

RENATO LEONI

**Esempi numerici riguardanti l'analisi
delle componenti principali**

**UNIVERSITÀ DI FIRENZE
DIPARTIMENTO DI STATISTICA "G. PARENTI"
FIRENZE, 2007**

Questo lavoro è destinato a un uso personale e ne è vietata la commercializzazione.

AVVERTENZA

Come è anticipato dal titolo, in questo lavoro sono presentati alcuni esempi numerici riguardanti l'analisi delle componenti principali (ACP).

Per una ampia esposizione dei concetti teorici ai quali si fa riferimento, si rinvia a:

Leoni, R., *Principal Component Analysis*, Department of Statistics, "G. Parenti", Firenze, 2007.

ESEMPIO 1

Questo semplicissimo esempio, basato su dati fittizi riferiti a 3 individui e 2 variabili, ha lo scopo di illustrare i principali concetti della ACP, seguendo l'ordine con cui questi sono presentati nel lavoro citato nella Avvertenza.

I DATI DI BASE E LA LORO STRUTTURA ALGEBRICA

Si consideri anzitutto la seguente matrice di dati grezzi

$$\mathbf{X} = \begin{bmatrix} 0 & 3 \\ -1 & 2 \\ 4 & -2 \end{bmatrix}.$$

Gli individui sono rappresentati dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$$

appartenenti a \mathbb{R}^2 , mentre le variabili sono rappresentate dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 3 \\ 2 \\ -2 \end{bmatrix}$$

appartenenti a \mathbb{R}^3 .

Ciò premesso, si definisca in \mathbb{R}^2 (spazio degli individui) e in \mathbb{R}^3 (spazio delle variabili) una metrica euclidea nel modo seguente.

In \mathbb{R}^2 la matrice rappresentativa della metrica – rispetto alla base canonica di \mathbb{R}^2 – è

$$\mathbf{Q} = \text{diag}(1, 1) = \mathbf{I}_2.$$

In \mathbb{R}^3 la matrice rappresentativa della metrica – rispetto alla base canonica di \mathbb{R}^3 – è

$$\mathbf{M} = \frac{1}{3} \text{diag}(1, 1, 1) = \frac{1}{3} \mathbf{I}_3.$$

La scelta della matrice $\mathbf{Q} = \mathbf{I}_2$, rappresentativa della metrica nello spazio degli individui, attribuisce la medesima importanza alle variabili $\mathbf{x}_1, \mathbf{x}_2$ e risulta giustificata dalla supposizione che quest'ultime siano espresse nella stessa unità di misura e dal fatto che le loro varianze, come verificheremo nel seguito, siano eguali.

A sua volta, la scelta della matrice $\mathbf{M} = (1/3) \mathbf{I}_3$, rappresentativa della metrica nello spazio delle variabili, attribuisce la medesima importanza (1/3) agli individui $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$.

Ciò premesso, poiché il baricentro o vettore delle medie è

$$\mathbf{g} = \mathbf{X}'\mathbf{M}\mathbf{u} = \begin{bmatrix} 0 & -1 & 4 \\ 3 & 2 & -2 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},$$

la matrice dei dati centrati risulta

$$\mathbf{Y} = \mathbf{X} - \mathbf{u}\mathbf{g}' = \begin{bmatrix} 0 & 3 \\ -1 & 2 \\ 4 & -2 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -2 & 1 \\ 3 & -3 \end{bmatrix}.$$

I vettori

$$\mathbf{y}_1 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} 3 \\ -3 \end{bmatrix}$$

e

$$\mathbf{y}_1 = \begin{bmatrix} -1 \\ -2 \\ 3 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 2 \\ 1 \\ -3 \end{bmatrix}$$

rappresentano, rispettivamente, gli individui e le variabili (misurati in termini di scarti dalle medie).

CONCETTI PRELIMINARI NELLO SPAZIO DEGLI INDIVIDUI

INERZIA RISPETTO AL BARICENTRO

La matrice di inerzia rispetto al baricentro è

$$\mathbf{V}_g = \mathbf{Y}'\mathbf{M}\mathbf{Y} = \begin{bmatrix} -1 & -2 & 3 \\ 2 & 1 & -3 \end{bmatrix} \frac{1}{3} \begin{bmatrix} -1 & 2 \\ -2 & 1 \\ 3 & -3 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix}.$$

Pertanto, l'inerzia rispetto al baricentro risulta

$$I_g = \text{tr} \{ \mathbf{V}_g \} = \text{tr} \left\{ \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} \right\} = \frac{1}{3} 28.$$

INERZIA SPIEGATA DA UNA VARIETÀ LINEARE PASSANTE PER IL BARI-CENTRO

Si consideri in \mathbb{R}^2 la varietà lineare $\mathbf{g} + C_1$ di direzione C_1 e traslazione \mathbf{g} dove C_1 è un sottospazio di \mathbb{R}^2 (di dimensione 1) generato dal vettore ($\mathbf{c}_1 \mathbf{c}_1 = 1$)

$$\mathbf{c}_1 = \frac{1}{5} \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Sia

$$\mathbf{P} = \mathbf{c}_1 \mathbf{c}_1' = \frac{1}{5} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \frac{1}{5} [3 \ 4] = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix}$$

la matrice di proiezione ortogonale in C_1 .

Ne consegue che la quantità

$$\begin{aligned} I_{\mathbf{g} + C_1} &= \text{tr} \{ \mathbf{V}_g \mathbf{P} \} &&= \text{tr} \left\{ \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \right\} \\ &= \text{tr} \left\{ \frac{1}{75} \begin{bmatrix} -30 & -40 \\ 51 & 68 \end{bmatrix} \right\} = \frac{38}{75} \end{aligned}$$

rappresenta l'inerzia spiegata da $\mathbf{g} + C_1$.

Analogamente, si consideri in \mathbb{R}^2 la varietà lineare $\mathbf{g} + C_1^\perp$ di direzione C_1^\perp e traslazione \mathbf{g} dove C_1^\perp è un sottospazio di \mathbb{R}^2 (di dimensione 1) generato dal vettore ($\mathbf{c}_2 \mathbf{c}_2 = 1$; $\mathbf{c}_1 \mathbf{c}_2 = 0$)

$$\mathbf{c}_2 = \frac{1}{5} \begin{bmatrix} -4 \\ 3 \end{bmatrix}.$$

Come facilmente si verifica, la quantità

$$I_{\mathbf{g}+C_1^\perp} = \text{tr} \left\{ \mathbf{V}_{\mathbf{g}} (\mathbf{I}_2 - \mathbf{P}) \right\} = \frac{662}{75}$$

denota l'inerzia non spiegata da $\mathbf{g} + C_1$.

Ovviamente,

$$I_{\mathbf{g}} = \frac{1}{3} 28 = \frac{38}{75} + \frac{662}{75} = I_{\mathbf{g}+C_1} + I_{\mathbf{g}+C_1^\perp}.$$

CONCETTI PRELIMINARI NELLO SPAZIO DELLE VARIABILI

MATRICE DELLE COVARIANZE E MATRICE DI CORRELAZIONE

Come è noto, la matrice delle covarianze coincide con la matrice di inerzia rispetto al baricentro.

Pertanto,

$$\mathbf{V} = \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix}.$$

Inoltre, posto

$$\mathbf{Q}_{1/\sigma} = \text{diag}(\sqrt{3/14}, \sqrt{3/14}),$$

la matrice di correlazione è data da

$$\begin{aligned} \mathbf{R} &= \mathbf{Q}_{1/\sigma} \mathbf{V} \mathbf{Q}_{1/\sigma} = \begin{bmatrix} \sqrt{3/14} & 0 \\ 0 & \sqrt{3/14} \end{bmatrix} \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} \begin{bmatrix} \sqrt{3/14} & 0 \\ 0 & \sqrt{3/14} \end{bmatrix} \\ &= \begin{bmatrix} 1 & -3/14 \\ -3/14 & 1 \end{bmatrix}. \end{aligned}$$

INTERPRETAZIONE DI ALCUNI CONCETTI DI INERZIA

Si noti anzitutto che

$$I_{\mathbf{g}} = \frac{1}{3} 28 = J_2,$$

vale a dire che l'inerzia rispetto al baricentro coincide con la variabilità glo-

bale di $\mathbf{x}_1, \mathbf{x}_2$.

Si considerino poi le variabili

$$\hat{\mathbf{x}}_1 = \mathbf{X} \mathbf{p}_1 = \begin{bmatrix} 0 & 3 \\ -1 & 2 \\ 4 & -2 \end{bmatrix} \frac{1}{25} \begin{bmatrix} 9 \\ 12 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 36 \\ 15 \\ 12 \end{bmatrix}$$

e

$$\hat{\mathbf{x}}_2 = \mathbf{X} \mathbf{p}_2 = \begin{bmatrix} 0 & 3 \\ -1 & 2 \\ 4 & -2 \end{bmatrix} \frac{1}{25} \begin{bmatrix} 12 \\ 16 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 48 \\ 20 \\ 16 \end{bmatrix}$$

dove \mathbf{p}_j ($j = 1, 2$) denota la j -esima colonna della trasposta della matrice di proiezione ortogonale \mathbf{P} in C_1 .

Tali variabili hanno varianze

$$\hat{\sigma}_1^2 = \mathbf{p}_1' \mathbf{V} \mathbf{p}_1 = \frac{1}{25} [9 \ 12] \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} \frac{1}{25} \begin{bmatrix} 9 \\ 12 \end{bmatrix} = \frac{342}{625 \times 3}$$

e

$$\hat{\sigma}_2^2 = \mathbf{p}_2' \mathbf{V} \mathbf{p}_2 = \frac{1}{25} [12 \ 16] \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} \frac{1}{25} \begin{bmatrix} 12 \\ 16 \end{bmatrix} = \frac{608}{625 \times 3}.$$

Risulta quindi che

$$I_{\mathbf{g}+C_1} = \frac{38}{75} = \frac{342}{625 \times 3} + \frac{608}{625 \times 3} = \hat{J}_2,$$

vale a dire che l'inerzia spiegata dalla varietà lineare $\mathbf{g} + C_1$ coincide con la variabilità globale di $\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2$.

UN APPROCCIO ALLA ACP

VETTORI PRINCIPALI, COMPONENTI PRINCIPALI, COMPONENTI PRINCIPALI STANDARDIZZATE

Come si verifica facilmente, dall'applicazione della ACP, nell'approccio che qui stiamo considerando, si ottengono i risultati seguenti.

- Una matrice

$$\tilde{\mathbf{D}}_2 = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2) = \begin{bmatrix} 9 & 0 \\ 0 & 1/3 \end{bmatrix}$$

i cui elementi sulla diagonale principale sono gli autovalori di \mathbf{V} .

- Una matrice

$$\tilde{\mathbf{C}}_2 = [\tilde{\mathbf{c}}_1 \quad \tilde{\mathbf{c}}_2] = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

costituita dai vettori principali.

- Una matrice

$$\tilde{\mathbf{Y}} = [\tilde{\mathbf{y}}_1 \quad \tilde{\mathbf{y}}_2] = \begin{bmatrix} -3/\sqrt{2} & -1/\sqrt{2} \\ -3/\sqrt{2} & 1/\sqrt{2} \\ 6/\sqrt{2} & 0 \end{bmatrix}$$

costituita dalle componenti principali.

- Una matrice

$$\tilde{\tilde{\mathbf{Y}}} = [\tilde{\tilde{\mathbf{y}}}_1 \quad \tilde{\tilde{\mathbf{y}}}_2] = \begin{bmatrix} -1/\sqrt{2} & -\sqrt{3}/\sqrt{2} \\ -1/\sqrt{2} & \sqrt{3}/\sqrt{2} \\ 2/\sqrt{2} & 0 \end{bmatrix}$$

costituita dalle componenti principali standardizzate.

RAPPRESENTAZIONE GRAFICA DEGLI INDIVIDUI E DELLE VARIABILI

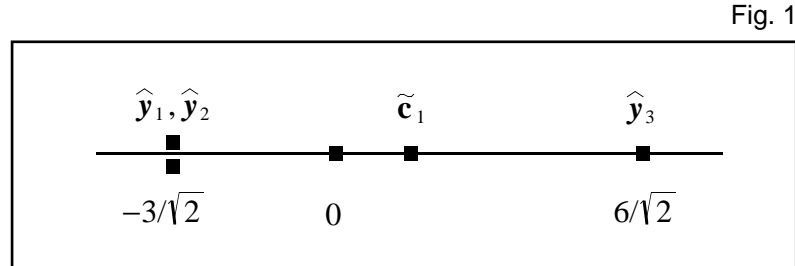
RAPPRESENTAZIONE GRAFICA DEGLI INDIVIDUI

Una rappresentazione grafica degli individui (misurati in termini di scarti dalle medie) si può ottenere considerando le loro proiezioni ortogonali $\hat{\mathbf{y}}_1$, $\hat{\mathbf{y}}_2$, $\hat{\mathbf{y}}_3$ nel sottospazio $\tilde{\mathbf{C}}_1$ generato dal vettore $\tilde{\mathbf{c}}_1$ (asse principale).

Poiché le coordinate, rispetto a $\tilde{\mathbf{c}}_1$, di tali proiezioni ortogonali sono

$$\tilde{y}_{11} = -3/\sqrt{2} \quad , \quad \tilde{y}_{21} = -3/\sqrt{2} \quad , \quad \tilde{y}_{31} = 6/\sqrt{2}$$

si ottiene il grafico di Fig. 1.



La qualità globale della rappresentazione degli individui si può misurare mediante l'indice GQRI che assume il valore

$$\text{GQRI} = \frac{\tilde{\lambda}_1}{\tilde{\lambda}_1 + \tilde{\lambda}_2} = \frac{27}{28}.$$

A sua volta, la qualità della rappresentazione di ciascun individuo si può misurare attraverso l'indice $\text{QR}(i; \tilde{c}_1)$ che assume i valori

$$\begin{aligned} \text{QR}(1; \tilde{c}_1) &= \frac{\tilde{y}_{11}^2}{\tilde{y}_{11}^2 + \tilde{y}_{12}^2} = \frac{9}{10} \\ \text{QR}(2; \tilde{c}_1) &= \frac{\tilde{y}_{21}^2}{\tilde{y}_{21}^2 + \tilde{y}_{22}^2} = \frac{9}{10} \\ \text{QR}(3; \tilde{c}_1) &= \frac{\tilde{y}_{31}^2}{\tilde{y}_{31}^2 + \tilde{y}_{32}^2} = 1. \end{aligned}$$

Come si può vedere, gli individui sono ben rappresentati sia globalmente sia singolarmente.

Infine, nell'interpretare i risultati dell'analisi è importante esaminare il contributo di ciascun individuo all'inerzia spiegata da \tilde{c}_1 , misurato dall'indice $C(i; \tilde{c}_1)$ il quale assume i valori

$$\begin{aligned} C(1; \tilde{c}_1) &= \frac{m_1 \tilde{y}_{11}^2}{\tilde{\lambda}_1} = \frac{1}{6} \\ C(2; \tilde{c}_1) &= \frac{m_2 \tilde{y}_{21}^2}{\tilde{\lambda}_1} = \frac{1}{6} \\ C(3; \tilde{c}_1) &= \frac{m_3 \tilde{y}_{31}^2}{\tilde{\lambda}_1} = \frac{2}{3}. \end{aligned}$$

Il contributo maggiore, come risulta del resto dall'esame diretto del grafico, è fornito dal terzo individuo.

RAPPRESENTAZIONE GRAFICA DELLE VARIABILI

Una rappresentazione grafica delle variabili (misurate in termini di scarti standardizzati) si può ottenere considerando le loro proiezioni ortogonali $\hat{\mathbf{y}}_1^*, \hat{\mathbf{y}}_2^*$ nel sottospazio generato dal vettore $\tilde{\mathbf{y}}_1$.

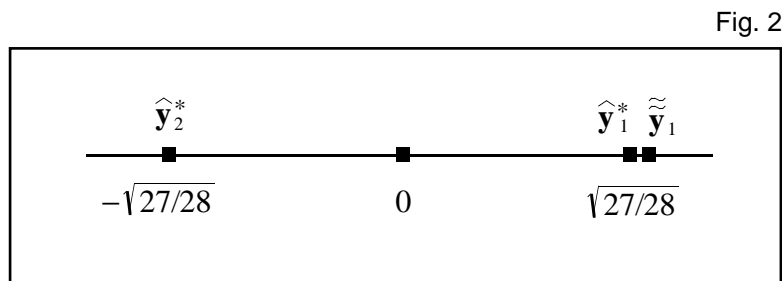
Poiché le coordinate, rispetto a $\tilde{\mathbf{y}}_1$, di tali proiezioni ortogonali sono

$$r_{11} = [-1 \quad -2 \quad 3] \sqrt{3/14} \frac{1}{3} \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \\ -2/\sqrt{2} \end{bmatrix} = \sqrt{27/28}$$

e

$$r_{21} = [2 \quad 1 \quad -3] \sqrt{3/14} \frac{1}{3} \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \\ 2/\sqrt{2} \end{bmatrix} = -\sqrt{27/28}$$

si ottiene il grafico di Fig. 2.



ALTRI APPROCCIA ALLA ACP

L'APPROCCIO IN TERMINI DI SOMMA PONDERATA DEI QUADRATI DELLE DISTANZE TRA COPPIE DI INDIVIDUI

Si considerino anzitutto i quadrati delle distanze tra coppie di individui

$$d_{1,2}^2 = \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right)' \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} -2 \\ 1 \end{bmatrix} \right) = [1 \quad 1] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 = d_{2,1}^2$$

$$\mathbf{d}_{1,3}^2 = \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \end{bmatrix} \right)' \left(\begin{bmatrix} -1 \\ 2 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \end{bmatrix} \right) = [-4 \quad 5] \begin{bmatrix} -4 \\ 5 \end{bmatrix} = 41 = \mathbf{d}_{3,1}^2$$

$$\mathbf{d}_{2,3}^2 = \left(\begin{bmatrix} -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \end{bmatrix} \right)' \left(\begin{bmatrix} -2 \\ 1 \end{bmatrix} - \begin{bmatrix} 3 \\ -3 \end{bmatrix} \right) = [-5 \quad 4] \begin{bmatrix} -5 \\ 4 \end{bmatrix} = 41 = \mathbf{d}_{3,2}^2$$

e la corrispondente matrice dei quadrati delle distanze

$$\mathbf{D}^2 = \begin{bmatrix} 0 & 2 & 41 \\ 2 & 0 & 41 \\ 41 & 41 & 0 \end{bmatrix}.$$

Come subito si verifica, risulta

$$\text{tr}(\mathbf{M}\mathbf{u}\mathbf{u}'\mathbf{M}\mathbf{D}^2) = \frac{1}{9} 168 = 2 \frac{28}{3},$$

vale a dire che la somma ponderata dei quadrati delle distanze tra coppie di individui è pari al doppio della inerzia rispetto al baricentro (28/3).

Successivamente, si considerino le proiezioni ortogonali di $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ nel sottospazio C_1 generato dal vettore $\mathbf{c}_1 = \frac{1}{5} [3 \quad 4]'$, vale a dire

$$\mathbf{P}\mathbf{y}_1 = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} 15 \\ 20 \end{bmatrix}$$

$$\mathbf{P}\mathbf{y}_2 = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} -6 \\ -8 \end{bmatrix}$$

$$\mathbf{P}\mathbf{y}_3 = \frac{1}{25} \begin{bmatrix} 9 & 12 \\ 12 & 16 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \frac{1}{25} \begin{bmatrix} -9 \\ -12 \end{bmatrix}.$$

I quadrati delle distanze tra coppie di tali proiezioni ortogonali risultano

$$\widehat{\mathbf{d}}_{1,2}^2 = \left(\frac{1}{25} \begin{bmatrix} 15 \\ 20 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -6 \\ -8 \end{bmatrix} \right)' \left(\frac{1}{25} \begin{bmatrix} 15 \\ 20 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -6 \\ -8 \end{bmatrix} \right) = \frac{49}{25} = \widehat{\mathbf{d}}_{2,1}^2$$

$$\widehat{\mathbf{d}}_{1,3}^2 = \left(\frac{1}{25} \begin{bmatrix} 15 \\ 20 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -9 \\ -12 \end{bmatrix} \right)' \left(\frac{1}{25} \begin{bmatrix} 15 \\ 20 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -9 \\ -12 \end{bmatrix} \right) = \frac{64}{25} = \widehat{\mathbf{d}}_{3,1}^2$$

$$\widehat{\mathbf{d}}_{2,3}^2 = \left(\frac{1}{25} \begin{bmatrix} -6 \\ -8 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -9 \\ -12 \end{bmatrix} \right)' \left(\frac{1}{25} \begin{bmatrix} -6 \\ -8 \end{bmatrix} - \frac{1}{25} \begin{bmatrix} -9 \\ -12 \end{bmatrix} \right) = \frac{1}{25} = \widehat{\mathbf{d}}_{3,2}^2$$

e la corrispondente matrice dei quadrati delle distanze è

$$\widehat{\mathbf{D}}^2 = \frac{1}{25} \begin{bmatrix} 0 & 49 & 64 \\ 49 & 0 & 1 \\ 64 & 1 & 0 \end{bmatrix}.$$

Pertanto, si ha che

$$\text{tr}(\mathbf{M}\mathbf{u}\mathbf{u}'\widehat{\mathbf{D}}^2) = \frac{1}{9} \frac{228}{25} = \frac{76}{75},$$

vale a dire che la somma ponderata dei quadrati delle distanze tra coppie di proiezioni ortogonali di $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$ nel sottospazio C_1 generato dal vettore $\mathbf{c}_1 = \frac{1}{5} [3 \ 4]'$ è pari al doppio della inerzia spiegata da $\mathbf{g} + C_1$ ($38/75$).

L'APPROCCIO IN TERMINI DEL MODELLO LINEARE MULTIVARIATO

Nel caso specifico, il modello a cui fare riferimento è della forma

$$\mathbf{Y} = \mathbf{Y}\mathbf{F}\mathbf{G} + \mathbf{E}$$

dove \mathbf{F} e \mathbf{G} – di ordine, rispettivamente, $(2, 1)$ e $(1, 2)$ – sono di rango pari a 1.

Pertanto, con i dati a nostra disposizione, risultando

$$\tilde{\mathbf{F}} = \tilde{\mathbf{c}}_1 \tilde{\mathbf{D}}_1^{-1/2} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \frac{1}{3} = \begin{bmatrix} 1/3\sqrt{2} \\ -1/3\sqrt{2} \end{bmatrix}$$

e

$$\tilde{\mathbf{G}} = \tilde{\mathbf{F}}' \mathbf{V} = [1/3\sqrt{2} \ -1/3\sqrt{2}] \frac{1}{3} \begin{bmatrix} 14 & -13 \\ -13 & 14 \end{bmatrix} = [3/\sqrt{2} \ -3/\sqrt{2}],$$

si ha

$$\tilde{\mathbf{H}} = \tilde{\mathbf{F}} \tilde{\mathbf{G}} = \begin{bmatrix} 1/3\sqrt{2} \\ -1/3\sqrt{2} \end{bmatrix} [3/\sqrt{2} \ -3/\sqrt{2}] = \begin{bmatrix} 1/2 & -1/2 \\ -1/2 & 1/2 \end{bmatrix}.$$

ESEMPIO 2

Questo esempio, basato su dati fittizi riferiti a 9 individui e 9 variabili, ha lo scopo di consentire al lettore di acquisire dimestichezza con i risultati che provengono dalla applicazione del programma, scritto in linguaggio Matlab, predisposto per la ACP.

Gli individui sono identificati con le etichette $1, 2, \dots, 9$ e le variabili con le etichette a, b, \dots, i .

COMMENTI

- La prima parte (*Preliminary Statistics*) è dedicata alla esposizione delle principali caratteristiche dei dati a nostra disposizione.
- La seconda parte (*Eigenvalues, Principal Vectors*) e la terza (*Principal Components and Standardized Principal Components*) mostrano alcuni risultati che si ottengono dalla ACP.

Sono altresì indicati i controlli effettuati sulla correttezza di tali risultati.

- La quarta parte (*Diagnostics*) è dedicata alla presentazione di alcuni indici che servono a interpretare correttamente i grafici contenuti nella quinta parte (*Graphics*).

Gli individui risultano ben rappresentati, sia globalmente sia singolarmente, nel piano principale. Gli individui 7, 8, 9 contribuiscono in misura preponderante all'inerzia spiegata dal primo asse, mentre i rimanenti individui 1, 2, 3, 4, 5, 6 contribuiscono in misura prevalente all'inerzia spiegata dal secondo asse.

La prima componente principale risulta fortemente correlata con le variabili g, h, i (con segno negativo), mentre la seconda componente principale è significativamente correlata con le variabili a, b, c (con segno negativo) e con le variabili d, e, f (con segno positivo). Tutte le variabili sono ben rappresentate nel cerchio di correlazione.

Come era da attendersi data la particolare struttura dei dati grezzi esaminati, sia gli individui sia le variabili si dispongono in ben distinti raggruppamenti, rispettivamente, nel piano principale e nel cerchio di correlazione.

RISULTATI NUMERICI E GRAFICI

PRELIMINARY STATISTICS

X: RAW DATA MATRIX

	a	b	c	d	e	f
1	2.0000	1.0000	3.0000	0.0000	0.0000	0.0000
2	3.0000	1.0000	2.0000	0.0000	0.0000	0.0000
3	1.0000	2.0000	3.0000	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000	2.0000	4.0000	6.0000
5	0.0000	0.0000	0.0000	6.0000	4.0000	2.0000
6	0.0000	0.0000	0.0000	4.0000	6.0000	2.0000
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

	g	h	i
1	0.0000	0.0000	0.0000
2	0.0000	0.0000	0.0000
3	0.0000	0.0000	0.0000
4	0.0000	0.0000	0.0000
5	0.0000	0.0000	0.0000
6	0.0000	0.0000	0.0000
7	6.0000	5.0000	7.0000
8	7.0000	5.0000	6.0000
9	5.0000	6.0000	7.0000

(n,p): ORDER OF X

n = 9 , p = 9

rX: RANK OF X

9

CHOICE OF METRIC M

(if c = 1, then M = (1/n)In; if c = 2, then M = Mn)

c = 1

g: BARYCENTRE

a	0.6667	f	1.1111
b	0.4444	g	2.0000
c	0.8889	h	1.7778
d	1.3333	i	2.2222
e	1.5556		

Y: MEAN CENTRED DATA MATRIX

	a	b	c	d	e	f
1	1.3333	0.5556	2.1111	-1.3333	-1.5556	-1.1111
2	2.3333	0.5556	1.1111	-1.3333	-1.5556	-1.1111
3	0.3333	1.5556	2.1111	-1.3333	-1.5556	-1.1111
4	-0.6667	-0.4444	-0.8889	0.6667	2.4444	4.8889
5	-0.6667	-0.4444	-0.8889	4.6667	2.4444	0.8889
6	-0.6667	-0.4444	-0.8889	2.6667	4.4444	0.8889
7	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111
8	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111
9	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111

	g	h	i
1	-2.0000	-1.7778	-2.2222
2	-2.0000	-1.7778	-2.2222
3	-2.0000	-1.7778	-2.2222
4	-2.0000	-1.7778	-2.2222
5	-2.0000	-1.7778	-2.2222
6	-2.0000	-1.7778	-2.2222
7	4.0000	3.2222	4.7778
8	5.0000	3.2222	3.7778
9	3.0000	4.2222	4.7778

r: RANK OF Y

8

V: COVARIANCE MATRIX

	a	b	c	d	e	f
a	1.1111	0.4815	1.0741	-0.8889	-1.0370	-0.7407
b	0.4815	0.4691	0.8272	-0.5926	-0.6914	-0.4938
c	1.0741	0.8272	1.6543	-1.1852	-1.3827	-0.9877
d	-0.8889	-0.5926	-1.1852	4.4444	4.1481	2.0741
e	-1.0370	-0.6914	-1.3827	4.1481	5.1358	3.1605
f	-0.7407	-0.4938	-0.9877	2.0741	3.1605	3.6543
g	-1.3333	-0.8889	-1.7778	-2.6667	-3.1111	-2.2222
h	-1.1852	-0.7901	-1.5802	-2.3704	-2.7654	-1.9753
i	-1.4815	-0.9877	-1.9753	-2.9630	-3.4568	-2.4691

	g	h	i
a	-1.3333	-1.1852	-1.4815
b	-0.8889	-0.7901	-0.9877
c	-1.7778	-1.5802	-1.9753
d	-2.6667	-2.3704	-2.9630
e	-3.1111	-2.7654	-3.4568
f	-2.2222	-1.9753	-2.4691
g	8.2222	7.0000	8.7778
h	7.0000	6.3951	7.9383
i	8.7778	7.9383	9.9506

Vars: VARIANCES

a	1.1111	f	3.6543
b	0.4691	g	8.2222
c	1.6543	h	6.3951
d	4.4444	i	9.9506
e	5.1358		

CHOICE OF METRIC Q

(if $c = 1$, then $Q = I_p$; if $c = 2$, then $Q = \text{diag}(1/\text{var}(x(j)))$; if $c = 3$, then $Q = Q_p$)

$c = 2$

R: CORRELATION MATRIX

	a	b	c	d	e	f
a	1.0000	0.6669	0.7922	-0.4000	-0.4341	-0.3676
b	0.6669	1.0000	0.9389	-0.4104	-0.4454	-0.3772
c	0.7922	0.9389	1.0000	-0.4371	-0.4744	-0.4017
d	-0.4000	-0.4104	-0.4371	1.0000	0.8682	0.5147
e	-0.4341	-0.4454	-0.4744	0.8682	1.0000	0.7295
f	-0.3676	-0.3772	-0.4017	0.5147	0.7295	1.0000

	g	h	i
a	-0.4411	-0.4446	-0.4455
b	-0.4526	-0.4562	-0.4571
c	-0.4820	-0.4858	-0.4869
d	-0.4411	-0.4446	-0.4455
e	-0.4788	-0.4825	-0.4836
f	-0.4054	-0.4086	-0.4095
g	1.0000	0.9653	0.9704
h	0.9653	1.0000	0.9951
i	0.9704	0.9951	1.0000

EIGENVALUES, PRINCIPAL VECTORS

Lambdas: EIGENVALUES

(1)	4.2225	(6)	0.0419
(2)	3.7510	(7)	0.0414
(3)	0.4981	(8)	0.0046
(4)	0.3566	(9)	0.0000
(5)	0.0839		

Sum of Lambdas: INERTIA I_g or GLOBAL VARIABILITY J_p

PV: PRINCIPAL VECTOR MATRIX

	(1)	(2)	(3)	(4)	(5)	(6)
(1)	0.2901	-0.3683	0.0007	0.8268	0.0007	-0.2720
(2)	0.1985	-0.2528	0.0010	-0.3932	0.0004	-0.4317
(3)	0.3927	-0.4995	0.0017	-0.2934	0.0005	0.9286
(4)	0.3162	0.9129	-1.2485	-0.0527	-1.1868	-0.0842
(5)	0.3691	1.0659	-0.2858	-0.0608	1.7481	-0.0968
(6)	0.2639	0.7623	1.5214	-0.0421	-0.5672	-0.0691
(7)	-1.3557	-0.2498	0.0031	-0.0623	0.0014	0.0850
(8)	-1.2056	-0.2222	0.0027	-0.0557	0.0013	-0.1853
(9)	-1.5066	-0.2776	0.0034	-0.0693	0.0016	-0.1996

	(7)	(8)	(9)
(1)	0.0186	-0.0172	0.3649
(2)	0.0333	-0.0070	0.1541
(3)	-0.0791	-0.0264	0.5433
(4)	0.0005	-0.0346	0.7298
(5)	0.0006	-0.0400	0.8433
(6)	0.0004	-0.0285	0.6000
(7)	2.3337	-0.2333	0.9000
(8)	-1.1624	-1.7359	0.7000
(9)	-1.0921	2.2759	1.0892

Check PV: Does $\text{transpose(PV)} * Q * PV$ equal I_p ?

ans =

	(1)	(2)	(3)	(4)	(5)	(6)
(1)	1.0000	-0.0000	0.0000	-0.0000	-0.0000	0.0000
(2)	-0.0000	1.0000	0.0000	0.0000	-0.0000	0.0000
(3)	0.0000	0.0000	1.0000	-0.0000	-0.0000	0.0000
(4)	-0.0000	0.0000	-0.0000	1.0000	0.0000	0.0000
(5)	-0.0000	-0.0000	-0.0000	0.0000	1.0000	-0.0000
(6)	0.0000	0.0000	0.0000	0.0000	-0.0000	1.0000
(7)	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000
(8)	0.0000	-0.0000	-0.0000	-0.0000	-0.0000	0.0000
(9)	0.0000	-0.0000	0.0000	0.0000	-0.0000	-0.0000

	(7)	(8)	(9)
(1)	-0.0000	0.0000	0.0000
(2)	-0.0000	-0.0000	-0.0000
(3)	-0.0000	-0.0000	0.0000
(4)	-0.0000	-0.0000	0.0000
(5)	-0.0000	-0.0000	-0.0000
(6)	0.0000	0.0000	-0.0000
(7)	1.0000	-0.0000	0.0000
(8)	-0.0000	1.0000	0.0000
(9)	0.0000	0.0000	1.0000

PRINCIPAL COMPONENTS (PC) AND STANDARDIZED PRINCIPAL COMPONENTS (SPC)

CHOICE OF NUMBER k OF PC AND SPC TO BE PRINTED (k = 2,3)

k = 2

FIRST k PC

	PC1	PC2		PC1	PC2
1	1.7988	-2.0228	6	1.0016	2.5691
2	1.8226	-2.0523	7	-2.8505	-0.4663
3	1.9608	-2.2303	8	-2.8639	-0.4688
4	1.0045	2.5775	9	-2.8741	-0.4707
5	1.0002	2.5648			

Reconstitution formula: Does PC*transpose(PV) equal Y ?

ans =

	a	b	c	d	e	f
1	1.3333	0.5556	2.1111	-1.3333	-1.5556	-1.1111
2	2.3333	0.5556	1.1111	-1.3333	-1.5556	-1.1111
3	0.3333	1.5556	2.1111	-1.3333	-1.5556	-1.1111
4	-0.6667	-0.4444	-0.8889	0.6667	2.4444	4.8889
5	-0.6667	-0.4444	-0.8889	4.6667	2.4444	0.8889
6	-0.6667	-0.4444	-0.8889	2.6667	4.4444	0.8889
7	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111
8	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111
9	-0.6667	-0.4444	-0.8889	-1.3333	-1.5556	-1.1111

	g	h	i
1	-2.0000	-1.7778	-2.2222
2	-2.0000	-1.7778	-2.2222
3	-2.0000	-1.7778	-2.2222
4	-2.0000	-1.7778	-2.2222
5	-2.0000	-1.7778	-2.2222
6	-2.0000	-1.7778	-2.2222
7	4.0000	3.2222	4.7778
8	5.0000	3.2222	3.7778
9	3.0000	4.2222	4.7778

FIRST k SPC

	SPC1	SPC2		SPC1	SPC2
1	0.8754	-1.0444	6	1.0016	2.5691
2	0.8870	-1.0597	7	-2.8505	-0.4663
3	0.9542	-1.1516	8	-2.8639	-0.4688
4	0.4888	1.3308	9	-2.8741	-0.4707
5	0.4867	1.3243			

DIAGNOSTICS

GQRI: GLOBAL QUALITY OF REPRESENTATION OF INDIVIDUALS ON PRINCIPAL PLANE

0.8859

QRi: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL ON PRINCIPAL PLANE

1	0.9594	6	0.9011
2	0.8380	7	0.9966
3	0.8290	8	0.9786
4	0.7233	9	0.9767
5	0.8369		

Ci;1,2: CONTRIBUTION OF EACH INDIVIDUAL TO INERTIA EXPLAINED BY AXES

	Ci;1	Ci;2		Ci;1	Ci;2
1	0.0851	0.1212	6	0.0264	0.1955
2	0.0874	0.1248	7	0.2138	0.0064
3	0.1012	0.1474	8	0.2158	0.0065
4	0.0266	0.1968	9	0.2174	0.0066
5	0.0263	0.1949			

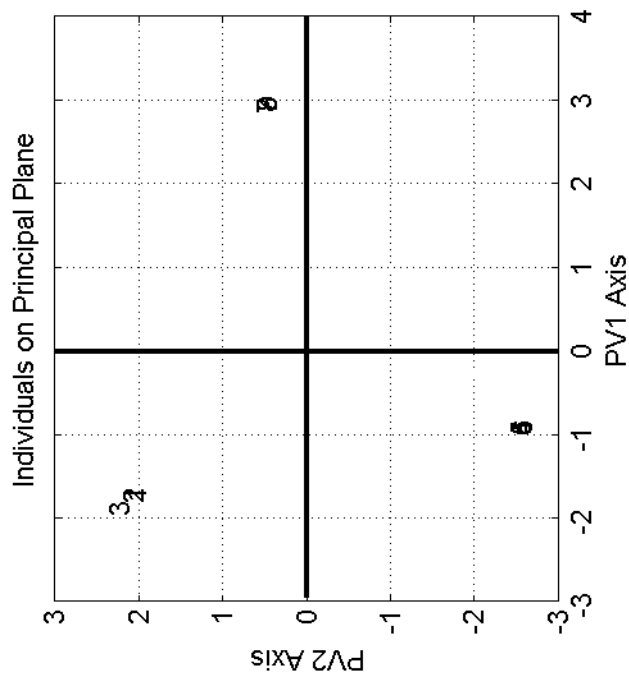
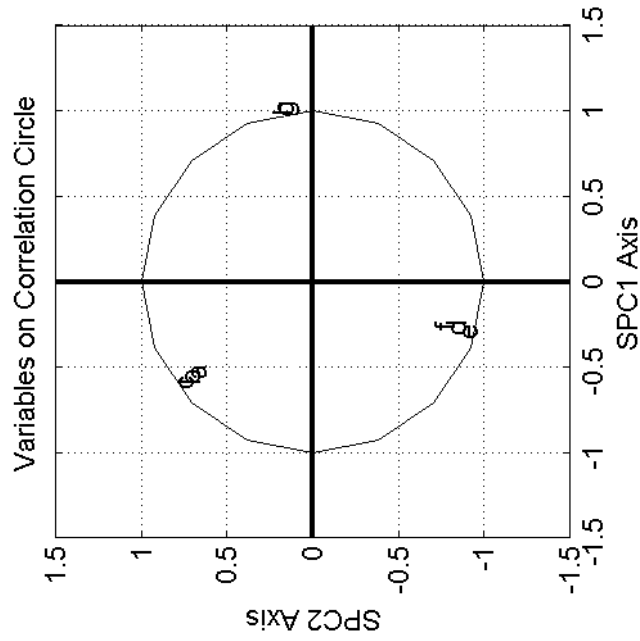
rj1,rj2: CORRELATIONS BETWEEN VARIABLES AND FIRST TWO PC

	rj1	rj2		rj1	rj2
a	0.5656	-0.6767	f	0.2837	0.7723
b	0.5955	-0.7150	g	-0.9715	-0.1687
c	0.6274	-0.7522	h	-0.9797	-0.1702
d	0.3082	0.8387	i	-0.9814	-0.1705
e	0.3347	0.9109			

QRj: QUALITY OF REPRESENTATION OF EACH VARIABLE ON CORRELATION CIRCLE

a	0.7778	f	0.6769
b	0.8658	g	0.9724
c	0.9595	h	0.9887
d	0.7984	i	0.9923
e	0.9418		

GRAPHICS



ESEMPIO 3

Questo esempio analizza i risultati conseguiti, durante i Giochi Olimpici del 1992, dai 28 atleti che hanno partecipato alle seguenti 5 prove di decathlon: 100, 400, 110h, peso, disco.

Gli atleti sono identificati con le etichette 1, 2, ..., 28, che riflettono anche l'ordinamento basato sul punteggio complessivo, e le prove con le etichette 100, 400, ..., disc.

Le unità di misura delle prove sono i secondi per i tempi e i metri per le lunghezze.

COMMENTI

- La prima parte (*Preliminary Statistics*) è dedicata alla esposizione delle principali caratteristiche dei dati a nostra disposizione.
- La seconda parte (*Eigenvalues, Principal Vectors*) e la terza (*Principal Components and Standardized Principal Components*) mostrano alcuni risultati che si ottengono dalla ACP.
- La quarta parte (*Diagnostics*) è dedicata alla presentazione di alcuni indici che servono a interpretare correttamente i grafici contenuti nella quinta parte (*Graphics*).

La qualità globale della rappresentazione degli atleti nel piano principale è ottima. Lo stesso può dirsi, salvo alcune eccezioni, della qualità di rappresentazione di ciascun atleta. Gli atleti 27 e 28 contribuiscono in misura preponderante all'inerzia spiegata dal primo asse, mentre gli atleti 5, 7, 27 contribuiscono in misura prevalente all'inerzia spiegata dal secondo asse.

La prima componente principale risulta la più significativa, in quanto assorbe oltre il 68% della variabilità globale ($3.4004/5 = 0.68008$). Tale componente è fortemente correlata sia con le prove 100, 400, 110h (con segno positivo), sia con le prove shot, disc (con segno negativo). Tutte le variabili sono ben rappresentate nel cerchio di correlazione.

RISULTATI NUMERICI E GRAFICI

PRELIMINARY STATISTICS

X: RAW DATA MATRIX

	100	400	110h		shot	disc
1	10.7800	48.6500	13.9500	1	14.5300	45.0000
2	11.0900	49.6600	14.5800	2	16.5000	49.6800
3	11.1600	49.7600	14.7600	3	15.2800	49.1200
4	11.0900	48.2400	14.8600	4	13.7300	39.2200
5	11.3600	50.0000	14.7500	5	16.0200	50.7400
6	10.7500	48.3300	15.2200	6	15.3400	42.1400
7	11.4200	50.4400	15.0200	7	15.4400	50.5800
8	10.9700	49.3000	14.7800	8	14.3500	45.0800
9	11.1100	48.9100	14.9700	9	14.3200	45.7600
10	11.0900	49.2400	14.8900	10	15.4800	48.3400
11	10.4900	47.9100	14.6400	11	14.5600	45.0600
12	11.3000	50.6200	14.7800	12	15.2100	44.8400
13	10.9500	47.9800	14.5100	13	13.5500	42.5600
14	11.0300	49.0300	14.6000	14	13.1100	44.5800
15	11.3100	49.1900	14.5000	15	13.5600	37.7800
16	10.9500	48.1400	15.1800	16	14.4100	43.5200
17	10.8000	48.3300	14.7600	17	15.2900	44.7000
18	11.1400	47.7300	14.9400	18	12.5600	37.4200
19	10.9700	50.0600	15.2900	19	13.6700	42.5400
20	11.1300	48.6900	14.8100	20	12.5600	41.2200
21	11.5600	50.0800	15.1700	21	12.9200	43.2600
22	11.4100	49.6000	14.8600	22	12.9900	40.1000
23	11.4000	51.5400	15.4200	23	13.4000	39.8000
24	11.4000	50.7600	15.0800	24	13.2100	39.9000
25	11.5500	51.7800	16.6800	25	13.5700	39.5200
26	11.1400	51.4500	15.7000	26	11.9700	40.6800
27	11.6300	51.4000	15.9800	27	8.6200	24.5800
28	11.7500	54.8100	16.2000	28	9.5400	30.2600

(n,p): ORDER OF X

n = 28 , p = 5

rX: RANK OF X

5

CHOICE OF METRIC M

(if c = 1, then M = (1/n)In; if c = 2, then M = Mn)

c = 1

g: BARYCENTRE

100 11.1689
 400 49.7011
 110h 15.0314
 shot 13.7746
 disc 42.4279

Y: MEAN CENTRED DATA MATRIX

	100	400	110h		shot	disc
1	-0.3889	-1.0511	-1.0814	1	0.7554	2.5721
2	-0.0789	-0.0411	-0.4514	2	2.7254	7.2521
3	-0.0089	0.0589	-0.2714	3	1.5054	6.6921
4	-0.0789	-1.4611	-0.1714	4	-0.0446	-3.2079
5	0.1911	0.2989	-0.2814	5	2.2454	8.3121
6	-0.4189	-1.3711	0.1886	6	1.5654	-0.2879
7	0.2511	0.7389	-0.0114	7	1.6654	8.1521
8	-0.1989	-0.4011	-0.2514	8	0.5754	2.6521
9	-0.0589	-0.7911	-0.0614	9	0.5454	3.3321
10	-0.0789	-0.4611	-0.1414	10	1.7054	5.9121
11	-0.6789	-1.7911	-0.3914	11	0.7854	2.6321
12	0.1311	0.9189	-0.2514	12	1.4354	2.4121
13	-0.2189	-1.7211	-0.5214	13	-0.2246	0.1321
14	-0.1389	-0.6711	-0.4314	14	-0.6646	2.1521
15	0.1411	-0.5111	-0.5314	15	-0.2146	-4.6479
16	-0.2189	-1.5611	0.1486	16	0.6354	1.0921
17	-0.3689	-1.3711	-0.2714	17	1.5154	2.2721
18	-0.0289	-1.9711	-0.0914	18	-1.2146	-5.0079
19	-0.1989	0.3589	0.2586	19	-0.1046	0.1121
20	-0.0389	-1.0111	-0.2214	20	-1.2146	-1.2079
21	0.3911	0.3789	0.1386	21	-0.8546	0.8321
22	0.2411	-0.1011	-0.1714	22	-0.7846	-2.3279
23	0.2311	1.8389	0.3886	23	-0.3746	-2.6279
24	0.2311	1.0589	0.0486	24	-0.5646	-2.5279
25	0.3811	2.0789	1.6486	25	-0.2046	-2.9079
26	-0.0289	1.7489	0.6686	26	-1.8046	-1.7479
27	0.4611	1.6989	0.9486	27	-5.1546	-17.8479
28	0.5811	5.1089	1.1686	28	-4.2346	-12.1679

r: RANK OF Y

5

V: COVARIANCE MATRIX

	100	400	110h	shot	disc
100	0.0819	0.3357	0.0929	-0.2515	-0.7129
400	0.3357	2.3165	0.6012	-1.3064	-3.5526
110h	0.0929	0.6012	0.3031	-0.5313	-1.7543
shot	-0.2515	-1.3064	-0.5313	2.9483	8.7387
disc	-0.7129	-3.5526	-1.7543	8.7387	31.1583

Vars: VARIANCES

100	0.0819
400	2.3165
110h	0.3031
shot	2.9483
disc	31.1583

CHOICE OF METRIC Q

(if $c = 1$, then $Q = I_p$; if $c = 2$, then $Q = \text{diag}(1/\text{var}(x(j)))$; if $c = 3$, then $Q = Q_p$)

$c = 2$

R: CORRELATION MATRIX

	100	400	110h	shot	disc
100	1.0000	0.7707	0.5893	-0.5117	-0.4462
400	0.7707	1.0000	0.7174	-0.4999	-0.4182
110h	0.5893	0.7174	1.0000	-0.5620	-0.5709
shot	-0.5117	-0.4999	-0.5620	1.0000	0.9118
disc	-0.4462	-0.4182	-0.5709	0.9118	1.0000

EIGENVALUES, PRINCIPAL VECTORS

Lambdas: EIGENVALUES

(1)	3.4004
(2)	0.9289
(3)	0.4013
(4)	0.1920
(5)	0.0774

Sum of Lambdas: INERTIA I_g or GLOBAL VARIABILITY J_p

PV: PRINCIPAL VECTOR MATRIX

	(1)	(2)	(3)	(4)	(5)
(1)	0.1244	-0.1187	0.1709	-0.1520	0.0038
(2)	0.6811	-0.7293	-0.0035	1.1167	-0.2716
(3)	0.2494	-0.0997	-0.4303	-0.1999	0.0766
(4)	-0.7889	-0.8445	-0.3050	-0.3583	-1.1796
(5)	-2.4617	-3.1752	-0.1747	0.3644	3.8540

PRINCIPAL COMPONENTS (PC) AND STANDARDIZED PRINCIPAL COMPONENTS (SPC)

CHOICE OF NUMBER k OF PC AND SPC TO BE PRINTED (k = 2,3)

k = 2

FIRST k PC

	PC1	PC2		PC1	PC2
1	-2.1949	0.7719	16	-0.9256	0.4667
2	-1.8055	-1.2439	17	-1.7719	0.3901
3	-1.1510	-1.0295	18	0.0220	1.5508
4	-0.4251	0.9705	19	0.0352	0.1089
5	-1.1109	-1.7688	20	-0.1181	0.9186
6	-1.2805	0.5579	21	0.9825	-0.5718
7	-0.5004	-1.9006	22	0.5894	0.2007
8	-0.9905	0.0623	23	1.5192	-0.6666
9	-0.7818	-0.1411	24	1.0531	-0.2650
10	-1.2952	-0.7849	25	2.8309	-1.3941
11	-2.2981	1.1836	26	1.6412	-0.0335
12	-0.3122	-1.0536	27	4.7696	1.7801
13	-1.2179	1.0816	28	5.4406	-0.3822
14	-0.7555	0.5256			
15	0.0515	0.6663			

FIRST k SPC

	SPC1	SPC2		SPC1	SPC2
1	-1.1903	0.8009	16	-0.5020	0.4842
2	-0.9791	-1.2906	17	-0.9609	0.4048
3	-0.6242	-1.0682	18	0.0119	1.6091
4	-0.2305	1.0069	19	0.0191	0.1130
5	-0.6024	-1.8352	20	-0.0641	0.9531
6	-0.6944	0.5789	21	0.5328	-0.5933
7	-0.2714	-1.9720	22	0.3196	0.2082
8	-0.5371	0.0646	23	0.8239	-0.6916
9	-0.4240	-0.1464	24	0.5711	-0.2750
10	-0.7024	-0.8143	25	1.5352	-1.4465
11	-1.2463	1.2281	26	0.8900	-0.0348
12	-0.1693	-1.0932	27	2.5865	1.8469
13	-0.6605	1.1222	28	2.9504	-0.3965
14	-0.4097	0.5453			
15	0.0279	0.6913			

DIAGNOSTICS

GQRI: GLOBAL QUALITY OF REPRESENTATION OF INDIVIDUALS ON PRINCIPAL PLANE

0.8659

QRI: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL ON PRINCIPAL PLANE

1	0.8217	16	0.5700
2	0.9699	17	0.8992
3	0.9728	18	0.7965
4	0.7875	19	0.0172
5	0.9336	20	0.7339
6	0.4996	21	0.5711
7	0.9469	22	0.3247
8	0.8960	23	0.9560
9	0.8069	24	0.8095
10	0.9793	25	0.7725
11	0.8406	26	0.6724
12	0.7238	27	0.9951
13	0.9549	28	0.9680
14	0.6309		
15	0.2237		

Ci;1,2: CONTRIBUTION OF EACH INDIVIDUAL TO INERTIA EXPLAINED BY AXES

	Ci;1	Ci;2		Ci;1	Ci;2
1	0.0506	0.0229	16	0.0090	0.0084
2	0.0342	0.0595	17	0.0330	0.0059
3	0.0139	0.0408	18	0.0000	0.0925
4	0.0019	0.0362	19	0.0000	0.0005
5	0.0130	0.1203	20	0.0001	0.0324
6	0.0172	0.0120	21	0.0101	0.0126
7	0.0026	0.1389	22	0.0036	0.0015
8	0.0103	0.0001	23	0.0242	0.0171
9	0.0064	0.0008	24	0.0116	0.0027
10	0.0176	0.0237	25	0.0842	0.0747
11	0.0555	0.0539	26	0.0283	0.0000
12	0.0010	0.0427	27	0.2389	0.1218
13	0.0156	0.0450	28	0.3109	0.0056
14	0.0060	0.0106			
15	0.0000	0.0171			

rj1,rj2: CORRELATIONS BETWEEN VARIABLES AND FIRST TWO PC

	rj1	rj2
100	0.8018	-0.3999
400	0.8252	-0.4618
110h	0.8352	-0.1745
shot	-0.8472	-0.4740
disc	-0.8132	-0.5482

QRj: QUALITY OF REPRESENTATION OF EACH VARIABLE ON CORRELATION CIRCLE

100	0.8027
400	0.8942
110h	0.7280
shot	0.9425
disc	0.9619

GRAPHICS

