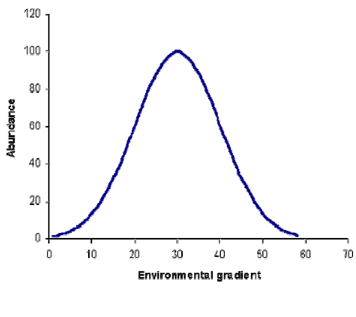
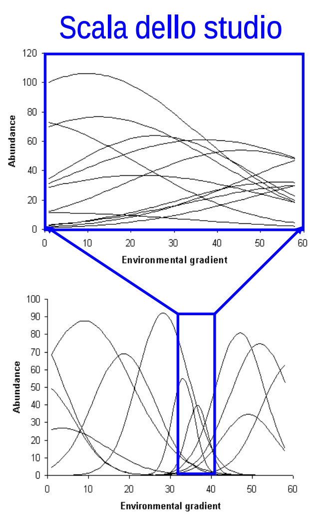
Caratteristiche dei dati ecologici

- I dati sono “sparsi”, cioè hanno molti valori nulli (a volte la maggioranza!)
- La gran parte delle specie presenti è rara.
- I fattori ambientali che influenzano la distribuzione delle specie sono molteplici e combinati fra loro,...
- ...ma quelli veramente importanti sono pochi (bassa dimensionalità intrinseca).
- I dati contengono molto “rumore” sia per eventi stocastici e contingenti, sia per l’errore di osservazione (anche in condizioni ideali le repliche sono diverse!)
- L’informazione è spesso ridondante (la specie A è associata alla specie B, ma questa può essere associata alla specie C, etc.): questo è un problema, ma è anche ciò che rende possibile interpretare i dati ecologici.

Gradienti ambientali e cenoclini





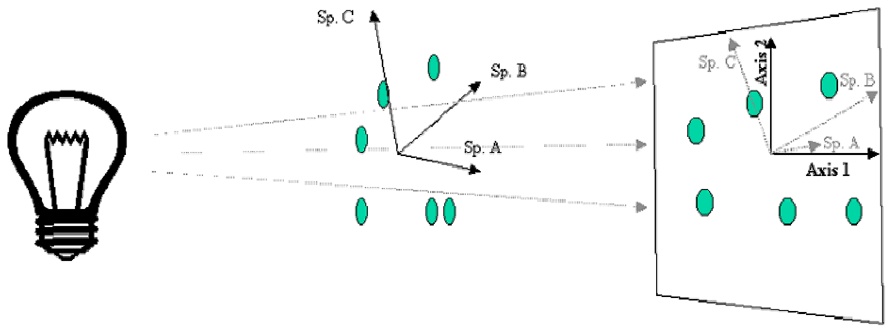
La cassetta degli attrezzi.

- Ordinamento (PCA, MDS, NMDS, CA, DCA, CCA, etc.)
- Classificazione (algoritmi gerarchici, k-means, reti neuronali, etc.)
- Analisi spaziale (correlogrammi, variogrammi, kriging, co-kriging, etc.)
- Analisi di serie (periodogrammi, runs tests, cross-correlation, cross-association, etc.)
- Confronti fra dati multivariati (MRPP, test di Mantel, INDVAL, etc.)
- Reti neurali
- ...

Tecniche di ordinamento

tre specie

due dimensioni

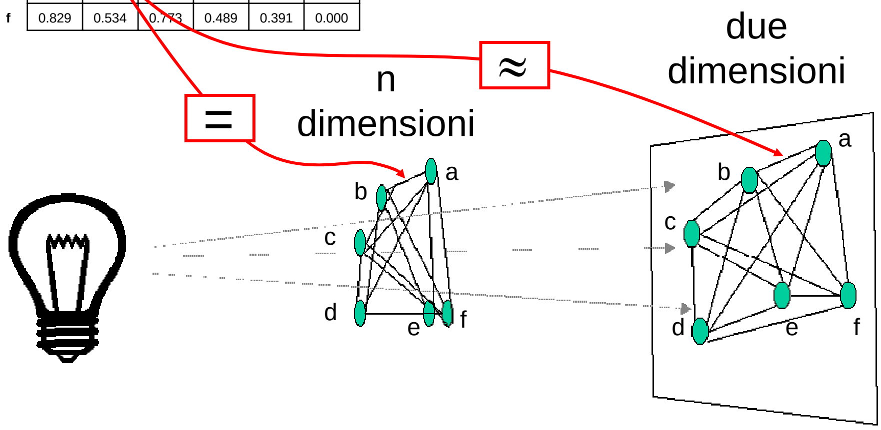


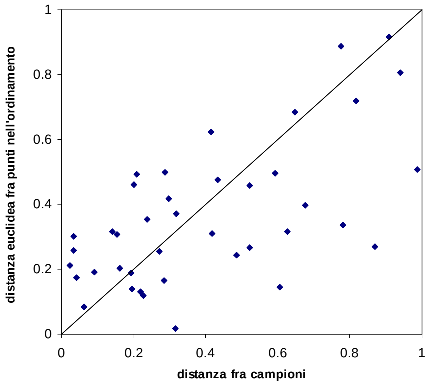
Analisi indiretta di gradiente

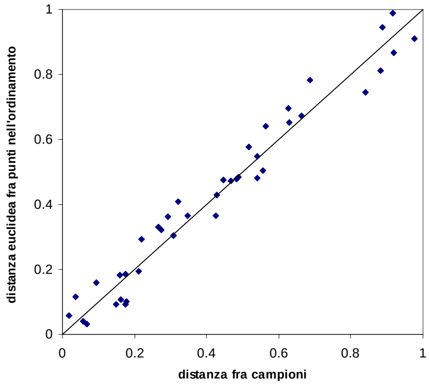
- Metodi basati su distanze
  - **Ordinamento polare (Bray-Curtis)**
  - **Analisi delle Coordinate Principali (PCoA)**
  - **Multidimensional Scaling Nonmetrico (NMDS)**
- Metodi basati su autovalori/autovettori
  - Modello lineare
    - **Analisi delle Componenti Principali (PCA)**
  - Modello unimodale
    - **Analisi delle Corrispondenze (CA)**
    - **Analisi delle Corrispondenze Detrendizzata (DCA)**

PCoA e NMDS

	a	b	c	d	e	f
a	0.000	0.317	0.405	0.982	0.923	0.829
b	0.317	0.000	0.184	0.617	0.695	0.534
c	0.405	0.184	0.000	0.571	0.614	0.773
d	0.982	0.617	0.571	0.000	0.092	0.489
e	0.923	0.695	0.614	0.092	0.000	0.391
f	0.829	0.534	0.773	0.489	0.391	0.000







Stress elevato: distanze nell'ordinamento diverse da quelle originali, quindi bassa qualità dell'ordinamento

Stress modesto: distanze nell'ordinamento simili a quelle originali, quindi alta qualità dell'ordinamento

Analisi delle Coordinate Principali.

$$X := \begin{pmatrix} 5 & 9 & 8 & 15 & 23 \\ 4 & 10 & 14 & 8 & 9 \end{pmatrix}$$

matrice dei dati (2 variabili x 5 osservazioni)

$$\Delta_{i,j} := \left(\sum_k \left| X_{k,i} - X_{k,j} \right| \right)$$

definizione di una misura di distanza
(es. metrica di Manhattan)

$$\Delta = \begin{pmatrix} 0 & 10 & 13 & 14 & 23 \\ 10 & 0 & 5 & 8 & 15 \\ 13 & 5 & 0 & 13 & 20 \\ 14 & 8 & 13 & 0 & 9 \\ 23 & 15 & 20 & 9 & 0 \end{pmatrix}$$

matrice distanze

$$A := \frac{-1}{2} \cdot \Delta$$

trasformazione

$$A = \begin{pmatrix} 0 & -5 & -6.5 & -7 & -11.5 \\ -5 & 0 & -2.5 & -4 & -7.5 \\ -6.5 & -2.5 & 0 & -6.5 & -10 \\ -7 & -4 & -6.5 & 0 & -4.5 \\ -11.5 & -7.5 & -10 & -4.5 & 0 \end{pmatrix}$$

matrice distanze trasformata

$$a_j := \frac{1}{n} \cdot \sum_i A_{i,j}$$

media della j-ma riga di A

$$a_m := \frac{1}{n} \cdot \sum_i a_i$$

media generale di A

$$A_{i,j} := A_{i,j} - a_j - a_i + a_m$$

il centroide degli oggetti è nell'origine
del nuovo sistema di assi

$$A = \begin{pmatrix} 6.8 & -0.4 & -0.6 & -1.8 & -4 \\ -0.4 & 2.4 & 1.2 & -1 & -2.2 \\ -0.6 & 1.2 & 5 & -2.2 & -3.4 \\ -1.8 & -1 & -2.2 & 3.6 & 1.4 \\ -4 & -2.2 & -3.4 & 1.4 & 8.2 \end{pmatrix}$$

matrice delle distanza dopo la seconda trasformazione (semidefinita positiva [t'At $\geq$0], ha un autovalore nullo)

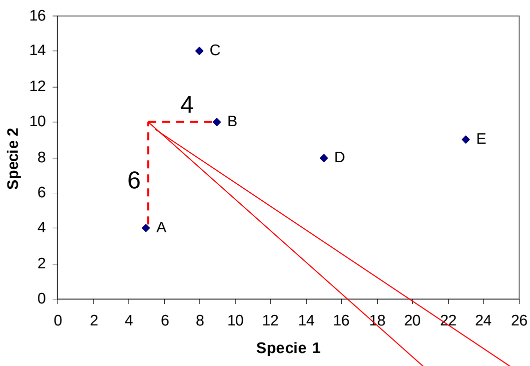
$\lambda_i := \text{reverse}(\text{sort}(\text{eigenvals}(A)))_i$ $\qquad \lambda^T = \begin{pmatrix} 13.5 & 6.9 & 3.6 & 2 & -3.6 \times 10^{-15} \end{pmatrix}$ autovalori

$U_{k,j} := \text{eigenvec}(A, \lambda_k)_j$ $\qquad U = \begin{pmatrix} -0.467 & -0.192 & -0.359 & 0.286 & 0.731 \\ -0.729 & 0.276 & 0.614 & -0.115 & -0.047 \end{pmatrix}$ autovettori (primi due)

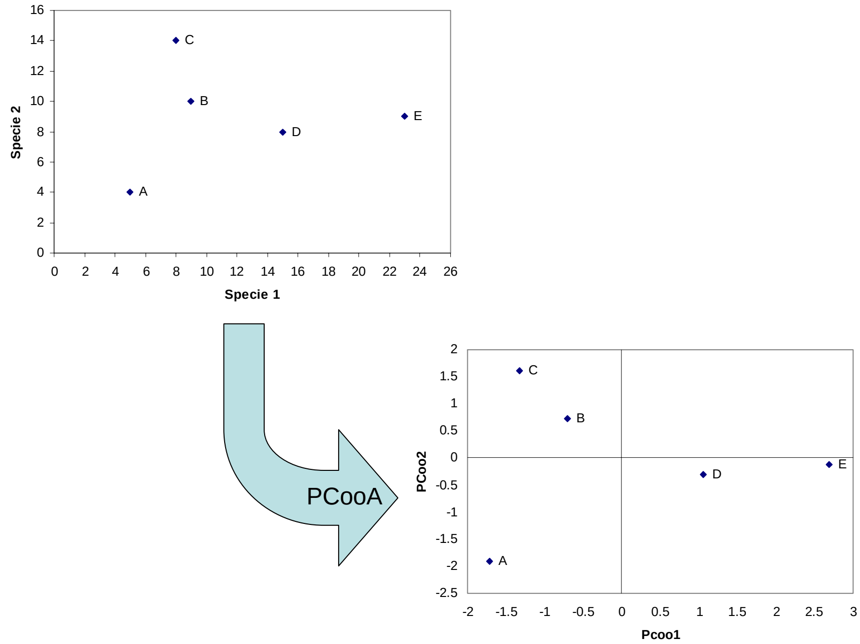
$C_{k,j} := \sqrt{\lambda_k} \cdot U_{k,j}$ $\qquad C = \begin{pmatrix} -1.718 & -0.706 & -1.32 & 1.053 & 2.692 \\ -1.917 & 0.725 & 1.616 & -0.302 & -0.122 \end{pmatrix}$ coordinate principali

$$\frac{\sum_k \left(\lambda_k + k \cdot |\max(\lambda)|\right)}{\sum_h \left[\lambda_h + (n-1) \cdot |\max(\lambda)|\right]} = 0.267 \qquad h := n - 1$$

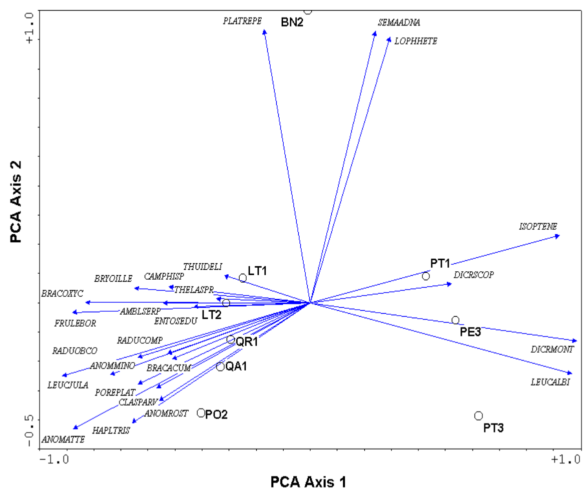
qualità della rappresentazione nel piano definito dai primi due assi



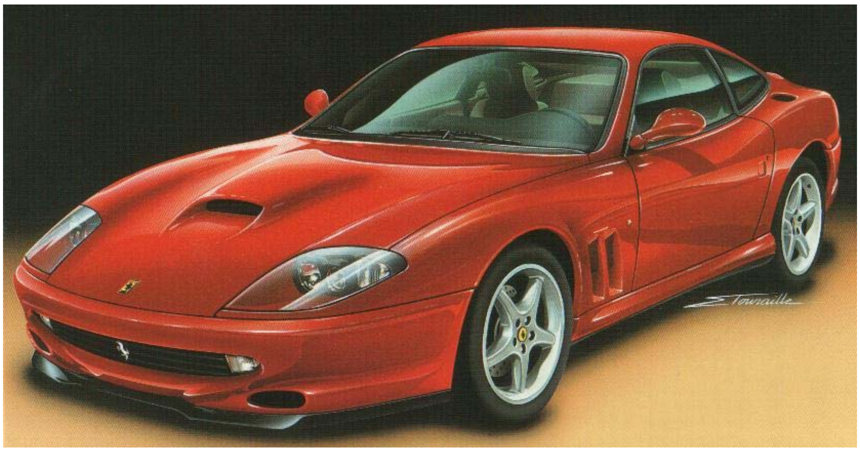
	A	B	C	D	E
A	0	10	13	14	23
B	10	0	5	8	15
C	13	5	0	13	20
D	14	8	13	0	9
E	23	15	20	9	0

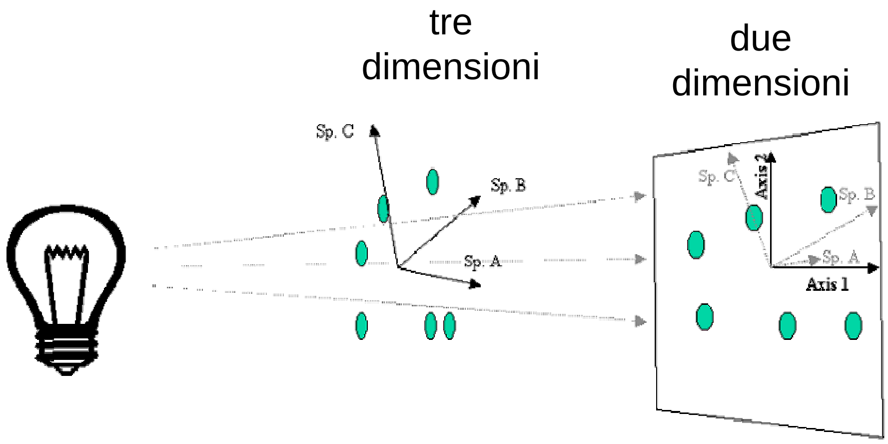

Specie 2
Specie 1
A
B
C
D
E
PCooA
PCoo2
Pcoo1

PCA


PCA Axis 2
PCA Axis 1
BN2
LT1
LT2
QR1
QA1
PO2
PT1
PE3
PT3
PLATREPE
SEMAADNA
LOPHHETE
ISOPTENE
DICRSCOP
DICRMONT
LEUCALBI
THUIDELI
CAMPHISP
BRYOILLE
THELASPR
BRACOXYC
AMBLSERP
ENTOSEDU
FRULEBOR
RADUCOMP
RADUOBCO
ANOMMINO
BRACACUM
LEUCJULA
POREPLAT
CLASPARV
ANOMROST
HAPLTRIS
ANOMATTE



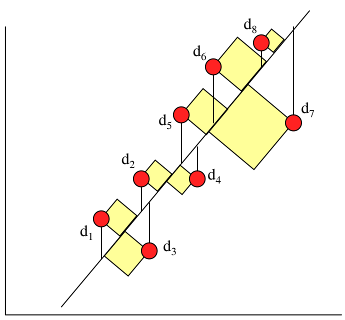





Asse Maggiore

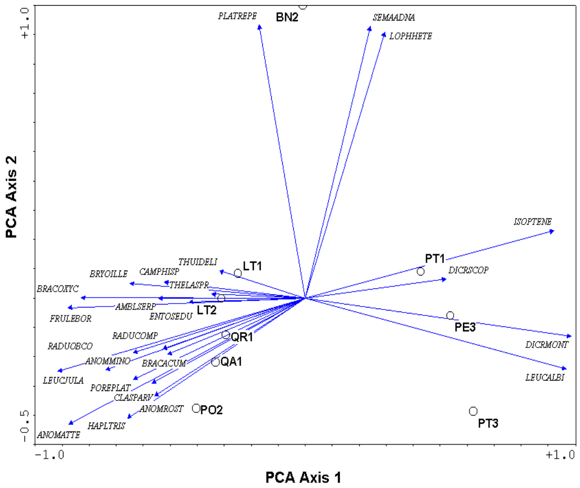
- Si minimizza la somma dei quadrati delle proiezioni dei punti sull'Asse Maggiore
- Il calcolo implica:
  - Estrazione di autovalori ed autovettori dalla matrice di covarianza

    oppure
  - Calcolo delle regressioni Y su X e X su Y e della bisettrice delle due rette



Asse maggiore

PCA



Perchè l’ordinamento?

"Ordination primarily endeavors to represent sample and variable relationships as faithfully as possible in a low-dimensional space.“

Gauch (1982)

- La PCA è una rotazione rigida degli assi: non cambia le posizioni degli oggetti nel loro spazio, ma ridefinisce il sistema di coordinate.
- Nella PCA gli assi sono definiti in modo che le distanze di ciascun oggetto dagli assi sia minimizzata (come nel caso dell'asse maggiore).
- Gli assi sono combinazioni lineari delle variabili originali.
- In queste combinazioni lineari ogni variabile ha un peso ("loading") noto e interpretabile.
- La PCA accetta valori negativi per le variabili analizzate.
- La PCA consente di proiettare nuovi punti in un ordinamento

- La PCA è adatta a trattare variabili dimensionalmente eterogenee, che possono essere standardizzate in modo da avere media nulla e varianza unitaria (in questo caso si lavora sulla matrice di correlazione)
- Gli autovalori hanno un significato legato alla varianza spiegata da ciascun asse e la loro somma corrisponde alla somma delle varianze di tutte le variabili (o al numero di variabili in caso di varianza unitaria).
- Gli assi sono linearmente indipendenti fra loro (ortogonali), cioè la somma dei prodotti dei pesi delle variabili che definiscono due diversi assi è nulla.
- La PCA ha seri problemi ad analizzare dati la cui distribuzione non sia normale, ma soprattutto non può rendere conto correttamente di relazioni fortemente non lineari o addirittura non monotone.

1901 Pearson sviluppa la PCA come una tecnica di regressione (quindi basata sulla covarianza)

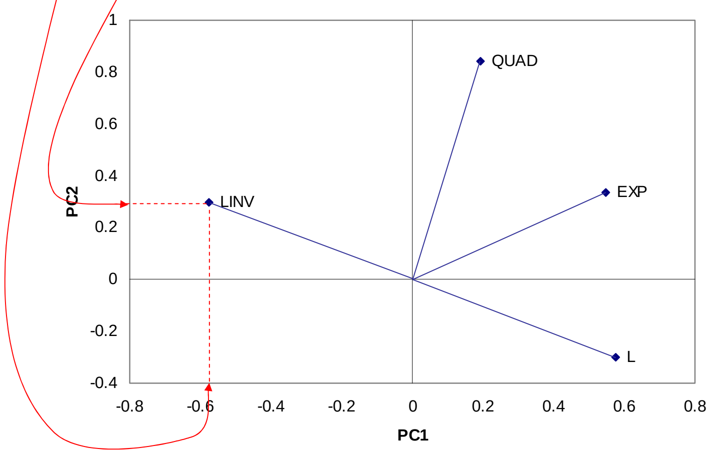
1933 Hotelling sviluppa la PCA come metodo per analizzare e comprendere il significato delle matrici di correlazione

1954 Goodall usa il termine “ordinamento” (“ordination”) per la PCA

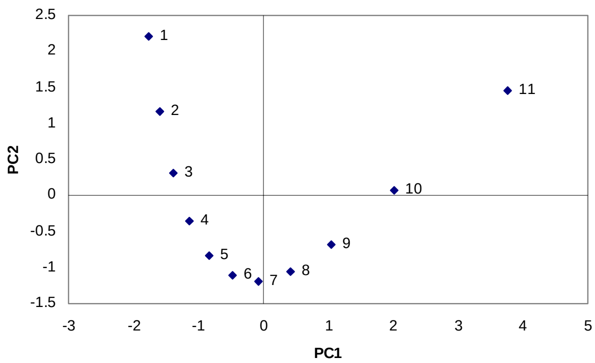
	L	**LINV**	**QUAD**	**EXP**
1	-5	10	25	0.01
2	-4	9	16	0.02
3	-3	8	9	0.05
4	-2	7	4	0.14
5	-1	6	1	0.37
6	0	5	0	1.00
7	1	4	1	2.72
8	2	3	4	7.39
9	3	2	9	20.09
10	4	1	16	54.60
11	5	0	25	148.41

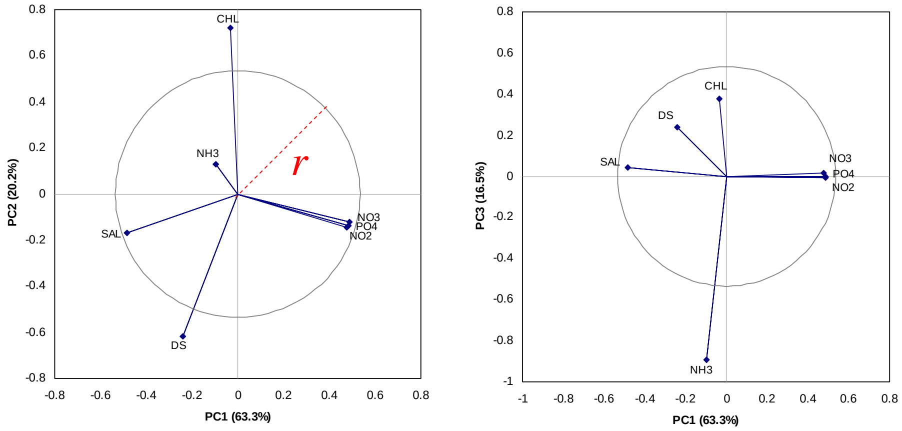
	PC1	PC2	PC3
L	0.575	-0.300	0.281
LINV	$\boxed{-0.575}$	$\boxed{0.300}$	-0.281
QUAD	0.193	0.842	0.504
EXP	0.548	0.334	-0.767

	autovalore	varianza spiegata
PC1	2.658	66.5%
PC2	1.232	30.8%
PC3	0.110	2.8%



	PC1	PC2	PC3
1	-1.763	2.214	0.348
2	-1.595	1.166	0.012
3	-1.384	0.310	-0.211
4	-1.129	-0.356	-0.320
5	-0.827	-0.830	-0.318
6	-0.477	-1.110	-0.209
7	-0.070	-1.192	-0.005
8	0.419	-1.059	0.261
9	1.054	-0.675	0.498
10	2.010	0.070	0.461
11	3.763	1.463	-0.517





$$= \sqrt{\frac{2}{p}}$$

Rotazione rigida di un insieme di punti.

$$X := \begin{pmatrix} -4.04 & -8.66 & 1.73 & -2.88 & 2.88 & -1.73 & 8.66 & 4.04 \\ 7.07 & 1.41 & 7.07 & 1.41 & -1.41 & -7.07 & -1.41 & -7.07 \\ 3.26 & 0 & -4.89 & -8.16 & 8.16 & 4.89 & 0 & -3.26 \end{pmatrix}$$

coordinate dei vertici di un parallepipedo (il baricentro coincide con l'origine degli assi) [vedi (a)]

$$R := X \cdot X^T$$

$$R = \begin{pmatrix} 205.209 & -65.206 & 3.741 \\ -65.206 & 207.892 & -46.059 \\ 3.741 & -46.059 & 202.251 \end{pmatrix}$$

matrice SSCP (sum of squares and cross products) , cioè: $R_{i,j} := \sum_k X_{i,k} \cdot X_{j,k}$

$$\Lambda_{i,i} := \text{reverse}(\text{sort}(\text{eigenvals}(R)))_i$$

$$\Lambda = \begin{pmatrix} 287.654 & 0 & 0 \\ 0 & 199.709 & 0 \\ 0 & 0 & 127.988 \end{pmatrix}$$

matrice degli autovalori (in ordine decrescente)

$$U_{i,j} := \text{eigenvec}\left(R, \Lambda_{i,i}\right)_j$$

$$U = \begin{pmatrix} 0.578 & -0.707 & 0.407 \\ -0.577 & -1.816 \cdot 10^{-3} & 0.817 \\ 0.577 & 0.707 & 0.409 \end{pmatrix}$$

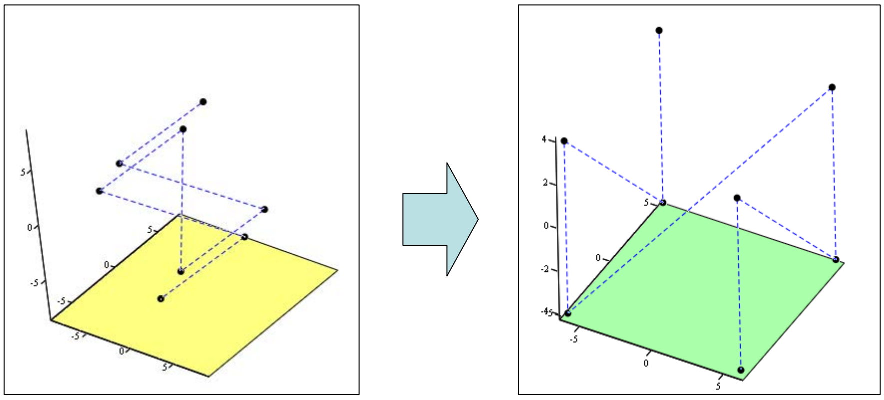
La matrice degli autovettori è ortogonale: in altre parole, gli autovettori (righe) sono linearmente indipendenti fra loro.

Quindi: $U \cdot U^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

$$Y := U \cdot X$$

$$Y = \begin{pmatrix} -6 & -6 & -6 & -6 & 6 & 6 & 6 & 6 \\ 5 & 5 & -5 & -5 & 5 & 5 & -5 & -5 \\ 4 & -4 & 4 & -4 & 4 & -4 & 4 & -4 \end{pmatrix}$$

Coordinate dei vertici del parallelepipedo dopo la rotazione che rende i suoi lati paralleli agli assi cartesiani [vedi (b)]



$\left[\left((X)^T\right)^{\langle 1\rangle}, \left((X)^T\right)^{\langle 2\rangle}, \left((X)^T\right)^{\langle 3\rangle}\right]$ $\quad$ $\left[\left((Y)^T\right)^{\langle 1\rangle}, \left((Y)^T\right)^{\langle 2\rangle}, \left((Y)^T\right)^{\langle 3\rangle}\right]$

$$X := \begin{pmatrix} 2 & 1 & 2 \\ 1 & 4 & 1.5 \\ 2 & 1.5 & 3 \end{pmatrix} \qquad i := 1..3 \qquad j := 1..3$$

XU=UΛ $\qquad$ XU-UΛ=0

$$\Lambda_{i,i} := \mathrm{eigenvals}(X)_i$$

$$\Lambda = \begin{pmatrix} 2.466 & 0 & 0 \\ 0 & 0.432 & 0 \\ 0 & 0 & 6.102 \end{pmatrix}$$

$$U\cdot\Lambda = \begin{pmatrix} 1.085 & -0.334 & 2.777 \\ -1.869 & -0.021 & 3.969 \\ 1.188 & 0.273 & 3.71 \end{pmatrix}$$

$$X\cdot U - U\cdot\Lambda = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\sum_i \Lambda_{i,i} = 9 \qquad \sum_i X_{i,i} = 9$$

$$X\cdot U = \begin{pmatrix} 1.085 & -0.334 & 2.777 \\ -1.869 & -0.021 & 3.969 \\ 1.188 & 0.273 & 3.71 \end{pmatrix}$$

$$U_{i,j} := \mathrm{eigenvec}\left(X, \Lambda_{j,j}\right)_i$$

$$U = \begin{pmatrix} 0.44 & -0.774 & 0.455 \\ -0.758 & -0.048 & 0.651 \\ 0.482 & 0.631 & 0.608 \end{pmatrix}$$

$$\sum_j U_{1,j}\cdot U_{2,j} = 0 \qquad \sum_j U_{1,j}\cdot U_{3,j} = 0 \qquad \sum_j U_{2,j}\cdot U_{3,j} = 0$$

Analisi delle Componenti Principali.

$$x := \begin{pmatrix} 632.53 & 267.5 & 310.549 & 172.835 & 211.882 & 120.379 \\ 26.919 & 9.325 & 13.07 & 5.647 & 6.338 & 3.92 \\ 79.126 & 63.657 & 48.223 & 213.082 & 69.371 & 44.174 \\ 41.908 & 13.137 & 20.448 & 8.48 & 7.474 & 10.821 \\ 24.552 & 29.273 & 28.43 & 31.266 & 33.766 & 31.879 \\ 0.737 & 0.702 & 0.644 & 0.791 & 2.452 & 0.852 \\ 1.289 & 13.875 & 1.59 & 2.528 & 1.253 & 1.858 \end{pmatrix}$$

dati bruti (7 variabili [righe] per 6 oggettii [colonne])

$n := \mathrm{cols}(x) \qquad p := \mathrm{rows}(x)$

$i := 1..\, p \qquad j := 1..\, n \qquad k := 1..\, p$

$$m_i := \frac{1}{n}\cdot\sum_j x_{i,j}$$

$$s_i := \mathrm{stdev}\left[\left(x^T\right)^{\langle i\rangle}\right]$$

$$y_{i,j} := \frac{x_{i,j} - m_i}{s_i} \longrightarrow y = \begin{pmatrix} 2.079 & -0.111 & 0.148 & -0.678 & -0.444 & -0.993 \\ 2.069 & -0.199 & 0.284 & -0.673 & -0.584 & -0.896 \\ -0.123 & -0.39 & -0.657 & 2.188 & -0.292 & -0.726 \\ 2.091 & -0.329 & 0.286 & -0.72 & -0.805 & -0.523 \\ -1.806 & -0.2 & -0.487 & 0.478 & 1.328 & 0.686 \\ -0.458 & -0.512 & -0.603 & -0.373 & 2.224 & -0.278 \\ -0.536 & 2.226 & -0.47 & -0.264 & -0.544 & -0.411 \end{pmatrix}$$

dati centrati e standardizzati (sulla riga)

$$y = \begin{pmatrix} 2.079 & -0.111 & 0.148 & -0.678 & -0.444 & -0.993 \\ 2.069 & -0.199 & 0.284 & -0.673 & -0.584 & -0.896 \\ -0.123 & -0.39 & -0.657 & 2.188 & -0.292 & -0.726 \\ 2.091 & -0.329 & 0.286 & -0.72 & -0.805 & -0.523 \\ -1.806 & -0.2 & -0.487 & 0.478 & 1.328 & 0.686 \\ -0.458 & -0.512 & -0.603 & -0.373 & 2.224 & -0.278 \\ -0.536 & 2.226 & -0.47 & -0.264 & -0.544 & -0.411 \end{pmatrix}$$

dati centrati e standardizzati (sulla riga)

$$S := \frac{1}{n}\cdot y\cdot y^T$$

$$S = \begin{pmatrix} 1 & 0.995 & -0.157 & 0.965 & -0.9 & -0.24 & -0.1 \\ 0.995 & 1 & -0.169 & 0.983 & -0.925 & -0.303 & -0.137 \\ -0.157 & -0.169 & 1 & -0.213 & 0.13 & -0.102 & -0.102 \\ 0.965 & 0.983 & -0.213 & 1 & -0.937 & -0.389 & -0.191 \\ -0.9 & -0.925 & 0.13 & -0.937 & 1 & 0.635 & -0.063 \\ -0.24 & -0.303 & -0.102 & -0.389 & 0.635 & 1 & -0.268 \\ -0.1 & -0.137 & -0.102 & -0.191 & -0.063 & -0.268 & 1 \end{pmatrix}$$

matrice di correlazione

$\Lambda_{i,i} := \text{reverse}(\text{sort}(\text{eigenvals}(S)))_i$

$$\Lambda = \begin{pmatrix} 4.086 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1.306 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1.065 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.534 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 8.167\times10^{-3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

matrice autovalori

$$\sum_j \Lambda_{j,j} = 7$$

N.B. Se il numero delle variabili (p) è maggiore del numero degli oggetti (n), la matrice avrà p-(n-1) autovalori nulli [nel caso in esame, 7-(6-1)=7-5=2

$j := 1..5 \qquad U_{i,j} := \text{eigenvec}(S, \Lambda_{j,j})_i$

$$U = \begin{pmatrix} -0.475 & -0.145 & -0.016 & 0.312 & -0.073 \\ -0.484 & -0.135 & 1.984\times10^{-3} & 0.188 & -0.254 \\ 0.096 & 0.128 & 0.894 & 0.409 & 0.064 \\ -0.488 & -0.123 & 4.606\times10^{-3} & -0.028 & 0.812 \\ 0.484 & -0.169 & -0.042 & 0.041 & 0.495 \\ 0.24 & -0.619 & -0.238 & 0.618 & -0.05 \\ 0.033 & 0.719 & -0.377 & 0.562 & 0.138 \end{pmatrix}$$

matrice autovettori (saturazioni o loadings delle variabili, in riga), norma = 1

(possono essere rappresentati come un istogramma per riga [cioè per PC] o come scatterplot)

$$\sqrt{\sum_i (U_{i,j})^2} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

$$U_{i,j} := U_{i,j} \cdot \frac{\sqrt{\Lambda_{j,j}}}{S_{i,i}}$$

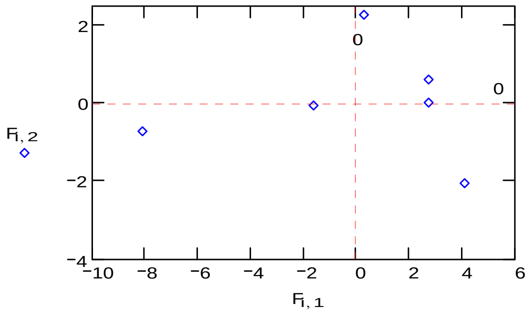
$$U = \begin{pmatrix} -0.959 & -0.165 & -0.016 & 0.228 & -6.614\times10^{-3} \\ -0.978 & -0.154 & 2.048\times10^{-3} & 0.138 & -0.023 \\ 0.194 & 0.147 & 0.923 & 0.299 & 5.75\times10^{-3} \\ -0.987 & -0.14 & 4.754\times10^{-3} & -0.02 & 0.073 \\ 0.979 & -0.193 & -0.044 & 0.03 & 0.045 \\ 0.484 & -0.708 & -0.246 & 0.452 & -4.481\times10^{-3} \\ 0.067 & 0.822 & -0.389 & 0.411 & 0.012 \end{pmatrix}$$

correlazioni fra variabili (righe) e PCs

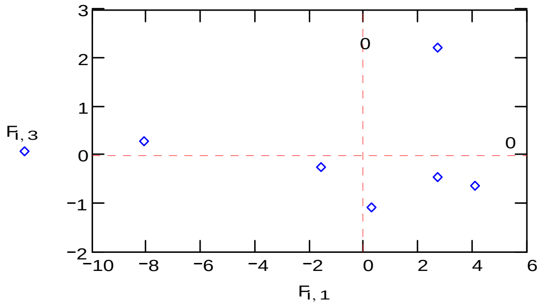
$F := y^T \cdot U$

$$F = \begin{pmatrix} -8.131 & -0.743 & 0.266 & 0.198 & 5.868\times10^{-3} \\ 0.255 & 2.269 & -1.091 & 0.515 & 1.157\times10^{-4} \\ -1.629 & -0.07 & -0.254 & -0.609 & -0.015 \\ 2.714 & 0.593 & 2.199 & 0.158 & -5.186\times10^{-4} \\ 4.076 & -2.044 & -0.66 & 0.568 & -1.656\times10^{-3} \\ 2.715 & -4.226\times10^{-3} & -0.46 & -0.83 & 0.011 \end{pmatrix}$$

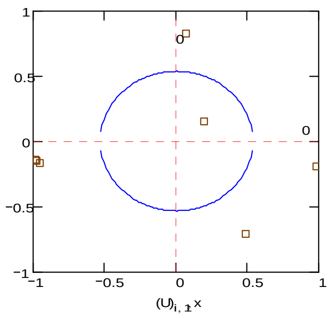
coordinate o scores oggetti (righe)



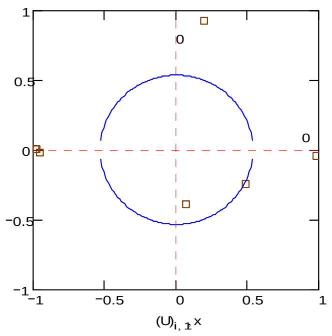
proiezione sulle prime due componenti principali



proiezione sulla prima e terza componente principale



punti-descrittore (norma=1) e cerchio del contributo equilibrato agli assi (dimensioni rappresentazione/n. descrittori=2/p): 1.o e 2.o asse.



idem, per 1.a e 3.a CP

Le quattro diverse versioni dell'Analisi delle Compenenti Principali.

$$X := \begin{pmatrix} 2 & 12 & 33 & 42 & 55 & 60 & 62 & 65 & 92 & 99 \\ 4 & 10 & 13 & 30 & 17 & 42 & 27 & 25 & 55 & 43 \end{pmatrix}$$

dati bruti (2 variabili per 10 osservazioni)

$$Xm_i := \frac{1}{n} \cdot \sum_j X_{i,j} \qquad Xm = \begin{pmatrix} 52.2 \\ 26.6 \end{pmatrix}$$

medie

$$Xs_i := \sqrt{\frac{1}{n} \cdot \sum_j \left(X_{i,j} - Xm_i\right)^2}$$

deviazioni standard

Dati standardizzati

Dati centrati

Si\Si x'=(x-m)/s	Si\No x'=x-m
No\Si x'=x/s	No\No x'=x

Soluzione #1: dati non centrati e non standardizzati

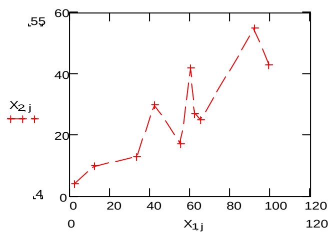
$$R1 := \frac{1}{n} \cdot X \cdot X^T \qquad R1 = \begin{pmatrix} 3596 & 1788.8 \\ 1788.8 & 946.6 \end{pmatrix}$$

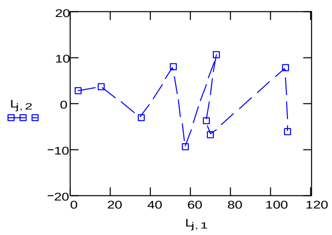
$$\Lambda 1_{i,i} := \mathrm{reverse}(\mathrm{sort}(\mathrm{eigenvals}(R1)))_i \qquad \Lambda 1 = \begin{pmatrix} 4497.201 & 0 \\ 0 & 45.399 \end{pmatrix}$$

$$U1_{i,k} := \mathrm{eigenvec}(R1, \Lambda 1_{k,k})_i \qquad U1 = \begin{pmatrix} 0.893 & -0.45 \\ 0.45 & 0.893 \end{pmatrix}$$

$$L := X^T \cdot U1$$

$$L = \begin{pmatrix} 3.586 & 2.672 \\ 15.216 & 3.532 \\ 35.32 & -3.238 \\ 51.007 & 7.895 \\ 56.767 & -9.564 \\ 72.481 & 10.513 \\ 67.518 & -3.783 \\ 69.297 & -6.919 \\ 106.908 & 7.725 \\ 107.76 & -6.141 \end{pmatrix}$$





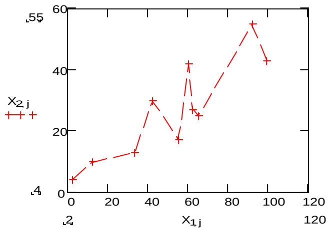
Soluzione #2: dati centrati e non standardizzati

$Xc_{i,j} := X_{i,j} - Xm_i$

$R_2 := \frac{1}{n} \cdot Xc \cdot Xc^T \qquad R_2 = \begin{pmatrix} 871.16 & 400.28 \\ 400.28 & 239.04 \end{pmatrix}$

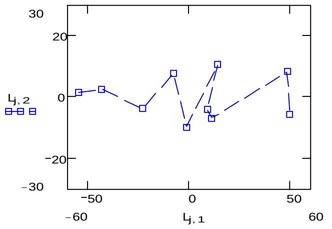
$\Lambda_{2_{i,i}} := \text{reverse}(\text{sort}(\text{eigenvals}(R_2)))_i \qquad \Lambda_2 = \begin{pmatrix} 1.065\cdot 10^3 & 0 \\ 0 & 45.082 \end{pmatrix}$

$U_{2_{i,k}} := \text{eigenvec}(R_2, \Lambda_{2_{k,k}})_i \qquad U_2 = \begin{pmatrix} 0.9 & -0.436 \\ 0.436 & 0.9 \end{pmatrix}$



$L := Xc^T \cdot U_2$

$$L = \begin{pmatrix} -55.031 & 1.552 \\ -43.415 & 2.591 \\ -23.209 & -3.867 \\ -7.697 & 7.508 \\ -1.666 & -9.86 \\ 13.735 & 10.457 \\ 8.994 & -3.913 \\ 10.821 & -7.021 \\ 48.201 & 8.202 \\ 49.268 & -5.649 \end{pmatrix}$$



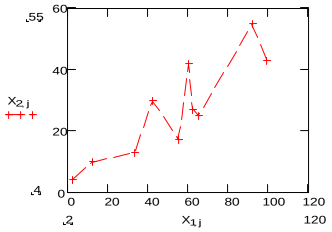
Soluzione #3: dati non centrati e standardizzati

$Z_{i,j} := \frac{X_{i,j}}{Xs_i}$

$R_3 := \frac{1}{n} \cdot Z \cdot Z^T \qquad R_3 = \begin{pmatrix} 4.128 & 3.92 \\ 3.92 & 3.96 \end{pmatrix}$

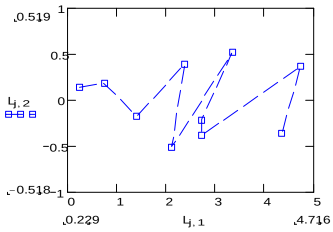
$\Lambda_{3_{i,i}} := \text{reverse}(\text{sort}(\text{eigenvals}(R_3)))_i \qquad \Lambda_3 = \begin{pmatrix} 7.965 & 0 \\ 0 & 0.123 \end{pmatrix}$

$U_{3_{i,k}} := \text{eigenvec}(R_3, \Lambda_{3_{k,k}})_i \qquad U_3 = \begin{pmatrix} 0.715 & -0.699 \\ 0.699 & 0.715 \end{pmatrix}$



$L := Z^T \cdot U_3$

$$L = \begin{pmatrix} 0.229 & 0.137 \\ 0.743 & 0.178 \\ 1.387 & -0.181 \\ 2.374 & 0.391 \\ 2.101 & -0.518 \\ 3.353 & 0.519 \\ 2.723 & -0.221 \\ 2.705 & -0.385 \\ 4.716 & 0.362 \\ 4.342 & -0.359 \end{pmatrix}$$



Soluzione #4: dati centrati e standardizzati [c

$$Zc_{i,j} := \frac{X_{i,j} - Xm_i}{Xs_i}$$

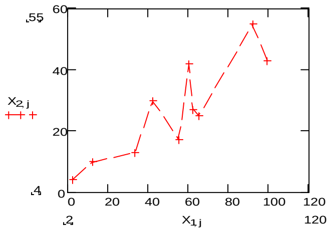
$$R_4 := \frac{1}{n} \cdot Zc \cdot Zc^T \qquad R_4 = \begin{pmatrix} 1 & 0.877 \\ 0.877 & 1 \end{pmatrix}$$

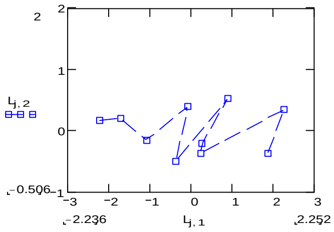
$$\Lambda_{4_{i,i}} := \text{reverse}(\text{sort}(\text{eigenvals}(R_4)))_i \qquad \Lambda_4 = \begin{pmatrix} 1.877 & 0 \\ 0 & 0.123 \end{pmatrix}$$

$$U_{4_{i,k}} := \text{eigenvec}(R_4, \Lambda_{4_{k,k}})_i \qquad U_4 = \begin{pmatrix} 0.707 & -0.707 \\ 0.707 & 0.707 \end{pmatrix}$$

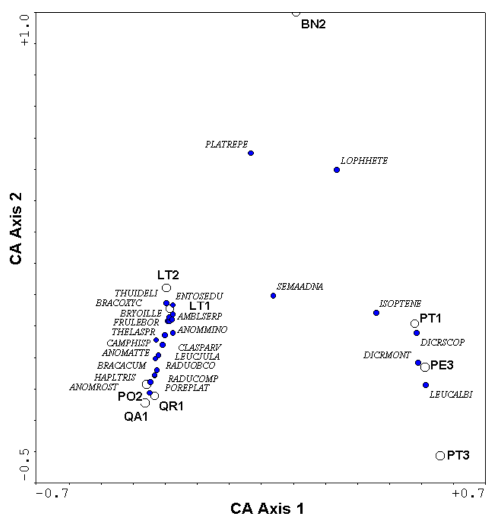
$$L := Zc^T \cdot U_4$$

$$L = \begin{pmatrix} -2.236 & 0.169 \\ -1.722 & 0.204 \\ -1.082 & -0.162 \\ -0.089 & 0.4 \\ -0.372 & -0.506 \\ 0.891 & 0.517 \\ 0.253 & -0.216 \\ 0.233 & -0.38 \\ 2.252 & 0.345 \\ 1.871 & -0.371 \end{pmatrix}$$





CA



Analisi diretta di gradiente

- Modello lineare
  - **Analisi di Ridondanza (RDA)**
- Modello unimodale
  - **Analisi Canonica delle Corrispondenze (CCA)**
  - **Analisi Canonica delle Corrispondenze Detrendizzata (DCCA)**

CCA