# Decomposizione Spettrale e Decomposizione in Valori Singolari Analisi Esplorativa

Aldo Solari



Decomposizione Spettrale

2 Matrice dei dati ortogonalizzati

- 3 Decomposizione in Valori Singolari
- 4 Appendice



### **Outline**

- Decomposizione Spettrale
- 2 Matrice dei dati ortogonalizzati
- 3 Decomposizione in Valori Singolari
- 4 Appendice



# Direzione e lunghezza di un vettore

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

• Lunghezza

$$\lambda = \|\underset{k \times 1}{a}\| = \sqrt{\underset{1 \times kk \times 1}{a'}} = \sqrt{\underset{j=1}{\overset{k}{\sum}}} a_j^2$$

• Direzione normalizzata

$$\underset{k\times 1}{v} = \frac{\overset{a}{k\times 1}}{\lambda}$$

$$\operatorname{con} \, \| \underset{k \times 1}{v} \| = 1$$



- Questa idea si può estendere ad una matrice simmetrica  $A\atop k\times k$  a valori reali, che è un insieme di vettori
- Gli autovalori (eigenvalues)  $\lambda$  e gli autovettori (eigenvectors) normalizzati  $v \in \mathbb{R} = 1$  sono definiti dall'equazione

$$\underset{k \times kk \times 1}{A} v = \lambda \underset{k \times 1}{v}$$

- Ci sono esattamente k coppie  $(\lambda_j,\ v_j\ ), j=1,\dots,k$  che soddisfano l'equazione
- Per convenzione, gli autovalori sono ordinati in maniera decrescente:

$$\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_k$$



- $A_{k \times k}$  una matrice simmetrica a valori reali
- A definita positiva  $\Leftrightarrow$  tutti gli autovalori di A sono positivi , i.e.  $\lambda_i>0, j=1,\ldots,k$
- A semidefinita positiva  $\Leftrightarrow$  tutti gli autovalori di A sono non negativi, i.e.  $\lambda_j \geq 0, \ j=1,\ldots,k$
- $A \atop k \times k$  con  $\mathrm{rango}(A \atop k \times k) = r \le k$ . Allora  $A \atop k \times k$  ha r autovalori non nulli, e i rimanenti k-r autovalori nulli
- Autovettori normalizzati  $v_j$  e  $v_l$  associati ad autovalori distinti  $\lambda_j \neq \lambda_l$  sono perpendicolari, i.e.  $v_j'$   $v_l = 0$   $1 \times k^{k \times 1}$



### Teorema di Decomposizione Spettrale

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

$$A = V \underset{k \times k}{\Lambda} V' = \sum_{j=1}^{k} \lambda_j \underbrace{v_j \ v'_j}_{k \times 1_1 \times k}$$

dove

- $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_k)$  è una matrice diagonale dove il j-simo elemento diagonale  $\lambda_j$  è il j-simo autovalore associato ad  $A = \sum_{k \times k} A_k$
- $\bullet \ \ \, \underset{k\times k}{V} = \left[ \begin{array}{ccc} v_1 & \cdots & v_k \\ k\times 1 & & k\times 1 \end{array} \right] \text{, dove la $j$-sima colonna } v_j \text{ è il $j$-simo autovettore normalizzato } \left( \parallel v_j \parallel = 1 \right) \text{ associato all'autovalore } \lambda_j$
- $V_{k \times k}$  è una matrice ortogonale:  $V_{k \times kk \times k} = V'_{k \times kk \times k} = I_{k \times k}$



### **Esempio**

$$\begin{array}{l} A \\ 2\times 2 \\ \end{array} = \left[ \begin{array}{cc} 2.2 & 0.4 \\ 0.4 & 2.8 \end{array} \right] \text{ simmetrica e a valori reali. Otteniamo gli autovalori} \\ \lambda_1 = 3 \text{ e } \lambda_2 = 2 \text{, a cui corrispondono gli autovalori (normalizzati)} \\ v_1 \\ z\times 1 \\ \end{array} = \left[ \begin{array}{c} 1/\sqrt{5} \\ 2/\sqrt{5} \end{array} \right] \text{ e } v_2 \\ z\times 1 \\ \end{array} = \left[ \begin{array}{c} 2/\sqrt{5} \\ -1/\sqrt{5} \end{array} \right] \text{, quindi} \\ \end{array}$$

$$\underset{2\times 2}{A} = 3 \left[ \begin{array}{c} 1/\sqrt{5} \\ 2/\sqrt{5} \end{array} \right] \left[ \begin{array}{c} 1/\sqrt{5} & 2/\sqrt{5} \end{array} \right] + 2 \left[ \begin{array}{c} 2/\sqrt{5} \\ -1/\sqrt{5} \end{array} \right] \left[ \begin{array}{c} 2/\sqrt{5} & -1/\sqrt{5} \end{array} \right]$$



### Lemma

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

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  $A \atop k imes k$  è definita positiva, vale anche per q < 0 razionali, ovvero per  $q \in \mathbb{Q}^{\backslash 0}$



## Applicazioni del Lemma

$$\bullet \ A^{-1} = \underset{k \times k}{V} \Lambda^{-1} V'$$

$$\bullet \ A^{-\frac{1}{2}} = V \Lambda^{-\frac{1}{2}} V'$$

$$k \times k \ k \times k \ k \times k$$

$$\bullet \ A^2 = \underset{k \times k}{V} \Lambda^2 V'$$



### Decomposizione Spettrale di S

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

$$S_{p \times p} = V \underset{p \times pp \times pp \times p}{\Lambda} V'$$

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

$$\operatorname{tr}(S) = \operatorname{tr}(V \underset{p \times p}{\Lambda} V') = \operatorname{tr}(N \underset{p \times pp \times pp \times p}{\Lambda} V') = \operatorname{tr}(N \underset{p \times p}{\Lambda} V$$

$$\det(S_{p\times p}) = \det(V \underset{p\times pp\times pp\times p}{\Lambda} V') = \det(V) \det(X_{p\times p}) \det(V')$$
$$= \det(X_{p\times p}) \det(V V') = \det(X_{p\times p}) \det(I) = \prod_{j=1}^{p} \lambda_{j}$$

Per le proprietà di traccia e determinante, vedi Appendice della lezione su varianza totale e generalizzata;



### **Outline**

- 1 Decomposizione Spettrale
- 2 Matrice dei dati ortogonalizzati
- 3 Decomposizione in Valori Singolari
- 4 Appendice



# Matrice dei dati ortogonalizzati $ilde{Z}$

Sia  $S_{p \times p}$ , la matrice di varianze/covarianze di  $S_{n \times p}$ , definita positiva. Possiamo allora ottenere la matrice dei dati ortogonalizzati

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

tale per cui

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

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



# Dati ortogonalizzati $ilde{Z}$

### Dimostrazione:

Il vettore delle medie di  $\tilde{\tilde{Z}}_{n\times p}$  è

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

#### ricordando che

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

$$(S_{p \times p}^{-\frac{1}{2}})' = (V_{p \times p} \Lambda_{p \times p}^{-\frac{1}{2}} V'_{p \times p})' = V_{p \times p} \Lambda_{p \times p}^{-\frac{1}{2}} V'_{p \times p} = S_{p \times p}^{-\frac{1}{2}}$$



# Dati ortogonalizzati $ilde{Z}$

#### Dimostrazione:

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

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

### ricordando che

- $V_{p \times p}$  è una matrice ortogonale:  $V_{p \times pp \times p} = V'_{p \times pp \times p} = I_{p \times pp \times p}$
- $\operatorname{diag}(a_1, \ldots, a_q)\operatorname{diag}(b_1, \ldots, b_q) = \operatorname{diag}(a_1b_1, \ldots, a_qb_q) = \operatorname{diag}(b_1, \ldots, b_q)\operatorname{diag}(a_1, \ldots, a_q)$

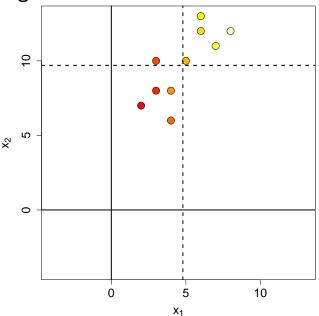


# Dati ortogonalizzati $\tilde{Z}$

Vettore delle medie	Matrice di varianze/covarianze	Matrice di correlazione
$\substack{ar{x} \\ p  imes 1}$	$\mathop{S}\limits_{p  imes p}$	$\mathop{R}_{p\times p}$
$\mathop{0}_{p\times 1}$	$S_{p \times p}^{\tilde{X}} = S_{p \times p}$	$R_{p\times p}^{\tilde{X}} = R_{p\times p}$
$_{p imes1}^{0}$	$S^Z_{p \times p} = \underset{p \times p}{R}$	$\underset{p\times p}{R^Z} = \underset{p\times p}{R}$
$\mathop{0}\limits_{p imes 1}$	$S^{\tilde{Z}}_{p \times p} = I_{p \times p}$	$R^{\tilde{Z}}_{p \times p} = I_{p \times p}$
	medie $\bar{x}$ $p \times 1$ $0$ $p \times 1$ $0$ $p \times 1$	medie varianze/covarianze $\frac{\bar{x}}{p \times 1}$ $S \\ p \times p$ $S \\ p \times p$ $S \\ N \\ $

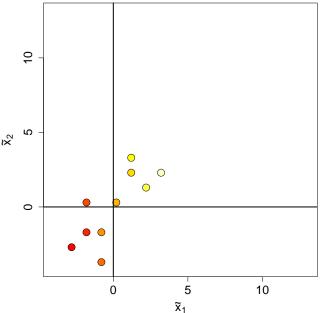


### Matrice originale X



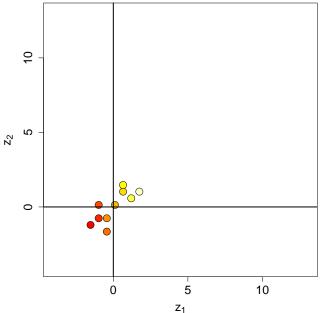


# Matrice centrata $\tilde{X}$ (traslazione)



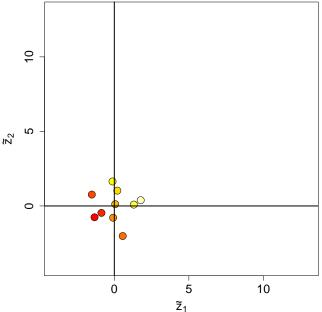


# Matrice standardizzata Z (compr./dilat.)





# Matrice ortogonalizzata $\tilde{Z}$ (decorrelazione)





### **Outline**

- 1 Decomposizione Spettrale
- 2 Matrice dei dati ortogonalizzati
- 3 Decomposizione in Valori Singolari
- 4 Appendice



- Si possono generalizzare le idee della Decomposizione Spettrale per ottenere la Decomposizione in Valori Singolari (Singular Value Decomposition) di una matrice rettangolare  $\underset{m \times k}{A}$ , dove:
- i vettori  $v_j$  e  $u_j$  della SVD di  $A \atop m \times k$  sono gli autovettori normalizzati delle matrici simmetriche  $K = A' A \atop k \times m \times k$  e  $M = A A' \atop m \times m = M A' \atop m \times k k \times m$ , rispettivamente
- i valori  $\delta_j$  della SVD di  $A \atop m \times k$  sono la radice quadrata degli autovalori > 0 della matrice K (o, equivalentemente, della matrice M)



Sia  $A \atop m imes k$  una matrice rettangolare a valori reali. Allora esiste una matrice ortogonale  $U \atop m imes m$  e una matrice ortogonale  $V \atop k imes k$  tali che

$$A = U \Delta V' = \sum_{m \times m} \delta_j u_j v'_j$$

$$M \times k = \sum_{j=1}^{\min(m,k)} \delta_j u_j v'_j$$

$$M \times 1_{1 \times k}$$

dove la matrice  $\Delta \atop m imes k$  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  $A \leq \ldots \leq \delta_{\min(m,k)}$ 



- Sia  $A \atop m \times k$  con  $\operatorname{rango}(A \atop m \times k) = r \le \min(m, k)$ .
- Le k colonne di V sono gli autovettori di K = A' A  $k \times m \times k$

$$A' A = V \Delta' U' U \Delta V' = V \Delta' \Delta V'$$

$$k \times mm \times k = V \sum_{k \times kk \times mm \times mm \times mm \times kk \times k} V' = V \sum_{k \times kk \times mm \times kk \times k} V'$$

$$= V \begin{bmatrix} \Delta_r^2 & 0 \\ r \times r & r \times (k-r) \\ 0 & 0 \\ (k-r) \times r & (k-r) \times (k-r) \end{bmatrix} V' = V \sum_{k \times kk \times mm \times kk \times k} V'$$



- Sia  $\underset{m \times k}{A}$  con rango $(\underset{m \times k}{A}) = r \le \min(m, k)$ .
- $\Delta_r^2 = \operatorname{diag}(\delta_1^2,\dots,\delta_r^2)$  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  $K = \underset{k \times mm \times k}{A'} \Lambda$  (o di  $M = \underset{m \times kk \times m}{A} \Lambda'$ )

Possiamo scrivere

$$A_{m \times k} = U_{m \times m} \Delta V'_{m \times kk \times k} = U_{r} \Delta_{r} V'_{r} = \sum_{j=1}^{r} \delta_{j} u_{j} v'_{j}$$

$$M \times k = U_{r} \Delta_{r} V'_{r} = \sum_{j=1}^{r} \delta_{j} u_{j} v'_{j}$$

con

$$\bullet \ \, \underset{m \times r}{U_r} = \left[ \begin{array}{ccc} u_1 & \cdots & u_r \\ m \times 1 & & m \times 1 \end{array} \right]$$

$$\bullet \ \ V_r = \left[ \begin{array}{ccc} v_1 & \dots & v_r \\ k \times r & k \times 1 \end{array} \right]$$

• 
$$\Delta_r = \operatorname{diag}(\delta_1, \dots, \delta_r) \operatorname{con} \delta_j > 0$$



### Matrice di varianze/covarianze S

 $\bullet$  Decomposizione spettrale di  $\underset{p\times p}{S}$ 

$$\underset{p \times p}{S} = \underset{p \times pp \times pp \times p}{N} V'$$

 $\bullet$  Decomposizione in valori singolari di  $\tilde{X}_{n\times p}$ 

$$\tilde{X}_{n \times p} = \underset{n \times n}{U} \underset{n \times p}{\Delta} V'$$

 $\bullet$  Decomposizione in valori singolari di  $\underset{p\times p}{S}$ 

$$S = \frac{1}{n}\tilde{X}'\tilde{X} = \frac{1}{n}V\Delta U'U\Delta V' = V\frac{1}{n}\Delta^2 V'$$



### **Outline**

- Decomposizione Spettrale
- 2 Matrice dei dati ortogonalizzati
- 3 Decomposizione in Valori Singolari
- 4 Appendice



Per ogni matrice quadrata  $\underset{k\times k}{A}$  , si possono trovare uno scalare  $\lambda$  e un vettore  $\underset{k\times 1}{v}$  tali che

$$\underset{k \times k}{A} v = \lambda v \\
 \underset{k \times 1}{v}
 \tag{1}$$

 $\lambda$  è detto autovalore (*eigenvalue*) di  $\underset{k \times k}{A}$  e  $\underset{k \times 1}{v}$  è l'autovettore (*eigenvector*) di  $\underset{k \times k}{A}$  correspondente a  $\lambda$ . Per risolvere l'equazione (1) scriviamo:

$$(A - \lambda I)_{k \times k} v = 0$$

$$k \times 1$$

$$k \times 1$$

$$k \times 1$$

$$k \times 1$$



- Se  $\det(\underset{k\times k}{A}-\lambda\underset{k\times k}{I})\neq 0$ ,  $(\underset{k\times k}{A}-\lambda\underset{k\times k}{I})$  è invertibile e  $\underset{k\times 1}{v}=\underset{k\times 1}{0}$  è l'unica soluzione. Per escludere le soluzioni banali, imponiamo quindi  $\det(\underset{k\times k}{A}-\lambda\underset{k\times k}{I})=0$  in modo tale da ottenere valori di  $\lambda$  che possano essere sostituiti in (1) e (2) per ottenere i corrispondenti valori di  $\underset{k\times 1}{v}$ .
- Quindi, affinchè (2) abbia una soluzione, è necessario che le colonne di  $A \lambda I = 0$  siano linearmente dipendenti. Concludendo,  $A \lambda I = 0$  deve essere singolare (determinate pari a 0) affinchè esista un A = 0 = 0 che sia soluzione di (1).
- L'equazione  $\det(A \atop k \times k} \lambda \atop k \times k)$  è detta equazione caratteristica. Visto che  $A \atop k \times k$  ha dimensioni  $k \times k$ , l'equazione caratteristica ha k radici; cioè,  $A \atop k \times k$  ha k autovalori  $\lambda_1, \lambda_2, \ldots, \lambda_k$ . I valori  $\lambda_j$  non sono tutti necessariamente distinti da zero.



- Però, se  $A \in \mathbb{R}$  è non singolare, gli autovalori saranno tutti diversi tra loro. Gli autovettori  $v_1$ ,  $v_1$ ,  $v_1$ ,  $v_2$ ,  $v_k$  associati a  $\lambda_1, \lambda_2, \dots, \lambda_k$  sono definiti da (2).
- Notiamo che se moltiplichiamo entrambi i membri della (1) per uno scalare k (e sfruttiamo la proprietà commutativa), otteniamo

$$(A - \lambda I)c v = c0 = 0.$$

Quindi, se  $v \in \mathbb{R}^n$  è un autovettore di  $v \in \mathbb{R}^n$ , anche  $v \in \mathbb{R}^n$  è un autovettore. In generale, ogni autovettore è unico a meno di valori scalari  $v \in \mathbb{R}^n$ . Si noti in ogni caso che la direzione indicata dal vettore  $v \in \mathbb{R}^n$  è unica (la moltiplicazione per  $v \in \mathbb{R}^n$  mantiene costanti i rapporti delle coordinate) e quindi la soluzione è essenzialmente unica.

• Per questo motivo è uso definire l'autovettore normalizzato  $\underset{k\times 1}{e} = \underset{k\times 1}{v}/\|\underset{k\times 1}{v}\| \text{ di lunghezza unitaria, cioè tale che } \underset{1\times kk\times 1}{e'} = 1.$ 

## Autovalori e autovettori: esempio

 Facciamo un esempio, troviamo gli autovalori e autovettori della matrice

$$\underset{2\times 2}{A} = \left[ \begin{array}{cc} 1 & 2\\ -1 & 4 \end{array} \right]$$

• L'equazione caratteristica è

$$\det({}_{2\times 2}^{A} - \lambda_{2\times 2}^{I}) = \det\begin{pmatrix} 1 - \lambda & 2 \\ -1 & 4 - \lambda \end{pmatrix} = (1 - \lambda)(4 - \lambda) + 2 = 0$$
$$\lambda^{2} - 5\lambda + 6 = (\lambda - 3)(\lambda - 2) = 0,$$

da cui si ricavano  $\lambda_1 = 3$  e  $\lambda_2 = 2$ .



# Autovalori e autovettori: esempio

• Per trovare l'autovettore  $e_1$  corrispondente a  $\lambda_1=3$ , usiamo la (2):

$$\begin{pmatrix} A - \lambda_{2 \times 2}^{I} e_{11} & 0 \\ 2 \times 2 & 2 \times 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \times 1 \end{pmatrix}$$

$$\begin{bmatrix} 1 - 3 & 2 \\ -1 & 4 - 3 \end{bmatrix} \begin{bmatrix} e_{11} \\ e_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{cases} -2e_{11} + 2e_{21} = 0 \\ -e_{11} + e_{21} = 0. \end{cases}$$



### Autovalori e autovettori: esempio

• Come atteso, la seconda equazione è ridondante in presenza della prima (e viceversa), rimane quindi una sola equazione con due parametri:  $e_{11}=e_{21}$ . Ora se all'unica equazione aggiungiamo la condizione di vettore di lunghezza unitaria ( $\parallel e_1 \parallel = 1$  cioè

$$\sqrt{e_{11}^2 + e_{21}^2} = 1$$
) otteniamo

$$e_1 = \left[ \begin{array}{c} e_{11} \\ e_{21} \end{array} \right] = \left[ \begin{array}{c} 1/\sqrt{2} \\ 1/\sqrt{2} \end{array} \right]$$

• In modo analogo, in corrispondenza di  $\lambda_2=2$ , si ha

$$e_2_{2\times 1} = \left[\begin{array}{c} 2/\sqrt{5} \\ 1/\sqrt{5} \end{array}\right].$$



### Autovalori e autovettori: ulteriori risultati

- Nonostante gli autovalori di una matrice A = A non siano necessariamente tutti reali, ciò è garantito nel caso A = A sia simmetrica e reale.
- Sia  $A \atop k \times k$  una matrice  $k \times k$  and  $B \atop k \times k$  una matrice  $k \times k$  con  $\mathrm{rango}(B \atop k \times k) = k. \ \ A \atop k \times k} \ \mathrm{e} \ B^{-1} \atop k \times k \times k \times k \times k \times k$  hanno gli stessi autovalori.
- una matrice  $\underset{k\times k}{A}$  con  $\mathrm{rango}(\underset{k\times k}{A}) < k$  ha almeno un autovalore pari a zero.
- una matrice idempotente ha solo autovalori pari a 1 o 0.
- Se  $A \atop k imes k$  è semidefinita positiva, allora  $\lambda_j \geq 0, j=1,\dots,k$ ;
  - Se  $\underset{k imes k}{A}$  è definita positiva, allora  $\lambda_j > 0, j = 1, \dots, k$
- Autovettori associati a distinti autovalori sono perpendicolari (se  $\lambda_j \neq \lambda_l, \ e'_j \ e_l = 0$ )

