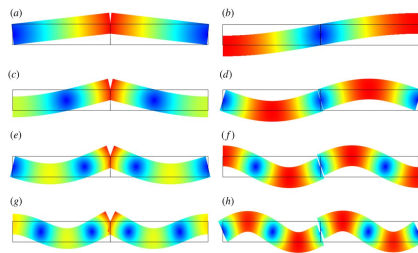
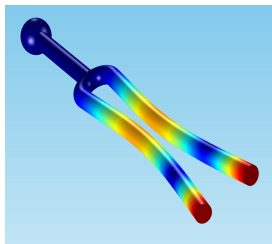
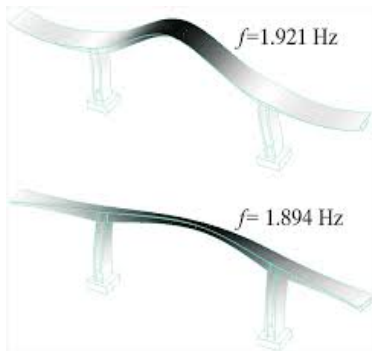


ICMC - Ramal 738176, gustavo.buscaglia@gmail.com

Cálculo de autovalores e autovetores

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



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 $v \in \mathbb{R}^n$, $v \neq \mathbf{0}$ tal que

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

ou seja:

$$A v = \lambda v$$

Definição - Polinômio característico

$$P(\lambda) = \det(A - \lambda \mathbb{I})$$

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

Matrizes semelhantes, matrizes diagonalizáveis

- **Teorema:** Se A e B , matrizes $n \times n$, são **semelhantes**, i.e., satisfazem

$$B = S^{-1} A S, \quad \text{ou} \quad A = S B S^{-1}$$

para alguma S não singular, e (λ, v) é auto-par de A , então $(\lambda, S^{-1}v)$ é auto-par de B .

- **Definição:** Uma matriz A é **diagonalizável** se ela é **semelhante a uma matriz diagonal**. Para isto, deverá existir uma matriz S não singular e uma matriz

$$\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$$

(os λ 's podem ser repetidos) tais que $A = S \Lambda S^{-1}$.

- Uma matriz $A \in \mathbb{R}^{n \times n}$ é **diagonalizável** se e só se existem n autovetores linearmente independentes.

Matrizes simétricas

- Toda matriz real simétrica é diagonalizável.
- Sejam v and w autovetores correspondentes a autovalores distintos λ e μ de uma matriz simétrica, i.e.

$$A v = \lambda v, \quad A w = \mu w$$

Então, v e w são ortogonais:

$$(v, w) = v^T w = 0$$

- Se A é simétrica, então $A = S \Lambda S^T$, sendo S ortogonal e Λ diagonal.

Mais propriedades importantes

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

$$\det(A) = \lambda_1 \lambda_2 \dots \lambda_n$$

Método das potências e as suas variantes

O método das potências é um método iterativo para achar o autovalor de maior módulo de uma matriz.

Vamos supor que temos uma matriz A diagonalizável cujo autovalor λ_1 é maior em valor absoluto que todos os autovalores restantes.

Escrevendo um vetor arbitrário z como combinação dos autovetores obtemos

$$z = c_1 v^{(1)} + c_2 v^{(2)} + \dots + c_n v^{(n)}$$

Agora multiplicamos por A em ambos lados

$$Az = c_1 A v^{(1)} + c_2 A v^{(2)} + \dots + c_n A v^{(n)}$$

e já que $A v^{(i)} = \lambda_i v^{(i)}$, $i = 1, \dots, n$, resulta

$$Az = c_1 \lambda_1 v^{(1)} + c_2 \lambda_2 v^{(2)} + \dots + c_n \lambda_n v^{(n)}$$

Tirando fator comum λ_1

$$Az = \lambda_1 \left[c_1 v^{(1)} + c_2 \left(\frac{\lambda_2}{\lambda_1} \right) v^{(2)} + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right) v^{(n)} \right]$$

Multiplicamos novamente por A em ambos lados

$$\begin{aligned}
 A^2 z &= \lambda_1^2 \left[c_1 v^{(1)} + c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^2 v^{(2)} + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^2 v^{(n)} \right] \\
 &\vdots \\
 &\vdots \\
 &\vdots \\
 A^k z &= \lambda_1^k \left[c_1 v^{(1)} + \underbrace{c_2 \left(\frac{\lambda_2}{\lambda_1} \right)^k v^{(2)} + \dots + c_n \left(\frac{\lambda_n}{\lambda_1} \right)^k v^{(n)}}_{\mathbf{r}} \right]
 \end{aligned}$$

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

Motivados por isto, vamos propor o **método das potências**:

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

Convergência do método:

$$\begin{aligned} |\eta^{(k)} - \lambda_1| &\leq C \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k \\ \|\mathbf{y}^{(k)} - \mathbf{v}^{(1)}\| &\leq C \left(\frac{|\lambda_2|}{|\lambda_1|} \right)^k \end{aligned}$$

Assim, $\eta^{(k)}$ tende ao autovalor **mais afastado do zero**, e $\mathbf{y}^{(k)}$ tende ao autovetor correspondente normalizado.

Método das potências inversas

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

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

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

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

Podemos 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 possui autovalores λ_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 $\mathbf{x}^{(0)} \in \mathbb{R}^n$, $\mathbf{y}^{(0)} = \frac{\mathbf{x}^{(0)}}{\|\mathbf{x}^{(0)}\|}$, $\lambda^{(0)} = \mathbf{y}^{(0)T} \mathbf{A}^{-1} \mathbf{y}^{(0)}$
- **for** $k = 1, 2, \dots$
 - $\mathbf{x}^{(k)} = \mathbf{A}^{-1} \mathbf{y}^{(k-1)}$
 - $\mathbf{y}^{(k)} = \frac{\mathbf{x}^{(k)}}{\|\mathbf{x}^{(k)}\|}$
 - $\mu^{(k)} = \mathbf{y}^{(k)T} \mathbf{A}^{-1} \mathbf{y}^{(k)}$
 - 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.

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

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:

$$\begin{aligned} \mathbf{z} &= \mathbf{L} \setminus \mathbf{y} \\ \mathbf{x} &= \mathbf{U} \setminus \mathbf{z} \end{aligned}$$

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 fazer se quisermos o autovalor mais próximo de um certo número q ? Neste caso podemos usar o **método das potências inversas com deslocamento**. Procedemos assim:

- Vamos supor uma matriz A com autovalores $\{\lambda_1, \lambda_2, \dots, \lambda_n\}$.
- Lembremos que a matriz $A_q = A - q\mathbb{I}$ tem autovalores $\{\lambda_1 - q, \lambda_2 - q, \dots, \lambda_n - q\}$ e seja λ_q o autovalor de A mais próximo de um número q .
- Aplicando 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 do Prof. Afonso Paiva).

O comando `eig`

O comando `eig` calcula a fatora  o $A = S \Lambda S^{-1}$:

`[S Lam]=eig(A)` $\longrightarrow$ `A=S*Lam*inv(S)`

- As colunas de S s  o os autovetores de A correspondentes a cada autovalor.
- Se A    sim  trica, Lam    real e S ortogonal.
- Em geral, tanto S como Lam s  o complexas, mas se A    real os autovalores aparecem em pares conjugados.
- Para isto, Octave utiliza o **m  todo de Francis** (ver slides do Prof. Afonso).

No cálculo de estruturas é usual encontrar o problema generalizado de autovalores

$$A v = \lambda M v$$

onde M é uma matriz definida positiva (“matriz de massas”).

O comando `eig` de Octave também resolve esse problema:

```
[S Lam]=eig(A,M)      --->    A=M*S*Lam*inv(S)
```

- As colunas de S são os autovetores de A correspondentes a cada autovalor, i.e.,

$A*S(:,k) - \text{Lam}(k,k)*M*S(:,k)$ é igual a zero.

- Se A é simétrica, Lam é real.

A decomposição SVD

É a generalização mais útil da diagonalização de matrizes.

Valores singulares e vetores singulares

- Um real não negativo σ é **valor singular** de uma matriz $M \in \mathbb{R}^{m \times n}$ se e só se existem vetores unitários $u \in \mathbb{R}^m$ e $v \in \mathbb{R}^n$ tais que

$$A v = \sigma u, \quad \text{e} \quad A^T u = \sigma v. \quad (1)$$

Os vetores u e v são chamados de vetores singulares para σ **a esquerda** e **a direita**, respectivamente.

- **Teorema:** Toda matriz real $A \in \mathbb{R}^{m \times n}$ pode ser fatorada como

$$A = U \Sigma V^T \quad \Leftrightarrow \quad [U \ S \ V] = \text{svd}(A) \quad (2)$$

onde $U \in \mathbb{R}^{m \times m}$ e $V \in \mathbb{R}^{n \times n}$ são matrizes ortogonais, e a matriz $\Sigma \in \mathbb{R}^{m \times n}$ é diagonal, satisfazendo $\Sigma_{11} \geq \Sigma_{22} \geq \dots \Sigma_{rr} > 0$, e todos os elementos restantes são zero. O inteiro $r \leq \min(m, n)$ é o posto de A . Usaremos também a notação $\sigma_i = \Sigma_{ii}$ para os valores singulares.

Exo.: Mostrar que, se se cumpre (2), então as colunas de U e de V são vetores singulares de A , a esquerda e direita respectivamente, e os elementos diagonais de Σ são valores singulares de A .

- Denotemos por $u^{(i)}$ a coluna número i de U , e por $v^{(i)}$ a coluna número i de V . Quando resultar conveniente e não gerar confusão, usaremos também u_i e v_i para os mesmos fins.

Exo. Ação sobre um vetor z :

Dizer se alguma das identidades abaixo está errada:

$$\begin{aligned}
 Az &= A [(v_1^T z)v_1 + \dots + (v_n^T z)v_n] \\
 &= (v_1^T z)Av_1 + \dots + (v_n^T z)Av_n \\
 &= \sigma_1(v_1^T z)u_1 + \dots + \sigma_r(v_r^T z)u_r \\
 &= [\sigma_1 u_1 v_1^T + \dots + \sigma_r u_r v_r^T] z
 \end{aligned}$$

Supondo que a última esteja correta, responda se ela é suficiente para poder afirmar que

$$A = \sum_{i=1}^r \sigma_i u^{(i)} (v^{(i)})^T$$

Exo.: Os seguintes itens são verdadeiros:

- As colunas $r + 1$ a n de V são uma **base ortogonal** do **núcleo** de A .
- As colunas 1 a r de U são uma **base ortogonal** da **imagem** de A .
- A diagonal de Σ tem comprimento $\min(m, n)$, quando A é matriz $m \times n$.
- Dada a decomposição $A = U\Sigma V^T$, o posto de A é r se e só se os elementos $r + 1$ a $\min(m, n)$ da diagonal de Σ são zero, e os primeiros r elementos positivos.

- **Teorema:** Seja $A \in \mathbb{R}^{m \times n}$ uma matriz cuja decomposição SVD é dada por $A = U\Sigma V^T$. A melhor aproximação a A com matrizes de posto $k < r$ é dada por

$$\tilde{A} = U \tilde{\Sigma} V^T, \quad (3)$$

onde $\tilde{\Sigma}$ é também diagonal, seus elementos diagonais 1 a k coincidem com os de Σ , e o resto são zero.

- **Teorema:** Seja $A \in \mathbb{R}^{n \times n}$ uma matriz quadrada, com decomposição $A = U\Sigma V^T$. A matriz **ortogonal** mais próxima a A é, então, a matriz $Q = UV^T$.

Pseudoinversa

- A **pseudoinversa** de uma matriz $A \in \mathbb{R}^{m \times n}$, que denotaremos por $A^\dagger$, é de $n \times m$ e definida por

$$A^\dagger = V \Sigma^\dagger U^T, \quad (4)$$

onde $\Sigma^\dagger \in \mathbb{R}^{n \times m}$ é a matriz diagonal cujos elementos valem $1/\Sigma_{ii}$ se $\Sigma_{ii} \neq 0$, e zero se $\Sigma_{ii} = 0$.

- O problema de **mínimos quadrados**

$$Ax = b$$

com A arbitrária (de posto incompleto, em particular), se resolve de maneira simples como

$$x = A^\dagger b \quad \Leftrightarrow \quad x = \text{pinv}(A) * b \quad (5)$$

Quando A é de posto completo a solução coincide com a obtida com a fatoração QR, que é solução de

$$A^T A x = A^T b. \quad (6)$$

Quando A é de posto incompleto, e a fatoração QR fornece então infinitas soluções, a solução obtida é a de mínima norma.

As seguintes afirmações são verdadeiras?

- $AA^\dagger A = A$
- $A^\dagger AA^\dagger = A^\dagger$
- $(AA^\dagger)^T = AA^\dagger$
- $(A^\dagger A)^T = A^\dagger A$
- $A^T A = V \Sigma^T \Sigma V^T$
- $AA^T = U \Sigma \Sigma^T U^T$
- As colunas de V são autovetores de $A^T A$.
- As colunas de U são autovetores de AA^T .