AULA 8

DECOMPOSIÇÃO EM VALORES SINGULARES

TEOREMA ESPECTRAL

Seja A uma matriz $n \times n$ simétrica. Então

(a) A tem n ~~autovalores~~ autovalores reais (contados com multiplicidade)

(b) Se λ é um autovalor de A com multiplicidade k, então o auto-espaço associado tem dimensão k.

(c) ~~Autovalores~~ ~~Autovetores~~ Autovetores correspondentes a autovalores diferentes são ortogonais

(d) A é ortogonalmente diagonalizável:

$$A = Q^{\cancel{T}} D Q^{T}$$

↳ ortogonal

↳ diagonal

→ melhor QDQ^T do que $Q^T D Q$

PROPOSIÇÃO 6 Seja A uma matriz $\boxed{\text{simétrica}}$ $n \times n$.

Defina

$$m = \min_{\|x\|=1} x^T A x \quad \text{e} \quad M = \max_{\|x\|=1} x^T A x$$

~~Então m é igual ao menor autovalor de A e o mínimo ocorre quando x é um autovetor unitário correspondente~~

~~M é igual ao maior autovalor de A e o máximo ocorre quando x for~~

Então: m é o menor autovalor de A e o mínimo ocorre quando x é um autovetor unitário associado,

M é o maior autovalor de A e o máximo ocorre quando x é um autovetor unitário associado.

Demonstração: Seja $\{v_1, \dots, v_n\}$ uma base ortonormal de autovetores associados aos autovalores $\lambda_1, \dots, \lambda_n$. Sem perda de generalidade, vamos supor que

$$\lambda_1 \geq \dots \geq \lambda_n .$$

~~Note que~~ $\cancel{\min_{\|x\|=1} \|Ax\|}$ ~~ocorre para o x ótimo~~

~~do problema de otimização~~ $\cancel{\min_{\|x\|^2=1} \|Ax\|^2}$~~. Agora.~~

~~$\|x\|^2 =$ Se~~ $\cancel{x = \alpha_1 v_1 + \dots + \alpha_n v_n}$~~, então~~

$$\cancel{\|x\|^2 = \alpha_1^2 + \dots + \alpha_n^2} \text{ e}$$

$$\cancel{\|Ax\|^2 = \lambda_1^2\alpha_1^2 + \dots + \lambda_n^2\alpha_n^2} .$$

~~Note que~~ $\cancel{\|x\| = 1 \iff \alpha_1^2 + \dots + \alpha_n^2 = 1}$.

Note que $\|x\|^2 = \alpha_1^2 + \dots + \alpha_n^2$. Logo

$$\|x\| = 1 \iff \alpha_1^2 + \dots + \alpha_n^2 = 1 .$$

Note também que
$$x^T A x = \left(\sum \alpha_i v_i\right)^T A \left(\sum \alpha_j v_j\right)$$
$$= \left(\sum \alpha_i v_i\right)^T \left(\sum \alpha_j \lambda_j v_j\right)$$
$$= \sum_i \sum_j \alpha_i \alpha_j \lambda_j v_i^T v_j$$

$$= \sum_i \lambda_i \alpha_i^2$$

Então $\max_{\|x\|=1} x^T A x$ é equivalente a

$$\max \sum \lambda_i \alpha_i^2$$
$$\text{s.a. } \alpha_1^2 + \dots + \alpha_n^2 = 1.$$

Sem perda de generalidade, suponha que $\lambda_1 \geq \dots \geq \lambda_n$.

Agora.

$$\left\{\begin{array}{l} \lambda_1 \alpha_1^2 = \lambda_1 \alpha_1^2 \\ \lambda_2 \alpha_2^2 \leq \lambda_1 \alpha_2^2 \\ \vdots \\ \lambda_n \alpha_n^2 \leq \lambda_1 \alpha_n^2 \end{array}\right. \Rightarrow \sum \lambda_i \alpha_i^2 \leq \lambda_1 \left(\sum \alpha_i^2\right) = \lambda_1$$

Assim, $M \leq \lambda_1$. Tomando $\alpha_2 = \dots = \alpha_n = 0$ e $\alpha_1 = 1$

vemos que o valor máximo $M = \lambda_1$ é atingido quando $x = 1v_1 + \dots + 0v_n = v_1$.

Um argumento análogo mostra que $m = \lambda_n$ e o mínimo ocorre quando $x = v_n$.

Se $(\alpha_1^*, \dots, \alpha_n^*)$ é outro ponto que maximiza $\sum \lambda_i \alpha_i^2$, com $(\alpha_1^*)^2 + \dots + (\alpha_n^*)^2 = 1$, então

PASSO IMPORTANTE

$$\lambda_1 = \lambda_1(\alpha_1^*)^2 + \cdots + \lambda_n(\alpha_n^*)^2$$

$$\|$$

$$\lambda_1\underbrace{\left((\alpha_1^*)^2 + \cdots + (\alpha_n^*)^2\right)}_{\substack{\| \\ 1}} \Rightarrow \underbrace{(\lambda_1-\lambda_2)}_{\geq 0}(\alpha_2^*)^2 + \cdots + \underbrace{(\lambda_1-\lambda_n)}_{\geq 0}(\alpha_n^*)^2 = 0$$

$$\cancel{(\lambda_1-\lambda_n)(\alpha_1^*)^2 + \cdots + (\lambda_{n-1}-\lambda_n)(\alpha_{n-1}^*)^2 = 0}$$

$$\Downarrow$$

para os autovalores λ_i diferentes de λ_1,
$\alpha_i^* = 0$.

Um argumento análogo mostra que um ponto de ~~máximo~~ mínimo de $x^T A x$ sujeito à $\|x\|^2 = 1$ deve ser um autovetor associado a λ_n. ∎

TEOREMA 7 Sejam A, v_1 e λ_1 como no enunciado e na demonstração do teorema precedente então o máximo de

$$x^T A x$$

sujeito às restrições

$$x^T x = 1 \text{ e } x^T v_1 = 0$$

é o segundo maior autovalor λ_2, e o máximo é atingido quando x é um autovetor unitário correspondente a λ_2.

Demonstração. Como antes, o problema é equivalente a

$$\max \sum \lambda_i \alpha_i^2$$

sujeito a

$$\alpha_1^2 + \dots + \alpha_n^2 = 1.$$

$$\alpha_1 = 0$$

$x = \sum_{i=1}^{n} \alpha_i \theta_i$

~~Agora~~ pois como $x^T v_1 = 0$, então $\alpha_1 = 0$. Então, como antes,

$A = Q^T D Q$

$$\lambda_2 \alpha_2^2 = \lambda_2 \alpha_2^2$$
$$\lambda_2 \alpha_3^2 \geq \lambda_3 \alpha_3^2$$
$$\vdots$$
$$\lambda_2 \alpha_n^2 \geq \lambda_n \alpha_n^2$$

$$\Rightarrow \lambda_2 \left(\sum_{j=2}^{n} \alpha_j^2\right) \geq \sum_{j=2}^{n} (\lambda_j \alpha_j)^2$$

$$\Downarrow$$

$$\lambda_2 \geq \sum_{j=2}^{n} (\lambda_j \alpha_j)^2.$$

O máximo ocorre quando $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = 0, \dots, \alpha_n = 0$. ∎

<u>TEOREMA 8</u> Seja A uma matriz $n \times n$ simétrica com uma diagonalização ortogonal $A = Q^T D Q$, onde os elementos de D estão ordenados de forma que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. As colunas de Q^T são autovetores unitários associados aos ~~respectivos~~ respectivos autovalores. Então, para $k = 2, \dots, n$,

o valor máximo de x^TAx sujeito às restrições

$$x^Tx = 1,\ x^Tv_1 = 0,\ \dots,\ x^Tv_{k-1} = 0$$

é o autovalor λ_k e este máximo ocorre quando $x = v_k$.

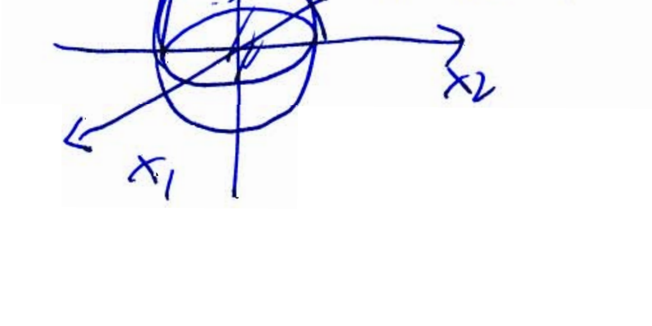
$$\max_{\|x\|=1} \|Ax\| = \max_{1 \le i \le n} \{|\lambda_i|\} = |\lambda_k|$$

simétrica e o máximo ocorre quando x é autovetor unitário associado ao autovalor λ_k

ISTO FUNCIONA PARA MATRIZES QUADRADAS

Esta descrição tem uma contrapartida para matrizes retangulares que conduzirá à decomposição em valores singulares.

EXEMPLO 1 $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$



A transformação linear $x \mapsto T(x) = Ax$ leva a esfera unitária na elipse indicada na figura. Encontre um vetor unitário x ~~[illegible]~~ para o qual $\|Ax\|$ é máximo. Calcule este comprimento máximo.

Solução: o ponto ótimo que maximiza $\|Ax\|$ é o mesmo que maximiza $\|Ax\|^2$, que é mais fácil de se estudar. Agora

$$\|Ax\|^2 = \langle Ax, Ax\rangle = x^T A^T A x$$

A matriz A^TA é simétrica. Assim o problema de maximizar $x^T(A^TA)x$ sujeito a $\|x\|=1$ pode ser resolvido através do teorema dado no início da aula:

$$A^TA = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$$

Os autovalores de A^TA são $\lambda_1 = 360$, $\lambda_2 = 90$ e $\lambda_3 = 0$. Os autovetores correspondentes são

$$v_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} \quad v_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad v_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

O valor máximo de $\|Ax\|^2$ é 360 que é obtido fazendo-se ~~[illegible]~~ $x = v_1$. O vetor

$$Av_1 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 18 \\ 6 \end{bmatrix}$$ é um dos

pontos da elipse mais afastados da origem.

OS VALORES SINGULARES DE UMA MATRIZ $m \times n$

Seja $A_{m\times n}$ uma matriz real.

A^TA é simétrica, logo ortogonalmente diagonalizável.

Seja $\{v_1, \dots, v_n\}$ uma base ortonormal formada pelos auto~~valores~~vetores de A^TA. Assim:

$$
\begin{aligned}
0 \leq \|Av_i\|^2 = v_i^T A^T A v_i &= v_i^T \lambda_i v_i \\
&= \lambda_i v_i^T v_i \\
&= \lambda_i \|v_i\|^2 \\
&= \lambda_i
\end{aligned}
$$

Isto mostra que todos os autovalores de A^TA são não negativos. Reordenando se necessário podemos assumir que $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$.

Os valores singulares de A são as raízes quadradas dos autovalores de A^TA:

$$\sigma_i = \sqrt{\lambda_i} \quad \left(= \sqrt{\|Av_i\|^2} = \|Av_i\| \right)$$

Os valores singulares são, portanto, os ~~compri~~ comprimentos dos vetores $Av_1, \dots, Av_n$

EXEMPLO 2

Seja A a matriz do exemplo anterior. Os autovalores de A^TA são 360, 90 e 0. Assim, os valores singulares de A são

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10} \text{ e } \sigma_3 = 0.$$

~~O primeiro valor singular de A é o máximo de ||Ax|| sobre a esfera unitária ||x||=1.~~

~~O segundo valor singular, pelo teorema enunciado no início da aula~~, é o ~~máximo~~ ~~de~~

$$\begin{cases} \sigma_1 = \max\limits_{\|x\|=1} \|Ax\| & \text{atingido em } x = v_1, \\[2ex] \sigma_2 = \max\limits_{\substack{\|x\|=1 \\ x \in \{v_1\}^\perp}} \|Ax\| & \text{atingido em } x = v_2. \end{cases}$$

Para o v_2 do exemplo anterior,

$$\cancel{Av_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & & \end{bmatrix}}$$

$$Av_2 = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3 \\ -9 \end{bmatrix}$$

Note então que os dois primeiros valores singulares determinam os comprimentos dos dois semi-eixos da elipse. ∎

TEOREMA 9 Se A é uma matriz $m \times n$ que possui r valores singulares não nulos:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$$
$$\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0,$$

então o posto de A é r.

Demonstração seja $\{v_1, \dots, v_n\}$ base ortonormal de autovetores de A^TA cujos autovalores satisfazem $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Assim,

$$\text{para } i \neq j, \ \langle Av_i, Av_j \rangle = v_j^T \underline{A^TAv_i}$$
$$= v_j^T \lambda_i v_i$$
$$= \lambda_i v_j^T v_i = 0.$$

Assim, $\{Av_1, \dots, Av_n\}$ é um conjunto ortogonal. (que gera $\mathcal{C}(A) = \mathrm{Im}(A)$)

Como ~~$\lambda_i = \|Av_i\|^2$~~ Com $\sigma_i = \|Av_i\|$, segue-se que $Av_{r+1} = Av_{r+2} = \dots = Av_n = 0$. Logo, o conjunto de vetores não nulos $\{Av_1, \dots, Av_r\}$ é LI e gera $\mathcal{C}(A)$.

Logo $\mathrm{rank}(A) = r$. ∎

A DECOMPOSIÇÃO EM VALORES SINGULARES

$$\Sigma_{m\times n} = \begin{array}{c} \left[\begin{array}{cc} D_{r\times r} & 0 \\ 0 & \underbrace{0}_{n-r \text{ colunas}} \end{array}\right] \end{array} \begin{array}{l} \\ \} \; m-r \text{ linhas} \end{array}$$

<u>TEOREMA 10</u> Seja A uma matriz $m\times n$ com posto r.
Existe uma matriz

$$\Sigma_{m\times n} = \begin{array}{c} \left[\begin{array}{cc} D_{r\times r} & 0 \\ 0 & \underbrace{0}_{n-r} \end{array}\right] \end{array} \begin{array}{l} \\ \} \; m-r \end{array}$$

onde os elementos da diagonal em D são os primeiros r ~~valores~~ valores singulares de A:

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0,$$

e existem matrizes ortogonais $U_{m\times m}$ e $V_{n\times n}$ tais que

$$A = U \Sigma V^T$$

- A decomposição não é única (colunas de U) → pois depende como a base $\{u_1, \dots, u_r\}$ é estendida ↳ ver (*) na página 12, mas se $A = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ com $\tilde{U}$ e $\tilde{V}$ ortogonais, então as entradas de $\tilde{\Sigma}$ são obrigatoriamente os valores singulares de A que estão na diagonal de $\tilde{\Sigma}$.

EXERCÍCIO: (H) $Q_{m\times m}$ ortogonal

$\Leftrightarrow$ A e QA têm os mesmos valores singulares.

Prova: $(QA)^T(QA)$
$$= A^TQ^TQA = A.$$

Prova da observação:

$$A = \tilde{U}\tilde{\Sigma}\tilde{V}^T$$
$$A^TA = (\tilde{U}\tilde{\Sigma}\tilde{V}^T)^T(\tilde{U}\tilde{\Sigma}\tilde{V})$$
$$= \tilde{V}^T\tilde{\Sigma}\tilde{U}^T\tilde{U}\tilde{\Sigma}\tilde{V}$$
$$= \tilde{V}^T\tilde{\Sigma}\tilde{\Sigma}\tilde{V}$$

Os autovalores de A^TA são os quadrados dos elementos da diagonal ~~de $\tilde{\Sigma}$~~ principal de $\tilde{\Sigma}$

$$A = Q^T D Q$$
$$\Downarrow$$
$$Q^T A = Q^T D$$
$$\Downarrow$$
$$(A^TQ)^T$$
$$AQ^T = Q^T D$$
$$A\begin{bmatrix} q_i \end{bmatrix} = \begin{bmatrix} \, \end{bmatrix} d\ldots$$

As colunas de U são denominadas de vetores singulares à ~~esq~~ esquerda e as colunas de V de vetores singulares a direita.

Demonstração. Como na demonstração do teorema 9, sejam λ_i e v_i, com

$$\sigma_i = \sqrt{\lambda_i} = \|Av_i\| > 0 \quad \text{para } 1 \leq i \leq r$$

e

$$\{Av_1, \ldots, Av_r\} \text{ base ortogonal de } \mathcal{C}(A).$$

Para $1 \leq i \leq r$, defina

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i$$

de modo que

$$Av_i = \sigma_i u_i \quad \text{para } i = 1, \ldots, r$$

Então $\{u_1, \ldots, u_r\}$ é uma base ortonormal de $\mathcal{C}(A)$. Estenda para uma base ortonormal de $\mathbb{R}^m$: $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$. Defina

(*)

$$U = \begin{bmatrix} \boxed{u_1} & \cdots & \boxed{u_m} \end{bmatrix} \text{ e } V = \begin{bmatrix} \boxed{v_1} & \cdots & \boxed{v_n} \end{bmatrix}.$$

As matrizes U e V são ortogonais e

$$AV = [Av_1 \ldots Av_r, 0 \ldots 0] = [\sigma_1 u_1 \ldots \sigma_r u_r 0 \ldots 0]$$

$$= \begin{bmatrix} u_1 & \cdots & u_r & u_{r+1} & \cdots & u_m \end{bmatrix} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \ddots & & 0 \\ & & \sigma_r & \\ & 0 & & 0 \end{bmatrix}}_{\Sigma}$$

Assim, $AV = U\Sigma$, de modo que

$$A = U\Sigma V^T.$$

$\blacksquare$

EXEMPLO 3 Encontre uma decomposição em valores singulares para a matriz $A = \begin{bmatrix} 4 & 11 & 4 \\ 8 & 7 & -2 \end{bmatrix}$

Solução. Do exemplo 1, temos que

$$u_1 = \frac{1}{\sigma_1} Av_1 = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} Av_2 = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

$\{u_1, u_2\}$ é base ortonormal de $\mathbb{R}^2$. Logo não é necessário fazer um completamento da base. Colocando $U = [u_1 \ u_2]$ $V = [v_1 \ v_2 \ v_3]$ e

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \ \Sigma = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

Então

$$A = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

$\blacksquare$

EXEMPLO 4 Calcule uma decomposição SVD de

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

Solução

- $A^TA = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$, com autovalores $\lambda = 18$

  e $\lambda_2 = 0$,

  Autovetores correspondentes:

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad v_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

- $Av_1 = \begin{bmatrix} \frac{2}{\sqrt{2}} \\ -\frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} \end{bmatrix} \qquad \sigma = \|Av_1\| = \sqrt{18} = 3\sqrt{2}$

$$u_1 = \frac{1}{\sigma_1} Av_1 = \frac{1}{3\sqrt{2}} Av_1 = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}$$

Note que $Av_2 = 0$.

• Estender $\{u_1\}$ para base ortonormal de $\mathbb{R}^3$.

$$\langle u, u_1 \rangle = 0$$

$$\Downarrow$$

$$x_1 - 2x_2 + 2x_3 = 0$$

$$\Downarrow$$

$$x_1 = -2x_2 + 2x_3$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -2x_2 + 2x_3 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= x_2 \begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, w_2 + x_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}, w_1$$

• Aplicando o processo de ortonormalização de Gram-Schmidt:

$$u_2 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \\ 0 \end{bmatrix} \quad u_3 = \begin{bmatrix} -2/\sqrt{45} \\ 4/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

• Finalmente, $U = [u_1 \; u_2 \; u_3]$, $V = [v_1 \; v_2]$

e $\Sigma = \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$

Então

$$A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 1/3 & 2/\sqrt{5} & -2/\sqrt{45} \\ -2/3 & 1/\sqrt{5} & 4/\sqrt{45} \\ 2/3 & 0 & 5/\sqrt{45} \end{bmatrix}$$

$$\begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$$

$\blacksquare$