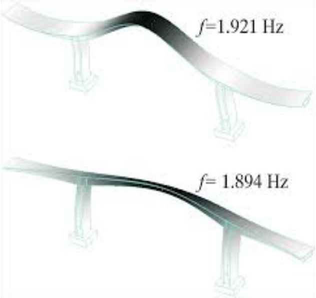
SME0305 - 2016

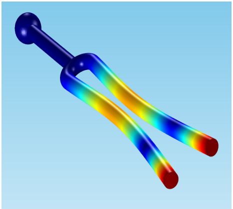
Roberto F. Ausas / Gustavo C. Buscaglia

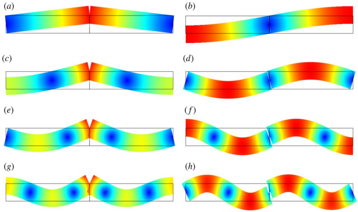
ICMC - Ramal 736628, rfausas@gmail.com
ICMC - Ramal 738176, gustavo.buscaglia@gmail.com

Cálculo de autovalores e autovetores

Existem vários problemas na engenharia em que precisamos calcular os autovalores e autovetores de uma matriz:


f=1.921 Hz
f= 1.894 Hz




(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)

Precisamos estudar o que há por traz dos métodos numéricos para calcular-los. A ideia é relembrar algumas noções de algebra linear sobre autovalores e autovetores de matrizes e depois introduzir o Método das potências.

Definição: Seja $A \in \mathbb{R}^{n\times n}$, $\lambda \in \mathbb{R}$ é autovalor de A se $A - \lambda\,\mathbb{I}$ é singular.

Por tanto, se λ é autovalor de A, existe $\mathbf{v} \in \mathbb{R}^n,\ \mathbf{v} \neq \mathbf{0}$ tal que

$$(A - \lambda\,\mathbb{I})\,\mathbf{v} = \mathbf{0}$$

ou seja:

$$A\,\mathbf{v} = \lambda\,\mathbf{v}$$

Se a matriz $A \in \mathbb{R}^{n \times n}$ é simétrica (i.e., $A = A^T$), então ela tem n autovalores (contados com a sua multiplicidade) e n autovetores linearmente independentes.

Definição - Polinômio característico: Se λ é autovalor da matriz $A \in \mathbb{R}^{n \times n}$, então:

$$P(\lambda) = \det\left(A - \lambda\,\mathbb{I}\right) = 0$$

$P(\lambda)$ é chamado de polinômio caracteristico de A e os autovalores de A são os zeros desse polinômio.

Observação: Se um autovalor tem multiplicidade maior do que 1, seja m, então, teremos m autovetores que geram o subespaço (de dimensão m) associado a esse autovalor.

Propriedades

- Sejam $\mathbf{v}$ and $\mathbf{w}$ autovetores correspondentes a autovalores distintos λ e μ de uma matriz simétrica, i.e.

$$A\,\mathbf{v} = \lambda\,\mathbf{v}$$
$$A\,\mathbf{w} = \mu\,\mathbf{w}$$

  Então, $\mathbf{v}$ e $\mathbf{w}$ são ortogonais:

$$(v, w) = \mathbf{v}^T\,\mathbf{w} = 0$$

- Se λ é autovalor de A, então $\lambda + q, \quad q \in \mathbb{R}$ é autovalor de $A + q\,\mathbb{I}$.

- Se λ é autovalor de A, então $\frac{1}{\lambda+q}, \quad q \in \mathbb{R}$ é autovalor de $(A + q\,\mathbb{I})^{-1}$.

- O traço de A é:

$$\text{trace}(A) = \sum_{i=1}^{n} A_{ii} = \sum_{i=1}^{n} \lambda_i$$

- O determinante de A é:

$$\det(A) = \lambda_1\,\lambda_2\,...\,\lambda_n$$

Definição: Matrizes semelhantes - Duas matrizes A e B dizem-se semelhantes, se existe uma matrix Q não singular tal que:

$$A = Q\,B\,Q^{-1}$$

ou

$$B = Q^{-1}\,A\,Q$$

Propiedade: Se $(\lambda, \mathbf{v})$ é **autopar** de A, então $(\lambda, Q^{-1}\mathbf{v})$ é autopar de B.

Definição: Uma matriz diz-se **diagonalizável** se ela for semelhante a uma matriz diagonal, ou seja, existe Q tal que

$$A = Q\,D\,Q^{-1}$$

em que D é uma matriz diagonal, i.e.,

$$D = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$$

Teorema: Seja $A \in \mathbb{R}^{n \times n}$ simétrica. Então, existe uma matriz Q ortogonal (i.e., $Q^{-1} = Q^T$), tal que

$$Q^T \, A \, Q = D = \begin{bmatrix} \lambda_1 & 0 & ... & 0 \\ 0 & \lambda_2 & ... & 0 \\ . & . & . & . \\ . & . & . & . \\ . & . & . & . \\ 0 & 0 & ... & \lambda_n \end{bmatrix}$$

em que os λ_i, $i = 1, \ldots, n$ são os autovalores de A. Notemos que as colunas de Q são de fato os autovetores de A:
Para ver isto, tomamos um vetor da base canónica, i.e., o vetor $\mathbf{e}_{(i)}$ está feito de zeros exceto na sua compenente i que vale 1:

$$\mathbf{e}_{(i)} = \begin{bmatrix} 0 \\ 0 \\ . \\ . \\ . \\ 1 \\ . \\ . \\ . \\ 0 \end{bmatrix}$$

Agora fazemos $D\,\mathbf{e}_{(i)} = (Q^T\,A\,Q)\,\mathbf{e}_{(i)} = \lambda_i\,\mathbf{e}_{(i)}$

Então:

$$(A\,Q\,)\,\mathbf{e}_{(i)} = Q\,(\lambda_i\,\mathbf{e}_{(i)}) = \lambda_i\,(Q\,\mathbf{e}_{(i)})$$

Agora chamemos $\mathbf{v}_{(i)} = Q\,\mathbf{e}_{(i)}$

$$A\,\mathbf{v}_{(i)} = \lambda_i\,\mathbf{v}_{(i)}$$

Isto significa que $\mathbf{v}_{(i)} = Q\,\mathbf{e}_{(i)}$ é o autovetor associado ao autovalor λ_i da matriz A. Mas, acontece que $Q\,\mathbf{e}_{(i)}$ é a coluna número i da matriz Q, i.e.,

$$Q = \begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \mathbf{v}_{(1)} & \mathbf{v}_{(2)} & \cdots & \mathbf{v}_{(n)} \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix}$$

Isto significa, que a matriz Q é feita de "pendurar" os autovetores de matriz A.

Método das potências e as suas variantes

Motivação: O método das potências é um método iterativo para achar o autovalor de maior módulo de uma matriz. Vamos motivar ele com um exemplo gráfico na louza.

Vamos supor que temos uma matriz com um autovalor λ_1 que é maior em valor absoluto que todos os autovalores restantes.

Podemos ver o que acontece escrevendo uma combinação linear dos autovetores da forma:

$$\mathbf{v} = c_1\,\mathbf{v}_1 + c_2\,\mathbf{v}_2 + \cdots + c_n\,\mathbf{v}_n$$

Agora multiplicamos por A em ambos lados

$$A\,\mathbf{v} = c_1\,A\,\mathbf{v}_1 + c_2\,A\,\mathbf{v}_2 + \cdots + c_n\,A\,\mathbf{v}_n$$

e já que $A\,\mathbf{v}_i = \lambda_i\,\mathbf{v}_i, \quad i = 1, \ldots, n$, resulta

$$A\mathbf{v} = c_1\,\lambda_1\,\mathbf{v}_1 + c_2\,\lambda_2\,\mathbf{v}_2 + \cdots + c_n\,\lambda_n\,\mathbf{v}_n$$

Tirando fator comum λ_1

$$A\mathbf{v} = \lambda_1\left[c_1\,\mathbf{v}_1 + c_2\,\left(\frac{\lambda_2}{\lambda_1}\right)\,\mathbf{v}_2 + \cdots + c_n\,\left(\frac{\lambda_2}{\lambda_1}\right)\,\mathbf{v}_n\right]$$

Multiplicamos novamente por A em ambos lados

$$A^2\mathbf{v} = \lambda_1^2 \left[c_1\,\mathbf{v}_1 + c_2\,\left(\tfrac{\lambda_2}{\lambda_1}\right)^2\,\mathbf{v}_2 + \cdots + c_n\,\left(\tfrac{\lambda_n}{\lambda_1}\right)^2\,\mathbf{v}_n \right]$$

$$\vdots$$

$$A^k\mathbf{v} = \lambda_1^k \left[c_1\,\mathbf{v}_1 + \underbrace{c_2\,\left(\frac{\lambda_2}{\lambda_1}\right)^k\,\mathbf{v}_2 + \cdots + c_n\,\left(\frac{\lambda_n}{\lambda_1}\right)^k\,\mathbf{v}_n}_{\mathbf{r}} \right]$$

Claramente, se $\left(\frac{\lambda_i}{\lambda_1}\right) < 1, \quad i = 2, 3, \dots, n$, então, quando $k \to \infty$, $\left(\frac{\lambda_i}{\lambda_1}\right)^k \to 0$. Isto é, quando $k \to \infty$, $A^k\mathbf{v}$ tende a ficar alinhado com $\mathbf{v_1}$, já que o resto $\mathbf{r}$ tende para zero.

Motivados por isto, vamos propor o seguinte método iterativo:

Vamos supor $A \in \mathbb{R}^{n \times n}$ que possui autovalores $\lambda_i, \ i = 1, \ldots, n$ e se verifica que

$$|\lambda_1| > |\lambda_2| \ \geq \ |\lambda_3| \ \geq \ \ldots \ \geq \ |\lambda_n|$$

O algoritmo é:

- Dado $\mathbf{x}^{(0)} \in \mathbb{R}^n$, $\mathbf{y}^{(0)} = \frac{\mathbf{x}^{(0)}}{\|\mathbf{x}^{(0)}\|}$, $\lambda^{(0)} = \mathbf{y}^{(0)\,T} A\,\mathbf{y}^{(0)}$
- **for** $k = 1, 2, ...$
  - $\mathbf{x}^{(k)} = A\,\mathbf{y}^{(k-1)}$
  - $\mathbf{y}^{(k)} = \frac{\mathbf{x}^{(k)}}{\|\mathbf{x}^{(k)}\|}$
  - $\lambda^{(k)} = \mathbf{y}^{(k)\,T} A\,\mathbf{y}^{(k)}$
  - Se $|\lambda^{(k)} - \lambda^{(k-1)}| < TOL\,|\lambda^{(k)}| \Rightarrow$ Sair
- **end for**
- **return** $\lambda^{(k)}$

Convergência do método:

$$\|\mathbf{y}^{(k)} - \mathbf{v}_{(1)}\| \leq C \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k$$

No exemplo, vimos que num algoritmo deste tipo $\mathbf{y}^{(k)}$ tende ao autovetor normalizado correspondente ao autovalor de maior módulo. Vejamos que efetivamente

$$\lambda^{(k)} = {\mathbf{y}^{(k)}}^T A\, \mathbf{y}^{(k)} \ \xrightarrow[k\to\infty]{} \ \lambda_1$$

Se $\mathbf{y}^{(k)} \ \xrightarrow[k\to\infty]{} \ \mathbf{v}_{(1)} \ \Rightarrow \ A\, \mathbf{y}^{(k)} \ \xrightarrow[k\to\infty]{} \ \lambda_1\, \mathbf{y}^{(k)}$

Então,

$${\mathbf{y}^{(k)}}^T A\, \mathbf{y}^{(k)} \ \xrightarrow[k\to\infty]{} \ \lambda_1\, {\mathbf{y}^{(k)}}^T \mathbf{y}^{(k)} = \lambda_1$$

O que de fato pode acontecer quando o outovalor dominante é negativo, é que $\mathbf{y}^k \to \pm\mathbf{v}_{(1)}$. Ilustremos isto com o programa `plot_vecs_norm.m`. Para ver isto, podemos fazer os cálculos de antes, mas normalizando a cada passo, como de fato fazemos no algoritmo. Nesse caso podemos provar que:

$$\lim_{k\to\infty} \mathbf{y}^{(k)} = \lim_{k\to\infty} \beta^{(k)} \lambda_1^k c_1 \mathbf{v}_{(1)}$$

em que $\beta^{(k)} = 1/\parallel A^k\,\mathbf{v}\parallel_2$ e por tanto

$$\beta^{(k)}\,\lambda_1\,c_1 = \frac{\lambda_1^k\,c_1}{\|\lambda_1^k(c_1\mathbf{v}_{(1)}+\mathbf{r})\|_2} \xrightarrow[k\to\infty]{} \pm 1$$

Exercício: Ver os detalhes completos nas slides 36 e 37 do professor Afonso Paiva.

Método das potências inversas

Lembremos que se $(\lambda, \mathbf{v})$ é autopar de $A \Rightarrow (\frac{1}{\lambda}, \mathbf{v})$ é autopar de A^{-1}, i.e.,

$$A^{-1}\,\mathbf{v} = A^{-1}\left(\frac{1}{\lambda}\,A\,\mathbf{v}\right) = \frac{1}{\lambda}\,\mathbf{v}$$

Então os autovalores de $B = A^{-1}$ serão: $\{\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}\}$, o que significa, que o autovalor de maior módulo da matriz $B = A^{-1}$, corresponderá ao autovalor de menor módulo de A

$$\lambda_{\mathrm{max}}(B) = \frac{1}{\lambda_{\mathrm{min}}(A)}$$

Poderiamos aplicar o método das potências que acabamos de introduzir na matriz $B(= A^{-1})$ para determinar o autovalor de menor módulo de A.

Vamos supor $A \in \mathbb{R}^{n \times n}$ que posui autovalores $\lambda_i, \; i = 1, \dots, n$ e se verifica que

$$|\lambda_1| \; \geq \; |\lambda_2| \; \geq \; |\lambda_3| \; \geq \; \dots > |\lambda_n|$$

O algoritmo é:

- Dado $\mathbf{x}^{(0)} \in \mathbb{R}^n$, $\mathbf{y}^{(0)} = \frac{\mathbf{x}^{(0)}}{\|\mathbf{x}^{(0)}\|}$, $\lambda^{(0)} = \mathbf{y}^{(0)\,T} A^{-1} \mathbf{y}^{(0)}$
- **for** $k = 1, 2, \ldots$
  - $\mathbf{x}^{(k)} = A^{-1} \mathbf{y}^{(k-1)}$
  - $\mathbf{y}^{(k)} = \frac{\mathbf{x}^{(k)}}{\|\mathbf{x}^{(k)}\|}$
  - $\mu^{(k)} = \mathbf{y}^{(k)\,T} A^{-1} \mathbf{y}^{(k)}$
  - Se $|\mu^{(k)} - \mu^{(k-1)}| < TOL\,|\mu^{(k)}| \;\Rightarrow$ Sair
- **end for**
- **return** $\frac{1}{\mu^{(k)}}$

Lembremos que em geral não precisamos calcular a inversa da matriz. Por exemplo no passo k teremos

$$A^{-1}\,\mathbf{y}^{(k)} = \mathbf{x}^{(k)} \;\Rightarrow\; A\,\mathbf{x}^{(k)} = \mathbf{y}^{(k)}$$

Este sistema pode ser resolvido para achar $\mathbf{x}^{(k)}$ usando a fatoração LU ou de Cholesky.

Uma vez que essa fatoração está pronta, ela pode ser usada cada vez que precisar dentro do algoritmo fazendo:

```
z = L \y
x = U \z
```

Potências inversas com deslocamento

Já vimos métodos para calcular:

- O autovalor de maior módulo, i.e., o autovalor mais afastado de zero;
- O autovalor de menor módulo, i.e., o autovalor mais perto de zero;

Que acontece se quisermos o autovalor mais próximo de um certo número q. Neste caso podemos usar o método das Potências inversas com deslocamento

Para chegar neste método procedemos assim:

- Vamos supor uma matriz A com autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$.
- Lembremos que a matriz $A - q\,\mathbb{I}$ tem autovalores $\{\lambda_1 - q, \lambda_2 - q, \dots, \lambda_q - q, \dots, \lambda_n - q\}$ em que λ_q é autovalor de A mais próximo de um número q.
- Aplicamos o método das potências inversas a A_q, i.e.

$$\lambda_{\max}(A_q^{-1}) = \frac{1}{\lambda_{\min}(A_q)} = \frac{1}{\lambda_q - q}$$

  Então:

$$\lambda_q = \frac{1}{\lambda_{\max}(A_q^{-1})} + q$$

Para ter uma ideia da região aonde procurar os autovalores podemos usar os chamados discos de Gershgorin (ver slides 44 e 45 do Prof. Afonso Paiva).