Scomposizione ai valori singolari.

Sia $A \in \mathbb{R}^{m \times n}$ una matrice, non necessariamente quadrata. Si potrebbe considerare anche una matrice complessa ottenendo risultati analoghi. Abbiamo:

Osservazione

$AA^T \in \mathbb{R}^{m \times m}$ e $A^T A \in \mathbb{R}^{n \times n}$ sono simmetriche semidefinite positive. Esse sono definite positive se e solo se A è invertibile. Inoltre: $rk(AA^T) = rk(A^T A) = rk(A)$. Infatti:

$$A^T Ax = 0 \implies x^T A^T Ax = (Ax)^T Ax =< Ax, Ax >= \|Ax\|_2^2 = 0 \implies Ax = 0.$$

Definizione

Due matrici $A, B \in \mathbb{R}^{m \times n}$, dello stesso tipo, si dicono equivalenti se esistono una matrice $Q \in \mathbb{R}^{m \times m}$ invertibile e una matrice $P \in \mathbb{R}^{n \times n}$ invertibile tali che

$$B = QAP.$$

Questa è una relazione di equivalenza.

Teorema (Scomposizione ai valori singolari)

Sia $A \in \mathbb{R}^{m\times n}$*. Allora esistono una matrice* $U \in O(m)$ *e una matrice* $V \in O(n)$*, tali che:*

$$U^T AV = \Sigma, \quad \textit{cioè} \quad A = U\Sigma V^T,$$

dove la matrice "diagonale" $\Sigma \in \mathbb{R}^{m\times n}$ *ha elementi*

$$\sigma_{ij} = \begin{cases} 0 & \textit{se } i \neq j, \\ \sigma_i & \textit{se } i = j, \end{cases}$$

con

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0, \quad p = \min\{m, n\}.$$

Teorema (Scomposizione ai valori singolari)

Sia $A \in \mathbb{R}^{m\times n}$. Allora esistono una matrice $U \in O(m)$ e una matrice $V \in O(n)$, tali che:

$$U^T AV = \Sigma, \quad \textit{cioè} \quad A = U\Sigma V^T,$$

dove la matrice "diagonale" $\Sigma \in \mathbb{R}^{m\times n}$ ha elementi

$$\sigma_{ij} = \begin{cases} 0 & \textit{se } i \neq j, \\ \sigma_i & \textit{se } i = j, \end{cases}$$

con

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = \sigma_p = 0, \quad p = \min\{m, n\}.$$

- Le colonne $u_1, \ldots, u_m$ di U sono detti vettori singolari sinistri di A. Essi sono autovettori di AA^T. (Infatti: $AA^T = U\Sigma\Sigma^T U^T$, cioè $U^T(AA^T)U = \Sigma\Sigma^T$, che è diagonale).
- Le colonne $v_1, \ldots, v_n$ di V sono detti vettori singolari destri di A. Essi sono autovettori di $A^T A$. (Infatti: $A^T A = V\Sigma^T\Sigma V^T$, cioè $V^T(A^T A)V = \Sigma^T\Sigma$, che è diagonale).
- I numeri reali $\sigma_1, \ldots, \sigma_p$ sono detti valori singolari di A. Essi sono le radice quadrate degli autovalori λ_j di $A^T A$:

$$\sigma_j = \sqrt{\lambda_j(AA^T)} = \sqrt{\lambda_j(A^T A)}.$$

Osservazione

- *I vettori singolari (sinistri e destri) e quindi le matrici U e V non sono univocamente determinati.*
- *I valori singolari e la matrice Σ sono univocamente determinati.*

Dimostrazione.

Essendo $A^T A$ simmetrica, per il teorema spettrale, esiste una matrice $V \in O(n)$ tale che $V^T(A^T A)V = \Sigma$ sia diagonale. Ricordiamo che le colonne di V sono date da autovettori di $A^T A$ e gli elementi diagonali di Σ sono i corrispondenti autovalori. Questi sono non negativi, essendo $A^T A$ semidefinita positiva. Possiamo ordinare gli autovettori (cioè le colonne di V) in modo da avere:

$$V^T A^T AV = diag(\sigma_1^2, \dots, \sigma_r^2, 0, \dots, 0)$$

con $\sigma_1 \geq \cdots \geq \sigma_r > 0$.

Se indichiamo con w_j il j-esimo vettore colonna della matrice AV, allora dalla relazione sopra segue che i vettori $w_1, \dots, w_r$ sono ortogonali fra loro. Si noti che:

$$w_j^T w_j = \begin{cases} \sigma_j^2 & \text{se } 1 \leq j \leq r, \\ 0 & \text{se } r < j \leq n. \end{cases}$$

Quindi se consideriamo i vettori

$$u_1 = \sigma_1^{-1} w_1, \dots, u_r = \sigma_r^{-1} w_r,$$

questi sono formano un insieme ortonormale di vettori. Lo possiamo completare fino ad ottenere una base ortonormale $u_1, \dots, u_r, \dots, u_m$ di R^m. Sia ora U la matrice le cui colonne sono date da questi vettori. Si ha $U \in O(m)$ e $UAV = diag(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ poiché

$$(UAV)_{ij} = (u_i)^T w_j = \begin{cases} \delta_{ij}\sigma_j^2 & \text{se } 1 \leq j \leq r, \\ 0 & \text{se } r < j \leq n. \end{cases}$$

Esempio

Sia

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}.$$

Allora:

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

La matrice $AA^T = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$ *è simmetrica (semi)definita positiva con autovalori* $\lambda_1 = \sigma_1^2 = 3, \lambda_2 = \sigma_2^2 = 1$ *e autovettori:*

$$u_1 = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad u_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}.$$

La matrice $A^T A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$ *ha autovettori:*

$$u_1 = \begin{pmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{pmatrix}, \quad u_2 = \begin{pmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{pmatrix}, \quad u_3 = \begin{pmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{pmatrix}.$$

Teorema di scomposizione polare

Dalla scomposizione ai valori singolari di una matrice si ottiene subito il seguente

Teorema (Scomposizione polare)

Ogni matrice $A \in \mathbb{R}^{n \times n}$ è scomponibile come prodotto $A = QS$ di una matrice ortogonale $Q \in O(n)$ e una matrice simmetrica S, semidefinita positiva.
Se A è invertibile, allora A è definita positiva e la scomposizione è unica.

(Il nome deriva dal fatto che per $m = n = 1$ si ottiene la forma polare di un numero).
Dalla scomposizione in valori singolari $A = U\Sigma V^T$ della matrice A infatti si ottiene:

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

in cui la matrice $Q = UV^T$ è ortogonale e la matrice $S = V\Sigma V^T$ è simmetrica.

Unicità:

$A = Q_1 S_1 = Q_2 S_2 \implies A^T A = S_1^T Q_1^T Q_1 S_1 = S_1^2 = S_2^2 \implies S_1 = S_2 \implies Q_1 = Q_2$.

Osservazione

Esiste un'analoga scomposizione polare $A = S'Q'$ con S' simmetrica e Q' ortogonale data da

$$A = U\Sigma V^T = (U\Sigma U^T)(UV^T) = S'Q'.$$

Si noti anche che $Q' = UV^T = Q$ mentre $S' = U\Sigma U^T = UV^T SVU^T = QSQ^T$ (che è simile a S ma in generale è diversa da questa).

Osservazione (Pseudoinversa di Moore-Penrose)

Sia $A \in \mathbb{R}^{m\times n}$. *Se la scomposizione ai valori singolari (SVD) di* A *è*

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

con

$$U \in O(m), \quad V \in O(n), \quad \Sigma = diag\,(\sigma_1, \dots, \sigma_r, 0, \dots, 0)\,,$$

allora la pseudoinversa di Moore-Penrose $A^\dagger$ *di* A *(che quando* $rk(A) = n$ *è data da* $(A^TA)^{-1}A^T$*) è uguale a*

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

dove

$$\Sigma^\dagger = diag\left(\frac{1}{\sigma_1}, \dots, \frac{1}{\sigma_r}, 0, \dots, 0\right).$$