}_j)^{1/2}} \quad (23.32)$$

para $j = 2, 3, \dots, N$. Este é o **processo de ortogonalização de Gram-Schmidt**.¹

Quando se está no âmbito do cálculo estrutural, da solução do autop problema (23.5), é mais útil realizar os produtos escalares com uma **matriz peso**, a matriz massa \mathbf{M} . Quando a função peso usada é a matriz identidade tem-se o chamado produto euclidiano. Os dois tipos de produtos tem as seguintes formas:

Produto interno Euclidiano	$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \sum_{i=1}^N u_i v_i,$	(23.33)
Produto interno ponderado	$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{M} \mathbf{v} = \sum_{i=1}^N \sum_{j=1}^N u_i M_{ij} v_j.$	

Então, usando o produto escalar ponderado, as eqs. (23.32) tomam a forma

$$\mathbf{e}_1 = \frac{\mathbf{v}_1}{(\mathbf{u}_1^T \mathbf{M} \mathbf{v}_1)^{1/2}}, \quad \bar{\mathbf{e}}_j = \mathbf{v}_j - \sum_{k=1}^{j-1} (\mathbf{v}_k^T \mathbf{M} \mathbf{e}_k) \mathbf{e}_k, \quad \mathbf{e}_j = \frac{\bar{\mathbf{e}}_j}{(\bar{\mathbf{e}}_j^T \mathbf{M} \bar{\mathbf{e}}_j)^{1/2}}. \quad (23.34)$$

¹Jørgen Pedersen Gram, dinamarques (1850 – 1916) e Erhard Schmidt, alemão (1876 – 1959) .

$$\mathbf{B} = \mathbf{L}\mathbf{U} \quad (23.41)$$

onde \mathbf{L} e \mathbf{U} são matrizes triangular inferior e superior, com $L_{jj} = 1$ para $j = 1, \dots, N$. Usando (23.40) e (15.17) tem-se o determinante de \mathbf{B} numa forma limpa e computacionalmente aceitável:

$$\det \mathbf{B} = \prod_{j=1}^N U_{jj} \quad (23.42)$$

Observa-se que se \mathbf{B} for singular, um ou mais dos últimos valores da diagonal de \mathbf{U} serão nulos. Além disso deve-se lembrar que, trocando duas linhas ou colunas de uma matriz, o sinal do determinante é trocado. Então se \mathbf{B} tiver um **elemento diagonal nulo**, \mathbf{B} não pode ser diretamente fatorado pelo método de Gauss. Neste caso torna-se necessária uma troca de linhas (ou coluna), o que gera uma matriz $\bar{\mathbf{B}}$ diferente de \mathbf{B} . $\bar{\mathbf{B}}$ é que será fatorada e terá seu determinante calculado. O determinante de \mathbf{B} deve ser obtido por $\det \mathbf{B} = (-1)^p \det \bar{\mathbf{B}}$, onde p é o número de permutações de linhas ou colunas realizados na definição de $\bar{\mathbf{B}}$.

Exemplo 23.3

Obter o determinante da matriz \mathbf{K} usando o método da diagonalização (eqs. (23.40) e (23.42)).

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

Solução:

Aplicando o método de fatoração de Gauss, a decomposição de \mathbf{K} na forma $\mathbf{K} = \mathbf{L}\mathbf{U}$ é:

$$\begin{bmatrix} 4 & 1 & -1 & 0 \\ 1 & 4 & 1 & 0 \\ -1 & 1 & 4 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & & & \\ \frac{1}{4} & 1 & & \\ -\frac{1}{4} & \frac{1}{3} & 1 & \\ 0 & 0 & \frac{3}{10} & 1 \end{bmatrix} \begin{bmatrix} 4 & 1 & -1 & 0 \\ & \frac{15}{4} & \frac{5}{4} & 0 \\ & & \frac{10}{3} & 1 \\ & & & \frac{17}{10} \end{bmatrix}.$$

Da equação (23.42), $\det \mathbf{K} = \det \mathbf{L} \det \mathbf{U}$. Para o determinante de \mathbf{L} usa-se (23.38) e chega-se a que $\det \mathbf{L} = \mathbf{1}^2$. Então, para o $\det \mathbf{U}$ pode-se usar diretamente a eq. (23.42): $\prod_{j=1}^4 U_{jj} = 4 \frac{15}{4} \frac{10}{3} \frac{17}{10} = 85$.

23.5.2 Método determinante por secante para autovalor

Voltemos ao polinômio característico em (23.36). O problema de achar os autovalores pode ser convertido no problema de achar as N raízes λ_j do polinômio característico. Este polinômio é uma função algébrica de uma única variável, como ilustrado na Figura 23.3. Dado um valor para λ , o cálculo de $p(\lambda)$ é feito calculando a matriz $\mathbf{A} = [\mathbf{K} - \lambda\mathbf{M}]$ e obtendo seu determinante. Por definição, eq. (23.36), o valor deste determinante é o valor do polinômio para o dado λ .

Tem-se também uma forma eficiente de computar o determinante de uma matriz de grande porte, que é pela sua fatoração, usando (23.41) e (23.42). Em resumo, apesar de não se ter $p(\lambda)$ em sua forma explícita, como por exemplo

$$p(\lambda) = a_0 + a_1\lambda + a_2\lambda^2 + \dots + a_N\lambda^N, \quad (23.43)$$

²De fato isto é uma propriedade geral, isto é, qualquer matriz real triangular com diagonal unitária tem determinante igual a um.

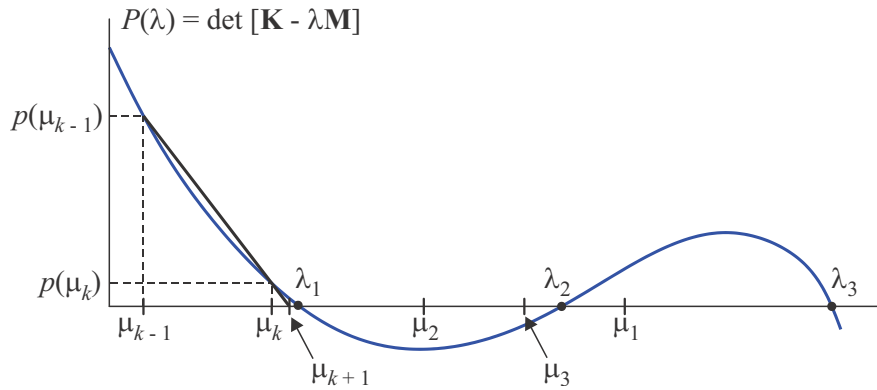


Figura 23.3: Esboço do polinômio característico e método da secante para obter λ_1 .

tem-se como calcular seu valor para um dado λ . O problema de como obter suas raízes torna-se então um problema conceitualmente simples, já estudado em métodos padrão de determinação numérica de raízes de polinômios. Ilustraremos aqui apenas um deles, o **método da secante**. Este é um método de interpolação linear. Observe-se a Figura 23.3, onde se tenta aproximar o primeiro autovalor, λ_1 . O procedimento pode ser aplicado para qualquer autovalor, uma vez que é realizado para aproximar cada autovalor, um por vez, de forma independente.

O processo iterativo do **método da secante** pode ser sintetizado no seguinte fluxograma:

Fluxograma do método da secante para determinação de autovalores.

Etapa 1 - Estabelece-se o autovalor alvo, λ_j , que se deseja aproximar. Definem-se duas estimativas iniciais para ele, valores μ_0 e μ_1 , o mais próximo possível de λ_j . Na próxima seção se comentará um processo que pode ser usado para estimar aproximadamente os valores dos λ_j numa certa faixa do espectro.

Etapa 2 - Para cada iteração em busca de λ_j tomam-se os dois últimos valores μ_{k-1} e μ_k da sequência definida no algoritmo.

Etapa 3 - Calcula-se

$$\begin{aligned} p_{k-1} &= p(\mu_{k-1}) = \det \underbrace{[\mathbf{K} - \mu_{k-1}\mathbf{M}]}_{\mathbf{A}_{k-1}} = \det(\mathbf{LU})_{k-1} = \prod_j (U_{jj})_{k-1}, \\ p_k &= p(\mu_k) = \det \underbrace{[\mathbf{K} - \mu_k\mathbf{M}]}_{\mathbf{A}_k} = \det(\mathbf{LU})_k = \prod_j (U_{jj})_k. \end{aligned} \quad (23.44)$$

Obtém-se então as coordenadas de dois pontos da curva, $(\mu_{k-1}; p_{k-1})$ e $(\mu_k; p_k)$, que podem ser unidos por uma reta como visualizado na Figura 23.3. Considera-se que esta reta aproxima naquela região da curva $p(\lambda)$. Espera-se que o ponto μ_{k+1} onde a reta intercepta o eixo das abscissas fique próximo ao ponto onde a curva intercepta o eixo, em λ_j .

Etapa 4 - Calcula-se em seguida a nova aproximação da raiz fazendo uma extrapolação linear, o que resulta em

$$\mu_{k+1} = \mu_{k-1} - \frac{p(\mu_{k-1})(\mu_k - \mu_{k-1})}{p(\mu_k) - p(\mu_{k-1})} \quad (23.45)$$

Etapa 5 - Define-se uma tolerância de erro, Tol , por exemplo 10^{-6} , para o autovalor. Como em vibrações $\lambda = \omega^2$, isto dará uma precisão de 10^{-3} na frequência natural. O valor de μ_{k+1} é considerado satisfatório se

23.5.3 Método da bisseção e teste de Sturm

Um dos princípios fundamentais da solução de autoprobemas é a **separação de autovalores** e a **sequência de Sturm** que são deriváveis do **quociente de Rayleigh**. Estes princípios não serão mostrados aqui e podem ser obtidos de forma clara em textos clássicos como o de Bathe [10].

Uma das consequências da sequência de Sturm tem uma utilização prática direta, como descrita a seguir. Considere-se o problema $[\mathbf{K} - \lambda\mathbf{M}]\phi = \mathbf{0}$, que possui N autovalores, que podem ser ordenados de forma crescente. Toma-se um número real qualquer μ e computa-se a matriz simétrica real $\mathbf{B} = \mathbf{K} - \mu\mathbf{M}$. Em seguida fatora-se \mathbf{B} pelo método de Gauss, em termos de uma matriz triangular inferior \mathbf{L} com diagonal unitária e uma diagonal \mathbf{D} :

$$\mathbf{B} = \mathbf{LDL}^T. \quad (23.49)$$

É provado que o número de elementos negativos na matriz diagonal \mathbf{D} é igual ao número de autovalores de $[\mathbf{K} - \lambda\mathbf{M}]\phi = \mathbf{0}$ menores que μ . Isto significa que, se μ for tomado entre λ_i e λ_{i+1} , isto é, $\lambda_i < \mu < \lambda_{i+1}$, a matriz \mathbf{D} terá exatamente i valores negativos. Este é o chamado **teste de Sturm**.

Uma das utilizações para esta propriedade é a seguinte. Nota-se que para usar o método do determinante para a obtenção dos autovalores, é necessário que se tenha uma estimativa prévia, mesmo que grosseira, da localização dos autovalores desejados. Para isso, pode-se usar o chamado **teste de Sturm** acoplado a um **método de bisseção**. Considere-se inicialmente o esboço de parte de um polinômio característico ilustrado na Figura 23.2. Toma-se um certo número real, por exemplo μ_1 , como indicado na figura. Calcula-se a matriz $\mathbf{A} = [\mathbf{K} - \mu_1\mathbf{M}]$ e a realiza-se sua fatoração, $\mathbf{A} = \mathbf{LDL}^T$. Conta-se o número n_1 de termos negativos em \mathbf{D} , que no exemplo seria $n_1 = 2$, isto é, no exemplo existem dois autovalores, λ_1 e λ_2 menores que μ_1 .

Se o processo for repetido para μ_k e se obtenha $n_k = 0$, se teria a indicação de não haver nenhum autovalor menor que μ_k . Deduz-se então que existem dois autovalores entre μ_k e μ_1 , e esses são os autovalores λ_1 e λ_2 . Usa-se então o método da bisseção, isto é, divide-se o intervalo μ_k a μ_1 ao meio, isto é, toma-se um valor μ_2 dado por $\mu_2 = (\mu_k + \mu_1)/2$ e repete-se o processo de calcular \mathbf{A} e fatorá-la. Caso se obtenha um termo negativo, $n_2 = 1$, se saberia que o autovalor λ_1 se encontra no intervalo $\mu_k - \mu_2$ e λ_2 no intervalo $\mu_2 - \mu_1$. Se for desejado um intervalo mais estreito em torno de um certo autovalor basta dividir o intervalo ao meio e repetir o processo. Uma síntese do **processo de bisseção** é mostrado no Fluxograma da Tabela 23.2.

Tabela 23.2: Fluxograma do método da biseção para autovalores.

Etapa 1	- Definir dois números reais μ_j e μ_k , tal que $\mu_j < \mu_k$.
Etapa 2	- Fatorar $\mathbf{K} - \mu_j\mathbf{M}$ e aplicar o teste de Sturm para determinar o número n_j de autovalores menores que μ_j .
Etapa 3	- Aplicar o teste de Sturm em $\mathbf{K} - \mu_k\mathbf{M}$ e determinar o número n_k de autovalores menores que μ_k . Determinar o número n_{jk} de autovalores entre μ_j e μ_k .
Etapa 4	- Se um intervalo contiver mais de um autovalor biseccionar este intervalo num ponto central $\mu_e = (\mu_j + \mu_k)/2$. Aplicar o teste de Sturm de forma sequencial até obter cada intervalo com apenas um autovalor.
Etapa 5	- Continuar o processo de bisseção para aproximar cada autovalor até o grau desejado de precisão.

Observação - Este método pode ser usado para determinar os autovalores porém apresenta **baixa eficiência computacional**. Seu melhor uso é apenas o de determinar regiões para cada autovalor para facilitar a aplicação do método do determinante pela secante.

Exemplo 23.4

Considere o autoproblema do Exemplo 23.1, equação (23.24). Use o teste de Sturm para determinar quantos autovalores existem abaixo de $\mu = 2,0$.

Solução:

Tomamos \mathbf{K} e \mathbf{M} do Exemplo 23.1 e calculamos a matriz $\mathbf{B} = \mathbf{K} - \mu\mathbf{M}$:

$$\mathbf{B} = \begin{bmatrix} -6 & -3 & 0 \\ -3 & -6 & -3 \\ 0 & -3 & -3 \end{bmatrix}.$$

A fatoração desta matriz na forma $\mathbf{B} = \mathbf{L}\mathbf{U}$ resulta no seguinte:

$$\mathbf{B} = \begin{bmatrix} 1 & & \\ \frac{1}{2} & 1 & \\ 0 & \frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} -6 & -3 & 0 \\ & -\frac{9}{2} & -3 \\ & & -1 \end{bmatrix}.$$

O número de termos negativos na diagonal é três, o que corresponde à solução analítica dos autovalores mostrados no exemplo da página 521, os três autovalores são inferiores a 2,0, ($\lambda_1 = 0,04675$, $\lambda_2 = 0,5$, $\lambda_3 = 1,6456$), isto é, $\lambda_j < \mu$. Observe que, para $\mu = 2,0$, o determinante de \mathbf{B} é $\det \mathbf{B} = \prod_{j=1}^3 U_{jj} = (-6)(-9/2)(-1) = -27$.

Exemplo 23.5

Use o método do determinante por secante para determinar uma aproximação ao terceiro autovalor do autoproblema do Exemplo 23.1.

Solução:

Do Exemplo 23.4 tem-se que, para $\mu = 2,0$, o determinante de $\mathbf{B} = \mathbf{K} - \mu\mathbf{M}$ é igual a -27 . Chamemos a este μ de μ_1 . Buscamos o determinante de \mathbf{B} para $\mu_2 = 1,0$. Então $\mathbf{B}_2 = \mathbf{K} - \mu_2\mathbf{M}$, que dá:

$$\mathbf{B}_2 = \begin{bmatrix} -2 & -2 & 0 \\ -2 & -2 & -2 \\ 0 & -2 & -1 \end{bmatrix}.$$

Calcula-se $\det \mathbf{B}_2 = +8,0$. Usando o teste de Sturm verifica-se que existem dois autovalores menores que $\mu_2 = 1,0$. Conclui-se que existe apenas um autovalor entre $\mu_1 = 2,0$ e $\mu_2 = 1,0$, precisamente o terceiro autovalor do sistema. Tem-se dois pontos, $(\mu_1; p(\mu_1))$ e $(\mu_2; p(\mu_2))$, onde $p(\mu_k)$ é o valor do polinômio característico, calculado como $p(\mu_k) = \det \mathbf{B}_k$, isto é, $p(\mu_1) = -27$ e $p(\mu_2) = +8,0$. Calcula-se uma nova estimativa μ_3 para o autovalor λ_3 usando a extrapolação linear (23.45):

$$\mu_3 = \mu_2 - \frac{p(\mu_2 - \mu_1)}{p(\mu_2) - p(\mu_1)} = 1,0 - \frac{8(1,0 - 2,0)}{8 - (-27)} = 1,228.$$

A matriz $\mathbf{B}_3 = \mathbf{K} - \mu_3 \mathbf{M}$ é

$$\mathbf{B}_3 = \begin{bmatrix} -2,912 & -2,228 & 0 \\ -2,228 & -2,912 & -2,228 \\ 0 & -2,228 & -1,456 \end{bmatrix}.$$

O valor do seu determinante é $p(\mu_3) = \det \mathbf{B}_3 = 9,3362$. Novamente, utilizam-se os pontos 2 e 3 na equação (23.45) para encontrar uma estimativa μ_4 para o terceiro autovalor λ_3 . Seguindo o processo obtém-se os valores mostrados nas colunas 2 e 3 da Tabela 23.3. Visivelmente as iterações divergem. Isto se explica pela má qualidade das estimativas iniciais μ_1 e μ_2 . A evolução das iterações pode ser vista com o auxílio da Figura 23.4 que mostra a variação do valor do polinômio característico $p(\mu_k)$ versus μ_k . Um melhor par de valores iniciais pode ser obtido utilizando o método

de bisseção algumas vezes. Feito isto consegue-se estreitar a região de localização do autovalor λ_3 para o intervalo de μ entre 1,5 a 2,0, (colunas 4 e 5) em vez de 1,0 a 2,0 (colunas 2 e 3) da Tabela 23.3. Observe que as iterações foram executadas até se obter convergência de 8 dígitos significativos, para tolerância $Tol = 10^{-6}$, o que resulta num erro relativo para a frequência, $w_3 = \sqrt{\mu}$ menor que 10^{-3} como desejado. Este é ainda um valor aproximativo, uma vez que o determinante não é nulo, mas apenas $3,2 \cdot 10^{-10}$.

Tabela 23.3: Determinação do terceiro autovalor da matriz do Exemplo 5 com o uso do método do determinante.

ponto k	Valores iniciais 1,0 e 2,0		Valores iniciais 1,0 e 1,5	
	μ_k	$p(\mu_k) = \det \mathbf{B}_k$	μ_k	$p(\mu_k) = \det \mathbf{B}_k$
1	2,0	-27,0	2,0	-27,0
2	1,0	+8,0	1,50	5,50
3	1,229	9,335	1,585	2,64
4	-0,369	18,96	1,663	-0,847
5	2,78	-183,3	1,644	0,079
6	-0,074 diverge		1,64552	0,00202
7			1,64556	$-5,05 \cdot 10^{-6}$
8			1,6455619	$3,2 \cdot 10^{-10}$
9			1,6455619	

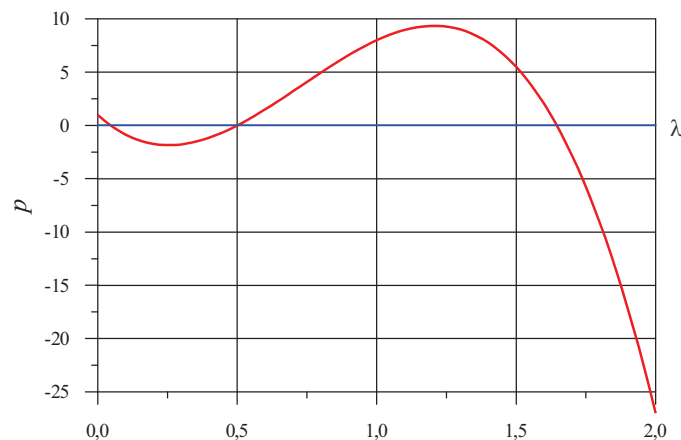


Figura 23.4: Plotagem de $p(\mu_k) = \det [\mathbf{K} - \mu_k \mathbf{M}]$ para o Exemplo 23.4. Os autovalores são os três zeros da função.

23.6 Método da iteração subespacial

O método da iteração subespacial, tanto quanto o de Lanczos, é um dos mais versáteis e eficientes métodos disponíveis para a solução de autoprobemas de grande porte envolvendo matrizes simétricas. Ele se fundamenta principalmente no método da potência e no método de Rayleigh-Ritz, sendo uma generalização destas. O primeiro já foi visto neste capítulo e o último será visto a seguir.

23.6.1 Quociente de Rayleigh

O quociente de Rayleigh é um dos pontos fundamentais da teoria de autoproblemas, e sua aplicação não se restringe ao método de iteração subespacial. Desta forma fazemos aqui sua definição e apresentamos uma de suas propriedades principais.

Considere-se o problema de autovalor (23.5), com \mathbf{K} e \mathbf{M} matrizes simétricas reais ordem N . Pode-se tomar um autopar $(\lambda_r; \phi^j)$ qualquer, para $j = 1, 2, \dots, N$, que satisfaz

$$\mathbf{K}\phi^j = \lambda_j \mathbf{M}\phi^j, \quad j = 1, 2, \dots, N. \quad (23.50)$$

Pré-multiplicando ambos os lados da igualdade por ϕ^{jT} pode-se fazer

$$\lambda_j = \frac{\phi^{jT} \mathbf{K} \phi^j}{\phi^{jT} \mathbf{M} \phi^j} \quad (\text{exato}) \quad (23.51)$$

Considera-se agora que, em vez de usar um autovetor ϕ^j no lado direito de (23.51) se tenha um vetor qualquer $\mathbf{v} \neq \phi^j$. O resultado daria um escalar, frequentemente denominado R , que seria função de \mathbf{v} , isto é, $R = R(\mathbf{v})$, e sem dúvida seria diferente de λ_j , isto é,

$$R(\mathbf{v}) \equiv \frac{\mathbf{v}^T \mathbf{K} \mathbf{v}}{\mathbf{v}^T \mathbf{M} \mathbf{v}} \quad (\text{aproximado}) \quad (23.52)$$

Este escalar é o chamado **quociente de Rayleigh**. Uma propriedade importante deste quociente pode ser vista como segue. Considere-se a expansão deste vetor arbitrário \mathbf{v} em termos dos autovetores ϕ^j do sistema (mesmo que estes não sejam ainda conhecidos), isto é, qualquer \mathbf{v} pode ser escrito como uma combinação linear dos ϕ^j , bastando determinar os coeficientes c_j :

$$\mathbf{v} = \sum_{j=1}^N c_j \phi^j = \Phi \mathbf{c}, \quad (23.53)$$

onde Φ é a matriz modal do sistema e \mathbf{c} é um vetor de ordem N . Considera-se que Φ seja um conjunto ortonormal pela massa, isto é, $\Phi^T \mathbf{M} \Phi = \mathbf{I}$ e $\Phi^T \mathbf{K} \Phi = \Lambda^2$, onde $\Lambda^2 = \text{Diag}[\lambda_1; \lambda_2; \dots; \lambda_N]$. Levando \mathbf{v} de (23.53) em (23.52) e simplificando tem-se

$$R(\mathbf{v}) = \frac{\mathbf{c}^T \Phi^T \mathbf{K} \Phi \mathbf{c}}{\mathbf{c}^T \Phi^T \mathbf{M} \Phi \mathbf{c}} = \frac{\mathbf{c}^T \Lambda^2 \mathbf{c}}{\mathbf{c}^T \mathbf{c}}.$$

Porém, como Λ^2 é diagonal, R fica

$$R(\mathbf{v}) = \frac{\sum_{j=1}^N \lambda_j c_j^2}{\sum_{j=1}^N c_j^2}. \quad (23.54)$$

Definem-se os números $\varepsilon_j = c_j/c_r$, para $r = 1, \dots, N$ arbitrário, onde c_r é um dos termos de \mathbf{c} . Então se $r = 2$, $\varepsilon_1 = c_1/c_2$, $\varepsilon_2 = 1$, $\varepsilon_3 = c_3/c_2$, etc. A equação (15.50) pode ser colocada em termos dos ε 's, colocando c_r em evidência:

$$R(\mathbf{v}) = \frac{\lambda_r + \sum_{j=1}^N (1 - \delta_{jr}) \lambda_j \varepsilon_j^2}{1 + \sum_{j=1}^N (1 - \delta_{jr}) \varepsilon_j^2}, \quad (23.55)$$

onde δ_{jr} é o operador delta de Kronecker. Por exemplo, para $r = 3$ e $N = 4$ esta equação fica

$$R(\mathbf{v}) = \frac{\lambda_3 + \lambda_1 \varepsilon_1^2 + \lambda_2 \varepsilon_2^2 + (1 - 1) \lambda_3 \varepsilon_3^2 + \lambda_4 \varepsilon_4^2}{1 + \varepsilon_1^2 + \varepsilon_2^2 + (1 - 1) \varepsilon_3^2 + \varepsilon_4^2}.$$

De fato, agora tem-se que R é uma função de $\boldsymbol{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_N)$, isto é, $R = R(\boldsymbol{\varepsilon})$. Pode-se fazer uma expansão de R em série de Taylor em torno de $\boldsymbol{\varepsilon} = \mathbf{0}$, retendo até os termos de segunda ordem:

$$\begin{aligned} \mathbf{R}(\varepsilon_j) = & R|_{\boldsymbol{\varepsilon}=\mathbf{0}} + \left. \frac{\partial R}{\partial \varepsilon_1} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}} \varepsilon_1 + \left. \frac{\partial R}{\partial \varepsilon_2} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}} \varepsilon_2 + \dots + \left. \frac{\partial R}{\partial \varepsilon_N} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}} \varepsilon_N \\ & + \frac{1}{2} \left. \frac{\partial^2 R}{\partial \varepsilon_1^2} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}} \varepsilon_1^2 + \dots + \frac{1}{2} \left. \frac{\partial^2 R}{\partial \varepsilon_N^2} \right|_{\boldsymbol{\varepsilon}=\mathbf{0}} \varepsilon_N^2 + 0(\varepsilon^3). \end{aligned}$$

As diferenciações podem ser feitas com a ajuda de um manipulador simbólico de equações e eliminados todos os termos de ordem superior a dois em ε_j . Isto resulta em

$$R(\mathbf{v}) = \lambda_r + \sum_{j=1}^N \underbrace{(\lambda_j - \lambda_r)}_{\geq 0} \underbrace{\varepsilon_j^2}_{\geq 0} + 0(\varepsilon^3). \quad (23.56)$$

Considere agora que $r = 1$. Se os autovalores estiverem ordenados de forma crescente, isto é, $\lambda_1 < \lambda_2 < \dots < \lambda_N$, então cada termo no somatório será maior ou igual a zero. A consequência é que sempre

$$\boxed{R(\mathbf{v}) = \frac{\mathbf{v}^T \mathbf{K} \mathbf{v}}{\mathbf{v}^T \mathbf{M} \mathbf{v}} \geq \lambda_1, \text{ qualquer que seja o vetor } \mathbf{v}.} \quad (23.57)$$

Isto significa que **o quociente de Rayleigh nunca é menor que o primeiro autovalor**. Uma outra maneira de ver esta equação é que **o mínimo de $R(\mathbf{v})$ dentre todos os possíveis valores de \mathbf{v} é o primeiro autovalor**:

$$\boxed{\text{Min}_{\mathbf{v}} R(\mathbf{v}) = \lambda_1} \quad (23.58)$$

Esta propriedade do quociente de Rayleigh pode ser usada em métodos de cálculo para os autovalores como visto na próxima subseção. Note que, quanto mais o vetor escolhido \mathbf{v} se aproximar do primeiro autovetor $\boldsymbol{\phi}_1$ mais $R(\mathbf{v})$ se aproximará de λ_1 . Isto pode ser visto em (23.55): se $\mathbf{v} \approx \boldsymbol{\phi}_1$, de (23.53) observa-se que $c_1 \approx 1$ e os outros c 's são pequenos. Como consequência, todos os ε 's são pequenos, com exceção de ε_1 que é ≈ 1 . Então, de (23.56), tem-se que $R(\mathbf{v}) \approx \lambda_1$.

23.6.2 Método de Rayleigh-Ritz

Na subseção anterior foi mostrado que o quociente de Rayleigh pode ser usado para obter uma estimativa do primeiro autovalor, se for disponível uma estimativa para o primeiro autovetor. Frequentemente o interesse é a obtenção de estimativas para os m primeiros autovalores do problema. Isto é conseguido como segue.

Considere-se novamente que qualquer vetor \mathbf{v} no espaço vetorial N dimensional gerado por \mathbf{M} ou \mathbf{K} pode ser escrito como uma combinação linear dos N autovetores $\boldsymbol{\phi}^j$ do autoproblema associado. Para problemas de alta ordem, é impossível obter todos os N autovetores do sistema. Em lugar disto se buscam apenas os m primeiros que serão utilizados no método de sobreposição modal ou em qualquer outra aplicação necessária.

O procedimento que segue é um processo inverso, baseado em que se **tenha um conjunto arbitrário de m , ($m \ll N$) vetores $\boldsymbol{\varphi}^j$, linearmente independentes. Deseja-se combinar linearmente estes vetores de forma a aproximar o melhor possível um dos autovetores $\boldsymbol{\phi}$** :

$$\boldsymbol{\phi} \approx \sum_{j=1}^m \boldsymbol{\varphi}^j z_j = \boldsymbol{\Psi} \mathbf{z} \quad (23.59)$$

onde \mathbf{z} é um vetor de dimensão m , ($m \ll N$) constituído pelas constantes a serem determinados, também chamadas **amplitudes de coordenadas generalizadas**. Tem-se então uma aproximação

$$\frac{\partial (z_k \bar{M}_{kl} z_l)}{\partial z_p} = 2 \bar{M}_{pj} z_j$$

Com isso (23.64) toma a forma (voltando à notação com o símbolo de somatório)

$$\frac{\partial R(\mathbf{z})}{\partial z_p} = \frac{(\mathbf{z}^T \bar{\mathbf{M}} \mathbf{z}) \left(2 \sum_j^m \bar{K}_{pj} z_j \right) - (\mathbf{z}^T \bar{\mathbf{K}} \mathbf{z}) \left(2 \sum_j^m \bar{M}_{pj} z_j \right)}{(\mathbf{z}^T \bar{\mathbf{M}} \mathbf{z})^2} = 0.$$

Em seguida escreve-se esta equação para todos os termos z_j :

$$\frac{\partial R(\mathbf{z})}{\partial \mathbf{z}} = \begin{Bmatrix} \partial R / \partial z_1 \\ \partial R / \partial z_2 \\ \vdots \\ \partial R / \partial z_m \end{Bmatrix} = 2 \frac{(\mathbf{z}^T \bar{\mathbf{M}} \mathbf{z}) \bar{\mathbf{K}} \mathbf{z} - (\mathbf{z}^T \bar{\mathbf{K}} \mathbf{z}) \bar{\mathbf{M}} \mathbf{z}}{(\mathbf{z}^T \bar{\mathbf{M}} \mathbf{z})^2} = \begin{Bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{Bmatrix} \quad (23.66)$$

Por outro lado, se se conseguir que ϕ aproxime um autovetor, o termo $\mathbf{z}^T \bar{\mathbf{K}} \mathbf{z}$ no numerador pode ser substituído por $\frac{\partial R(\mathbf{z})}{\partial z_p} = \lambda \mathbf{z}^T \bar{\mathbf{M}} \mathbf{z}$, e simplificando $\mathbf{z}^T \bar{\mathbf{M}} \mathbf{z}$, a segunda igualdade resulta em

$$\boxed{\bar{\mathbf{K}} \mathbf{z} - \lambda \bar{\mathbf{M}} \mathbf{z} = \mathbf{0}} \quad (23.67)$$

Observa-se que agora se tem um problema de autovalor de dimensões reduzidas $m \ll n$. Sua solução dá simultaneamente os coeficientes \mathbf{z} que minimizam R e ao mesmo tempo dá a estimativa do λ associado ao autovetor minimizante. Este autop problema reduzido tem matrizes também simétricas e possui m autopares $(\lambda_j; \mathbf{z}^j)$.

Então as etapas de cálculo são mostradas no Fluxograma da Tabela 23.4.

Tabela 23.4: Fluxograma do Método de Reyleigh-Ritz para autopares.

Etapa 1 - Escolher uma base de m vetores φ^j linearmente independentes, formando o conjunto Ψ
Etapa 2 - Calcular as matrizes reduzidas $\bar{\mathbf{K}}$ e $\bar{\mathbf{M}}$ de (23.61)
Etapa 3 - Resolver o problema reduzido (23.67) obtendo todos os m autopares $(\lambda_j; \mathbf{z}^j)$
Etapa 4 - Obter a estimativa do conjunto Φ dos m autovetores, eq. (23.59): $\Phi \approx \Psi \mathbf{z}$

Observação - Se os vetores escolhidos Ψ fossem exatamente os autovetores Φ do autop problema (23.50), as matrizes reduzidas (23.61) seriam diagonais. Já na situação usual em Ψ é próximo de Φ , estas matrizes terão os termos fora da diagonal não nulos, mas menores que os termos das diagonais.

23.6.3 Método da iteração subespacial

Este método pode ser visto como uma generalização do método da potência, onde se iterava com um único vetor, alterado para iterar simultaneamente com um conjunto de m vetores. Entretanto, o simples uso de um conjunto de vetores nas iterações resulta num método pobre, que converge apenas de forma muito lenta. O que dá eficiência ao método da iteração sub espacial é que cada uma das iterações incorpora o método de Rayleigh-Ritz. As etapas do método são as listadas no fluxograma a seguir.

Fluxograma para o processo iterativo pelo método da iteração subespacial.

Etapa 0 - Inicializações. Definir os vetores iniciais Ψ_1 (sua determinação é vista no fim da seção). Fatorar \mathbf{K} .

Etapa 1 - Para iterações $k = 1, 2, \dots, n_k$ calcular:

Etapa 8 - Teste de Sturm - Devido às deficiências no processo de convergência, é algo comum que alguns dos últimos autopares sejam saltados. Por exemplo, para $p = 4$ autopares desejados, o algoritmo pode não convergir $\lambda_1, \lambda_2, \lambda_3$ e λ_4 mas para $\lambda_1, \lambda_2, \lambda_3$ e λ_5 ou $\lambda_1, \lambda_2, \lambda_3$ e λ_8 , por exemplo. Para verificar se isso ocorreu ou não, após a convergência realiza-se o chamado teste de Sturm. Para isso toma-se o p -ésimo autovalor calculado em $\mathbf{\Lambda}_{k+1}$, isto é, $\lambda_p^{(k+1)}$. Toma-se um número real μ um pouco maior que $\lambda_p^{(k+1)}$, por exemplo,

$$\mu = 1,01\lambda_p^{(k+1)}. \quad (23.76)$$

O teste de Sturm dirá quantos autovalores menores que μ existem no sistema matricial. Se houver p autovalores menores que μ então conclui-se que o p -ésimo valor em $\mathbf{\Lambda}_{k+1}$ é realmente uma estimativa do p -ésimo autovalor do sistema, e nenhum autovalor foi pulado. Por outro lado, se, por exemplo, se $p = 5$ e o teste de Sturm mostra que existem 6 autovalores menores que $\lambda_5^{(k+1)}$, então $\lambda_5^{(k+1)}$ deve estar aproximando o sexto autovalor e algum dos outros foi pulado. Deve-se lembrar, de (23.49), que o teste de Sturm é feito primeiro calculando a matriz \mathbf{A} como

$$\mathbf{A} = \mathbf{K} - \mu\mathbf{M} \quad (23.77)$$

e depois fatorando-a pelo método de Gauss:

$$\mathbf{A} = \mathbf{L}\mathbf{S}. \quad (23.78)$$

onde \mathbf{L} é uma matriz triangular inferior com diagonal unitária, e \mathbf{S} triangular superior. O número de termos negativos da diagonal de \mathbf{S} é o número de autovalores de $(\mathbf{K} - \lambda\mathbf{M})\phi = \mathbf{0}$ menores que μ . Com isto termina o fluxograma do método.

Vetores iniciais

Dentre as formas simples de definir os vetores iniciais Ψ_1 necessários em (23.68), aparentemente a mais eficiente é a seguinte. Devem-se gerar m vetores φ^j . O primeiro, φ^1 , é gerado com 1's em todas as N posições, isto é,

$$\varphi^1 = \{1; 1; \dots; 1\}^T. \quad (23.79)$$

o vetor φ^2 é definido como um vetor nulo, com um 1 na posição k , sendo k o número da linha em que a relação entre os termos das diagonais das matrizes, K_{ii}/M_{ii} , seja mínima:

$$\varphi^2 = \{0; 0; \dots; 1; \dots; 0\}^T. \quad (23.80)$$

Os vetores 3 a m são gerados seguindo o mesmo procedimento de geração do vetor 2 tomando k 's iguais ao número da linha onde K_{ii}/M_{ii} seja a segunda menor, a terceira menor, e assim por diante. Desta forma geram-se todos os m vetores iniciais do conjunto Ψ_1 .

Prova-se que, se o conjunto inicial for uma combinação linear dos primeiros autovetores o processo converge em uma única iteração. Como em geral não é possível estimar previamente esse conjunto, o vetor em (23.79) é usado para excitar todos os modos do sistema, e os demais vetores tentam excitar os primeiros modos.

Exemplo 23.6

Construir o conjunto inicial de $m = 3$ vetores para as seguintes matrizes:

23.7 Método de Lanczos

Atualmente o método de Lanczos para solução de problemas de autovalores de sistemas de grande porte provenientes do método de elementos finitos é um dos mais utilizados, tanto quanto o método de iteração subespacial. De fato, atualmente existe o que se pode chamar de uma família de métodos de Lanczos em vez de um único método. A partir da proposição inicial enunciada por Lanczos [56] em 1950, diversas variações tem sido desenvolvidas para lidar com os mais diversos tipos de sistemas matriciais, como sistemas de uma ou duas matrizes, com uma ou ambas simétricas, sistemas amortecidos, sistemas com uma das matrizes antisimétrica (como ocorre em sistemas giroscópicos [7]), dentre outros. Compilações tem sido reunidas em livros totalmente dedicados ao assunto [30]. O método foi lançado como uma forma de determinar uma aproximação para os maiores autovalores do sistema, mas logo foi percebido como um procedimento de transformar uma matriz simétrica de grande porte em uma tridiagonal de pequeno porte, o que, de fato, é a etapa principal do método. Entretanto, em 1955 o método Householder-QR estabeleceu-se como um método mais eficiente para realizar a tridiagonalização da matriz e cálculo de todos os autopares, especialmente em sistemas de pequeno porte. O método de Lanczos permanece como mais eficiente em sistemas de grande porte, quando o objetivo é a determinação de um número pequeno de autopares. Para cumprir essa mesma tarefa, o método mais concorrente é o de iteração subespacial. A referência [76] faz uma comparação entre os dois métodos.

Inicialmente considera-se um problema de autovalor na forma

$$\mathbf{K}\mathbf{u} - \lambda\mathbf{M}\mathbf{u} = \mathbf{0}, \quad (23.81)$$

onde \mathbf{K} e \mathbf{M} são matrizes reais e simétricas, **positiva definidas**, de ordem N , e λ e \mathbf{u} são um autovalor e autovetor do problema. Uma vez que as matrizes são positiva-definidas, a condição (19.22) é satisfeita, de forma que se garante que todos os autopares do problema são reais. Existem outras deduções para o método de Lanczos que são adequadas a sistemas com autopares complexos. A dedução é feita usando a forma padrão do problema de autovalor correspondente a (23.81):

$$\left[\underbrace{\mathbf{K}^{-1}\mathbf{M}}_{\mathbf{A}} - \mu\mathbf{I} \right] \mathbf{u} = \mathbf{0}. \quad (23.82)$$

onde $\mu = 1/\lambda$. A matriz $\mathbf{A} \equiv \mathbf{K}^{-1}\mathbf{M}$ é não simétrica mas seus autovalores reais e são os mesmos de (23.81) e os autovalores os inversos daqueles. Nota-se que a inversa de \mathbf{K} é indicada aqui e durante a dedução, porém o algoritmo final não exigirá sua determinação. Em vez disso será exigido apenas sua fatoração de Cholesky ou solução de sistema linear por outro meio.

Inicialmente considera-se a chamada **sequência de Krilov**. Dada uma matriz \mathbf{A} quadrada de ordem N , não singular, e um vetor \mathbf{v}_0 de ordem N , a sequência de j vetores

$$\{\mathbf{v}_0, \mathbf{A}\mathbf{v}_0, \mathbf{A}^2\mathbf{v}_0, \dots, \mathbf{A}^{j-1}\mathbf{v}_0\} \quad (23.83)$$

forma a sequência de Krilov, e o conjunto de j vetores é a base de um subespaço de ordem j . **Conforme $j \rightarrow \infty$ a sequência converge para o autovetor correspondente ao autovalor de modulo mais próximo de zero** (positivo ou negativo). A notação \mathbf{A}^3 , por exemplo, significa $\mathbf{A}\mathbf{A}\mathbf{A}$. Diversos métodos iterativos de determinação de autovalores podem ser vistos como relacionados a esse subespaço, como o método de Lanczos e o de gradientes conjugados. De forma geral pode-se dizer que **o método de Lanczos toma os vetores da sequência de Krilov como vetores teste de Ritz**, para obter um problema reduzido cuja matriz é tridiagonal, portanto mais fácil de ser resolvida em seus autopares.

O processo iterativo é centrado em determinar um conjunto de vetores de Lanczos, um a cada iteração, de forma que se parte de um vetor arbitrário \mathbf{v}_0 , e numa iteração arbitrária j se tem conhecidos os vetores de Lanczos $\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j\}$ e se busca \mathbf{q}_{j+1} . Esse conjunto foi determinado de forma a garantir a ortonormalidade pela massa, isto é, definindo a matriz \mathbf{Q}_j de dimensões $N \times j$ formada pelas colunas $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_j$, isto é,

$$\mathbf{q}_j^T \mathbf{M} \mathbf{r}_j = \underbrace{\mathbf{q}_j^T \mathbf{M} \mathbf{p}_j}_0 + a_j \underbrace{\mathbf{q}_j^T \mathbf{M} \mathbf{q}_j}_1 + \underbrace{b_j \mathbf{q}_j^T \mathbf{M} \mathbf{q}_{j-1} + c_j \mathbf{q}_j^T \mathbf{M} \mathbf{q}_{j-2} + \cdots + z_j \mathbf{q}_j^T \mathbf{M} \mathbf{q}_1}_0.$$

Por definição, \mathbf{p}_j é ortonormal aos vetores anteriores, logo o primeiro termo à direita é nulo. O produto matricial no segundo termo a direita é igual a 1, e os demais termos à direita são nulos devido à ortogonalidade, resultando a componente de \mathbf{r}_j na direção \mathbf{q}_j :

$$\boxed{a_j = \mathbf{q}_j^T \mathbf{M} \mathbf{r}_j} \quad (23.92)$$

Determinação de b_j

A componente de \mathbf{r}_j na direção \mathbf{q}_{j-1} , isto é, b_j , é obtida pré-multiplicando (23.91) por $\mathbf{q}_{j-1}^T \mathbf{M}$, o que resulta

$$b_j = \mathbf{q}_{j-1}^T \mathbf{M} \mathbf{r}_j. \quad (23.93)$$

Em seguida utiliza-se (23.90) para eliminar \mathbf{r}_j em termos de \mathbf{q}_j e em seguida usa-se a mesma expressão para representar $\mathbf{r}_{j-1} = (\mathbf{K}^{-1} \mathbf{M}) \mathbf{q}_{j-1}$. Isso resulta em

$$\boxed{b_j = \mathbf{r}_{j-1}^T \mathbf{M} \mathbf{q}_j} \quad (23.94)$$

Uma **forma alternativa** pode ser construída. O vetor \mathbf{r}_{j-1} pode ser representado em termos dos vetores de Lanczos precedentes, usando uma expressão similar a (23.91):

$$\mathbf{r}_{j-1} = \mathbf{p}_{j-1} + a_{j-1} \mathbf{q}_{j-1} + b_{j-1} \mathbf{q}_{j-2} + c_{j-1} \mathbf{q}_{j-3} + \cdots + z_{j-1} \mathbf{q}_1. \quad (23.95)$$

Pré-multiplicando essa relação pelo transposto de $\mathbf{M} \mathbf{q}_j$ e levando o resultado a (23.94) obtém-se

$$\begin{aligned} b_j &= \mathbf{q}_j^T \mathbf{M} \mathbf{r}_{j-1}, \quad \text{usando (23.95),} \\ &= \mathbf{q}_j^T \mathbf{M} \mathbf{p}_{j-1} + \underbrace{a_{j-1} \mathbf{q}_j^T \mathbf{M} \mathbf{q}_{j-1} + b_{j-1} \mathbf{q}_j^T \mathbf{M} \mathbf{q}_{j-2} + c_{j-1} \mathbf{q}_j^T \mathbf{M} \mathbf{q}_{j-3} + \cdots + z_{j-1} \mathbf{q}_j^T \mathbf{M} \mathbf{q}_1}_0. \end{aligned} \quad (23.96)$$

A ortonormalidade da base anula todos os termos à direita exceto o primeiro. Pelo processo de construção da base ortonormal, o vetor obtido na iteração anterior, \mathbf{q}_j , foi obtido como a parte unitária de \mathbf{p}_{j-1} (lembrando que esse último é a parte de \mathbf{r}_{j-1} ortogonal à base anterior), isto é,

$$\mathbf{q}_j = \frac{\mathbf{p}_{j-1}}{\sqrt{\mathbf{p}_{j-1}^T \mathbf{M} \mathbf{p}_{j-1}}}. \quad (23.97)$$

Levando \mathbf{q}_j a (23.96) obtém-se a expressão final para b_j :

$$\boxed{b_j = \sqrt{\mathbf{p}_{j-1}^T \mathbf{M} \mathbf{p}_{j-1}}} \quad (23.98)$$

De fato, tanto essa quanto a expressão (23.94) podem ser usadas para calcular b_j , embora experimentação numérica mostre que a segunda apresenta melhor resultado geral, por lidar melhor com a perda natural de ortonormalidade da base que caracteriza o método de Lanczos.

As expressões (23.97) e (23.98) podem ser usadas para representar

$$\mathbf{q}_j = \frac{\mathbf{p}_{j-1}}{b_j} \quad \text{e} \quad \mathbf{q}_{j+1} = \frac{\mathbf{p}_j}{b_{j+1}}. \quad (23.99)$$

Determinação de c_j e demais componentes

Impõe-se a condição de \mathbf{M} -ortonormalidade de \mathbf{r}_j , pré-multiplicando (23.91) por $\mathbf{q}_{j-2}^T \mathbf{M}$:

Substituindo essa expressão em (23.82) e pré-multiplicando-a por $\mathbf{Q}_j^T \mathbf{M}$ obtém-se

$$\mathbf{T}_j \mathbf{s} = \mu \mathbf{s}. \quad (23.117)$$

onde $\mu = 1/\lambda$. Esse problema possui j autopares. Os autovalores são aproximações daqueles do problema original, e seus autovetores geram os autovetores no espaço físico através de (23.116). **O método fornecerá aproximações para os maiores autovalores de (23.82), o que significa que se terão os menores autovalores λ do problema generalizado original $[\mathbf{K} - \lambda \mathbf{M}] \mathbf{u} = \mathbf{0}$.**

A Tabela 23.5 mostra um fluxograma do algoritmo de Lanczos para o problema generalizado de autovalor, $\mathbf{K}\mathbf{u} - \lambda \mathbf{M}\mathbf{u} = \mathbf{0}$, onde \mathbf{K} e \mathbf{M} são matrizes reais e simétricas, **positiva-definidas**, de ordem N . Buscam-se os m primeiros autopares λ e \mathbf{u} do problema com um dado erro de convergência ε pré-definido. Para obter m autopares, é usual fazer as iterações até uma quantidade maior de valores, \bar{m} , no caso de problemas de grandes ordens N . Para N pequeno o método pode ser usado para obter todos os autopares, com $m = \bar{m} = N$.

Tabela 23.5: Fluxograma do método de Lanczos para matrizes \mathbf{K} e \mathbf{M} simétricas e positiva-definidas.

Etapa 1 -	Inicialização:
1	$\mathbf{q}_0 = \mathbf{0}$ e vetor arbitrário \mathbf{p}_0 Arbitrar $m < \bar{m}$ e a tolerância ε
2	$b_1 = \sqrt{(\mathbf{p}_0^T \mathbf{M} \mathbf{p}_0)}$ eq. (23.98)
3	$\mathbf{q}_1 = \frac{\mathbf{p}_0}{b_1}$ eq. (23.99)
4	$\mathbf{g}_1 = \mathbf{M} \mathbf{q}_1$
5	Fatoração de Cholesky $\mathbf{K} = \mathbf{C} \mathbf{C}^T$
Etapa 2 - DO	$j = 1, \bar{m}$
1	$r_j = \mathbf{K}^{-1} \mathbf{g}_j$ eq. (23.90)
2	$\bar{r}_j = r_j - b_j \mathbf{q}_{j-1}$ eq. (23.108)
3	$a_j = \mathbf{g}_j^T \bar{r}_j$ eq. (23.92)
4	$\mathbf{p}_j = \bar{r}_j - a_j \mathbf{q}_j$ eq. (23.108)
5	$\bar{\mathbf{g}}_j = \mathbf{M} \mathbf{p}_j$
6	$b_{j+1} = \sqrt{(\mathbf{p}_j^T \mathbf{M} \mathbf{p}_j)} = \sqrt{(\bar{\mathbf{g}}_j^T \mathbf{p}_j)}$ eq. (23.98)
7	$\mathbf{q}_{j+1} = \frac{\mathbf{p}_j}{b_{j+1}}$
8	M -ortonorm. de \mathbf{q}_{j+1} em relação aos q 's anteriores.
9	$\mathbf{g}_{j+1} = \frac{\bar{\mathbf{g}}_j}{b_{j+1}}$
10	Montar \mathbf{T}_j , de ordem j
11	Resolver o autoproblema simétrico real $\mathbf{T} \mathbf{s} = \mu \mathbf{s}$, eq. (23.117)
12	Calcular $\lambda_i^{(j)} = 1/\mu_i^{(j)}$, para $i = 1, \dots, j$.
13	Ordenar os λ 's e \mathbf{s} , em matrizes $\mathbf{\Lambda}$ e \mathbf{Q} .
14	Se $j > m$, teste de convergência para os m los termos. Eqs. (23.118) ou. (23.125). Se convergiu, ir à Etapa 2.15 .
	enddo
15	Solução final $\mathbf{u} = \mathbf{Q} \mathbf{s}$ e $\mathbf{\Lambda}$.
16	Teste de Sturm

resta o custo de obter o denominador $\|\mathbf{u}_m^{(j)}\|_{\mathbf{M}}$. Isso pode ser evitado considerando que o denominador em (23.120) é apenas um fator de normalização e pode ser substituído por outro de mesma ordem de grandeza. Por exemplo, pode-se tomar apenas os termos da diagonal de \mathbf{M} , de forma que o teste de convergência pode ser feito usando

$$\boxed{\frac{b_{j+1}q}{\sqrt{\sum_{t=1}^N M_{tt}u_t^2}} < \varepsilon_r} \quad (23.125)$$

Perda de ortonormalidade e autovalores repetidos

Toda a descrição mostrada nessa seção trata de operações aritméticas exatas. Nesse caso, o método de Lanczos apresenta um comportamento peculiar quando o problema possui **autovalores repetidos**. Nesse caso, em aritmética exata, o método só obtém um autovetor, isto é, apenas um dos autovetores correspondentes. Entretanto essa deficiência termina sendo contornada porque na computação as operações algébricas são feitas com erro de truncamento. Se existem dois autovetores para o autovalor μ_p , \mathbf{u}_1 e \mathbf{u}_2 , eles formam um subespaço $\bar{\mathbf{u}} = a\mathbf{u}_1 + b\mathbf{u}_2$, que é ortogonal aos demais autovetores. Em aritmética exata, se poderia ter um conjunto de vetores \mathbf{Q}_j completamente ortogonal a \mathbf{u}_2 . Entretanto, em precisão finita, os vetores \mathbf{Q}_j possuirão pequenas componentes na direção \mathbf{u}_2 , que crescerão durante as k iterações seguintes, até levar a um conjunto \mathbf{Q}_{j+k} que contenha $\bar{\mathbf{u}}$.

Quanto à **perda de ortonormalidade**. Em aritmética exata, as operações indicadas no Fluxograma da Tabela 23.5 não necessitariam a Etapa 2.8, para a M -ortonormalização de \mathbf{q}_{j+1} em relação aos q 's anteriores, pois as etapas anteriores, que visam cumprir a expressão (23.108), já seria suficiente para garantir essa ortogonalização. Entretanto, em precisão finita, em poucas iterações os vetores obtidos até a Etapa 2.7 são progressivamente não ortogonais. Esse é um fenômeno já detectado no trabalho original de Lanczos [56]. Sua recomendação foi exatamente a de realizar o que se chama de ortogonalização completa, em todas as iterações, como indicado na Etapa 2.8 do Fluxograma. Essa ortogonalização é feita pelo processo clássico de Gram-Schmidt, descrito na seção 23.4. Quando o número de autopares desejados, m , é pequeno, o custo da ortogonalização completa não é pesado. Quando m é grande, outros procedimentos são disponíveis, de menor custo.

Observações:

1. A quantidade de iterações \bar{m} é, de forma arbitrária, estimada como

$$\bar{m} = 1,5m \text{ ou } 2m. \quad (23.126)$$

onde m é o número de autopares que se deseja com uma precisão definida ε .

2. O comportamento do método de Lanczos, tanto quanto de outros métodos, é definido pelo número de iterações, se \bar{m} é “grande” ou “pequeno”. De forma usual, se considera que \bar{m} é grande por comparação com a largura de banda teórica de \mathbf{K} . A largura de banda é usada aqui apenas como indicativo do custo das operações de produto matriz-vetor e de triangularização de Cholesky, mesmo se o processo de armazenamento seja outro, o de matriz esparsa.
3. Em diversas etapas do método aparece a notação \mathbf{K}^{-1} . Entretanto deve ficar claro que uma matriz de grande porte como de rigidez obtida por uma modelagem de elementos finitos, não é recomendável a inversão. Em vez disso, a Etapa 2.1 do Fluxograma, $r_j = \mathbf{K}^{-1}\mathbf{g}_j$, o vetor r_j é obtido como a solução do problema algébrico linear

$$\mathbf{K}r_j = \mathbf{g}_j. \quad (23.127)$$

De fato, a fatoração de Cholesky (seção 12.2) é feita antes das iterações, na Etapa 1.5, gerando a forma $\mathbf{K} = \mathbf{C}\mathbf{C}^T$ onde apenas a matriz triangular inferior \mathbf{C} é armazenada, usando uma forma adequada (em meia banda, skyline, esparsa, etc.). Então, operação feita na Etapa 2.1

consiste apenas nas operações de substituição progressiva e regressiva, que são de menor custo que a fatoração. Em outros pontos da dedução do método de Lanczos a matriz \mathbf{K}^{-1} aparece, mas ao final ela está ausente das etapas de cálculo.

23.8 Sistemas lineares com matriz quadrada ou retangular, singular ou quase singular

Em muitas situações se tem a necessidade de resolver o sistema linear padrão $\mathbf{K}\mathbf{U} = \mathbf{F}$, onde \mathbf{K} é singular ou quase-singular. A singularidade é caracterizada pela presença de um ou mais autovalores de \mathbf{K} nulos, enquanto que ela é dita quase singular em diversas situações. Por exemplo, quando um ou mais autovalores são quase nulos, ou ainda quando a matriz é mal condicionada. O mal condicionamento, por sua vez, é determinado pelo **número de condição** da matriz, $c(\mathbf{K})$. Numericamente, o número de condição é determinado pela relação entre o maior e o menor autovalor da matriz, como definido em (12.95).

Quando $c(\mathbf{K})$ é maior que o número de algarismos significativos usados na computação (usualmente 15 em precisão dupla), os cálculos perdem precisão e logo tornam-se lixo numérico. Modelos mecânicos podem apresentar matrizes mal condicionadas por diversos motivos:

1. condições de contorno com pouca redundância em sistemas muito flexíveis,
2. condições de contato irregulares,
3. combinações de valores de propriedades de material de diferentes ordens de grandezas,
4. sistemas bem vinculados, porém muito longos e delgados,
5. cargas coplanares compressivas próximas ao valor crítico de instabilidade, e muitas outras situações.

Em suma, mesmo que do ponto de vista analítico a matriz devesse ser não singular, existem combinações especiais de parâmetros do modelo que podem tornar o menor autovalor muito próximo de zero.

Note que o problema não é apenas a proximidade de um ou mais autovalores do zero. Se todos eles tiverem ordem de grandeza igualmente pequenas, não se apresenta nenhuma dificuldade numérica. **O problema vem da presença de valores pequenos e grandes ao mesmo tempo.** Por exemplo, nas operações de soma de dois termos de ordens de grandeza extremas, a influência do menor valor tende a desaparecer do resultado.

Existem procedimentos de melhorar o condicionamento da matriz, os chamados **métodos de pré-condicionamento**, que consistem em fazer uma transformação

$$\mathbf{U} = \mathbf{T}\bar{\mathbf{U}}, \quad (23.128)$$

onde \mathbf{T} é uma matriz $N \times N$ não singular, conhecida, cuja definição produza os benefícios desejados. Na seção 12.3.3 descrevemos um dos métodos mais eficientes de condicionamento, o chamado **pré-condicionamento pela fatorização incompleta de Cholesky**. Uma versão simplificada consiste em definir \mathbf{T} simplesmente como uma diagonal onde $T_{jj} = 1/\sqrt{K_{jj}}$. Isso geralmente produz uma redução de 1 ou 2 no número de condição.

Qualquer que seja o método de condicionamento, se ele tiver a forma (23.128), sua substituição no sistema linear, e pré-multiplicação do resultado por \mathbf{T}^T , gera

$$\begin{aligned} \mathbf{K}\mathbf{U} &= \mathbf{F}, \\ \mathbf{K}\mathbf{T}\bar{\mathbf{U}} &= \mathbf{F}, \\ \underbrace{\mathbf{T}^T\mathbf{K}\mathbf{T}}_{\bar{\mathbf{K}}}\bar{\mathbf{U}} &= \underbrace{\mathbf{T}^T\mathbf{F}}_{\bar{\mathbf{F}}} \quad \rightarrow \quad \bar{\mathbf{K}}\bar{\mathbf{U}} = \bar{\mathbf{F}}. \end{aligned} \quad (23.129)$$

Observemos o resultado da operação $\mathbf{K}\mathbf{K}^+$, que é igual a alguma matriz quadrada simétrica \mathbf{A} :

$$\mathbf{K}\mathbf{K}^+ = \mathbf{A}. \quad (23.138)$$

Substituindo as expressões temos

$$\begin{aligned} \mathbf{K}\mathbf{K}^+ &= \mathbf{A}, \\ \mathbf{K}\mathbf{X}\bar{\Lambda}^{-1}\mathbf{X}^T &= \mathbf{A}, \quad (\text{pré-multiplicando por } \mathbf{X}\mathbf{X}^T) \\ \underbrace{\mathbf{X}\mathbf{X}^T\mathbf{K}\mathbf{X}\bar{\Lambda}^{-1}\mathbf{X}^T}_{\mathbf{A}} &= \mathbf{X}\mathbf{X}^T\mathbf{A}, \\ \mathbf{X}\mathbf{X}^T &= \mathbf{X}\mathbf{X}^T\mathbf{A}. \end{aligned} \quad (23.139)$$

A ortonormalidade da base, (23.132), implica que $\mathbf{X}\mathbf{X}^T = \mathbf{I}$, isto é, $\mathbf{K}^+ = \mathbf{K}^{-1}$. Como consequência temos que, **caso \mathbf{K} seja quadrada e de posto completo**,

$$\boxed{\mathbf{K}\mathbf{K}^+ = \mathbf{K}^+\mathbf{K} = \mathbf{I}} \quad (23.140)$$

23.8.2 Matriz singular ou de posto deficiente

Consideremos o espectro dos N autovalores de \mathbf{K} , ordenados de forma crescente. Podemos identificar três grupos distintos:

1. **Modos de corpo rígido.** Correspondem aos primeiros c_r autovalores são nulos, correspondentes aos r movimentos de corpo rígido do sistema que não tenham sido vinculados. O posto da matriz é $R(\mathbf{K}) = N - c_r$. Se $c_r > 0$, diz-se que o **posto é deficiente**, do contrário, a matriz é de **posto completo**, o que significa não singular.
2. **Modos de autovalor quase nulos.** Correspondem a p autovalores pequenos em relação aos autovalores superiores.
3. **Modos de deformação.** Correspondem aos d autovalores restantes, não nulos e nem incorretamente próximos de zero. Em geral são os modos de interesse na matriz.

Se \mathbf{K} for obtido pelo MEF, grande parte dos modos de deformação não possuem precisão e também são de pouco interesse na maioria das aplicações. Assim, pode ser interessante representar \mathbf{K} por uma matriz que contenha apenas a influência de um conjunto selecionado de modos, $\bar{m} \ll N$, os menores dentro do conjunto de modos de deformação.

Uma vez que os principais métodos de obter autovalores identificam um bloco único dos menores deles, faz-se a determinação dos menores m apenas, onde m deve ser calculado de forma a incluir os r modos de corpo rígido (se houverem), e os p modos quase nulos, mais uma quantidade adequada dos primeiros modos de deformação.

Consideremos a matriz modal truncada, $\bar{\mathbf{X}}$, agora de dimensões $N \times m$, isto é, constituída apenas pelos autovetores correspondentes aos m **autovalores selecionados**. Todas as etapas desenvolvidas nas eqs.(23.132)-(23.137), adaptadas para $\bar{\mathbf{X}}$ em lugar de \mathbf{X} , continuam formalmente iguais. Entretanto, algumas consequências críticas surgem, como por exemplo,

1. A ortonormalização (23.132), $\bar{\mathbf{X}}^T\bar{\mathbf{X}} = \mathbf{I}_m$, onde \mathbf{I}_m é uma identidade de **ordem reduzida** $m \times m$.
2. Entretanto, $\bar{\mathbf{X}}\bar{\mathbf{X}}^T$ é igual a uma matriz quadrada e simétrica, de dimensões $N \times N$. Entretanto, $\mathbf{B} \equiv \bar{\mathbf{X}}\bar{\mathbf{X}}^T \neq \mathbf{I}_N$. A matriz \mathbf{B} é quadrada, mas possui apenas posto $m < N$, isto é, $\bar{\mathbf{X}}$ não mais é capaz de representar qualquer vetor N -dimensional.

23.8.3 Sistemas lineares com matriz retangular

Consideramos sistemas algébricos $\mathbf{AU} = \mathbf{F}$ onde a matriz \mathbf{A} seja retangular real, de dimensões $a \times b$. Sistemas onde $a > b$ são sobre determinados, isto é, possuem mais restrições que incógnitas, por isso mesmo podem não possuir nenhuma solução. Já no caso $a < b$, não existem restrições suficientes e pode existirem infinitas soluções. Alguns casos notáveis são sumarizados a seguir, **apenas para o caso em que \mathbf{A} tenha posto completo, isto é**, $\text{posto}(\mathbf{A}) = \min\{a, b\}$. Casos em que $r = \text{posto}(\mathbf{A}) < \min\{a, b\}$, a pseudo inversa será obtida via Decomposição Singular (SVD), vista na seção seguinte e na eq. (23.163).

Caso 1 - $a \leq b$, com $\text{posto}(\mathbf{A}) = a$. Nesse caso as **linhas** de \mathbf{A} são linearmente independentes e formam uma base para \mathbf{U} . Então se pode escrever $\mathbf{U} = \mathbf{A}^T \mathbf{V}$. Então sistema fica:

$$\boxed{\phantom{\mathbf{A}}}_{a \times b} \rightarrow \begin{cases} \mathbf{AU} = \mathbf{F}, \\ \mathbf{AA}^T \mathbf{V} = \mathbf{F} \end{cases} \quad (23.146)$$

Agora \mathbf{AA}^T tem dimensões $a \times a$ e posto a . Logo é não singular e pode ser invertida (ou fatorada), resultando

$$\begin{aligned} \mathbf{V} &= (\mathbf{AA}^T)^{-1} \mathbf{F}, \text{ o que resulta em} \\ \mathbf{U} &= \mathbf{A}^T \mathbf{V}, \\ &= \underbrace{\mathbf{A}^T (\mathbf{AA}^T)^{-1}}_{\mathbf{A}^+} \mathbf{F}, \text{ isto é, a pseudo inversa é:} \end{aligned} \quad (23.147)$$

$$\boxed{\mathbf{A}^+ = \mathbf{A}^T (\mathbf{AA}^T)^{-1}} \quad (23.148)$$

: **Caso 2** - $a \geq b$, com $\text{posto}(\mathbf{A}) = b$. Nesse caso as **colunas** de \mathbf{A} são linearmente independentes e formam uma base para \mathbf{U} . O sistema fica

$$\boxed{\phantom{\mathbf{A}}}_{a \times b} \rightarrow \begin{cases} \mathbf{AU} = \mathbf{F}, \\ \mathbf{A}^T \mathbf{AU} = \mathbf{A}^T \mathbf{F}. \end{cases} \text{ que pré-multiplicado por } \mathbf{A}^T \text{ fica}$$

A matriz $\mathbf{A}^T \mathbf{A}$ tem dimensões $b \times b$ e posto b . Logo ela é não singular e pode ser invertida, resultando

$$\mathbf{U} = \underbrace{(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T}_{\mathbf{A}^+} \mathbf{F}, \text{ isto é, a pseudo inversa é:} \quad (23.149)$$

$$\boxed{\mathbf{A}^+ = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T} \quad (23.150)$$

Caso 3 - Casos em que $r = \text{posto}(\mathbf{A}) < \min\{a, b\}$, as matrizes \mathbf{AA}^T e $\mathbf{A}^T \mathbf{A}$ dos dois casos acima serão singulares. As quantidades de autovalores nulos são $(a - r)$ e $(b - r)$, respectivamente. De fato, o cálculo dessa quantidade de autovalores não nulos é uma das muitas maneiras de identificar o posto de uma matriz retangular. Assim essas matrizes quadradas não possuem inversa, mas sempre existe uma pseudo inversa, que pode ser obtida via Decomposição Singular (SVD), como visto na próxima seção, e eq. (23.163).

23.8.4 Decomposição singular - SVD (*Singular value decomposition*)

Inicialmente listamos algumas definições:

1. **Matriz normal.** É uma matriz complexa quadrada tal que

$$\mathbf{A}^* \mathbf{A} = \mathbf{A} \mathbf{A}^* \quad (23.151)$$

onde:

- A. \mathbf{U} é uma matriz de dimensões $a \times a$, **real ou complexa unitária**, cujas colunas são os chamados **vetores singulares a esquerda**. Se provará que \mathbf{U} são os autovetores ortonormalizados de \mathbf{AA}^* , isto é,

$$[\mathbf{AA}^*]_{a \times a} \mathbf{U}_{a \times a} = \mathbf{U}\mathbf{S}. \quad (23.158)$$

\mathbf{S} é uma matriz quadrada diagonal de ordem $a \times a$, com $r = \min\{a, b\}$ autovalores não negativos. (Note que \mathbf{AA}^* é uma matriz quadrada e simétrica). Esses autovalores são ordenados em ordem decrescente e as colunas em \mathbf{U} são ordenados de forma correspondentes.

- B. \mathbf{V} é uma matriz de dimensões $b \times b$, **real ou complexa unitária**, cujas colunas são os chamados **vetores singulares a direita**. Se provará que \mathbf{V} são os autovetores ortonormalizados de $\mathbf{A}^*\mathbf{A}$, isto é,

$$[\mathbf{A}^*\mathbf{A}]_{b \times b} \mathbf{V}_{b \times b} = \mathbf{V}\mathbf{s}. \quad (23.159)$$

\mathbf{s} é uma matriz quadrada diagonal de ordem $b \times b$, com $r = \min\{a, b\}$ autovalores não negativos. (Note que \mathbf{AA}^* é uma matriz quadrada e simétrica). Esses autovalores são ordenados em ordem decrescente e as colunas em \mathbf{V} são ordenados de forma correspondentes. Prova-se que os valores não nulos em \mathbf{S} e \mathbf{s} são os mesmos.

- C. \mathbf{D} é uma matriz retangular diagonal, $a \times b$, cujos valores na diagonal são **reais não negativos**. A quantidade de valores não nulos em \mathbf{D} é igual a r . Os termos não nulos da diagonal de \mathbf{D} é igual à raiz quadrada dos autovalores na diagonal de \mathbf{S} ou \mathbf{s} .

As aplicações da SVD são as mesmas da representação espectral, porém acrescidas de outras que permitem o uso em matrizes retangulares, por exemplo:

- determinação de pseudo inversa de problemas lineares com matriz retangular,
- ajuste de curva pelo método dos mínimos quadrados,
- aproximação e truncamento de matrizes,
- determinação do posto (igual a r), dentre muitas outras aplicações.

A demonstração de (23.158) e (23.159) pode ser feita da seguinte forma. Dada a SVD (23.157), $\mathbf{A} = \mathbf{UDV}^*$, e usando a propriedade de \mathbf{V} e \mathbf{U} de serem unitárias, ($\mathbf{U}^*\mathbf{U} = \mathbf{UU}^* = \mathbf{I}_a$, e $\mathbf{V}^*\mathbf{V} = \mathbf{VV}^* = \mathbf{I}_b$),

$$\begin{aligned} \mathbf{A}^*\mathbf{A} &= (\mathbf{VD}^*\mathbf{U}^*)(\mathbf{UDV}^*), \\ &= \mathbf{VD}^*\mathbf{DV}^*, \\ \mathbf{AA}^* &= (\mathbf{UDV}^*)(\mathbf{VD}^*\mathbf{U}^*), \\ &= \mathbf{UDD}^*\mathbf{U}^*. \end{aligned} \quad (23.160)$$

$\mathbf{VD}^*\mathbf{DV}^*$ (e $\mathbf{UDD}^*\mathbf{U}^*$) é a decomposição modal de $\mathbf{A}^*\mathbf{A}$ (e de \mathbf{AA}^*), onde $\mathbf{D}^*\mathbf{D}$ e \mathbf{DD}^* são matrizes diagonais reais positivas semidefinidas. Pós-multiplicando a primeira igualdade por \mathbf{V} e a segunda por \mathbf{U} se chega aos dois autoproblemas (23.158) e (23.159), com $\mathbf{s} = \mathbf{D}^*\mathbf{D}$ e $\mathbf{S} = \mathbf{DD}^*$. As matrizes \mathbf{U} e \mathbf{V} são então as bases ortonormais próprias de $\mathbf{A}^*\mathbf{A}$ e de \mathbf{AA}^* .

Relações do SVD e decomposição modal

Caso \mathbf{A} seja simétrica e real, a decomposição modal, proveniente do teorema espectral, tem a forma $\mathbf{A} = \mathbf{UDU}^T$. Na SVD a representação da matriz retangular tem a forma $\mathbf{A} = \mathbf{UDV}^*$. No primeiro caso existe apenas uma base ortonormal para definir a representação, e no segundo caso

$$\mathbf{A}^+ = \mathbf{VD}^+\mathbf{U}^*, \quad (23.164)$$

onde \mathbf{D}^+ é a pseudo inversa de \mathbf{D} , que é obtida simplesmente tomando o de \mathbf{D} e fazendo a inversa dos valores não nulos da diagonal. Deve-se observar que esse procedimento é válido qualquer que seja o posto de \mathbf{A} , mesmo que ele seja menor que o $\min\{a, b\}$, o que é diferente do procedimento mostrado nas eqs. (23.148) e (23.150).

Exemplo 23.8 - Determinação de decomposição singular e pseudo inversa

Determine a SVD e a pseudo inversa das seguintes matrizes. Nas matrizes quadradas obter também os autopares de sua representação espectral e sua inversa.

$$\begin{aligned} 1. \quad \mathbf{A}_1 &= \begin{bmatrix} 1 & 3 \\ -4 & 3 \end{bmatrix}; & 2. \quad \mathbf{A}_2 &= \begin{bmatrix} 1-i & 3 \\ -4 & 3 \end{bmatrix}; \\ 3. \quad \mathbf{A}_3 &= \begin{bmatrix} 1 & 0 & 3 \\ -4 & 0 & 3 \end{bmatrix}; & 4. \quad \mathbf{A}_4 &= \begin{bmatrix} 1 & 0 & 3 \\ -4 & 0 & 3 \\ 0 & 10^{-10} & 0 \end{bmatrix}; \end{aligned}$$

Solução do caso 1

Essa é uma matriz real não simétrica e seus autopares podem ser complexos. No presente caso a solução do autoproblema associado $\mathbf{A}_1\mathbf{X} = \mathbf{X}\mathbf{\Lambda}$ é realmente complexa:

$$\mathbf{\Lambda} = \begin{bmatrix} 2 + 3,317i & \\ & 2 - 3,317i \end{bmatrix} \text{ e } \mathbf{X} = \begin{bmatrix} 0,18898 - 0,6268i & 0,18898 + 0,6268i \\ 0,75593 & 0,75593 \end{bmatrix} \quad (23.165)$$

Os autovetores não estão normalizados. A SVD de \mathbf{A}_1 , tal que $\mathbf{A}_1 = \mathbf{UDV}^*$ é

$$\mathbf{D} = \begin{bmatrix} 5,14916 & \\ & 2,91309 \end{bmatrix}, \quad \mathbf{U} = \begin{bmatrix} 0,28978 & 0,95709 \\ 0,95709 & -0,28978 \end{bmatrix}, \quad \mathbf{V} = \begin{bmatrix} -0,68722 & 0,72645 \\ 0,72645 & 0,68722 \end{bmatrix}. \quad (23.166)$$

Nota-se que toda a decomposição é real. Os vetores são tais que $\mathbf{UU}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$ (o mesmo para \mathbf{V}). A pseudo inversa \mathbf{A}_1^+ é obtida fazendo a inversa de \mathbf{D}^* (como ela é quadrada e real trata-se aqui apenas da inversa):

$$\mathbf{D}^+ = \begin{bmatrix} 1/5,1492 & \\ & 1/2,9131 \end{bmatrix} \text{ e } \mathbf{A}_1^+ = \mathbf{VD}^+\mathbf{U}^* = \begin{bmatrix} 0,2 & -0,2 \\ 0,26667 & 0,06667 \end{bmatrix}, \quad (23.167)$$

Um cálculo direto mostra que \mathbf{A}_1^+ é igual à própria inversa de \mathbf{A}_1 , como esperado.

Solução do caso 2

A matriz \mathbf{A}_2 é quadrada porém complexa. A matriz \mathbf{AA}^* em (23.158) fica

$$\mathbf{AA}^* = \begin{bmatrix} 11 & 5 + 4i \\ 5 + 4i & 25 \end{bmatrix}$$

É uma matriz Hermitiana, cujos autovalores são

$$\mathbf{S} = \begin{bmatrix} 27,4868 & \\ & 8,5132 \end{bmatrix},$$

e os autovetores são \mathbf{U} , que são os vetores a direita da SVD de \mathbf{A}_2 , tal que $\mathbf{A}_2 = \mathbf{UDV}^*$.

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & & \\ & 1.22475 \cdot 10^{-5} & \\ & & -1.22475 \cdot 10^{-5} \end{bmatrix} \quad (23.174)$$

Então seu número de condição é igual a $1/1.22475 \cdot 10^{-5} = 8,16 \cdot 10^4$.

A SVD de \mathbf{A}_4 , tal que $\mathbf{A}_4 = \mathbf{UDV}^*$ é

$$\begin{aligned} \mathbf{D} &= \begin{bmatrix} 5,14916 & & & \\ & 2,91309 & & \\ & & 10^{-11} & \\ & & & \end{bmatrix}, & \mathbf{U} &= \begin{bmatrix} -0,28978 & 0,95709 & 0 & \\ -0,95709 & -0,28978 & 0 & \\ 0 & 0 & 0 & -1 \end{bmatrix}, \\ \mathbf{V} &= \begin{bmatrix} 0,687215 & 0,72645 & 0 & \\ 0 & 0 & -1 & \\ -0,72645 & 0,687215 & 0 & \end{bmatrix}. \end{aligned} \quad (23.175)$$

Como obrigatório, \mathbf{D} é real, e aqui os vetores também são reais, tais que $\mathbf{UU}^* = \mathbf{U}^*\mathbf{U} = \mathbf{I}$ (o mesmo para \mathbf{V}). \mathbf{D} mostra que o posto de \mathbf{A}_4 é 3, igual ao número de linhas. A pseudo inversa é obtida a partir da inversa transposta de \mathbf{D} :

$$\begin{aligned} \mathbf{D}^+ &= \begin{bmatrix} 1/5,14916 & & & \\ & 1/2,91309 & & \\ & & 10^{11} & \\ & & & \end{bmatrix} \text{ e} \\ \mathbf{A}_4^+ &= \mathbf{VD}^+\mathbf{U}^* = \begin{bmatrix} 0,2 & 0,2 & 0 & \\ 0 & 0 & 10^{11} & \\ 0,266667 & 0,066667 & 0 & \end{bmatrix}, \end{aligned} \quad (23.176)$$

que é a mesma já obtida em (23.173), uma vez que a pseudo inversa é única.

Agora, buscamos identificar uma matriz aproximada de \mathbf{A}_4 , expurgando seu valor singular quase nulo. Assim a matriz aproximada $\tilde{\mathbf{A}}_4$ será bem condicionada. Para isso, truncaremos as matrizes da SVD, anulando os termos associados a D_{33} em todas as matrizes:

$$\mathbf{D} = \begin{bmatrix} 5,14916 & & \\ & 2,91309 & \\ & & 0 \end{bmatrix}, \quad (23.177)$$

tal que

$$\mathbf{D}^+ = \begin{bmatrix} 1/5,14916 & & \\ & 1/2,91309 & \\ & & 0 \end{bmatrix}. \quad (23.178)$$

Assim, a representação da aproximação de \mathbf{A}_4 e de sua pseudo inversa são dadas por

$$\begin{aligned} \tilde{\mathbf{A}}_4 &= \mathbf{UDV}^* = \begin{bmatrix} 1 & 0 & 3 \\ -4 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix} \text{ e} \\ \tilde{\mathbf{A}}_4^+ &= \mathbf{VD}^+\mathbf{U}^* = \begin{bmatrix} 0,2 & 0,2 & 0 \\ 0 & 0 & 0 \\ 0,266667 & 0,066667 & 0 \end{bmatrix} \end{aligned} \quad (23.179)$$

Note que agora o número de condição de $\tilde{\mathbf{A}}_4$ é infinito pois possui um autovalor nulo. $\tilde{\mathbf{A}}_4$ é perfeitamente singular, logo, não invertível. Entretanto a pseudo inversa é obtida sem que esse mal condicionamento afete qualquer das operações.