Varianza totale e generalizzata

Analisi Esplorativa

Aldo Solari

1 Varianza totale

2 Varianza generalizzata

3 Appendice

Variabilità

- Nel caso $p = 1$, la variabilità (o dispersione) presente nelle misurazioni della variabile considerata è descritta da un singolo numero, la varianza $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$
- Nel caso $p > 1$, la variabilità presente nelle misurazioni delle p variabili considerate è descritta da p varianze $s_{jj}, j = 1, \ldots, p$ e $\frac{1}{2}p(p-1)$ covarianze $s_{jk}, j \neq k = 1, \ldots, p$, ovvero

  $$p + \frac{1}{2}p(p-1)$$

  numeri, contenuti nella matrice di varianze/covarianze $\underset{p \times p}{S}$
- Possiamo riassumere la variabilità descritta da $\underset{p \times p}{S}$ in un singolo numero (senza perdere troppa informazione)?

Outline

1. Varianza totale
2. Varianza generalizzata
3. Appendice

Varianza totale

$$\boxed{\text{Varianza totale} = \mathrm{tr}(\underset{p\times p}{S}) = \sum_{j=1}^{p} s_{jj}}$$

Data una matrice quadrata $\underset{q\times q}{A}$, la traccia di $\underset{q\times q}{A}$ è definita come $\mathrm{tr}(\underset{q\times q}{A}) = \sum_{j=1}^{q} a_{jj}$, ovvero la somma dei suoi elementi diagonali; vedi Appendice;

Varianza totale: spazio delle osservazioni

Nello spazio delle osservazioni, la varianza totale può essere interpretata come la somma delle lunghezze al quadrato dei p vettori scarto dalla media $\underset{p\times 1}{\tilde{x}_j}$, $j = 1, \ldots, p$, divisa per n

Dimostrazione:

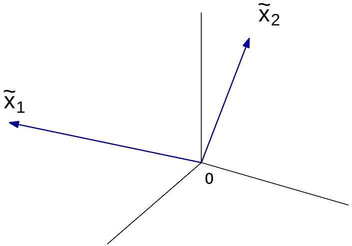
$$\mathrm{tr}(\underset{p\times p}{S}) = \frac{1}{n}\sum_{j=1}^{p} n s_{jj} = \frac{1}{n}\sum_{j=1}^{p} \| \underset{p\times 1}{\tilde{x}_j} \|^2$$

□

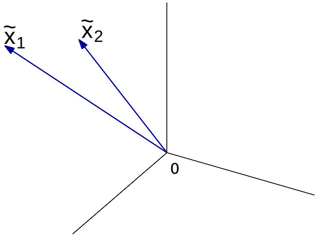
$$\underset{3\times 2}{\tilde{X}} = \begin{bmatrix} 2 & -2 \\ -3 & 0 \\ 1 & 2 \end{bmatrix}$$

$$\mathrm{tr}(\underset{2\times 2}{S}) = \frac{14}{3} + \frac{8}{3}$$

$$r_{12} = -0.19$$



$$\underset{3\times 2}{\tilde{X}} = \begin{bmatrix} 1 & 0 \\ -3 & -2 \\ 2 & 2 \end{bmatrix}$$



$$\mathrm{tr}(\underset{2\times 2}{S}) = \tfrac{14}{3} + \tfrac{8}{3}$$

$$r_{12} = 0.95$$

Distanza Euclidea tra 2 punti p-dimensionali

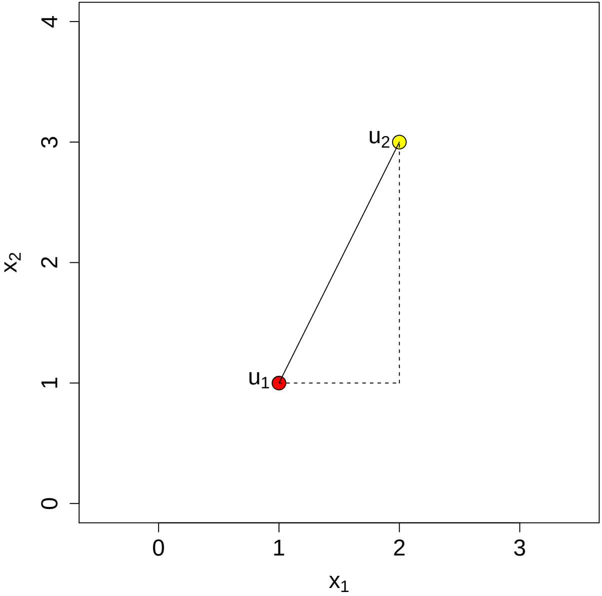
- tra due unità statistiche $\underset{1\times p}{u'_i}$ e $\underset{1\times p}{u'_l}$:

$$d(u_i, u_l) = \sqrt{\underset{1\times p}{(u_i - u_l)'}\underset{p\times 1}{(u_i - u_l)}} = \sqrt{\sum_{j=1}^{p}(x_{ij} - x_{lj})^2}$$

- tra l'i-sima unità statistica $\underset{1\times p}{u'_i}$ e il baricentro $\underset{1\times p}{\bar{x}'}$:

$$d(u_i, \bar{x}) = \sqrt{\underset{1\times p}{(u_i - \bar{x})'}\underset{p\times 1}{(u_i - \bar{x})}} = \sqrt{\sum_{j=1}^{p}(x_{ij} - \bar{x}_j)^2}$$

ricordando che $\underset{1\times p}{u'_i} = \underset{1\times p}{x'_i}$ e $u_{ij} = x_{ij}$



$$\underset{1\times 2}{u_1'} = [1\ 1],\ \underset{1\times 2}{u_2'} = [2\ 3],\ d(u_1, u_2) = \sqrt{(1-2)^2 + (1-3)^2} = \sqrt{5}$$

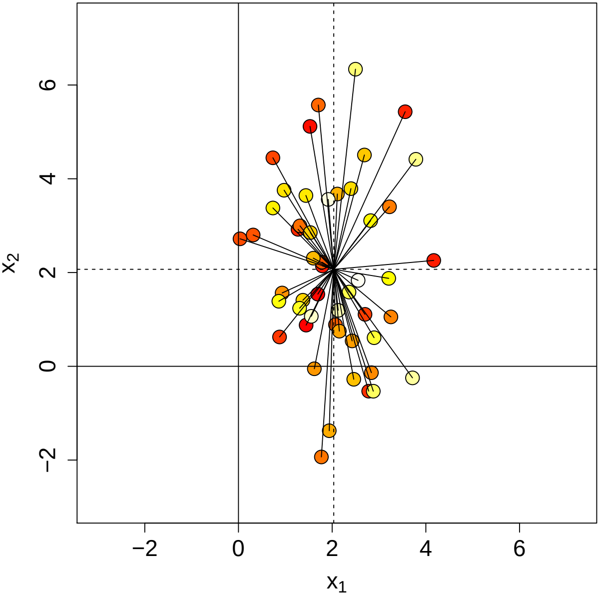
Varianza totale: spazio delle variabili

Nello spazio delle variabili, la varianza totale può essere interpretata come la media aritmetica delle distanze Euclidee al quadrato delle n unità statistiche $\underset{1\times p}{u_i'}$, $i = 1, \dots, n$, dal baricentro $\underset{1\times p}{\bar{x}'}$

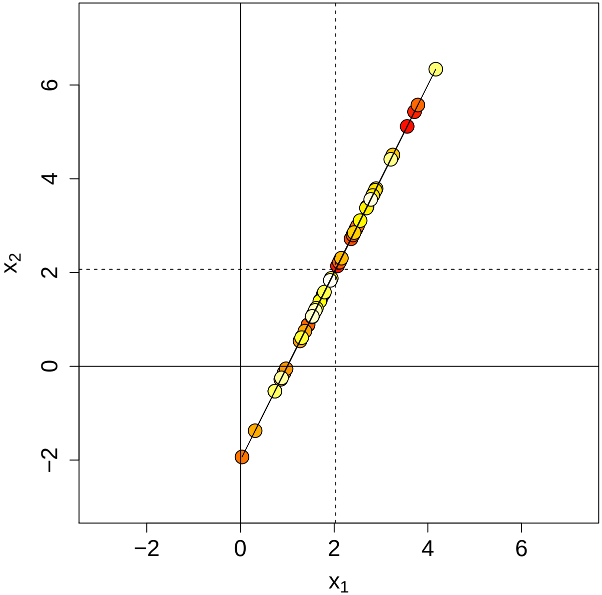
Dimostrazione:

$$\operatorname{tr}(\underset{p\times p}{S}) = \sum_{j=1}^{p} \frac{1}{n} \sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{p} (u_{ij} - \bar{x}_j)^2 = \frac{1}{n} \sum_{i=1}^{n} d^2(u_i, \bar{x})$$

□



$$\mathrm{tr}(\mathrm{S}) = s_{11} + s_{22} = 0.84 + 3.36 = 4.2,\ r_{12} = -0.07$$



$$\mathrm{tr}(\mathrm{S}) = s_{11} + s_{22} = 0.84 + 3.36 = 4.2,\ r_{12} = 1$$

Varianza totale: cosa perdiamo

Sintetizzando la matrice di varianze/covarianze con un singolo numero dato dalla varianza totale, perdiamo tutta l'informazione sulla struttura di correlazione (di covarianza) tra le p variabili

Outline

1. Varianza totale
2. Varianza generalizzata
3. Appendice

Varianza generalizzata

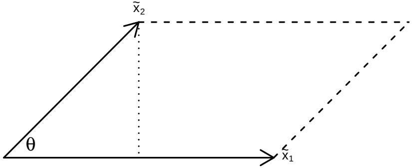
$$\boxed{\text{Varianza generalizzata} = \det(\underset{p\times p}{S})}$$

$\det(\underset{q\times q}{A})$, indica il determinante di una matrice quadrata $\underset{q\times q}{A}$; vedi Appendice;

Varianza generalizzata: spazio delle oss.

Consideriamo geometricamente l'area generata da $p = 2$ vettori scarto dalla media $\underset{n\times 1}{\tilde{x}_1}$ e $\underset{n\times 1}{\tilde{x}_2}$ nello spazio n-dimensionale

$$\begin{aligned}\text{Area parallelogramma} &= \text{base paral.} \cdot \text{altezza paral.}\\ &= \| \underset{n\times 1}{\tilde{x}_1} \| \cdot \| \underset{n\times 1}{\tilde{x}_2} \| \sqrt{1-\cos^2(\theta)}\\ &= n\sqrt{s_{11}s_{22}(1-r_{12}^2)}\end{aligned}$$



Varianza generalizzata: spazio delle oss.

$$\begin{aligned}\det(\underset{2\times 2}{S}) &= \det\left(\begin{bmatrix} s_{11} & \sqrt{s_{11}}\sqrt{s_{22}}r_{12} \\ \sqrt{s_{11}}\sqrt{s_{22}}r_{12} & s_{22}\end{bmatrix}\right) \\ &= s_{11}s_{22} - s_{11}s_{22}r_{12}^2 \\ &= s_{11}s_{22}(1 - r_{12}^2)\end{aligned}$$

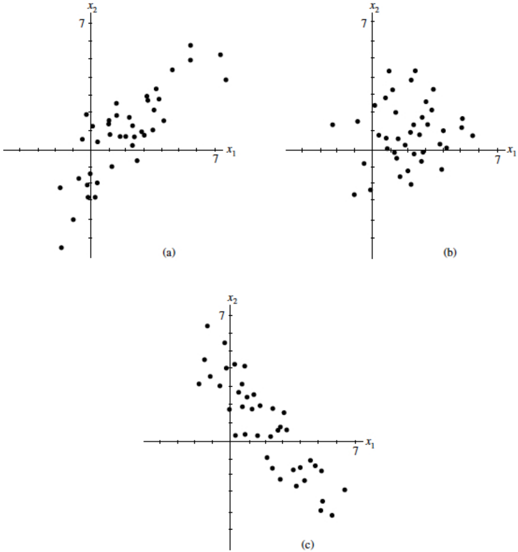
Quindi

$$\det(\underset{2\times 2}{S}) = \frac{(\text{Area parallelogramma})^2}{n^2}$$

In generale, per p vettori n-dimensionali $\underset{n\times 1}{\tilde{x}_j}, j = 1, \ldots, p$:

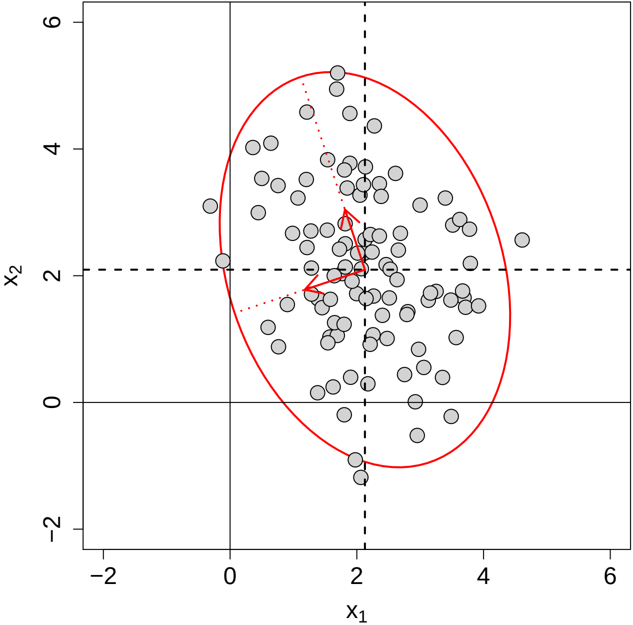
$$\det(\underset{p\times p}{S}) = \frac{(\text{Volume parallelepipedo } p\text{−dimensionale})^2}{n^p}$$

Varianza generalizzata: spazio delle variabili



Qual è la configurazione più variabile (in termini di varianza generalizzata)?

Varianza generalizzata: spazio delle variabili



Autovalori e autovettori di $\underset{p \times p}{S}$

Alla matrice di varianze/covarianze $\underset{p \times p}{S}$ possiamo associare p coppie di autovalori e autovettori

$$(\lambda_1, \underset{p \times 1}{v_1}), (\lambda_2, \underset{p \times 1}{v_2}), \ldots, (\lambda_p, \underset{p \times 1}{v_p})$$

dove gli autovalori (*eigenvalues*) sono ordinati in maniera decrescente, ovvero

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \geq 0$$

e dove gli autovettori (*eigenvectors*) $\underset{p \times 1}{v_1}, \ldots, \underset{p \times 1}{v_p}$ sono tali che

- hanno lunghezza unitaria $\| \underset{1 \times p}{v_1} \| = \ldots = \| \underset{1 \times p}{v_p} \| = 1$
- sono mutualmente perpendicolari: $\underset{1 \times p}{v_j'} \underset{p \times 1}{v_k} = 0$ per $j \neq k$

Iper-ellissoide

L'equazione

$$\underset{1\times p}{(x-\bar{x})'}\underset{p\times p}{S^{-1}}\underset{p\times 1}{(x-\bar{x})} = c^2$$

definisce l'iper-elissoide

- centrato sul baricentro $\underset{1\times p}{\bar{x}'}$
- con il j-simo asse orientato secondo il j-simo autovettore $\underset{p\times 1}{v_j}$ di $\underset{p\times p}{S}$
- la (metà) lunghezza del j-simo asse é $\frac{c}{\sqrt{\lambda_j}}$, proporzionale al j-simo autovalore λ_j di $\underset{p\times p}{S}$,

dove stiamo assumendo che $\underset{p\times p}{S}$ è una matrice definita positiva in modo da garantire l'esistenza di $\underset{p\times p}{S^{-1}}$

Varianza generalizzata: spazio delle variabili

Il volume dell'iper-ellissoide è funzione della varianza generalizzata:

$$\text{Volume di } \left\{ \underset{1\times p}{x'} : (\underset{1\times p}{x - \bar{x}})' \underset{p\times p}{S^{-1}} (\underset{p\times 1}{x - \bar{x}}) \leq c^2 \right\} = k_p c^p \sqrt{\det(S)}$$

dove $k_p = \dfrac{2\pi^{p/2}}{p\Gamma(p/2)}$ e $\Gamma(\cdot)$ è la funzione Gamma.

Quindi

$$(\text{Volume iperellissoide})^2 = (\text{costante})(\text{varianza generalizzata})$$

Varianza generalizzata: spazio delle variabili

(a) $\underset{2\times 1}{\bar{x}'} = \begin{bmatrix} 2 & 1 \end{bmatrix}$, $\underset{2\times 2}{S} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}, r_{12} = 0.8$, $\det(\underset{2\times 2}{S}) = 9$,

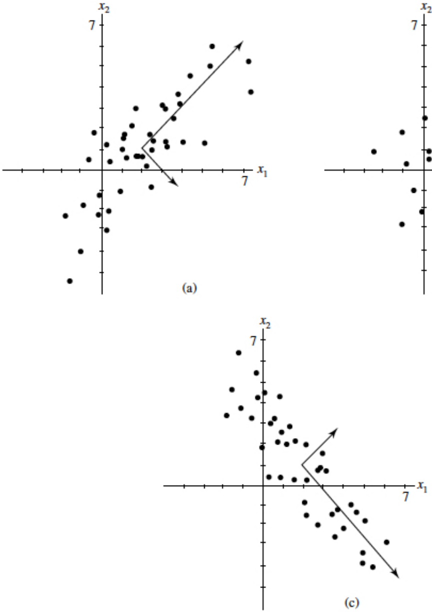
$\underset{2\times 1}{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \lambda_1 = 9$, $\underset{2\times 1}{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \lambda_2 = 1$

(b) $\underset{2\times 1}{\bar{x}'} = \begin{bmatrix} 2 & 1 \end{bmatrix}$, $\underset{2\times 2}{S} = \begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix}, r_{12} = 0$, $\det(\underset{2\times 2}{S}) = 9$,

$\underset{2\times 1}{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \lambda_1 = 3$, $\underset{2\times 1}{v_2} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \lambda_2 = 3$

(c) $\underset{2\times 1}{\bar{x}'} = \begin{bmatrix} 2 & 1 \end{bmatrix}$, $\underset{2\times 2}{S} = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}, r_{12} = -0.8$, $\det(\underset{2\times 2}{S}) = 9$

$\underset{2\times 1}{v_1} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \lambda_1 = 9$, $\underset{2\times 1}{v_2} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \lambda_2 = 1$



x_2
7
x_1
7
(a)
x_2
7
x_1
7
(b)
x_2
7
x_1
7
(c)

Varianza generalizzata: cosa perdiamo

Sintetizzando la matrice di varianze/covarianze con un singolo numero dato dalla varianza generalizzata, perdiamo l'informazione riguardante l'orientamento della nuvola di punti p-dimensionale formata dalle n unità statistiche

Quando la varianza generalizzata è 0?

La varianza generalizzata è 0 se e solo se le colonne di $\underset{n\times p}{\tilde{X}}$ sono linearmente dipendenti.

Si ricordi che le colonne di $\underset{n\times p}{\tilde{X}}$, ovvero i vettori $\underset{n\times 1}{\tilde{x}_j}$, $j = 1, \ldots, p$, sono linearmente dipendenti se esiste un vettore non nullo $\underset{p\times 1}{c} \neq \underset{p\times 1}{0}$ tale che

$$\underset{n\times p}{\tilde{X}}\underset{p\times 1}{c} = c_1 \underset{n\times 1}{\tilde{x}_1} + \ldots + c_p \underset{n\times 1}{\tilde{x}_p} = \underset{n\times 1}{0}$$

Per la definizione di vettori linearmente dipendenti, vedi Appendice;

Quando la varianza generalizzata è 0?

Dimostrazione $\Leftarrow$:

Vedi lavagna e/o libro di testo (la dimostrazione verrà resa disponibile alla fine del corso)

Quando la varianza generalizzata è 0?

Dimostrazione $\Rightarrow$:

Vedi lavagna e/o libro di testo (la dimostrazione verrà resa disponibile alla fine del corso)

Esempio

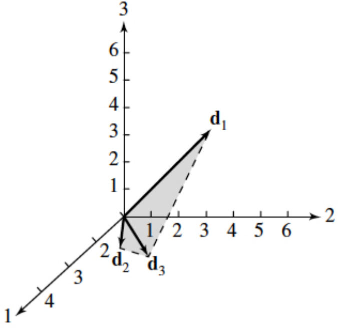
$$\underset{3\times 3}{X} = \begin{bmatrix} 1 & 2 & 5 \\ 4 & 1 & 6 \\ 4 & 0 & 4 \end{bmatrix}, \ \underset{3\times 3}{\tilde{X}} = \begin{bmatrix} -2 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & -1 \end{bmatrix}, \text{ quindi poichè}$$

$$\underset{3\times 1}{\tilde{x}_3} = \underset{3\times 1}{\tilde{x}_1} + 2\,\underset{3\times 1}{\tilde{x}_2}$$

le colonne $\underset{3\times 3}{\tilde{X}}$ sono linearmente dipendenti, ovvero $\underset{3\times 3}{\tilde{X}}\underset{3\times 1}{c} = \underset{3\times 1}{0}$ con $\underset{3\times 1}{c} = \begin{bmatrix} 1 \\ 2 \\ -1 \end{bmatrix} \neq \underset{3\times 1}{0}$.

Esempio

Geometricamente questo significa che uno dei vettori scarto dalla media, ad esempio $\underset{3\times 1}{\tilde{x}_3}$, giace nel piano generato da $\underset{3\times 1}{\tilde{x}_1}$ e $\underset{3\times 1}{\tilde{x}_2}$. Di conseguenza, il volume del parallelepipedo tridimensionale è 0.



Se $n \leq p$, allora $\det(S) = 0$

Se $n \leq p$, ovvero

$$(\text{numerosità}\ \ \text{dei dati}) \leq (\text{dimensionalità}\ \ \text{dei dati})$$

allora $\det(\underset{p\times p}{S}) = 0$.

Se $n \leq p$, allora $\det(S) = 0$

Dimostrazione:

Vedi lavagna e/o libro di testo (la dimostrazione verrà resa disponibile alla fine del corso)

Varianza generalizzata per dati standardizzati

$$\boxed{\begin{array}{c}\text{Varianza generalizzata per} \\ \text{dati standardizzati } \underset{n\times p}{Z}\end{array} = \det(\underset{p\times p}{S^Z}) = \det(\underset{p\times p}{R})}$$

Relazione tra $\det(S)$ e $\det(R)$

Dimostrazione:

$$
\begin{aligned}
\det(\underset{p\times p}{S}) &= \det(\underset{p\times p}{D^{1/2}}\,\underset{p\times p}{R}\,\underset{p\times p}{D^{1/2}}) \\
&= \det(\underset{p\times p}{D^{1/2}})\det(\underset{p\times p}{R})\det(\underset{p\times p}{D^{1/2}}) \\
&= (s_{11}s_{22}\cdots s_{pp})\det(\underset{p\times p}{R}) \\
&= \left(\prod_{j=1}^{p} s_{jj}\right)\det(\underset{p\times p}{R})
\end{aligned}
$$

dove $\underset{p\times p}{D^{1/2}} = \mathrm{diag}(\sqrt{s_{11}}, \ldots, \sqrt{s_{pp}})$

$\square$

Per le proprietà del determinante, vedi Appendice;

Varianza generalizzata per dati originali e dati standardizzati

- Se cambiamo l'unità di misura per la prima variabile $\underset{n\times 1}{x_1}$, ad esempio da Kg a gr, e quindi moltiplicando $\underset{n\times 1}{x_1}$ per 1000, abbiamo che la varianza s_{11} aumenta di un fattore moltiplicativo pari a 1000^2
- Questo cambio di unità di misura da Kg a gr influenza la varianza generalizzata:

$$\det(\underset{p\times p}{S^{gr}}) = \Big((1000^2 s_{11}) s_{22} \cdots s_{pp}\Big) \det(\underset{p\times p}{R}) = 1000^2 \det(\underset{p\times p}{S^{Kg}})$$

- Per questo motivo, spesso è conveniente calcolare la varianza generalizzata considerando i dati standardizzati $\underset{n\times p}{Z}$

Indice relativo di variabilità

$$0 \leq \begin{array}{c}\text{Indice relativo}\\ \text{di variabilità}\end{array} = \det(\underset{p\times p}{R}) = \frac{\det(\underset{p\times p}{S})}{\prod_{j=1}^{p} s_{jj}} \leq 1$$

Esempio

$$\underset{3\times 3}{S} = \begin{bmatrix} 4 & 3 & 1 \\ 3 & 9 & 2 \\ 1 & 2 & 1 \end{bmatrix}, \ \underset{3\times 3}{R} = \begin{bmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 & \frac{2}{3} \\ \frac{1}{2} & \frac{2}{3} & 1 \end{bmatrix},$$

$\det(\underset{3\times 3}{S}) = 14$, $\det(\underset{3\times 3}{R}) = \frac{7}{18}$,

$$14 = \det(\underset{3\times 3}{S}) = s_{11}s_{22}s_{33}\det(\underset{3\times 3}{R}) = (4)(9)(1)\frac{7}{18} = 14$$

Outline

1 Varianza totale

2 Varianza generalizzata

3 Appendice

Traccia di una matrice

La traccia di una matrice quadrata $\underset{q\times q}{A}$ è la somma degli elementi diagonali:

$$\mathrm{tr}(\underset{q\times q}{A}) = \sum_{j=1}^{p} a_{jj}$$

Sia $\underset{q\times q}{B}$ una matrice quadrata e c una costante.

- $\mathrm{tr}(c\underset{q\times q}{A}) = c\,\mathrm{tr}(\underset{q\times q}{A})$
- $\mathrm{tr}(\underset{q\times q}{A} \pm \underset{q\times q}{B}) = \mathrm{tr}(\underset{q\times q}{A}) \pm \mathrm{tr}(\underset{q\times q}{B})$
- $\mathrm{tr}(\underset{q\times q}{A}\underset{q\times q}{B}) = \mathrm{tr}(\underset{q\times q}{B}\underset{q\times q}{A})$
- $\mathrm{tr}(\underset{q\times q}{B^{-1}}\underset{q\times q}{A}\underset{q\times q}{B}) = \mathrm{tr}(\underset{q\times q}{A})$
- $\mathrm{tr}(\underset{q\times q}{A}\underset{q\times q}{A'}) = \sum_{j=1}^{q}\sum_{k=1}^{q} a_{jk}^2$

Vettori linearmente dipendenti

I vettori $\underset{m\times 1}{a_1}, \ldots, \underset{m\times 1}{a_q}$ sono detti linearmente dipendenti se esistono delle costanti $c_1, \ldots, c_q$ non tutte nulle tali che

$$c_1 \underset{m\times 1}{a_1} + \ldots + c_q \underset{m\times 1}{a_q} = \underset{m\times 1}{0}$$

Se $\underset{m\times 1}{a_1}, \ldots, \underset{m\times 1}{a_q}$ sono linearmente dipendenti, questo implica che almeno un vettore può essere scritto come combinazione lineare dei restanti.

Vettori che non sono linearmente dipendenti sono detti linearmente indipendenti

Rango di una matrice

Il rango di una matrice $\underset{m\times q}{A}$ è il numero massimo di colonne (o di righe) linearmente indipendenti. Abbiamo

$$\text{rango}(\underset{m\times q}{A}) \leq \min(m, q)$$

Inoltre vale

$$\begin{aligned}\text{rango}(\underset{m\times q}{A}) &= \text{rango}(\underset{q\times m}{A'}) \\ &= \text{rango}(\underset{q\times m}{A'}\underset{m\times q}{A}) \\ &= \text{rango}(\underset{m\times q}{A}\underset{q\times m}{A'})\end{aligned}$$

Matrice non singolare

Una matrice quadrata $\underset{q\times q}{A}$ è non singolare se le colonne della matrice sono $\underset{q\times q}{A}$ linearmente indipendenti, ovvero

$$\underset{q\times q}{A}\underset{q\times 1}{c} = c_1 \underset{q\times 1}{a_1} + \ldots + c_q \underset{q\times 1}{a_q} = \underset{q\times 1}{0}$$

implica che $\underset{q\times 1}{c} = \underset{q\times 1}{0}$.
Se le colonne della matrice $\underset{q\times q}{A}$ sono linearmente dipendenti, la matrice $\underset{q\times q}{A}$ è detta singolare.

Equivalentemente,
Una matrice quadrata $\underset{q\times q}{A}$ è non singolare se è a pieno rango, ovvero

$$\mathrm{rango}(\underset{q\times q}{A}) = q$$

Se $\mathrm{rango}(\underset{q\times q}{A}) < q$, la matrice $\underset{q\times q}{A}$ è detta singolare.

Determinante di una matrice

Il determinante di una matrice quadrata $\underset{q\times q}{A}$ con $q > 1$ è

$$\det(\underset{q\times q}{A}) = \sum_{j=1}^{q} a_{ij}\det(\underset{(q-1)\times(q-1)}{A_{ij}})(-1)^{i+j}$$

dove $\underset{(q-1)\times(q-1)}{A_{1j}}$ è la matrice che si ottiene eliminando l'i-sima riga e la j-sima colonna di $\underset{q\times q}{A}$

Per $q = 2$:
$\det(\underset{2\times 2}{A}) = \det\left(\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}\right) = a_{11}a_{22} - a_{12}a_{21}$

Determinante di una matrice

Sia $\mathrm{diag}(a_1, \ldots, a_q)$ una matrice diagonale. Vale

- $\det(\mathrm{diag}(a_1, \ldots, a_q)) = \prod_{j=1}^{q} a_j$

Siano $\underset{q\times q}{A}$ e $\underset{q\times q}{B}$ due matrici quadrate e c una costante. Vale

- $\det(\underset{q\times q}{A}) = \det(\underset{q\times q}{A'})$
- Se $\underset{q\times q}{A}$ è non singolare, allora $\det(\underset{q\times q}{A}) = 1/\det(\underset{q\times q}{A^{-1}})$
- $\det(\underset{q\times q}{A}\underset{q\times q}{B}) = \det(\underset{q\times q}{A})\det(\underset{q\times q}{B})$
- $\det(c\underset{q\times q}{A}) = c^q \det(\underset{q\times q}{A})$

Proprietà equivalenti

Per una matrice quadrata $\underset{q\times q}{A}$, le seguenti proprietà sono equivalenti

- $\underset{q\times q}{A}$ è non singolare
- Esiste una matrice $\underset{q\times q}{A}^{-1}$ tale che $\underset{q\times q}{A}\,\underset{q\times q}{A^{-1}} = \underset{q\times q}{A^{-1}}\,\underset{q\times q}{A} = \underset{q\times q}{I}$
- $\det(\underset{q\times q}{A}) \neq 0$

Matrice ortogonale

Una matrice quadrata $\underset{q\times q}{Q}$ è detta ortogonale se e solo se

$$\underset{q\times q}{Q^{-1}} = \underset{q\times q}{Q'}$$

Per una matrice ortogonale $\underset{q\times q}{Q}$, abbiamo $\underset{q\times q}{Q}\,\underset{q\times q}{Q'} = \underset{q\times q}{Q'}\,\underset{q\times q}{Q} = \underset{q\times q}{I}$ quindi sia le righe che le colonne considerate come vettori sono mutualmente ortogonali e di lunghezza unitaria:

$$\underset{1\times q}{q_j'}\,\underset{q\times 1}{q_j} = 1,\ \underset{1\times q}{q_j'}\,\underset{q\times 1}{q_k} = 0, j \neq k$$