Decomposizione Matriciale SVD
(Singular Value Decomposition)

Federico Riva
Ingegneria Matematica
Politecnico di Milano

Relatore: Professor Marco Verani

Contents

1 Introduzione

La decomposizione SVD è una fattorizzazione che consente di rappresentare una matrice come il prodotto di tre matrici ciascuna con caratteristiche specifiche. Le applicazioni di questo metodo possono essere ritrovate in una vasta gamma di campi tra cui l'analisi dei dati, il machine learning e l'elaborazione delle immagini.

2 Definizione

Sia $A \in \mathbb{R}^{m \times n}$, oppure equivalentemente $A \in \mathbb{C}^{m \times n}$, allora la decomposizione SVD della una matrice nel caso $m > n$ è definita come:

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

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\mathbf{u_1}\,\mathbf{u_2}\,\ldots\,\mathbf{u_m}] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} [\mathbf{v_1}\,\mathbf{v_2}\,\ldots\,\mathbf{v_n}]^T$$

Mentre per $m < n$:

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = [\mathbf{u_1}\,\mathbf{u_2}\,\ldots\,\mathbf{u_m}] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \end{bmatrix} [\mathbf{v_1}\,\mathbf{v_2}\,\ldots\,\mathbf{v_n}]^T$$

Dove:

- $U \in \mathbb{R}^{m \times m}$ è una matrice ortogonale contenente i vettori singolari sinistri di A,
- $\Sigma \in \mathbb{R}^{m \times n}$ è una matrice diagonale contenente i valori singolari di A ordinati in modo decrescente,
- $V \in \mathbb{R}^{n \times n}$ è una matrice ortogonale contenente i vettori singolari destri di A.

Definizione 1. ***Matrice Ortogonale***
Una matrice quadrata $Q \in \mathbb{R}^{n \times n}$ si dice ortogonale se $Q^T Q = I$.

Definizione 2. ***Matrice Simmetrica***
Una matrice quadrata $Q \in \mathbb{R}^{n \times n}$ si dice simmetrica se $Q^T = Q$.

Definizione 3. ***Forma Quadratica***
Si definisce forma quadratica un polinomio di secondo grado in n variabili $p : \mathbb{R}^n \to \mathbb{R}$ t.c.

$$p(x) = p(x_1, \dots, x_n) = \sum_{1 \leq i,j \leq n} a_{ij} x_i x_j = a_{11}x_1^2 + a_{12}x_1x_2 + \cdots + a_{nn-1}x_n x_{n-1} + a_{nn}x_n^2 = \mathbf{x}^T A \mathbf{x}$$

dove $\mathbf{x} = (x_1, \dots, x_n)$ e A è una matrice simmetrica di dimensione $n \times n$.

Proposizione 1.
Hp: Data una qualsiasi $M \in \mathbb{R}^{m \times n}$
Ts: $D = M^T M$ e $S = MM^T$ sono matrici simmetriche e definite positive.

Proof. 1) Vogliamo dimostrare che $D = D^T$. Segue quindi dalle proprietà di trasposizione del prodotto di due matrici: $D^T = (M^T M)^T = MM^T = D$. Si dimostra in maniera analoga il risultato per S.
2) Per dimostrare la semipositività della matrice $A^T A$ dal momento che è una matrice quadrata possiamo studiare il segno della forma quadratica associata. Se quest'ultima sarà positiva allora lo saranno anche tutti gli autovalori.

$$\forall \mathbf{x} \neq 0 : \mathbf{x} \in \mathbb{R}^n$$
$$\mathbf{x}^T A^T A \mathbf{x} = (A\mathbf{x})^T (A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$$

□

Osservazione: Gli autovalori che otterremo dalle due matrici saranno ≥ 0 e potremo quindi calcolarne la radice quadrata.

Proposizione 2.
Hp: Sia $A \in \mathbb{R}^{m \times n}$
Ts: $A^T A$ e AA^t hanno gli stessi autovalori non nulli.

Proof. Ipotizziamo che la coppia autovalor' autovettore $(\mathbf{x}, \lambda)$ sia associata alla matrice $A^T A$. Possiamo quindi scrivere

$$A^T A \mathbf{x} = \lambda \mathbf{x}$$

e moltiplicando ambo i membri per A otteniamo

$$AA^T A \mathbf{x} = \lambda A \mathbf{x}$$

Ma allora definendo $\mathbf{y} = A\mathbf{x}$ possiamo riscrivere

$$AA^T \mathbf{y} = \lambda \mathbf{y}$$

Il che conclude la dimostrazione poichè prova che λ è un autovalore anche della matrice AA^T □

Teorema 1. ***Teorema Spettrale***
Hp: Data $Q \in \mathbb{R}^{n\times n}$ una matrice simmetrica
Ts: È ortognalmente digonalizzabile $Q = U^{-1}DU = U^T DU$

Prendendo in cosiderazione le **Proposizioni 1**, **2** ed il **Teorema 1** possiamo osss ervare che data una qualsiasi matrice A è possibile costruire a partire da essa due matrici simmetriche ortgonalmente diagonalizzabili. Possiamo quindi fornire le seguenti definizioni.

Definizione 4. ***Vettori Singolari di sinistra***
I vettori singolari di sinstra di A sono gli autovettori $\mathbf{u_1}, \dots, \mathbf{u_m}$ di $AA^T = U^T D_l U$ e compongono la Matrice Singolare di Sinistra U.

Definizione 5. ***Vettori Singolari di destra***
I vettori singolari di destra di A sono gli autovettori $\mathbf{v_1}, \dots, \mathbf{v_n}$ di $A^T A = V^T D_r V$ e compongono la Matrice Singolare di destra V.

Definizione 6. ***Valori Singolari***
I valori singolari sono le radici quadrate degli autovalori non nulli di $A^T A$ e $A^T A$
$\sigma_1 \geq \dots \geq \sigma_{r=\min n,m} \geq 0$. Inchicheremo con r il numero di valori singolari che nel caso di assenza di autovalori nulli sarà $r = \min\{m, n\}$.

3 Legame fra la decomposizione SVD ed il Teorema Spettrale

Da un punto di vista intuitivo la farrotizzazion SVD può essere considerata come la generalizzaizone del concetto di diagonalizzione di una matrice quadrata; infatti se la seconda può essere applicata solamente a matrici quadrate $(n \times n)$ risuulta valida per qualsiasi tipo di matrice $(m \times n)$ senza limite alcuno sulla sua forma.
Se A ammette la fattorizzazione SVD allora $A = U\Sigma V^T$ valgono le seguenti relazioni:

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

$$AA^T = U\Sigma V^T (U\Sigma V^T)^T = \dots = U\Sigma\Sigma^T U^T$$

Vale duque un risultato ancora più specifico della semplice dagonalizzazione di una matricre ovvero il Teorema Spettrale

4 Interpretazione geometrica

Data una matrice $A \in \mathbb{R}^{m\times n}$ consideriamo l'applicazione lineare ad essa associata

$$A : \mathbb{R}^n \to \mathbb{R}^m$$

La fattorizzazione SVD scopone la trasformazione lineare in 3 trasformazioni geometriche rappresentate dalle 3 matici U,V e Σ dove le matrici ortonormali rapppresentano una rotazinone o una simmetria assiale mentre la matrice Σ è una dilatazione di intestità σ_i lungo la componente $\mathbf{x}_i$.Di seguito illustiamo visimaente lo specifico caso $A \in \mathbb{R}^{2\times 2}$ nel caso di una circonferenza di raggio unitario che viene mappata in un'ellise.

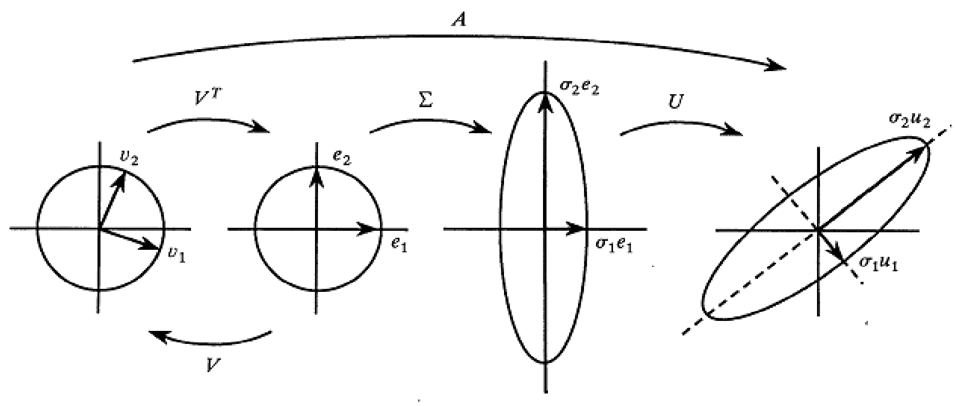

Figure 1: Illustrazione in $\mathbb{R}^2$

Possiamo estendere questo ragionemento ad un caso più generale considerando una sfera di raggio unitario in $\mathbb{R}^n$ e l'ellissoide ad essa associata da $A \in \mathbb{R}^{m\times n}$ in $\mathbb{R}^m$.

Notiamo che le matrici U e V sono unitarie e le colonne di ciscuna formano una mase

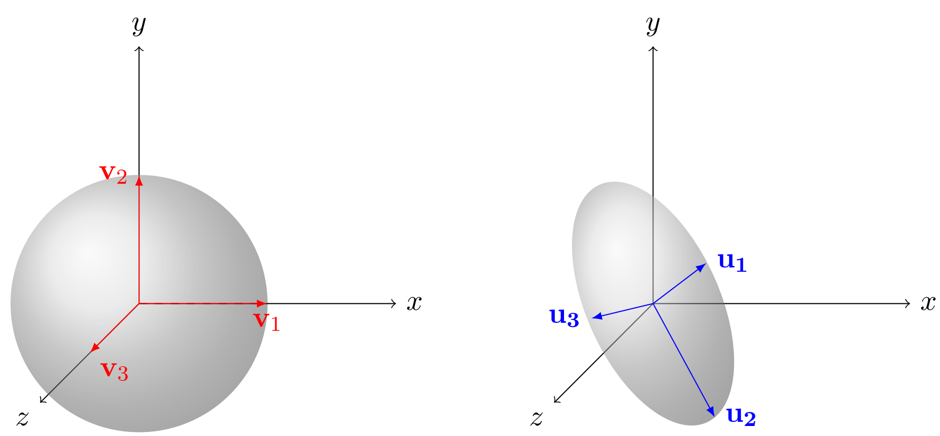

Figure 2: Mappa da una sfera in $\mathbb{R}^n$ ad un'ellisoide in $\mathbb{R}^m$.

di vettori ortonormali. Inoltre la matrice A mappa il v_i elemento della base in nel vettore dilatao $\sigma_i u_i$.

Verifichiamo lo:

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

$$AV = U\Sigma$$

$$[A\mathbf{v_1}\, A\mathbf{v_2}\, \dots\, A\mathbf{v_n}] = [U\boldsymbol{\sigma_1}\, U\boldsymbol{\sigma_2}\, \dots\, U\boldsymbol{\sigma_n}]$$

Tenendo conto che

$$\boldsymbol{\sigma_i} = [0\, \dots\, 0\, \sigma_i\, 0\, \dots\, 0]$$

Otteniamo

$$\forall i = 1, \dots, n \;\; A\mathbf{v_i} = \sigma_i \mathbf{u_i}$$

Possiamo quindi scivere più in generale che: $A\mathbf{v} = \sigma\mathbf{u}$ e $A^T\mathbf{u} = \sigma\mathbf{v}$.
Dal punto di vista degli spazi associati allo *Span* delle colonne di ciascuna matrice se la matrice A non ha rango massimo, $\mathrm{Rg}(A) = r < \min\{m, n\}$, allora sussiste la seguente relazione:

- Le prime r colonne di U sono una base dello $\mathrm{Span}(A)$;
- Le ultime $m - r$ colonne di U sono una base del $\mathrm{Ker}(A^T)$;
- Le prime r colonne di V sono una base dello $\mathrm{Span}(A^T)$;
- Le ultime $n - r$ colonne di V sonouna base del $\mathrm{Ker}(A)$

Osservazione: $\mathrm{Span}(A) = \mathrm{Col}(A)$ e quindi $\mathrm{Span}(A^T) = \mathrm{Col}(A^T) = \mathrm{Row}(A)$.
La dimostrazione dell'elenco delle proprietà segue dalla **Proposizione 3** unita al **Teorema 1** e alla definzione di decomposizione SVD.

Teorema 2. ***Teorema di Nullità più rango***
Hp: Sia $A \in \mathbb{R}^{m\times n}$ con Rg(A)=r
Ts: dim Ker(A)=$n - r$

Proposizione 3.
Hp: $\forall A \in \mathbb{R}^{m\times n}$ valogno le seguenti uguaglianze
Ts: $Ker(A) = Ker(A^TA)$, $Col(A^T) = Col(A^TA)$ e di conseguenza $Rg(A) = Rg(A^TA)$

Proof. Osserviamo che

$$\|A\mathbf{x}\|^2 = (A\mathbf{x})^T(A\mathbf{x}) = \mathbf{x}^TA^TA\mathbf{x}$$

Ne segue che, se $\mathbf{x}$ appartiene al Nucleo di A^TA allora $\|A\mathbf{x}\|^2 = 0$ e quindi $A\mathbf{x} = 0$, cioè $\mathbf{x} \in \mathrm{Ker}(A)$. Viceversa se $\mathbf{x} \in \mathrm{Ker}(A)$ ripercorrendo la catena di uguaglianze a retroso $\mathbf{x} \in \mathrm{Ker}(A^TA)$ e per arbitrarietà della scelta del vettore i due nuclei sono uguali. Dall'ugualianza dei nuclei pèr il **Teorema 2** segue l'uquaglianza dei ranghi.
$\mathrm{Col}(A^T)$=$\mathrm{Row}(A)$ è il complemento ortogonale di $\mathrm{Ker}(A)$. Ma $\mathrm{Ker}(A) = \mathrm{Ker}(A^TA)$ il cui complemento ortogonale è $\mathrm{Row}(A^TA)$. Quindi $\mathrm{Col}(A^T) = \mathrm{Row}(A^TA)$ ma per simmetria di A^TA $\mathrm{Row}(A^TA)$=$\mathrm{Col}(A^TA) \Rightarrow \mathrm{Col}(A)$=$\mathrm{Col}(A^TA)$. □

5 Dimostrazione della decomposizione SVD e Teorema Eckart-Young

5.1 Dimostrazione

Data una matrice $A \in \mathbb{R}^{m\times n}$ poichè $A^TA \in \mathbb{R}^{n\times n}$ è simmetrica e semidefinita positiva segue dal Teorema Spettrale che $\exists$ una matrice ortonormale $V \in \mathbb{R}^{n\times n}$ tale che

$$A^TA = V\bar{D}V^T$$

e a sua volta poichè $V^{-1} = V^T$

$$V^TA^TAV = \bar{D} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

dove $D \in \mathbb{R}^{\ell\times\ell}$ è una matrice diagonale e positiva di dimensione con $\ell \leq \min\{m, n\}$ il numero di autovalori non nulli della matrice A^TA. Osserviamo che per definizione di autovalori e autovettori la $i-esima$ colonna di V corrisponde all' $i-esimo$ autovalore $\bar{D}_{ii}$. Di consguenza possiamo indivuare $V_1 = [\mathbf{v}_1 \ldots \mathbf{v}_\ell] \in \mathbb{R}^{n\times\ell}$ e $V_2 = [\mathbf{v}_{\ell+1} \ldots \mathbf{v}_n] \in \mathbb{R}^{n\times n-\ell}$ tali che $V = [V_1\, V_2]$ che sono rispettivamente gli autovettori associati ad autovalori non nulli e nulli. Possiamo quindi riscrivere l'equazione come:

$$\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} A^TA \begin{bmatrix} V_1 & V_2 \end{bmatrix} = \begin{bmatrix} V_1^TA^TAV_1 & V_1^TA^TAV_2 \\ V_2^TA^TAV_1 & V_2^TA^TAV_2 \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

Da cui ricaviamo le seguenti equazioni:

$$V_1^TA^TAV_1 = D \ \text{ e } \ V_2^TA^TAV_2 = 0$$

Siamo quindi in grado di risolvere la seconda

$$\|AV_2\|_2 = 0 \Rightarrow AV_2 = 0$$

Costruiamo inolte le seguenti matrici di diverse dimensioni che torneranno utili in seguito:

- $V_1^TV_1 = I_\ell$ dove $(\ell \times n) \times (n \times \ell) = (\ell \times \ell)$
- $V_2^TV_2 = I_{n-\ell}$ dove $(n - \ell \ \times n) \times (n \times \ n - \ell) = (n - \ell \times n - \ell)$
- $[V_1V_2][V_1V_2]^T = V_1V_1^T + V_2V_2^T = I_n$ dove

$$(n \times \ell) \times (\ell \times n) = (n \times n) \text{ e } (n \times n - \ell) \times (n - \ell \times n) = (n \times n)$$

Definiamo ora la matrice U_1 come:

$$U_1 = AV_1D^{-\frac{1}{2}} \in \mathbb{R}^{m\times\ell} \ \{(m\times\ell) = (m\times n)\times(n\times\ell)\times(\ell\times\ell)\}$$

sapendo che $D^{-\frac{1}{2}}$ è ben definita poiché $D_{ii} > 0 \ \forall \ i = 1,\dots,\ell$.
Invertendo la relazione e sostuendo U_1:

$$A = U_1D^{\frac{1}{2}}V_1^T = AV_1D^{-\frac{1}{2}}D^{\frac{1}{2}}V_1^T = AV_1V_1^T$$

e sfruttando $V_1V_1^T + V_2V_2^T = I_n$ e $AV_2 = 0$ otteniamo:

$$A = A\left(I - V_2V_2^T\right) = A - AV_2V_2^T = A - 0 = A$$

il che sabilisce la veridicità dalla definizione di U_1 proposta.
Abbiamo quindi quasi dimostrato il risltato completo, non ci resta che verificare che le colonne della matrice U_1 formino una base ortonormale:

$$U_1^TU_1 = D^{-\frac{1}{2}}V_1^TA^TAV_1D^{-\frac{1}{2}} = D^{-\frac{1}{2}}DD^{-\frac{1}{2}} = I_\ell \in \mathbb{R}^{\ell\times\ell}$$

Possiamo quindi costruire il complemento orgonale di U_1 ovvero $U_2 = \text{Span}(\text{Col}(U_1)^\perp) \in \mathbb{R}^{m\times n-\ell}$ in modo che $U = [U_1 \ U_2] \in \mathbb{R}^{m\times m}$ sia una matrice ortogonale.
Per concludere aggiungiamo un numero opportune di righe e colonne nulle alla matrice $D^{\frac{1}{2}} \in \mathbb{R}^{\ell\times\ell}$ in modo da creare:

$$\Sigma = \begin{bmatrix} \begin{bmatrix} D^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix} \in \mathbb{R}^{m\times n}$$

Allora sfruttando le proprietà del prodotto matriciale fra matrici a blocchi:

$$U\Sigma V^T = [U_1 \ U_2] \begin{bmatrix} \begin{bmatrix} D^{\frac{1}{2}} & 0 \\ 0 & 0 \end{bmatrix} \\ 0 \end{bmatrix} [V_1 \ V_2]^T = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} D^{\frac{1}{2}}V_1^T \\ 0 \end{bmatrix} = U_1D^{\frac{1}{2}}V_1^T = A$$

□

5.2 Norme Matriciali

Dal momento che $\mathbb{R}^{m\times n}$ è isomorgf a $\mathbb{R}^{mn}$ la norma di una matrice è equivalente alla norma di un vettore.

Definizione 7. ***Norma***
Una funzione f: $\mathbb{R}^{m\times n} \to \mathbb{R}$ è detta norma se sooddisfa le seguenti proprietà

- $f(A) \geq 0 \ \forall A \in \mathbb{R}^{m\times n}$ *e* $f(A) = 0 \iff A = 0$

- $f(A+B) \leq f(A) + f(B) \;\; \forall A, B \in \mathbb{R}^{m\times n}$
- $f(\alpha A) = |\alpha| f(A), \quad \alpha \in \mathbb{R}, A \in \mathbb{R}^{m\times n}$

Useremo quindi la notazione $\|A\| = f(A)$ per indicare la 'norma della matrice A'.

Fra le numerose norme esistenti le due più frequenti sono:

Definizione 8. ***Norma di Frobenius***

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}|a_{ij}|^2}$$

Definizione 9. ***Norma p-esima***

$$\|A\|_p = \sup_{\mathbf{x}\neq 0}\frac{\|A\mathbf{x}\|_p}{\|\mathbf{x}\|_p} = \max_{\|\mathbf{x}\|_p=1}\|A\mathbf{x}\|_p$$

Definizione 10. ***Norma*** ∞

$$\|A\|_\infty = \max_{1\leq i\leq m}\sum_{j=1}^{n}|a_{ij}|$$

Proposizione 4.
Hp: Data una matrice ortogonale $Q \in \mathbb{R}^{n\times n}$
Ts:$\forall \mathbf{x} \in \mathbb{R}^n \;\; \|Q\mathbf{x}\| = \|\mathbf{x}\|$

Proof. $\|Q\mathbf{x}\|^2 = (Q\mathbf{x})^T(Q\mathbf{x}) = \mathbf{x}^T Q^T Q \mathbf{x} = \mathbf{x}^T I \mathbf{x} = \mathbf{x}^T\mathbf{x} = \|\mathbf{x}\|^2$ □

Enunciamo ora una serie di corollari che seguono dalla dimostrazione della fattorizzazione SVD.

Corollario 1.
Hp: Sia $U^T AV = \Sigma$ la decomposizione SVD della matrice $A \in \mathbb{R}^{m\times n}$ con $m \geq n$.
Ts: $\forall i = 1, \ldots, n \; A\mathbf{v}_i = \sigma_i \mathbf{u}_i e A^T \mathbf{u}_i = \sigma_i \mathbf{v}_i$

Proof. Come abbiamo già discusso precedentemente nell'intepretazione geometrica:

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

$$AV = U\Sigma$$

$$[A\mathbf{v_1}\, A\mathbf{v_2}\, \ldots\, A\mathbf{v_n}] = [U\boldsymbol{\sigma_1}\, U\boldsymbol{\sigma_2}\, \ldots\, U\boldsymbol{\sigma_n}]$$

Tenendo conto che

$$\boldsymbol{\sigma_i} = [0 \ldots 0\, \sigma_i\, 0 \ldots 0]$$

Otteniamo

$$\forall i = 1, \dots, n \; A\mathbf{v_i} = \sigma_i \mathbf{u_i}$$

Analogamente si dimostra il risultato per $A^T\mathbf{u}_i = \sigma_i \mathbf{v}_i$ $\square$

Corollario 2.
Hp: Sia $A \in \mathbb{R}^{m\times n}$ *con* $m \geq n$
Ts: $\|A\|_2 = \sigma_1$ *e* $\|A\|_F = \sqrt{\sigma_1^2 + \dots + \sigma_r^2}$

Proof. 1)

$$\|A\|_2 = \max_{\|\mathbf{x}\|_2=1} \|A\mathbf{x}\|_2 = \max_{\|\mathbf{x}\|_2=1} \|U\Sigma V^T\mathbf{x}\|_2$$

Dalla **Proposizione 4** sappiamo che le due matrici ortognoali U e V non modificano la norma. Possiamo quindi riscrivere l'uguaglianza come

$$\max_{\|\mathbf{x}\|_2=1} \|\Sigma\mathbf{x}\|_2$$

cioè

$$\max_{\|\mathbf{x}\|_2=1} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

e poichè $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$ otteniamo il massimo per $\mathbf{x} = [1\, 0 \dots 0]^T$ ovvero $\|A\|_2 = \sigma_1$
2)

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |(U\Sigma V^T)_{ij}|^2}$$

scrivendo il prodotto matriciale elemento per elemento otteniamo che

$$(U\Sigma V^T)_{ij} = \sum_{k=1}^{r} U_{ik}\Sigma_{kk}(V^T)_{kj} = \sum_{k=1}^{r} U_{ik}\Sigma_{kk}V_{jk}$$

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{r} U_{ik}\Sigma_{kk}V_{jk}}$$

Tuttavia per costruzione della matrice Σ

$\Sigma_{kk} = \sigma_k \geq 0 \; \forall k = 1, \dots, r$ con $r \leq n = \min\{m, n\}$ e $\Sigma_{ij} = 0 \; \forall i = 1, \dots, m$ e $\forall j = 1, \dots, n$

Otteniamo la tesi

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2}$$

□

Corollario 3.
Hp:
Ts:

Proof. □

Corollario 4.
Hp:
Ts:

Proof. □

Teorema 3. ***Teorema di Eckart-Young***
Hp: Sia A una matrice $m \times n$ di rango r, e sia $k \in \mathbb{N}$ tale che $1 \leq k \leq r$.
Ts: $\forall$ matrice B di rango k, la soluzione del problema di minimo quadrato:

$$\min_{rank(X)=k} \|A - X\|_F^2$$

è data dalla matrice X data dalla decomposizione ai valori singolari troncata di A, ovvero:

$$X = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

dove $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0$ sono i valori singolari di A, e u_i e v_i sono rispettivamente la i-esima colonna di U e la i-esima colonna di V nella decomposizione ai valori singolari di A, ovvero $A = U\Sigma V^T$, con U una matrice $m \times r$, V una matrice $n \times r$, e Σ una matrice diagonale $r \times r$ contenente i valori singolari di A sulla diagonale.

6 Algoritmi per il Calcolo

7 Applicazioni Teoriche

7.1 Pseudoinversa di una Matrice

7.2 Soluzione ai Minimi Quadrati

8 Applicazioni Pratiche

8.1 PCA - Principal Component Analysis

8.2 Compressione Immagini

8.3 DMD - Dyniamic Mode Decomposition

8.3.1 Esempio