

# Autovalores e autovetores



Ricardo Biloti

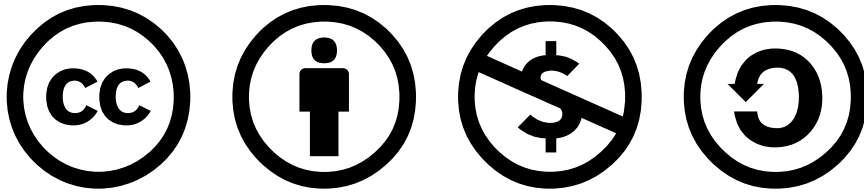
biloti@unicamp.br

Álgebra Linear Computacional

2S/2021

[www.ime.unicamp.br/~biloti](http://www.ime.unicamp.br/~biloti)

A large area of the page is filled with horizontal lines for writing, with a vertical red margin line on the left side.



Este trabalho é licenciado sob os termos da Licença Internacional Creative Commons Atribuição-NãoComercial-Compartilhalgal 4.0.

Para ver uma cópia desta licença, visite <http://creativecommons.org/licenses/by-nc-sa/4.0/>.

## Seus direitos e deveres são:

- Você é livre para copiar e redistribuir este material, em qualquer meio ou formato, para adaptá-lo, transformá-lo ou utilizá-lo para construir seu próprio material.
- Você deve dar os créditos apropriados, fornecendo link para a licença e indicando se alterações foram feitas. Você pode fazer isto de qualquer forma razoável, porém sem tentar passar a ideia ou sugerir que o autor endosse suas alterações ou seu uso do material.
- Você não pode utilizar este material para fins comerciais.
- Se você alterar, transformar ou construir seu próprio material com base neste trabalho, você deverá distribuí-lo sob a mesma licença usada no original.

Seja  $A$  uma matriz quadrada. Um vetor não nulo  $v$  tal que

$$Av = \lambda v,$$

para algum escalar  $\lambda$ , é dito um **autovetor** de  $A$ , associado ao **autovalor**  $\lambda$ .

Neste caso  $(A - \lambda I)$  é singular, uma vez que

$$(A - \lambda I)v = 0, v \neq 0 \iff \det(A - \lambda I) = 0.$$

Os autovalores de  $A$  são as raízes do **polinômio característico**

$$p(\lambda) = \det(A - \lambda I).$$

Mesmo que  $A$  seja real, seus autovalores e autovetores podem ser complexos.

$$\text{Se } A = \begin{bmatrix} -4 & 1 & 0 \\ 1 & 0 & -1 \\ 2 & 1 & 1 \end{bmatrix}, \text{ então } p(\lambda) = \lambda(\lambda + 4)(1 - \lambda) - 7$$

$$p(\lambda) = 0 \iff \lambda \in \{-4.3063, 0.6532 \pm 1.0949i\}$$

Dado que não é possível no problema de determinar autovalores e autovetores nos restringirmos a trabalhar no corpo dos reais, daqui para frente vamos considerar matrizes sobre o corpo dos complexos.

Considere  $A \in \mathbb{C}^{n \times n}$ .

- ▶ Se  $\lambda$  é autovalor de  $A$ , então  $\mathcal{L}_\lambda \equiv \{v \in \mathbb{C}^n \mid Av = \lambda v\}$  é um subespaço vetorial de  $\mathbb{C}^n$  ( $\mathcal{L}_\lambda$  é o **autoespaço** de  $A$  associado a  $\lambda$ ).
- ▶  $\mathcal{L}_\lambda = \mathcal{N}(A - \lambda I)$ .
- ▶  $A$  e  $A^T$  têm os mesmos autovalores.
- ▶ Se  $\lambda$  é autovalor de  $A$ , uma **matriz real**, então  $\bar{\lambda}$  também é autovalor de  $A$ .
- ▶ Se  $v_1$  e  $v_2$  são autovetores associados aos autovalores  $\lambda_1 \neq \lambda_2$ , então  $v_1$  e  $v_2$  são LI.

Assuma que cada uma dessas propriedades é um exercício e demonstre-as. Se estiver com dificuldade será um indicativo de que você deve revisar conteúdos de Álgebra Linear.

Note que todos os elementos **não nulos** de  $\mathcal{L}_\lambda$  são autovetores de  $A$ . Por outro lado,  $\mathcal{L}_\lambda$  é um subespaço vetorial de  $\mathbb{C}^n$  mesmo que  $\lambda$  não seja autovalor de  $A$  (por quê?).

Se  $A \in \mathbb{C}^{n \times n}$  tem  $n$  autovetores LI, então  $A$  é dita uma matriz **simples**.

## Corolário

Se  $A$  tem  $n$  autovalores distintos, então  $A$  é simples.

O corolário sai da propriedade que vimos antes de que autovetores associados a autovalores distintos são sempre LI.

## Exemplo



$$A = \begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 2 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix} \text{ é simples?}$$

- ▶  $p(A) = (\lambda - \frac{3}{2})^2(\lambda - 2) - \frac{1}{4}(\lambda - 2) = (\lambda - 1)(\lambda - 2)^2$
- ▶ Não é tão *simples* dizer se  $A$  é simples...

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

- ▶  $A$  é simples, apesar de não ter 3 autovalores distintos. 😊

Neste exemplo, apesar de  $A$  não ter 3 autovalores distintos, ainda é possível obter 3 autovetores LI.

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \text{ é simples?}$$

▶  $p(A) = (\lambda - 2)^3$

▶  $A$  tem apenas um autovetor LI

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

▶  $A$  não é simples. 😞

Neste exemplo, 2 é um autovalor de multiplicidade algébrica 3, mas tem apenas um autovetor LI associado a ele. Logo, a matriz  $A$  não é simples, mas sim *defective* ou *defeituosa*.



$$A = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} + \epsilon B,$$

onde  $B$  é uma matriz aleatória, é simples?

```
>> M = [2 1 0; 0 2 1; 0 0 2];
>> for k=1:100
    B = 2 * (rand(3,3) - 0.5); # Matriz aleatória com -1 <= M(i,j) <= 1
    A = M + 1.0e-12 * B;
    [V,D] = eig(A);           # Calcula os autovalores e autovetores de A
    posto(k) = rank(V);       # Conta quantos autovetores são LI
endfor
>> length(find(posto == 3))   # Quantas vezes o posto foi 3?
ans = 100
```

*Para toda matriz quadrada, sempre há um matriz simples, arbitrariamente próxima.*

Neste exemplo, fizemos um experimento numérico para testar o quão “estável” é a propriedade da matriz ser defeituosa. A uma matriz defeituosa, somamos uma pequena perturbação aleatória. Com o auxílio do comando `eig` do Octave, computamos seus autovetores e depois perguntamos qual o posto da matriz que tem os autovetores como coluna. Repetimos isto 100 vezes. Ao final descobrimos quantas dessas vezes o posto foi 3 indicando que havia três autovetores LI. A conclusão foi que em todas as vezes isso aconteceu.

Isto não *demonstra* nada, mas dá indícios de que ao perturbar uma matriz defeituosa o usual é que ela torne-se simples.

De fato, é um resultado que pode ser demonstrado com o auxílio da decomposição SVD que matrizes simples são densas no espaço das matrizes, ou seja, para qualquer matriz sempre haverá uma matriz simples arbitrariamente próxima.

Seja  $A \in \mathbb{C}^{n \times n}$ , uma matriz **simples**, com autovalores  $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|$ , respectivamente associados aos autovetor  $v_1, v_2, \dots, v_n$ .

Se  $|\lambda_1| > |\lambda_2|$ , dizemos que  $\lambda_1$  é o autovalor **dominante** e  $v_1$  ou autovetor **dominante**.

Neste caso, utilizaremos o **método da potência** para determinar  $\lambda_1$  e  $v_1$ .

Como vimos que as matrizes simples são muito mais abundantes que as defeituosas e que pequenas perturbações (como por exemplo as causadas por erros de contas em precisão finita) tendem a tornar a matriz uma matriz simples, vamos nos restringir daqui para frente em lidar com este caso.

Como  $A$  é simples, existe uma base de autovetores para  $\mathbb{C}^n$ . Se  $q \in \mathbb{C}^n$  é um vetor qualquer, então

$$q = c_1 v_1 + c_2 v_2 + \dots + c_n v_n.$$

No método da potência, construímos a sequência

$$q, Aq, A^2q, A^3q, \dots$$

$$q = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

$$\begin{aligned} Aq &= c_1 Av_1 + c_2 Av_2 + \cdots + c_n Av_n \\ &= c_1 \lambda_1 v_1 + c_2 \lambda_2 v_2 + \cdots + c_n \lambda_n v_n \end{aligned}$$

$$\begin{aligned} A^2 q &= c_1 A^2 v_1 + c_2 A^2 v_2 + \cdots + c_n A^2 v_n \\ &= c_1 \lambda_1^2 v_1 + c_2 \lambda_2^2 v_2 + \cdots + c_n \lambda_n^2 v_n \end{aligned}$$

$$A^k q = c_1 \lambda_1^k v_1 + c_2 \lambda_2^k v_2 + \cdots + c_n \lambda_n^k v_n$$

O que aconteceria se  $c_1 = 0$ , na expansão de  $q$  em combinação linear de autovetores de  $A$ ?

$$A^k q = \lambda_1^k \left[ c_1 v_1 + c_2 (\lambda_2/\lambda_1)^k v_2 + \cdots + c_n (\lambda_n/\lambda_1)^k v_n \right]$$

- ▶  $|\lambda_j/\lambda_1|^k \rightarrow 0$ , para  $j = 2, \dots, n$ , quando  $k \rightarrow \infty$
- ▶ Todo múltiplo de um autovetor também é autovetor ( $\lambda_1^k$  é irrelevante)
- ▶  $q_k \equiv A^k q / \lambda_1^k \rightarrow c_1 v_1$
- ▶ A velocidade de convergência depende de  $|\lambda_2/\lambda_1|$

- ▶ Na prática, não conhecemos  $\lambda_1$
- ▶ Dado  $q_0$ , construímos a sequência

$$q_{k+1} = Aq_k / \sigma_{k+1},$$

para algum fator de escala  $\sigma_{k+1}$ .

- ▶ Por exemplo,  $\sigma_{k+1} = \pm \|Aq_k\|_\infty$ , tal que o maior valor de  $q_{k+1}$  seja sempre 1.



$$A = \begin{bmatrix} 4 & 6 \\ 6 & 4 \end{bmatrix}, \quad p(\lambda) = (4 - \lambda)^2 - 36 = 0 \iff \lambda \in \{10, -2\}$$

- ▶  $\hat{q}_1 = A[1, 0]^T = [4, 6]^T$ ,  
 $q_1 = [4/6, 1]^T$
- ▶  $\hat{q}_2 = Aq_1 = [8.6667, 8]^T$ ,  
 $q_2 = [1, 0.9231]^T$
- ▶  $\hat{q}_3 = Aq_2 = [9.5385, 9.6923]^T$ ,  
 $q_3 = [0.9841, 1]^T$
- ▶  $\hat{q}_4 = Aq_3 = [9.9365, 9.9048]^T$ ,  
 $q_4 = [1, 0.9968]^T$
- ▶  $\hat{q}_5 = Aq_4 = [9.9808, 9.9872]^T$ ,  
 $q_5 = [0.9994, 1]^T$
- ▶  $\hat{q}_6 = Aq_5 = [9.9974, 9.9962]^T$ ,  
 $q_6 = [1, 0.9999]^T$

A convergência é linear se houver  $0 < r < 1$  tal que

$$\lim_{k \rightarrow \infty} \frac{\|q_{k+1} - v_1\|}{\|q_k - v_1\|} = r$$

No caso do exemplo, essa razão foi

0.3333, 0.2308, 0.2063, 0.2013, 0.2003, 0.2001

Lembre que  $|\lambda_2/\lambda_1| = 10/2 = 0.2$ .

**Exercício:** Mostre que o método da potência converge linearmente com  $r = |\lambda_2/\lambda_1|$ , se  $|\lambda_2| > |\lambda_3|$ .



Seja  $A \in \mathbb{C}^{n \times n}$ , uma matriz simples, e seja  $q_0 \in \mathbb{C}^n$  dado.

- ▶ Para  $k = 0, 1, 2, \dots$ 
  - ▶  $\tilde{q} \leftarrow Aq_k$
  - ▶  $\sigma_{k+1} \leftarrow \tilde{q}_s$ , onde  $|\tilde{q}_s| = \|\tilde{q}\|_\infty$
  - ▶  $q_{k+1} \leftarrow \tilde{q}/\sigma_{k+1}$

Exercício: Mostre que  $\sigma_k \rightarrow \lambda_1$

O que aconteceria se  $\sigma_{k+1} \rightarrow \tilde{q}_1$ , ou seja, se o vetor  $\tilde{q}$  fosse sempre normalizado para ter sua primeira componente igual a 1?

Seja  $A \in \mathbb{C}^{n \times n}$  uma matriz simples. Suponha que  $A$  seja não-singular.

**Exercício:** Se  $v$  é autovetor de  $A$ , associado ao autovalor  $\lambda$ , então  $v$  também será autovetor de  $A^{-1}$ , associado ao autovalor  $1/\lambda$ .

- ▶ Se  $A$  é simples,  $A^{-1}$  também será.
- ▶ Se  $|\lambda_n^{-1}| > |\lambda_{n-1}^{-1}|$ , então podemos aplicar o método da potência a  $A^{-1}$ .
- ▶ A velocidade de convergência dependerá do quociente  $|\lambda_{n-1}^{-1}/\lambda_n^{-1}| = |\lambda_n/\lambda_{n-1}|$ .



# Exemplo



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

$q_0 = [1, 2, 1]^T$	$\rho$	$k$	$\lambda$
m. potência	—	69	3
	2.70	35	3
	2.90	14	3
	2.99	8	3
m. pot. inversa	2.30	34	2
	1.90	14	2
	0.50	27	1
	0.80	17	1

Critério de parada

$$\|q_k - q_{k-1}\|_\infty \leq 10^{-12}$$





Seja  $q$  uma aproximação para um autovetor de  $A$ . Qual seria a melhor aproximação, no sentido de quadrados mínimos, para o autovalor associado?

$$q\lambda = Aq$$

O sistema normal é

$$q^* q \lambda = q^* A q \Rightarrow \lambda = \frac{q^* A q}{q^* q}$$

Esse é um sistema sobredeterminado, uma vez que a incógnita é  $\lambda$  e há  $n$  equações. Note que  $Aq$  é conhecido e faz o papel do vetor independente em um sistema linear.

Como o vetor  $q \in \mathbb{C}$ , o sistema normal é obtido multiplicando-se ambos os lados da equação por  $q^*$  (transposto conjugado de  $q$ ).

## Exemplo



$$v = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \text{ é autovetor de } A = \begin{bmatrix} 3/2 & 0 & -1/2 \\ 0 & 2 & 0 \\ -1/2 & 0 & 3/2 \end{bmatrix}, \text{ associado a } \lambda = 1$$

Se  $q = [1.0, 0.1, 0.9]^T$ , então

$$\frac{q^* A q}{q^* q} = 1.0082 \approx 1$$



# Iteração de Rayleigh



Seja  $A \in \mathbb{C}^{n \times n}$ , uma matriz simples, e seja  $q_0 \in \mathbb{C}^n$  dado.

▶ Para  $k = 0, 1, 2, \dots$

▶  $\rho \leftarrow (q_k^* A q_k) / (q_k^* q_k)$

▶  $\tilde{q} \leftarrow (A - \rho I)^{-1} q_k$

na prática, resolve-se  $(A - \rho I)\tilde{q} = q_k$

▶  $\sigma_{k+1} \leftarrow \tilde{q}_1$

▶  $q_{k+1} \leftarrow \tilde{q} / \sigma_{k+1}$

Neste algoritmo escolhi outro tipo de normalização. O vetor  $\tilde{q}$  está sendo normalizado sempre pelo valor de sua primeira componente. Isto significa que, dentre todos os autovetores associados a um autovalor, estamos selecionando aquele cuja primeira componente vale 1.

Claro que isso pode dar problema se justamente essa componente fosse nula...



$$\begin{bmatrix} 5 & 0 & -1 \\ 0 & 2 & 0 \\ -1 & 0 & 5 \end{bmatrix} \sim \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{bmatrix}$$

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}, \quad \lambda = 1 \pm i$$

Duas matrizes,  $A, B \in \mathbb{C}^{n \times n}$  são **similares** se existe  $P \in \mathbb{C}^{n \times n}$  tal que

$$B = P^{-1}AP.$$

## Teorema

Matrizes similares têm os mesmos autovalores.

$$\det(B - \lambda I) = \det(P^{-1}AP - \lambda P^{-1}P) = \det(P^{-1}) \det(A - \lambda I) \det(P) = \det(A - \lambda I)$$

**Exercício:** Mostre que se  $v$  é autovetor de  $A$ , então  $P^{-1}v$  é autovetor de  $B$ .

Handwritten notes area with a red vertical margin line on the left and horizontal ruling lines.

# Multiplicidade algébrica $\times$ geométrica



Seja  $\lambda$  autovalor de  $A$ .

Sua **multiplicidade algébrica** é sua multiplicidade como raiz de  $p_A(\lambda) = \det(A - \lambda I)$ .

Sua **multiplicidade geométrica** é a dimensão do autoespaço

$$\mathcal{L}_\lambda = \{v \in \mathbb{C}^n \mid Av = \lambda v\}$$

$$\begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}, \quad p(\lambda) = (2 - \lambda)^3, \quad v = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \text{ é o único autovetor}$$

A large area of the page is filled with horizontal lines, resembling a notebook page, intended for student notes or exercises.

## Teorema

Uma matriz é simples se e somente se é similar a uma matriz diagonal.

Seja  $V = [v_1, v_2, \dots, v_n]$  uma matriz com autovetores como coluna, então

$$V^{-1}AV = V^{-1}[Av_1, Av_2, \dots, Av_n] = V^{-1}[\lambda_1 v_1, \lambda_2 v_2, \dots, \lambda_n v_n] = V^{-1}VD = D,$$

onde  $D$  é matriz com os autovalores na diagonal.

Se  $A$  é hermitiana ( $A^* = A$ ), então seus autovetores são todos ortogonais entre si. Neste caso, tomando autovetores de norma 1, a matriz  $V = [v_1, v_2, \dots, v_n]$  é **unitária** ( $V^* V = I$ ).

Se  $A$  é real e simétrica, então seus autovetores são todos ortogonais entre si. Neste caso, tomando autovetores de norma 1, a matriz  $V = [v_1, v_2, \dots, v_n]$  é **ortogonal** ( $V^T V = I$ ).

## Teorema espectral

Seja  $A \in \mathbb{R}^{n \times n}$  uma matriz simétrica. Então existe  $V$  uma matriz ortogonal e  $D$  uma matriz diagonal, tais que  $D = V^T A V$ .

$A$  e  $D$  são ortogonalmente similares.



# Decomposição de Schur



Para toda  $A \in \mathbb{R}^{n \times n}$  é possível encontrar  $V$ , ortogonal e real, tal que  $V^T A V$  é diagonal?

## Teorema de Schur Real

Se  $A \in \mathbb{R}^{n \times n}$ , então existem  $U \in \mathbb{R}^{n \times n}$ , ortogonal, e  $T \in \mathbb{R}^{n \times n}$ , triangular superior por blocos, tais que  $U^T A U = T$ . Especificamente,

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1n} \\ & T_{22} & & \vdots \\ & & \ddots & \vdots \\ & & & T_{nn} \end{bmatrix},$$

onde  $T_{ii} \in \mathbb{R}^{1 \times 1}$  ou  $\mathbb{R}^{2 \times 2}$ . Cada bloco  $1 \times 1$  tem um autovalor real de  $A$  e cada bloco  $2 \times 2$  tem como autovalor um par complexo conjugado de autovalores de  $A$ .

Handwritten notes area with a vertical red margin line on the left and horizontal grey lines for writing.

$$U^T \begin{bmatrix} 6 & 3 & 1 & 4 \\ 2 & 2 & 2 & 5 \\ 1 & 5 & 2 & 4 \\ 7 & 5 & 4 & 6 \end{bmatrix} U = \begin{bmatrix} 15.497 & -1.522 & 0.932 & -2.368 \\ 0 & 3.279 & 1.882 & 0.963 \\ 0 & 0 & -1.388 & 1.433 \\ 0 & 0 & -0.963 & -1.388 \end{bmatrix}$$

$$\Lambda(A) = \{15.497, 3.279, -1.388 \pm 1.174i\}$$

- ▶ Seria ótimo ter um algoritmo que, através de transformações de similaridade, convertesse a matriz em uma matriz triangular.
- ▶ Este algoritmo deveria forçosamente ser iterativo (e não finito). Por quê?
- ▶ Existe algoritmo finito que converte a matriz, através de transformações de similaridade, a uma **matriz Hessenberg superior**.

$$\begin{bmatrix} * & * & * & \cdots & * & * \\ * & * & * & \cdots & * & * \\ & * & * & \cdots & * & * \\ & & \ddots & & \vdots & \vdots \\ & & & \ddots & & \vdots \\ & & & & * & * \end{bmatrix}$$