Revisão de Álgebra Linear

Abel Soares Siqueira

10/08/2018

Machine Learning

Vetores

- Elementos do $\mathbb{R}^n$;
- Características;

Vetores

$$v = (v_1, \ldots, v_n)$$

$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$$

Vetores

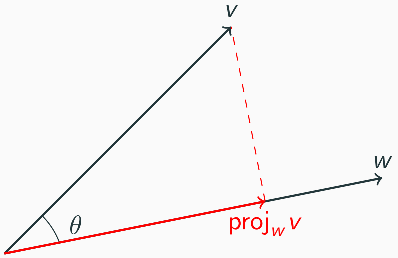
$$v^T w = v_1 w_1 + \cdots + v_n w_n = \sum_{i=1}^{n} v_i w_i$$

$$\|v\|_2 = \left(\sum_{i=1}^{n} v_i^2 \right)^{1/2} = \sqrt{v^T v}$$

$$\|v\|_1 = \sum_{i=1}^{n} |v_i|$$

$$\|v\|_\infty = \max_{i=1,\ldots,n} |v_i|$$

Vetores



$$\cos\theta = \frac{v^T w}{\|v\| \, \|w\|}.$$

$$\mathrm{proj}_w v = \frac{v^T w}{w^T w} w.$$

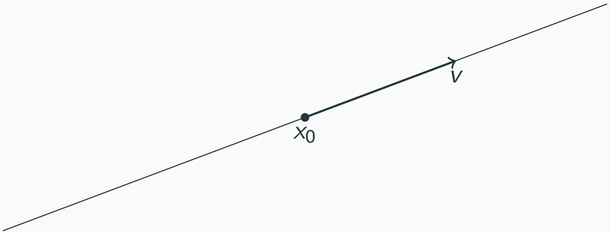
Retas

$$\mathbb{R}^2: ax + by = c, \quad a, b, c \in \mathbb{R}$$

$$\mathbb{R}^3: \frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}, \quad x_0, y_0, z_0 \in \mathbb{R}, a, b, c \in \mathbb{R}^*$$

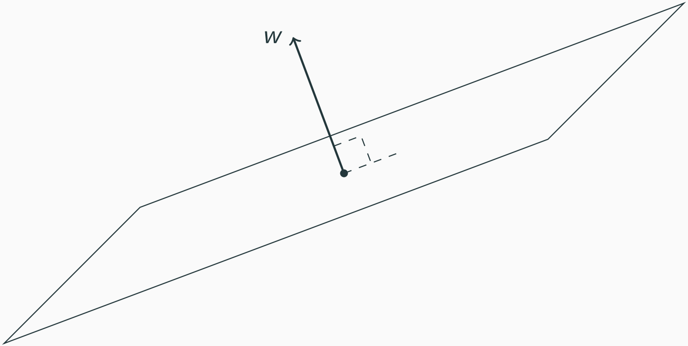
Parametrizado:

$$x = x_0 + tv, \quad v, x_0 \in \mathbb{R}^n, t \in \mathbb{R}.$$



Planos

$$\mathbb{R}^2 : ax + by = c, \quad a, b, c \in \mathbb{R}$$
$$\mathbb{R}^3 : ax + by + cz = d, \quad a, b, c, d \in \mathbb{R}$$
$$\mathbb{R}^n : w^T x + b = 0, \quad w \in \mathbb{R}^n, b \in \mathbb{R}.$$



Matrizes e sistemas lineares

$$\left\{\begin{array}{rcl} a_{11}x_1 + \cdots + a_{1n}x_n & = & b_1 \\ \vdots \quad \ldots \quad \vdots & = & \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n & = & b_m \end{array}\right.$$

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}$$

$$Ax = b, \qquad A \in \mathbb{R}^{m\times n},\ x \in \mathbb{R}^n,\ b \in \mathbb{R}^m.$$

Matrizes e sistemas lineares

- SL aparecem em toda parte;
- Maior parte das linguagens tem um jeito fácil de resolver SL;
- SL especiais precisam de métodos especiais;

Espaço vetorial

- V é um EV, então $v \in V$ é um vetor;
- Vamos usar só $V = \mathbb{R}^n$;
- $\langle v_1, \ldots, v_k \rangle = \{\alpha_1 v_1 + \cdots + \alpha_k v_k : \alpha_i \in \mathbb{R}\}$;
- $\{v_1, \ldots, v_k\} \subset \mathbb{R}^n$ é LI se

  $$\alpha_1 v_1 + \cdots + \alpha_k v_k = 0 \Rightarrow \alpha_i = 0,$$

  que é um sistema

  $$\underbrace{[v_1 \cdots v_k]}_{n \times k} \alpha = 0;$$

- v e w são Linearmente Independentes se não são paralelos;

Subespaços vetoriais

- Espaço vetorial $\mathbb{R}^n$;
- Subspaço vetorial $E \subset \mathbb{R}^n$;
- $\alpha, \beta \in \mathbb{R}$ e $v, w \in E$, temos $\alpha v + \beta w \in E$;
- $\{0\}$ e $\mathbb{R}^n$ são subespaços do $\mathbb{R}^n$;

Subespaços vetoriais

$$\langle v_1, \ldots, v_k \rangle = \{\alpha_1 v_1 + \cdots + \alpha_k v_k : \alpha_i \in \mathbb{R}\}.$$

$$\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\rangle = \{(x, y) \in \mathbb{R}^2 : y = x\}$$

$$\left\langle \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right\rangle = \{(x, y) \in \mathbb{R}^2 : y = x\}$$

Subespaços vetoriais

- $\langle v_1, \ldots, v_k \rangle$ é um subespaço;
- Se E é subespaço de $\mathbb{R}^n$ é gerado por $v_1, \ldots, v_k$;
- Se $\beta = \{v_1, \ldots, v_k\}$ é LI, e $E = \langle v_1, \ldots, v_k \rangle$, então β é uma base para E, e $\dim(E) = k$;

Matrizes e subespaços

- Sistema homogêneo: $Ax = 0$;
- Núcleo de A: $\mathrm{Nu}(A) = \{x : Ax = 0\} \subset \mathbb{R}^n$;
- Imagem de A: $\mathrm{Im}(A) = \{Ax,\ x \in \mathbb{R}^n\} \subset \mathbb{R}^m$;
- Posto de A é $p = \dim(\mathrm{Im}(A))$;
- $Ax = b$ quer dizer que $b \in \mathrm{Im}(A)$;

Eliminação Gaussiana

$$[A \mid b] : \text{Matriz aumentada}$$

$$L_i : \text{linha } i$$

- $L_i \leftarrow \alpha L_i$, $\alpha \neq 0$;
- $L_i \leftrightarrow L_j$;
- $L_i \leftarrow L_i - \alpha L_j$.

Eliminação Gaussiana - Forma escada

- Cada linha não nula tem mais zeros à esquerda que a linha de cima;
- As linhas nulas estão todas abaixo;
- O primeiro elemento não nulo é chamado de pivô;
- Se alguma linha é nula, exceto o valor mais à direita (b_j), o sistema é **impossível**, i.e., **não tem solução**;
- O número de linhas não nulas é o **posto** p da matriz;
- Se $p = n$, o sistema é **possível e determinado**, i.e., **existe uma única solução**;
- Se o posto for menor que n, o sistema é **possível e indeterminado**, i.e., **existem infinitas soluções**.

Eliminação Gaussiana - Forma escada reduzida

- Pivôs 1, zerar acima do pivô, e mudar a ordem das colunas de A obtendo algo tipo

$$\begin{bmatrix} 1 & 0 & \cdots & 0 & \times & \cdots & \times & c_1 \\ 0 & 1 & \cdots & 0 & \times & \cdots & \times & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & \times & \cdots & \times & c_p \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 \end{bmatrix},$$

ou $\begin{bmatrix} I & B & c \\ 0 & 0 & 0 \end{bmatrix}$.

- Quanto o posto é menor que n, existem $n - p$ liberdades;
- Um (hiper-)plano é um sistema linear com 1 equação;
- Uma reta é um sistema linear com $n - 1$ equações;

Projeções

- Dado um conjunto $\Omega \subset \mathbb{R}^n$, a projeção de y em Ω é um vetor $z \in \Omega$, se existir, tal que

$$\|z - y\| \leq \|x - y\|, \qquad \forall x \in \Omega.$$

- Isso pode ser visto como encontrar o ponto em Ω mais próximo de y.
- Se o conjunto Ω for um espaço vetorial, a projeção sempre existe, e é única; Nesse caso também vale

$$z - y \perp z - x, \qquad \forall x \in \Omega.$$

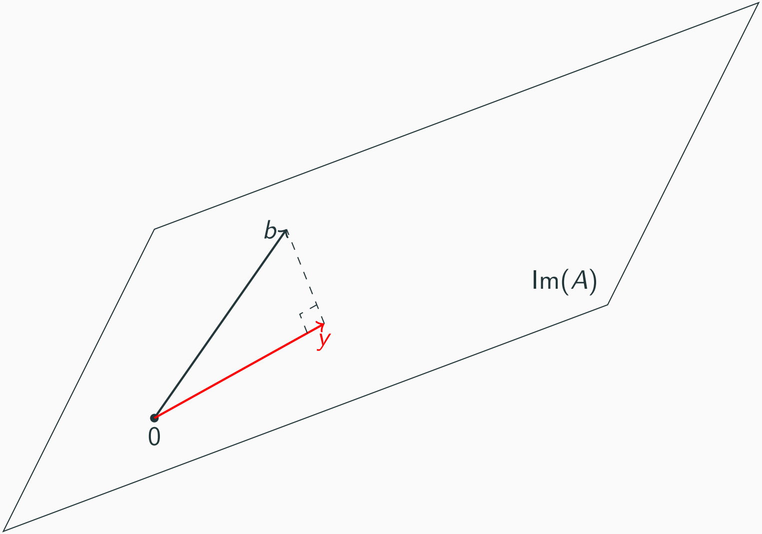
- Se Ω é EV e $\{v_1, \ldots, v_k\}$ é uma base ortogonal de Ω, então

$$z = \frac{y^T v_1}{v_1^T v_1} + \cdots + \frac{y^T v_k}{v_k^T v_k}.$$

Quadrados Mínimos

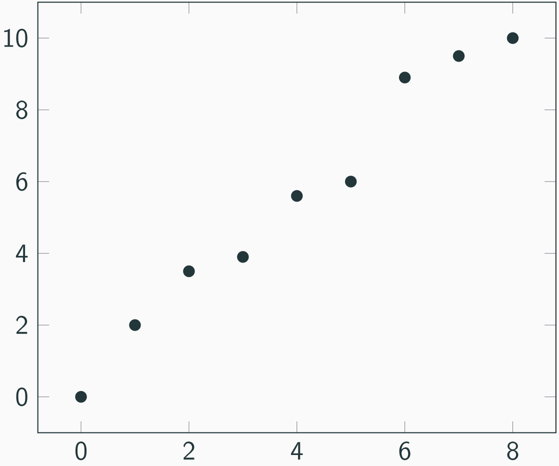
- Considere $A \in \mathbb{R}^{m \times n}$, com $m > n$, e $b \in \mathbb{R}^m$.
- O sistema $Ax = b$ pode não ter solução mesmo que as colunas de A sejam LI.
- Uma maneira de obter uma solução aproximada é considerar o problema $A^T(Ax - b) = 0$, que sempre terá solução.
- Essa solução pode ser obtida considerando o problema de encontrar x tal que o erro $\|Ax - b\|$ é mínimo.
- Esse problema pode ser interpretado como encontrar y mais próximo de b tal que $y = Ax$ para algum x.
- Em outras palavras, y é a projeção de b em $\mathrm{Im}(A)$.

Quadrados Mínimos



Quadrados Mínimos

- Conjunto de dados $\{(x_i, y_i) : i = 1, \ldots, m\}$, onde $y_i \approx \beta_0 + \beta_1 x_i$.



Quadrados Mínimos

$$\underbrace{\begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}}_{\beta} \approx \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}}_{y}$$

$$A\beta \approx y \qquad \Rightarrow \qquad \underbrace{A^T A}_{M} \beta = \underbrace{A^T y}_{c}$$

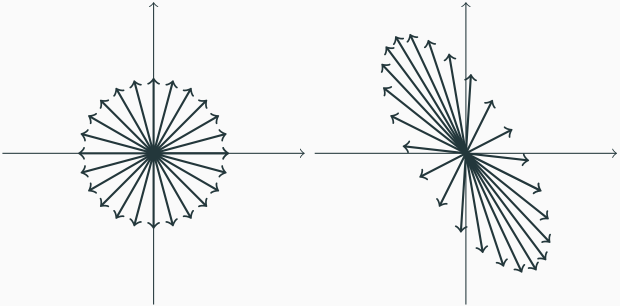
$$M = \begin{bmatrix} m & \sum_{i=1}^m x_i \\ \sum_{i=1}^m x_i & \sum_{i=1}^m x_i^2 \end{bmatrix} \qquad c = \begin{bmatrix} \sum_{i=1}^m y_i \\ \sum_{i=1}^m x_i y_i \end{bmatrix}$$

Autovalores e autovetores

- Se existem λ e $v \neq 0$ tais que $Av = \lambda v$ então λ é dito um autovalor e v um autovetor associado à λ.
- Se λ é autovalor, existem infinitos autovetores. Escolhemos com norma 1 em geral.
- Se A é simétrica, existe uma base $\{v_1, \ldots, v_n\}$ ortonormal de autovetores.

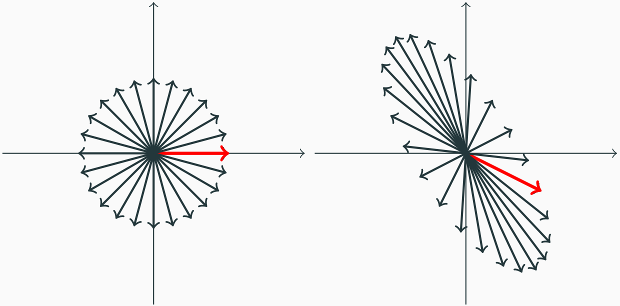
Autovalores e autovetores

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



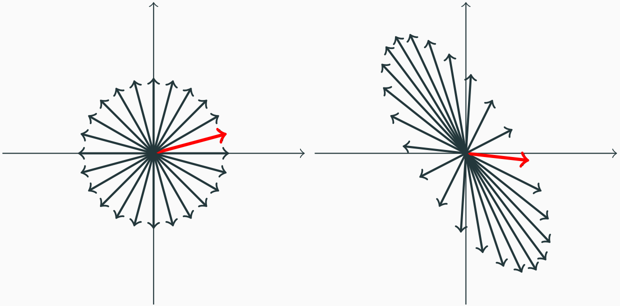
Autovalores e autovetores

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



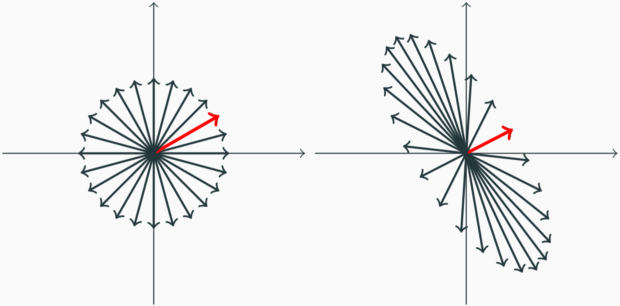
Autovalores e autovetores

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



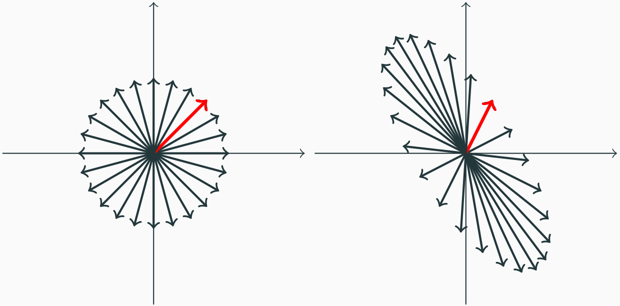
Autovalores e autovetores

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



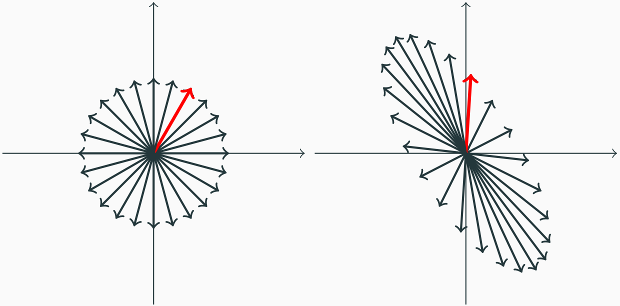
Autovalores e autovetores

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



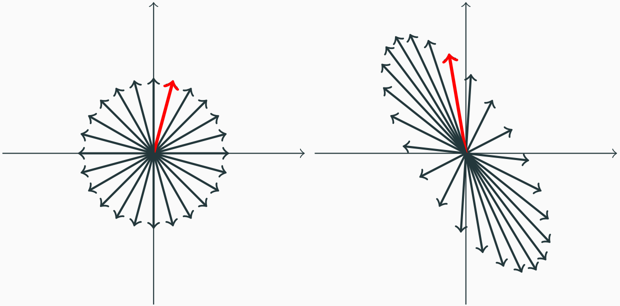
Autovalores e autovetores

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



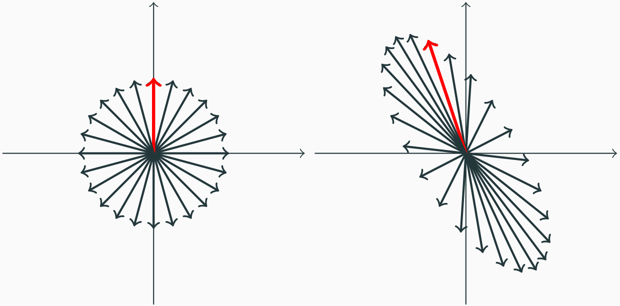
Autovalores e autovetores

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



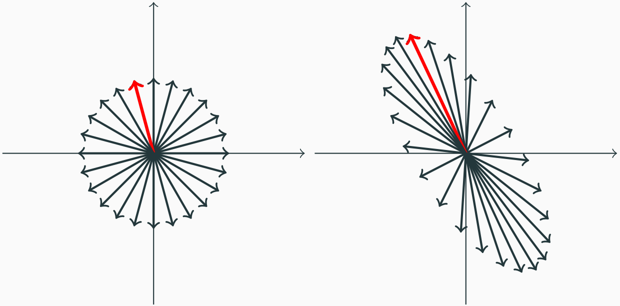
Autovalores e autovetores

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



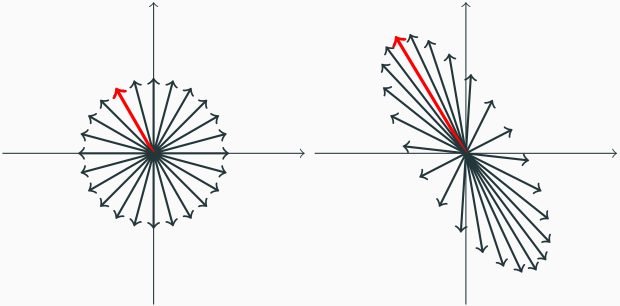
Autovalores e autovetores

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



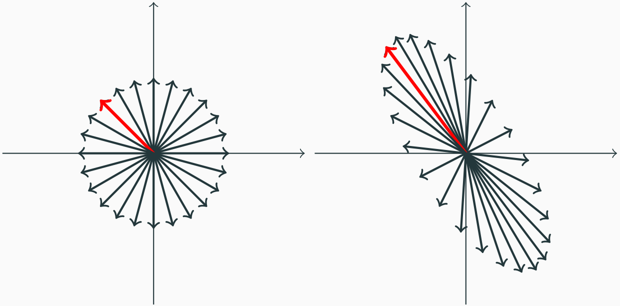
Autovalores e autovetores

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



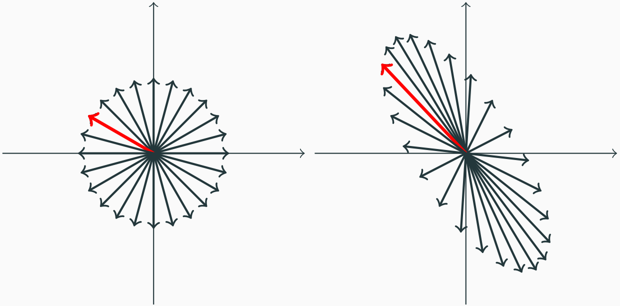
Autovalores e autovetores

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



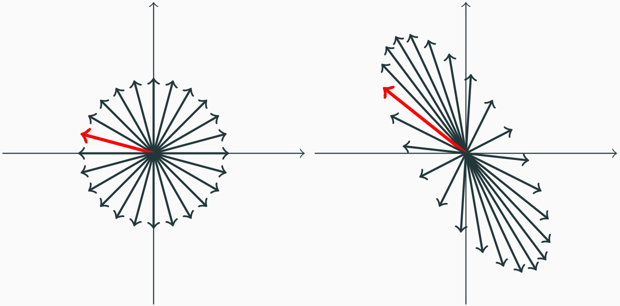
Autovalores e autovetores

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



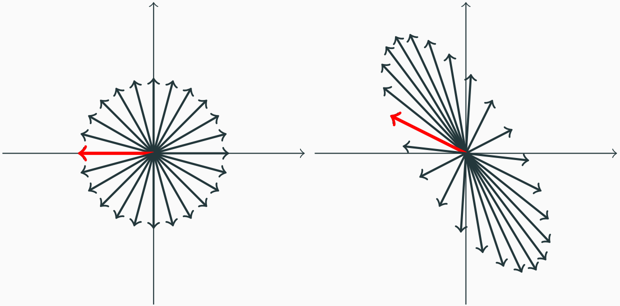
Autovalores e autovetores

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



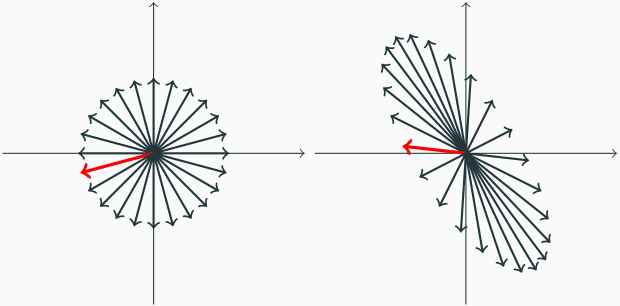
Autovalores e autovetores

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



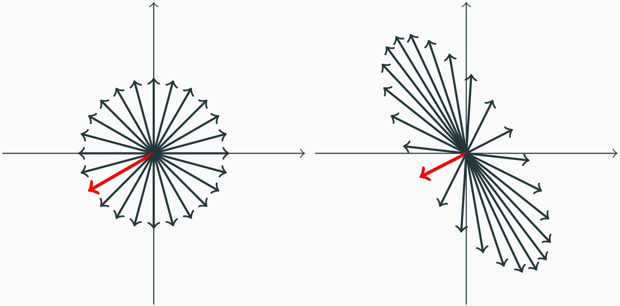
Autovalores e autovetores

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



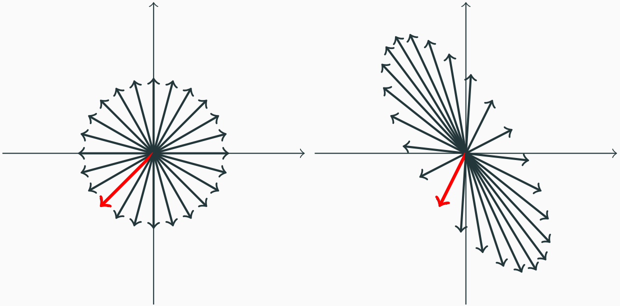
Autovalores e autovetores

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



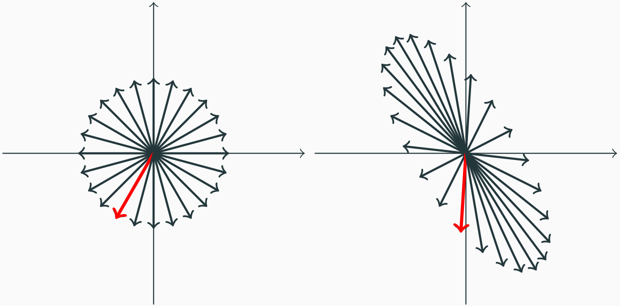
Autovalores e autovetores

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



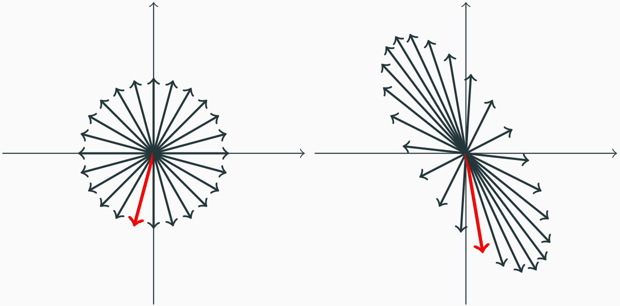
Autovalores e autovetores

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



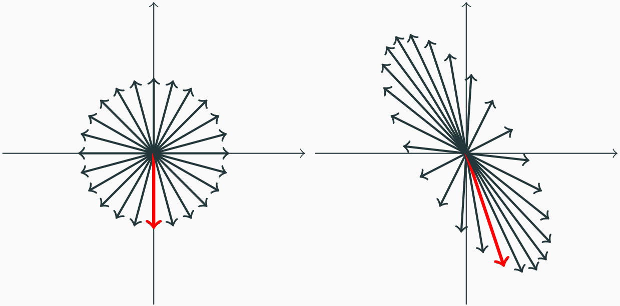
Autovalores e autovetores

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



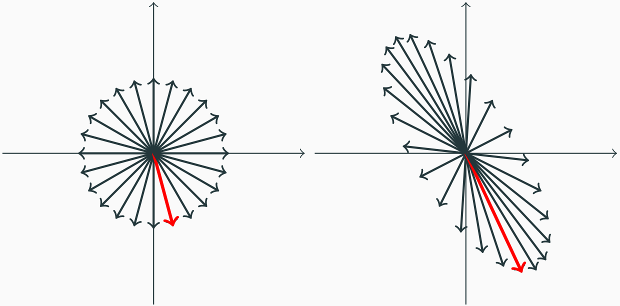
Autovalores e autovetores

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



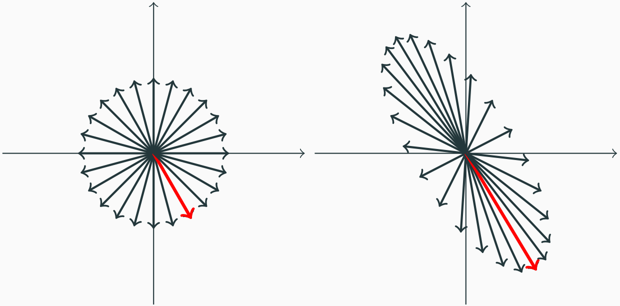
Autovalores e autovetores

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



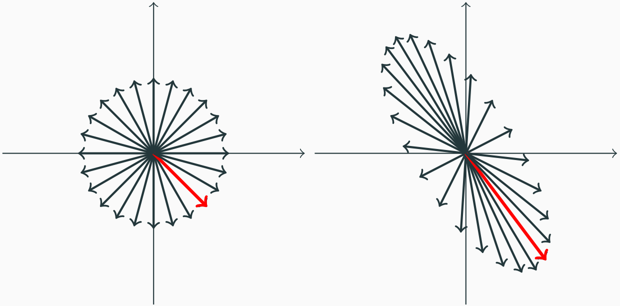
Autovalores e autovetores

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



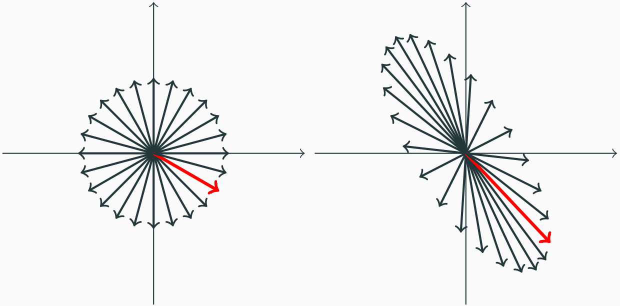
Autovalores e autovetores

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



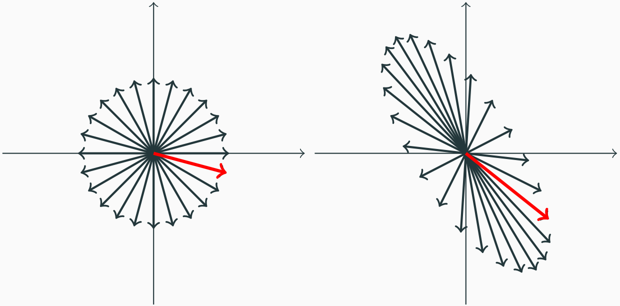
Autovalores e autovetores

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



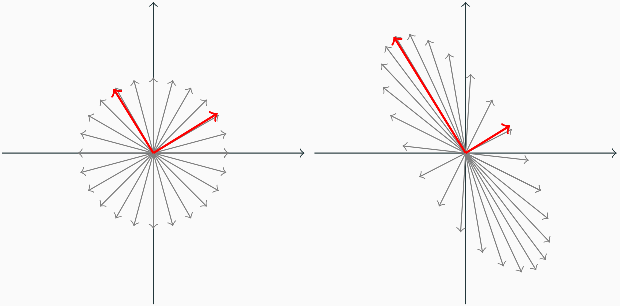
Autovalores e autovetores

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



Autovalores e autovetores

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



FIM