CdL in Scienze Statistiche ed Economiche - Università degli Studi di Milano-Bicocca

Lezione : Teorema di Decomposizione Spettrale

Docente: Aldo Solari

1 Direzione e lunghezza di un vettore

Sia $\underset{k\times 1}{a}$ un vettori a valori reali. Allora $\underset{k\times 1}{a}$ si può decomporre in due componenti,

- Lunghezza
$$\lambda = \|\underset{k\times 1}{a}\| = \sqrt{\underset{1\times k}{a'}\underset{k\times 1}{a}} = \sqrt{\sum_{j=1}^{k} a_j^2}$$

- Direzione *normalizzata*
$$\underset{k\times 1}{v} = \frac{\underset{k\times 1}{a}}{\lambda}$$
con $\|\underset{k\times 1}{v}\| = 1$

Questa idea si può estendere ad una matrice simmetrica $\underset{k\times k}{A}$ a valori reali, che è un insieme di vettori. Gli autovalori (*eigenvalues*) λ e gli autovettori (*eigenvectors*) normalizzati $\underset{k\times 1}{v}$ con $\|\underset{k\times 1}{v}\| = 1$ sono definiti dall'equazione
$$\underset{k\times k}{A}\underset{k\times 1}{v} = \lambda \underset{k\times 1}{v}$$
Ci sono esattamente k coppie $(\lambda_j, \underset{k\times 1}{v_j}), j = 1, \ldots, k$ che soddisfano l'equazione. Per convenzione, gli autovalori sono ordinati in maniera decrescente: $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_k$

1.1 Autovalori e autovettori

- $\underset{k\times k}{A}$ una matrice simmetrica a valori reali
- $\underset{k\times k}{A}$ definita positiva $\Leftrightarrow$ tutti gli autovalori di $\underset{k\times k}{A}$ sono positivi , i.e. $\lambda_j > 0, j = 1, \ldots, k$
- $\underset{k\times k}{A}$ semidefinita positiva $\Leftrightarrow$ tutti gli autovalori di $\underset{k\times k}{A}$ sono non negativi, i.e. $\lambda_j \geq 0$, $j = 1, \ldots, k$
- $\underset{k\times k}{A}$ con $\mathrm{rango}(\underset{k\times k}{A}) = r \leq k$. Allora $\underset{k\times k}{A}$ ha r autovalori non nulli, e i rimanenti $k - r$ autovalori nulli

- Autovettori normalizzati $\underset{k\times 1}{v_j}$ e $\underset{k\times 1}{v_l}$ associati ad autovalori distinti $\lambda_j \neq \lambda_l$ sono perpendicolari, i.e. $\underset{1\times k}{v_j'}\underset{k\times 1}{v_l} = 0$

2 Teorema di decomposizione spettrale

Theorem 2.1. *Sia $\underset{k\times k}{A}$ una matrice simmetrica a valori reali. Allora $\underset{k\times k}{A}$ si può esprimere come*

$$\boxed{\underset{k\times k}{A} = \underset{k\times k}{V}\underset{k\times k}{\Lambda}\underset{k\times k}{V'} = \sum_{j=1}^{k} \lambda_j \underset{k\times 1}{v_j}\underset{1\times k}{v_j'}}$$

dove

- $\underset{k\times k}{\Lambda} = \operatorname{diag}(\lambda_1, \dots, \lambda_k)$ *è una matrice diagonale dove il j-simo elemento diagonale λ_j è il j-simo autovalore associato ad $\underset{k\times k}{A}$*
- $\underset{k\times k}{V} = \begin{bmatrix} \underset{k\times 1}{v_1} & \cdots & \underset{k\times 1}{v_k} \end{bmatrix}$*, dove la j-sima colonna $\underset{k\times 1}{v_j}$ è il j-simo autovettore normalizzato ($\| \underset{k\times 1}{v_j} \| = 1$) associato all'autovalore λ_j*
- $\underset{k\times k}{V}$ *è una matrice ortogonale:* $\underset{k\times k}{V}\underset{k\times k}{V'} = \underset{k\times k}{V'}\underset{k\times k}{V} = \underset{k\times k}{I}$

Example 2.2. $\underset{2\times 2}{A} = \begin{bmatrix} 2.2 & 0.4 \\ 0.4 & 2.8 \end{bmatrix}$ *simmetrica e a valori reali. Otteniamo gli autovalori $\lambda_1 = 3$ e $\lambda_2 = 2$, a cui corrispondono gli autovalori (normalizzati)* $\underset{2\times 1}{v_1} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$ *e* $\underset{2\times 1}{v_2} = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$*, quindi*

$$\underset{2\times 2}{A} = 3 \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} + 2 \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \end{bmatrix}$$

Lemma 2.3.

$$\boxed{\underset{k\times k}{A^q} = \underset{k\times k}{V}\underset{k\times k}{\Lambda^q}\underset{k\times k}{V'}}$$

dove $\underset{k\times k}{\Lambda}{}^q = \operatorname{diag}(\lambda_1^q, \dots, \lambda_k^q)$

- *Se $\underset{k\times k}{A}$ è semidefinita positiva, vale solo per $q > 0$ razionali, ovvero per $q \in \mathbb{Q}^+$*
- *Se $\underset{k\times k}{A}$ è definita positiva, vale anche per $q < 0$ razionali, ovvero per $q \in \mathbb{Q}^{\setminus 0}$*

Esempi di applicazioni del Lemma:

- $\underset{k\times k}{A^{-1}} = \underset{k\times k}{V}\,\underset{k\times k}{\Lambda^{-1}}\,\underset{k\times k}{V'}$
- $\underset{k\times k}{A^{-\frac{1}{2}}} = \underset{k\times k}{V}\,\underset{k\times k}{\Lambda^{-\frac{1}{2}}}\,\underset{k\times k}{V'}$
- $\underset{k\times k}{A^{2}} = \underset{k\times k}{V}\,\underset{k\times k}{\Lambda^{2}}\,\underset{k\times k}{V'}$
- $\underset{k\times k}{A^{\frac{1}{2}}} = \underset{k\times k}{V}\,\underset{k\times k}{\Lambda^{\frac{1}{2}}}\,\underset{k\times k}{V'}$

2.1 Decomposizione Spettrale di S

La decomposizione spettrale della matrice simmetrica $\underset{p\times p}{S}$ è

$$\underset{p\times p}{S} = \underset{p\times p}{V}\,\underset{p\times p}{\Lambda}\,\underset{p\times p}{V'}$$

Ricordando che $\underset{p\times p}{V}\,\underset{p\times p}{V'} = \underset{p\times p}{V'}\,\underset{p\times p}{V} = \underset{p\times p}{I}$, otteniamo

$$\mathrm{tr}(\underset{p\times p}{S}) = \mathrm{tr}(\underset{p\times p}{V}\,\underset{p\times p}{\Lambda}\,\underset{p\times p}{V'}) = \mathrm{tr}(\underset{p\times p}{\Lambda}\,\underset{p\times p}{V'}\,\underset{p\times p}{V}) = \mathrm{tr}(\underset{p\times p}{\Lambda}) = \sum_{j=1}^{p}\lambda_j$$

$$\begin{aligned}\det(\underset{p\times p}{S}) &= \det(\underset{p\times p}{V}\,\underset{p\times p}{\Lambda}\,\underset{p\times p}{V'}) = \det(\underset{p\times p}{V})\det(\underset{p\times p}{\Lambda})\det(\underset{p\times p}{V'}) \\ &= \det(\underset{p\times p}{\Lambda})\det(\underset{p\times p}{V}\,\underset{p\times p}{V'}) = \det(\underset{p\times p}{\Lambda})\det(\underset{p\times p}{I}) = \prod_{j=1}^{p}\lambda_j\end{aligned}$$

3 Matrice dei dati ortogonalizzati

Proposition 3.1. *Sia $\underset{p\times p}{S}$, la matrice di varianze/covarianze di $\underset{n\times p}{X}$, definita positiva. Possiamo allora ottenere la matrice dei dati ortogonalizzati*

$$\boxed{\underset{n\times p}{\tilde{Z}} = \underset{n\times n}{H}\,\underset{n\times p}{X}\,\underset{p\times p}{S^{-\frac{1}{2}}}}$$

tale per cui

- $\underset{n\times p}{\tilde{Z}}$ *ha vettore delle medie nullo* $\underset{p\times 1}{0}$
- $\underset{n\times p}{\tilde{Z}}$ *ha matrice di varianze/covarianze* $\underset{p\times p}{S^{\tilde{Z}}} = \underset{p\times p}{I}$

Questa trasformazione lineare dei dati originali $\underset{n\times p}{X}$ si chiama trasformazione di Mahalanobis

Dimostrazione. Il vettore delle medie di $\underset{n\times p}{\tilde{Z}}$ è

$$\frac{1}{n}\underset{p\times n}{\tilde{Z}'}\underset{n\times 1}{1} = \frac{1}{n}(\underset{p\times p}{S^{-\frac{1}{2}}})'\underset{p\times n}{\tilde{X}'}\underset{n\times 1}{1} = \underset{p\times p}{S^{-\frac{1}{2}}}\underset{p\times 1}{0} = \underset{p\times 1}{0}$$

ricordando che

- il vettore delle medie di $\underset{n\times p}{\tilde{X}}$ è nullo: $\frac{1}{n}\underset{p\times n}{\tilde{X}'}\underset{n\times 1}{1} = \underset{p\times 1}{0}$
- la decomposizione spettrale di $\underset{p\times p}{S}$ è $\underset{p\times p}{S} = \underset{p\times p}{V}\underset{p\times p}{\Lambda}\underset{p\times p}{V'}$
- per il Lemma, abbiamo $\underset{p\times p}{S^{-\frac{1}{2}}} = \underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'}$ simmetrica: $(\underset{p\times p}{S^{-\frac{1}{2}}})' = (\underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'})' = \underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'} = \underset{p\times p}{S^{-\frac{1}{2}}}$

La matrice di varianze/covarianze di $\underset{n\times p}{\tilde{Z}}$ è

$$\begin{aligned}
\underset{p\times p}{S^{\tilde{Z}}} &= \frac{1}{n}\underset{n\times p}{\tilde{Z}'}\underset{n\times n}{H}\underset{p\times n}{\tilde{Z}} = \frac{1}{n}(\underset{p\times p}{S^{-\frac{1}{2}}})'\underset{n\times p}{\tilde{X}'}\underset{n\times p}{\tilde{X}}\underset{p\times p}{S^{-\frac{1}{2}}} = \underset{p\times p}{S^{-\frac{1}{2}}}\underset{p\times p}{S}\underset{p\times p}{S^{-\frac{1}{2}}} \\
&= (\underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'})(\underset{p\times p}{V}\underset{p\times p}{\Lambda}\underset{p\times p}{V'})(\underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'}) = \underset{p\times p}{V}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{\Lambda}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'} \\
&= \underset{p\times p}{V}\underset{p\times p}{\Lambda}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{\Lambda^{-\frac{1}{2}}}\underset{p\times p}{V'} = \underset{p\times p}{V}\underset{p\times p}{\Lambda}\underset{p\times p}{\Lambda^{-1}}\underset{p\times p}{V'} = \underset{p\times p}{V}\underset{p\times p}{I}\underset{p\times p}{V'} = \underset{p\times p}{I}
\end{aligned}$$

ricordando che

- $\underset{p\times p}{V}$ è una matrice ortogonale: $\underset{p\times p}{V}\underset{p\times p}{V'} = \underset{p\times p}{V'}\underset{p\times p}{V} = \underset{p\times p}{I}$
- $\mathrm{diag}(a_1,\ldots,a_q)\mathrm{diag}(b_1,\ldots,b_q) = \mathrm{diag}(a_1b_1,\ldots,a_qb_q) = \mathrm{diag}(b_1,\ldots,b_q)\mathrm{diag}(a_1,\ldots,a_q)$

□

Riassumendo:

Matrice dei dati	Vettore delle medie	Matrice di varianze/covarianze	Matrice di correlazione
$\underset{n\times p}{X}$	$\underset{p\times 1}{\bar{x}}$	$\underset{p\times p}{S}$	$\underset{p\times p}{R}$
$\underset{n\times p}{\tilde{X}}$	$\underset{p\times 1}{0}$	$\underset{p\times p}{S^{\tilde{X}}} = \underset{p\times p}{S}$	$\underset{p\times p}{R^{\tilde{X}}} = \underset{p\times p}{R}$
$\underset{n\times p}{Z}$	$\underset{p\times 1}{0}$	$\underset{p\times p}{S^{Z}} = \underset{p\times p}{R}$	$\underset{p\times p}{R^{Z}} = \underset{p\times p}{R}$
$\underset{n\times p}{\tilde{Z}}$	$\underset{p\times 1}{0}$	$\underset{p\times p}{S^{\tilde{Z}}} = \underset{p\times p}{I}$	$\underset{p\times p}{R^{\tilde{Z}}} = \underset{p\times p}{I}$

4 Decomposizione in Valori Singolari

Si possono generalizzare le idee della Decomposizione Spettrale per ottenere la Decomposizione in Valori Singolari (*Singular Value Decomposition*) di una matrice rettangolare $\underset{m\times k}{A}$.

Theorem 4.1. *Sia $\underset{m\times k}{A}$ una matrice rettangolare a valori reali. Allora esiste una matrice ortogonale $\underset{m\times m}{U}$ e una matrice ortogonale $\underset{k\times k}{V}$ tali che*

$$\boxed{\underset{m\times k}{A} = \underset{m\times m}{U}\,\underset{m\times k}{\Delta}\,\underset{k\times k}{V'} = \sum_{j=1}^{\min(m,k)} \delta_j \underset{m\times 1}{u_j}\,\underset{1\times k}{v_j'}}$$

dove la matrice $\underset{m\times k}{\Delta}$ ha elemento di posizione (j,j) pari a $\delta_j \geq 0$ per $j = 1,\ldots,\min(m,k)$ e gli altri elementi pari a 0.
Le costanti $\delta_1 \geq \ldots \geq \delta_{\min(m,k)}$ sono dette valori singolari *di $\underset{m\times k}{A}$.*

Observation 4.2. *Abbiamo che*

- *i vettori $\underset{k\times 1}{v_j}$ e $\underset{m\times 1}{u_j}$ della SVD di $\underset{m\times k}{A}$ sono gli autovettori normalizzati delle matrici simmetriche $\underset{k\times k}{K} = \underset{k\times m}{A'}\,\underset{m\times k}{A}$ e $\underset{m\times m}{M} = \underset{m\times k}{A}\,\underset{k\times m}{A'}$, rispettivamente*
- *i valori δ_j della SVD di $\underset{m\times k}{A}$ sono la radice quadrata degli autovalori > 0 della matrice K (o, equivalentemente, della matrice M)*

Sia $\underset{m\times k}{A}$ con $\text{rango}(\underset{m\times k}{A}) = r \leq \min(m,k)$.
Le k colonne di $\underset{k\times k}{V}$ sono gli autovettori di $K = \underset{k\times m}{A'}\,\underset{m\times k}{A}$

$$\begin{aligned}
\underset{k\times m}{A'}\,\underset{m\times k}{A} &= \underset{k\times k}{V}\,\underset{k\times m}{\Delta'}\,\underset{m\times m}{U'}\,\underset{m\times m}{U}\,\underset{m\times k}{\Delta}\,\underset{k\times k}{V'} = \underset{k\times k}{V}\,\underset{k\times m}{\Delta'}\,\underset{m\times k}{\Delta}\,\underset{k\times k}{V'} \\
&= \underset{k\times k}{V}\begin{bmatrix} \underset{r\times r}{\Delta_r^2} & \underset{r\times(k-r)}{0} \\ \underset{(k-r)\times r}{0} & \underset{(k-r)\times(k-r)}{0} \end{bmatrix}\underset{k\times k}{V'} = \underset{k\times k}{V^K}\,\underset{k\times k}{\Lambda^K}\,\underset{k\times k}{(V^K)'}
\end{aligned}$$

Le m colonne di $\underset{m\times m}{U}$ sono gli autovettori di $M = \underset{m\times k}{A}\,\underset{k\times m}{A'}$

$$\begin{aligned}
\underset{m\times k}{A}\,\underset{k\times m}{A'} &= \underset{m\times m}{U}\,\underset{m\times k}{\Delta}\,\underset{k\times k}{V'}\,\underset{k\times k}{V}\,\underset{k\times m}{\Delta'}\,\underset{m\times m}{U'} = \underset{m\times m}{U}\,\underset{m\times k}{\Delta}\,\underset{k\times m}{\Delta'}\,\underset{m\times m}{U'} \\
&= \underset{m\times m}{U}\begin{bmatrix} \underset{r\times r}{\Delta_r^2} & \underset{r\times(m-r)}{0} \\ \underset{(m-r)\times r}{0} & \underset{(m-r)\times(m-r)}{0} \end{bmatrix}\underset{m\times m}{U'} = \underset{m\times m}{V^M}\,\underset{m\times m}{\Lambda^M}\,\underset{m\times m}{(V^M)'}
\end{aligned}$$

Example 4.3. *Sia* $\underset{2\times 3}{A} = \begin{bmatrix} 3 & 1 & 1 \\ -1 & 3 & 1 \end{bmatrix}$*. Allora* $M = AA' = \begin{bmatrix} 11 & 1 \\ 1 & 11 \end{bmatrix}$ *con autovalori* $\lambda_1^M = 12$ *e* $\lambda_2^M = 10$ *e autovettori* $v_1^M = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$ *e* $v_2^M = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$*. Inoltre* $K = A'A = \begin{bmatrix} 10 & 0 & 2 \\ 0 & 10 & 4 \\ 2 & 4 & 2 \end{bmatrix}$ *con autovalori* $\lambda_1^K = 12$*,* $\lambda_2^K = 10$ *e* $\lambda_3^K = 0$ *e autovettori* $v_1^K = \begin{bmatrix} 1/\sqrt{6} \\ 2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$, $v_2^K = \begin{bmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \\ 0 \end{bmatrix}$ *e* $v_3^K = \begin{bmatrix} 1/\sqrt{30} \\ 2/\sqrt{30} \\ -5/\sqrt{30} \end{bmatrix}$*. Quindi abbiamo*

$$A = \sqrt{\lambda_1^M} v_1^M (v_1^K)' + \sqrt{\lambda_2^M} v_2^M (v_2^K)' = d_1 u_1 v_1' + d_2 u_2 v_2'$$

La decomposizione in valori singolari può essere anche espressa in funzione del rango r di A.

- Sia $\underset{m\times k}{A}$ con $\text{rango}(\underset{m\times k}{A}) = r \leq \min(m,k)$.
- $\underset{r\times r}{\Delta_r^2} = \text{diag}(\delta_1^2, \ldots, \delta_r^2)$ i cui elementi diagonali sono gli r autovalori non nulli $\lambda_1^K \geq \lambda_2^K \geq \ldots \geq \lambda_r^K > 0$ di $K = \underset{k\times m}{A'} \underset{m\times k}{A}$ (o di $M = \underset{m\times k}{A} \underset{k\times m}{A'}$)

Possiamo scrivere

$$\boxed{\underset{m\times k}{A} = \underset{m\times m}{U} \underset{m\times k}{\Delta} \underset{k\times k}{V'} = \underset{m\times r}{U_r} \underset{r\times r}{\Delta_r} \underset{r\times k}{V_r'} = \sum_{j=1}^{r} \delta_j \underset{m\times 1}{u_j} \underset{1\times k}{v_j'}}$$

con

- $\underset{m\times r}{U_r} = \begin{bmatrix} \underset{m\times 1}{u_1} & \cdots & \underset{m\times 1}{u_r} \end{bmatrix}$
- $\underset{k\times r}{V_r} = \begin{bmatrix} \underset{k\times 1}{v_1} & \cdots & \underset{k\times 1}{v_r} \end{bmatrix}$
- $\underset{r\times r}{\Delta_r} = \text{diag}(\delta_1, \ldots, \delta_r)$ con $\delta_j > 0$