

Corso di Analisi Numerica per Ingegneria

Complementi sulla SVD

ANGELES MARTINEZ CALOMARDO  
Dipartimento di Matematica e Geoscienze  
Università di Trieste  
e-mail: [amartinez@units.it](mailto:amartinez@units.it)  
pagina web: [www.dmi.units.it/~acalomar](http://www.dmi.units.it/~acalomar)

## Norma 2 e norma di Frobenius in termini dei valori singolari

Il seguente teorema mette in relazione  $\|A\|_2$  e  $\|A\|_F$  con i valori singolari di  $A$

### Teorema

Sia  $A \in \mathbb{R}^{m \times n}$  e siano  $\sigma_1, \sigma_2, \dots, \sigma_p$  tutti i valori singolari di  $A$ . Si ha

$$\begin{aligned}\|A\|_2 &= \sigma_1 \\ \|A\|_F &= \sqrt{\sigma_1^2 + \dots + \sigma_p^2}\end{aligned}$$

NOTA: compaiono tutti i valori singolari (anche quelli nulli) per considerare il caso particolare in cui tutti i valori singolari siano nulli ( $r = 0$ ,  $\text{rank}(A) = 0 \Rightarrow A = 0$ )

### Dimostrazione.

Sia  $A = V\Sigma U^T$  una SVD di  $A$ .

Dato che la norma 2 e la norma di Frobenius sono ortogonalmente invarianti si ha che

$$\begin{aligned}\|A\|_2 &= \|V\Sigma U^T\|_2 = \|\Sigma\|_2 \\ \|A\|_F &= \|V\Sigma U^T\|_F = \|\Sigma\|_F.\end{aligned}$$

Calcoliamo dunque queste norme per  $\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}$ .



## Dimostrazione (segue).

Per definizione di norma 2

$$\|\Sigma\|_2 = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|\Sigma x\|_2}{\|x\|_2} = \sigma_1.$$

Infatti

$$\|\Sigma x\|_2 = \sqrt{\sum_{i=1}^p (\sigma_i x_i)^2} = \sqrt{\sum_{i=1}^p \sigma_i^2 x_i^2} \leq \sqrt{\sum_{i=1}^p \sigma_1^2 x_i^2} = \sigma_1 \sqrt{\sum_{i=1}^p x_i^2} \leq \sigma_1 \sqrt{\sum_{i=1}^n x_i^2} = \sigma_1 \|x\|_2.$$

Segue che

$$\|\Sigma\|_2 = \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\|\Sigma x\|_2}{\|x\|_2} \leq \max_{x \in \mathbb{R}^n \setminus \{0\}} \frac{\sigma_1 \|x\|_2}{\|x\|_2} = \sigma_1.$$

Vale l'uguaglianza ovvero  $\|\Sigma\|_2 = \sigma_1$ , poiché c'è almeno un vettore per cui questo limite superiore viene raggiunto.

Per esempio scegliendo  $x = e^{(1)}$  si ha  $\|x\|_2 = 1$ , e  $\|\Sigma x\|_2 = \|\Sigma(:, 1)\|_2 = \sqrt{\sigma_1^2} = \sigma_1$ .

Risulta poi per la norma di Frobenius

$$\|\Sigma\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^m \Sigma_{ij}^2} = \sqrt{\sigma_1^2 + \dots + \sigma_p^2}.$$

□

## SVD di matrici simmetriche

Una SVD di matricie simmetrica si ottiene facilmente a partire dalla diagonalizzazione (mediante trasformazione ortogonale) della matrice:

$A \in \mathbb{R}^{n \times n}$  simmetrica  $\implies U^T A U = \text{diag}(\lambda_1, \dots, \lambda_n)$ , con  $U = [u^{(1)}, \dots, u^{(n)}]$  ortogonale.

Ordiniamo autovalori e corrispondenti autovettori in modo che

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_n|, \quad \text{con} \quad |\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_r| > 0.$$

Si noti che  $A = U D U^T$  NON è una SVD di  $A$  perché gli elementi diagonali di  $D$  non sono, in generale, tutti  $\geq 0$ .

Scriviamo pertanto

$$\begin{aligned} A &= U \text{diag}(\lambda_1, \dots, \lambda_n) U^T \\ &= U \text{diag}(\text{sgn}(\lambda_1), \dots, \text{sgn}(\lambda_r), 1, \dots, 1) \cdot \text{diag}(|\lambda_1|, \dots, |\lambda_r|, 0, \dots, 0) U^T \\ &= \underbrace{[\text{sgn}(\lambda_1)u^{(1)}, \dots, \text{sgn}(\lambda_r)u^{(r)}, u^{(r+1)}, \dots, u^{(n)}]}_V |D| U^T. \end{aligned}$$

Concludendo:

- I valori singolari di  $A$  sono i **valori assoluti degli autovalori**.
- $U$  è la matrice di trasformazione,  $V$  è la matrice descritta sopra.

**Esercizio.** Determinare una SVD della matrice simmetrica  $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$ .

Diagonalizziamo  $A$  con trasformazione ortogonale

$$P_A(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 - 4 = (-1-\lambda)(3-\lambda) = (\lambda+1)(\lambda-3)$$

quindi  $D = \text{diag}(3, -1)$ , autovalori ordinati per modulo decrescente  $|3| > |-1|$ .

L'autospazio relativo a  $\lambda = 3$ :

$$\text{Ker}(A - 3I) = \text{Ker} \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} = \{-x_1 + x_2 = 0, (x_1, x_2) \in \mathbb{R}^2\} = \langle (1, 1) \rangle \rightarrow u^{(1)} = \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)^T.$$

L'altro autovettore è ortogonale e vale pertanto  $u^{(2)} = \left( -\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right)^T$ . Abbiamo quindi

$$U = \begin{bmatrix} \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix}, \quad V = [u^{(1)} \quad -u^{(2)}] = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

infine

$$A = V|D|U^T = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \end{bmatrix},$$

come è facile verificare direttamente.

## SVD leggera (*Thin SVD*)

Sia  $A = V\Sigma U^T$  una SVD di  $A \in \mathbb{R}^{m \times n}$  e siano  $\sigma_1, \sigma_2, \dots, \sigma_r$  i valori singolari non nulli di  $A$ . Siano poi  $u^{(1)}, \dots, u^{(r)}$  e  $v^{(1)}, \dots, v^{(r)}$  i corrispondenti vettori singolari destri e sinistri, rispettivamente.

La scomposizione

$$A = V_r \Sigma_r U_r^T = \sum_{i=1}^r \sigma_i \underbrace{v^{(i)} u^{(i)T}}_{E_i}.$$

con  $E_i$  matrici di rango 1, è detta **SVD leggera di  $A$**  in quanto contiene solo le colonne delle matrici  $U$  e  $V$  corrispondenti ai valori singolari non nulli.

Queste sono le informazioni indispensabili per la ricostruzione della matrice  $A$ .

**Si noti che**  $V_r = [v^{(1)} \dots v^{(r)}] \in \mathbb{R}^{m \times r}$ ;  $U_r = [u^{(1)} \dots u^{(r)}] \in \mathbb{R}^{n \times r}$ ;  $\Sigma_r \in \mathbb{R}^{m \times r}$ .

Quindi con  $r$  vettori lunghi  $m$ ,  $r$  vettori lunghi  $n$  e  $r$  scalari si ricostruisce una matrice  $m \times n$ .

## SVD leggera

**Dimostriamo ora che**  $V\Sigma U^T = V_r \Sigma_r U_r^T$ .

Si ha

$$V\Sigma U^T = [V_r | Z] \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_r^T \\ W^T \end{bmatrix} = [V_r | Z] \begin{bmatrix} \Sigma_r U_r^T \\ 0 \end{bmatrix} = V_r \Sigma_r U_r^T.$$

Inoltre

$$\begin{aligned} [v^{(1)} \dots v^{(r)}] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} u^{(1)T} \\ \dots \\ u^{(r)T} \end{bmatrix} &= [v^{(1)} \dots v^{(r)}] \begin{bmatrix} \sigma_1 u^{(1)T} \\ \dots \\ \sigma_r u^{(r)T} \end{bmatrix} \\ &= \sigma_1 v^{(1)} u^{(1)T} + \dots + \sigma_r v^{(r)} u^{(r)T} \\ &= \sum_{i=1}^r \sigma_i v^{(i)} u^{(i)T}. \end{aligned}$$

## SVD troncata

Se invece di usare tutti gli  $r$  addendi nella SVD leggera come

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

se ne usano solo alcuni ( $l \leq r$ ) si ottiene una ricostruzione approssimata di  $A$  che chiameremo  $A_l$  dove

$$A_l = \sum_{i=1}^l \sigma_i v^{(i)} u^{(i)T}, \quad 0 \leq l \leq r.$$

Si ha che

$$A_l = \begin{bmatrix} v^{(1)} & \dots & v^{(l)} \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_l \end{bmatrix} \begin{bmatrix} u^{(1)T} \\ \dots \\ u^{(l)T} \end{bmatrix}$$

è un'approssimazione di  $A$  di rango  $l$ .

Più addendi si aggiungono più  $A_l$  si avvicina ad  $A$ .



## SVD troncata

Le matrici

$$A_0, A_1, \dots, A_r \in \mathbb{R}^{m \times n}$$

di rango, rispettivamente,  $0, 1, \dots, r$  risultano essere sempre migliori approssimazioni della matrice  $A$ .

Si ha il seguente

**Teorema (13, pag. 40 sulle dispense)**

*Sia  $l \in \{0, 1, \dots, r\}$ . Si ha che  $\text{rank}(A_l) = l$  e*

$$\min_{\substack{B \in \mathbb{R}^{m \times n} \\ \text{rank}(B) = l}} \|B - A\|_2 = \|A_l - A\|_2 = \begin{cases} \sigma_{l+1} & \text{se } l < r \\ 0 & \text{se } l = r \end{cases}$$

Questo risultato ci dice che  $A_l$  ha rango  $l$  ed è la miglior approssimazione di rango  $l$  della matrice  $A$ . L'errore di approssimazione nella metrica dedotta dalla norma 2 è pari a  $\sigma_{l+1}$  se  $l < r$  o a zero se  $l = r$  perchè in tale caso  $A_l = \sum_{i=1}^r \sigma_i v^{(i)} u^{(i)T} = A$ .

Un'importante applicazione: la SVD (troncata) viene ampiamente utilizzata nell'elaborazione di immagini (si veda la descrizione di questa applicazione alla fine della dispensa) e nell'elaborazione dei segnali.