

RENATO LEONI

**Esempi numerici riguardanti l'analisi
della correlazione canonica**

**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 della correlazione canonica (ACC).

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

Leoni, R., *Canonical Correlation Analysis*, Department of Statistics "G. Parenti", Florence, 2007.

ESEMPIO 1

Questo esempio, basato su dati fittizi riferiti a 5 individui e 4 variabili, ha lo scopo di illustrare i principali concetti della ACC, 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

$$\mathbf{X} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

Le variabili sono rappresentate dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

appartenenti a \mathbb{R}^5 , mentre gli individui sono rappresentati dai vettori

$$\mathbf{x}_1 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{x}_4 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{x}_5 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

appartenenti a \mathbb{R}^4 .

Ciò premesso, in \mathbb{R}^5 (spazio delle variabili) e in \mathbb{R}^4 (spazio degli individui) è definita una metrica euclidea nel modo seguente.

- In \mathbb{R}^5 la matrice rappresentativa della metrica – rispetto alla base canonica di \mathbb{R}^5 – è

$$\mathbf{M} = \frac{1}{5} \text{diag}(1, 1, 1, 1, 1) = \frac{1}{5} \mathbf{I}_5.$$

Ne consegue che, poiché il baricentro o vettore delle medie è

$$\mathbf{g} = \mathbf{X}'\mathbf{M}\mathbf{u} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \frac{1}{5} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.4 \\ 0.4 \\ 0.4 \end{bmatrix},$$

la matrice dei dati centrati risulta

$$\mathbf{Y} = \begin{bmatrix} -0.2 & 0.6 & -0.4 & -0.4 \\ -0.2 & -0.4 & 0.6 & -0.4 \\ -0.2 & -0.4 & 0.6 & -0.4 \\ 0.8 & -0.4 & -0.4 & 0.6 \\ -0.2 & 0.6 & -0.4 & 0.6 \end{bmatrix}.$$

I vettori

$$\mathbf{y}_1 = \begin{bmatrix} -0.2 \\ -0.2 \\ -0.2 \\ 0.8 \\ -0.2 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} 0.6 \\ -0.4 \\ -0.4 \\ -0.4 \\ 0.6 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} -0.4 \\ 0.6 \\ 0.6 \\ -0.4 \\ -0.4 \end{bmatrix}, \quad \mathbf{y}_4 = \begin{bmatrix} -0.4 \\ -0.4 \\ -0.4 \\ 0.6 \\ 0.6 \end{bmatrix}$$

e

$$\mathbf{y}_1 = \begin{bmatrix} -0.2 \\ 0.6 \\ -0.4 \\ -0.4 \end{bmatrix}, \quad \mathbf{y}_2 = \begin{bmatrix} -0.2 \\ -0.4 \\ 0.6 \\ -0.4 \end{bmatrix}, \quad \mathbf{y}_3 = \begin{bmatrix} -0.2 \\ -0.4 \\ 0.6 \\ -0.4 \end{bmatrix}, \quad \mathbf{y}_4 = \begin{bmatrix} 0.8 \\ -0.4 \\ -0.4 \\ 0.6 \end{bmatrix}, \quad \mathbf{y}_5 = \begin{bmatrix} -0.2 \\ 0.6 \\ -0.4 \\ 0.6 \end{bmatrix}$$

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

Infine, eseguita una ripartizione (per colonne) della matrice \mathbf{Y} nel seguente modo ($p_1 = p_2 = 2$)

$$\mathbf{Y} = [\mathbf{y}_1 \mathbf{y}_2 \mid \mathbf{y}_3 \mathbf{y}_4] = [\mathbf{Y}_1 \mathbf{Y}_2],$$

possiamo scrivere la matrice delle covarianze \mathbf{V} nella forma della matrice a blocchi

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_{11} & \mathbf{V}_{12} \\ \mathbf{V}_{21} & \mathbf{V}_{22} \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0.16 & -0.08 \\ -0.08 & 0.24 \end{bmatrix} & \begin{bmatrix} -0.08 & 0.12 \\ -0.16 & 0.04 \end{bmatrix} \\ \begin{bmatrix} -0.08 & -0.16 \\ 0.12 & 0.04 \end{bmatrix} & \begin{bmatrix} 0.24 & -0.16 \\ -0.16 & 0.24 \end{bmatrix} \end{bmatrix}.$$

Si noti che

$$r(\mathbf{V}_{11}) = p_1 = 2, \quad r(\mathbf{V}_{22}) = p_2 = 2, \quad r(\mathbf{V}_{12}) = r(\mathbf{V}_{21}) = k = 2 = p_1 = p_2.$$

- In \mathbb{R}^4 la matrice rappresentativa della metrica – rispetto alla base canonica di \mathbb{R}^4 – è

$$\mathbf{Q} = \text{diag}(\mathbf{V}_{11}^{-1}, \mathbf{V}_{22}^{-1})$$

dove

$$\mathbf{V}_{11}^{-1} = \begin{bmatrix} 7.5 & 2.5 \\ 2.5 & 5.0 \end{bmatrix}, \quad \mathbf{V}_{22}^{-1} = \begin{bmatrix} 7.5 & 5.0 \\ 5.0 & 7.5 \end{bmatrix}.$$

UN APPROCCIO ALLA ACC

FATTORI CANONICI, VARIABILI CANONICHE, COEFFICIENTI DI CORRELAZIONE CANONICA

Il primo passo

Si consideri l'equazione

$$\begin{aligned} & \det(-r^2 \mathbf{I}_{p_1} + \mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \\ &= \det(-r^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 7.5 & 2.5 \\ 2.5 & 5.0 \end{bmatrix} \begin{bmatrix} -0.08 & 0.12 \\ -0.16 & 0.04 \end{bmatrix} \begin{bmatrix} 7.5 & 5.0 \\ 5.0 & 7.5 \end{bmatrix} \begin{bmatrix} -0.08 & -0.16 \\ 0.12 & 0.04 \end{bmatrix}) \\ &= \det(-r^2 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.5 \\ 0.25 & 0.75 \end{bmatrix}) = 0 \end{aligned}$$

il cui primo autovalore è $\tilde{r}_1^2 = 1$.

Successivamente, si consideri l'equazione

$$(-(1) \mathbf{I}_{p_1} + \mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{a}_{(1)1} = \mathbf{0}$$

la quale ammette l'autovettore

$$\tilde{\mathbf{a}}_{(1)1} = \begin{bmatrix} 2.0412 \\ 2.0412 \end{bmatrix},$$

tale che $\tilde{\mathbf{a}}_{(1)1}' \mathbf{V}_{11} \tilde{\mathbf{a}}_{(1)1} = 1$.

A sua volta, l'equazione

$$\mathbf{a}_{(2)1} = \frac{1}{1} \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{a}_{(1)1}$$

dà luogo alla soluzione

$$\tilde{\mathbf{a}}_{(2)1} = \begin{bmatrix} -2.0412 \\ 0.0000 \end{bmatrix},$$

tale che $\tilde{\mathbf{a}}_{(2)1}' \mathbf{V}_{22} \tilde{\mathbf{a}}_{(2)1} = 1$.

Inoltre, come si verifica facilmente, risulta

$$\tilde{\mathbf{a}}_{(1)1}' \mathbf{V}_{12} \tilde{\mathbf{a}}_{(2)1} = \tilde{\mathbf{a}}_{(2)1}' \mathbf{V}_{21} \tilde{\mathbf{a}}_{(1)1} = \tilde{r}_1 = 1.$$

I vettori $\tilde{\mathbf{a}}_{(1)1}$ e $\tilde{\mathbf{a}}_{(2)1}$ di cui sopra costituiscono i primi due fattori canonici.

Infine, il vettore

$$\tilde{\mathbf{z}}_{(1)1} = \mathbf{Y}_1 \tilde{\mathbf{a}}_{(1)1} = \begin{bmatrix} -0.2 & 0.6 \\ -0.2 & -0.4 \\ -0.2 & -0.4 \\ 0.8 & -0.4 \\ -0.2 & 0.6 \end{bmatrix} \begin{bmatrix} 2.0412 \\ 2.0412 \end{bmatrix} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix},$$

tale che $\tilde{\mathbf{z}}_{(1)1}' \mathbf{M} \tilde{\mathbf{z}}_{(1)1} = 1$, e il vettore

$$\tilde{\mathbf{z}}_{(2)1} = \mathbf{Y}_2 \tilde{\mathbf{a}}_{(2)1} = \begin{bmatrix} -0.4 & -0.4 \\ 0.6 & -0.4 \\ 0.6 & -0.4 \\ -0.4 & 0.6 \\ -0.4 & 0.6 \end{bmatrix} \begin{bmatrix} -2.0412 \\ 0.0000 \end{bmatrix} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix},$$

tale che $\tilde{\mathbf{z}}_{(2)1}' \mathbf{M} \tilde{\mathbf{z}}_{(2)1} = 1$, costituiscono le prime due variabili canoniche corrispondenti al coefficiente di correlazione canonica $\tilde{r}_1 = 1$.

Ovviamente, risulta che

$$\tilde{\mathbf{z}}_{(1)1}' \mathbf{M} \tilde{\mathbf{z}}_{(2)1} = \tilde{\mathbf{z}}_{(2)1}' \mathbf{M} \tilde{\mathbf{z}}_{(1)1} = \tilde{r}_1 = 1.$$

Il secondo passo

Si consideri nuovamente l'equazione

$$\det(-r^2 \mathbf{I}_{p_1} + \mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) = 0$$

il cui secondo autovalore è $\tilde{r}_2^2 = 0.25$.

Successivamente, si consideri l'equazione

$$(-(0.25) \mathbf{I}_{p_1} + \mathbf{V}_{11}^{-1} \mathbf{V}_{12} \mathbf{V}_{22}^{-1} \mathbf{V}_{21}) \mathbf{a}_{(1)2} = \mathbf{0}$$

la quale ammette l'autovettore

$$\tilde{\mathbf{a}}_{(1)2} = \begin{bmatrix} -1.8257 \\ 0.9129 \end{bmatrix},$$

tale che $\tilde{\mathbf{a}}'_{(1)2} \mathbf{V}_{11} \tilde{\mathbf{a}}_{(1)2} = 1$, $\tilde{\mathbf{a}}'_{(1)2} \mathbf{V}_{11} \tilde{\mathbf{a}}_{(1)1} = 0$.

A sua volta, l'equazione

$$\mathbf{a}_{(2)2} = \frac{1}{0.5} \mathbf{V}_{22}^{-1} \mathbf{V}_{21} \mathbf{a}_{(1)2}$$

dà luogo alla soluzione

$$\tilde{\mathbf{a}}_{(2)2} = \begin{bmatrix} -1.8257 \\ -2.7386 \end{bmatrix},$$

tale che $\tilde{\mathbf{a}}'_{(2)2} \mathbf{V}_{22} \tilde{\mathbf{a}}_{(2)2} = 1$, $\tilde{\mathbf{a}}'_{(2)2} \mathbf{V}_{22} \tilde{\mathbf{a}}_{(2)1} = 0$.

Inoltre, come si verifica facilmente, risulta

$$\tilde{\mathbf{a}}'_{(1)1} \mathbf{V}_{12} \tilde{\mathbf{a}}_{(2)2} = \tilde{\mathbf{a}}'_{(2)1} \mathbf{V}_{21} \tilde{\mathbf{a}}_{(1)2} = 0$$

e

$$\tilde{\mathbf{a}}'_{(1)2} \mathbf{V}_{12} \tilde{\mathbf{a}}_{(2)2} = \tilde{\mathbf{a}}'_{(2)2} \mathbf{V}_{21} \tilde{\mathbf{a}}_{(1)2} = \tilde{r}_2 = 0.5 .$$

I vettori $\tilde{\mathbf{a}}_{(1)2}$ e $\tilde{\mathbf{a}}_{(2)2}$ di cui sopra costituiscono i secondi due fattori canonici.

Infine, il vettore

$$\tilde{\mathbf{z}}_{(1)2} = \mathbf{Y}_1 \tilde{\mathbf{a}}_{(1)2} = \begin{bmatrix} -0.2 & 0.6 \\ -0.2 & -0.4 \\ -0.2 & -0.4 \\ 0.8 & -0.4 \\ -0.2 & 0.6 \end{bmatrix} \begin{bmatrix} -1.8257 \\ 0.9129 \end{bmatrix} = \begin{bmatrix} 0.9129 \\ 0.0000 \\ 0.0000 \\ -1.8257 \\ 0.9129 \end{bmatrix},$$

tale che $\tilde{\mathbf{z}}'_{(1)2} \mathbf{M} \tilde{\mathbf{z}}_{(1)2} = 1$, $\tilde{\mathbf{z}}'_{(1)2} \mathbf{M} \tilde{\mathbf{z}}_{(1)1} = 0$, e il vettore

$$\tilde{\mathbf{z}}_{(2)2} = \mathbf{Y}_2 \tilde{\mathbf{a}}_{(2)2} = \begin{bmatrix} -0.4 & -0.4 \\ 0.6 & -0.4 \\ 0.6 & -0.4 \\ -0.4 & 0.6 \\ -0.4 & 0.6 \end{bmatrix} \begin{bmatrix} -1.8257 \\ -2.7386 \end{bmatrix} = \begin{bmatrix} 1.8257 \\ 0.0000 \\ 0.0000 \\ -0.9129 \\ -0.9129 \end{bmatrix},$$

tale che $\tilde{\mathbf{z}}'_{(2)2} \mathbf{M} \tilde{\mathbf{z}}_{(2)2} = 1$, $\tilde{\mathbf{z}}'_{(2)2} \mathbf{M} \tilde{\mathbf{z}}_{(2)1} = 0$, rappresentano le variabili canoniche corrispondenti al coefficiente di correlazione canonica $\tilde{r}_2 = 0.5$.

Ovviamente, risulta che

$$\tilde{\mathbf{z}}'_{(1)1} \mathbf{M} \tilde{\mathbf{z}}_{(2)2} = \tilde{\mathbf{z}}'_{(2)1} \mathbf{M} \tilde{\mathbf{z}}_{(1)2} = 0$$

e

$$\tilde{\mathbf{z}}'_{(1)2} \mathbf{M} \tilde{\mathbf{z}}_{(2)2} = \tilde{\mathbf{z}}'_{(2)2} \mathbf{M} \tilde{\mathbf{z}}_{(1)2} = \tilde{r}_2 = 0.5.$$

RAPPRESENTAZIONE GRAFICA DELLE VARIABILI E DEGLI INDIVIDUI

RAPPRESENTAZIONE GRAFICA DELLE VARIABILI

Una rappresentazione grafica delle variabili (misurate in termini di scarti standardizzati) si può ottenere considerando le loro proiezioni ortogonali $\hat{\mathbf{y}}_1^*$, $\hat{\mathbf{y}}_2^*$, $\hat{\mathbf{y}}_3^*$, $\hat{\mathbf{y}}_4^*$ nel sottospazio generato dal vettore $\tilde{\mathbf{z}}_{(1)1}$.

Poiché le coordinate, rispetto a $\tilde{\mathbf{z}}_{(1)1}$, di tali proiezioni ortogonali sono date dai coefficienti di correlazione lineare di $\tilde{\mathbf{z}}_{(1)1}$ con le variabili in questione, e questi, come facilmente si verifica, sono

$$r_{11} = 0.4082 \quad , \quad r_{12} = 0.6667 \quad , \quad r_{13} = -1 \quad , \quad r_{14} = 0.6667 \quad ,$$

si ha il grafico di Fig. 1.

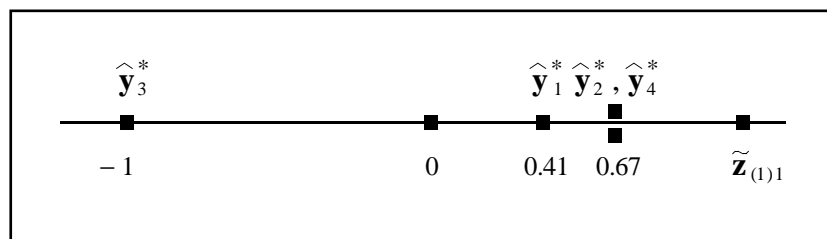


Fig. 1

Nel caso specifico, una misura della qualità della rappresentazione ottenuta è fornita dagli indici

$$\begin{aligned} \text{QR}(1; \tilde{\mathbf{z}}_{(1)1}) = r_{11}^2 = 0.1667 & \quad , \quad \text{QR}(2; \tilde{\mathbf{z}}_{(1)1}) = r_{12}^2 = 0.4444 , \\ \text{QR}(3; \tilde{\mathbf{z}}_{(1)1}) = r_{13}^2 = 1 & \quad , \quad \text{QR}(4; \tilde{\mathbf{z}}_{(1)1}) = r_{14}^2 = 0.4444 . \end{aligned}$$

Pertanto, l'unica variabile ben rappresentata è la terza.

RAPPRESENTAZIONE GRAFICA DEGLI INDIVIDUI

Si considerino anzitutto gli individui (misurati in termini di scarti dalle medie)

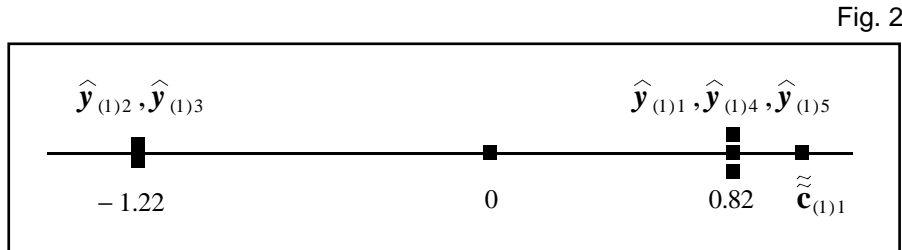
$$\mathbf{y}_{(1)1} = \begin{bmatrix} -0.2 \\ 0.6 \end{bmatrix} , \mathbf{y}_{(1)2} = \begin{bmatrix} -0.2 \\ -0.4 \end{bmatrix} , \mathbf{y}_{(1)3} = \begin{bmatrix} -0.2 \\ -0.4 \end{bmatrix} , \mathbf{y}_{(1)4} = \begin{bmatrix} 0.8 \\ -0.4 \end{bmatrix} , \mathbf{y}_{(1)5} = \begin{bmatrix} -0.2 \\ 0.6 \end{bmatrix} .$$

Una loro rappresentazione grafica si può ottenere considerando le proiezioni ortogonali $\hat{\mathbf{y}}_{(1)1}, \hat{\mathbf{y}}_{(1)2}, \hat{\mathbf{y}}_{(1)3}, \hat{\mathbf{y}}_{(1)4}, \hat{\mathbf{y}}_{(1)5}$ nel sottospazio generato dal vettore $\tilde{\tilde{\mathbf{c}}}_{(1)1}$.

Poiché le coordinate, rispetto a $\tilde{\tilde{\mathbf{c}}}_{(1)1}$, di tali proiezioni ortogonali sono date dagli elementi del vettore canonico

$$\tilde{\mathbf{z}}_{(1)1} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix} ,$$

si ha il grafico della Fig. 2.



Nel caso specifico, una misura della qualità della rappresentazione ottenuta è fornita dagli indici

$$\begin{aligned} \text{QR}(1; \tilde{\mathbf{c}}_{(1)1}) &= \frac{(\mathbf{y}'_{(1)1} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)1})^2}{(\mathbf{y}'_{(1)1} \mathbf{V}_{11}^{-1} \mathbf{y}_{(1)1})(\hat{\mathbf{y}}'_{(1)1} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)1})} = 0.4444, \\ \text{QR}(2; \tilde{\mathbf{c}}_{(1)1}) &= \frac{(\mathbf{y}'_{(1)2} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)2})^2}{(\mathbf{y}'_{(1)2} \mathbf{V}_{11}^{-1} \mathbf{y}_{(1)2})(\hat{\mathbf{y}}'_{(1)2} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)2})} = 1, \\ \text{QR}(3; \tilde{\mathbf{c}}_{(1)1}) &= \frac{(\mathbf{y}'_{(1)3} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)3})^2}{(\mathbf{y}'_{(1)3} \mathbf{V}_{11}^{-1} \mathbf{y}_{(1)3})(\hat{\mathbf{y}}'_{(1)3} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)3})} = 1, \\ \text{QR}(4; \tilde{\mathbf{c}}_{(1)1}) &= \frac{(\mathbf{y}'_{(1)4} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)4})^2}{(\mathbf{y}'_{(1)4} \mathbf{V}_{11}^{-1} \mathbf{y}_{(1)4})(\hat{\mathbf{y}}'_{(1)4} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)4})} = 0.1667, \\ \text{QR}(5; \tilde{\mathbf{c}}_{(1)1}) &= \frac{(\mathbf{y}'_{(1)5} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)5})^2}{(\mathbf{y}'_{(1)5} \mathbf{V}_{11}^{-1} \mathbf{y}_{(1)5})(\hat{\mathbf{y}}'_{(1)5} \mathbf{V}_{11}^{-1} \hat{\mathbf{y}}_{(1)5})} = 0.4444. \end{aligned}$$

Pertanto, gli unici individui ben rappresentati sono il secondo e il terzo.

Successivamente, si considerino gli individui

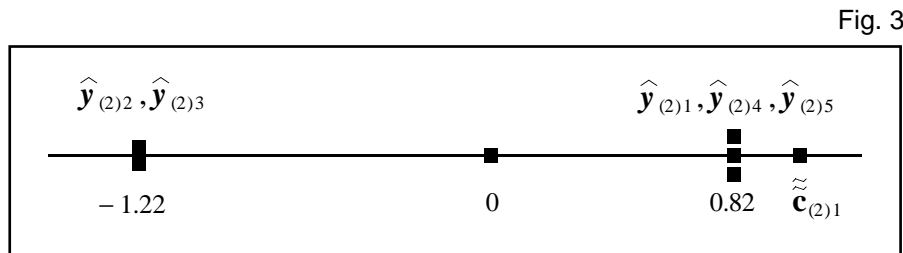
$$\mathbf{y}_{(2)1} = \begin{bmatrix} -0.4 \\ -0.4 \end{bmatrix}, \mathbf{y}_{(2)2} = \begin{bmatrix} 0.6 \\ -0.4 \end{bmatrix}, \mathbf{y}_{(2)3} = \begin{bmatrix} 0.6 \\ -0.4 \end{bmatrix}, \mathbf{y}_{(2)4} = \begin{bmatrix} -0.4 \\ 0.6 \end{bmatrix}, \mathbf{y}_{(2)5} = \begin{bmatrix} -0.4 \\ 0.6 \end{bmatrix}.$$

Una loro rappresentazione grafica si può ottenere considerando le proiezioni ortogonali $\hat{\mathbf{y}}_{(2)1}, \hat{\mathbf{y}}_{(2)2}, \hat{\mathbf{y}}_{(2)3}, \hat{\mathbf{y}}_{(2)4}, \hat{\mathbf{y}}_{(2)5}$ nel sottospazio generato dal vettore $\tilde{\mathbf{c}}_{(2)1}$.

Poiché le coordinate, rispetto a $\tilde{\mathbf{c}}_{(2)1}$, di tali proiezioni ortogonali sono date dagli elementi del vettore canonico

$$\tilde{\mathbf{z}}_{(2)1} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix},$$

si ha il grafico della Fig. 3.



Nel caso specifico, una misura della qualità della rappresentazione ottenuta è fornita dagli indici

$$\begin{aligned} \text{QR}(1; \tilde{\mathbf{c}}_{(2)1}) &= \frac{(\mathbf{y}'_{(2)1} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)1})^2}{(\mathbf{y}'_{(2)1} \mathbf{V}_{22}^{-1} \mathbf{y}_{(2)1})(\hat{\mathbf{y}}'_{(2)1} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)1})} = 0.1667, \\ \text{QR}(2; \tilde{\mathbf{c}}_{(2)1}) &= \frac{(\mathbf{y}'_{(2)2} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)2})^2}{(\mathbf{y}'_{(2)2} \mathbf{V}_{22}^{-1} \mathbf{y}_{(2)2})(\hat{\mathbf{y}}'_{(2)2} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)2})} = 1, \\ \text{QR}(3; \tilde{\mathbf{c}}_{(2)1}) &= \frac{(\mathbf{y}'_{(2)3} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)3})^2}{(\mathbf{y}'_{(2)3} \mathbf{V}_{22}^{-1} \mathbf{y}_{(2)3})(\hat{\mathbf{y}}'_{(2)3} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)3})} = 1, \\ \text{QR}(4; \tilde{\mathbf{c}}_{(2)1}) &= \frac{(\mathbf{y}'_{(2)4} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)4})^2}{(\mathbf{y}'_{(2)4} \mathbf{V}_{22}^{-1} \mathbf{y}_{(2)4})(\hat{\mathbf{y}}'_{(2)4} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)4})} = 0.4444, \\ \text{QR}(5; \tilde{\mathbf{c}}_{(2)1}) &= \frac{(\mathbf{y}'_{(2)5} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)5})^2}{(\mathbf{y}'_{(2)5} \mathbf{V}_{22}^{-1} \mathbf{y}_{(2)5})(\hat{\mathbf{y}}'_{(2)5} \mathbf{V}_{22}^{-1} \hat{\mathbf{y}}_{(2)5})} = 0.4444. \end{aligned}$$

Pertanto, gli unici individui ben rappresentati sono il secondo e il terzo.

ALTRI APPROCCI ALLA ACC

L'APPROCCIO IN TERMINI DELLA ACP

Si consideri l'equazione

$$(\mathbf{V}\mathbf{Q} - \lambda\mathbf{I}_p)\mathbf{c} = \mathbf{0}$$

dove

$$\mathbf{Q} = \text{diag}(\mathbf{V}_{11}^{-1}, \mathbf{V}_{22}^{-1}) = \begin{bmatrix} 7.5 & 2.5 & 0 & 0 \\ 2.5 & 5.0 & 0 & 0 \\ 0 & 0 & 7.5 & 5.0 \\ 0 & 0 & 5.0 & 7.5 \end{bmatrix}.$$

I primi due autovalori di tale equazione sono $\tilde{\lambda}_1 = 2$, $\tilde{\lambda}_2 = 1.5$ e a questi corrispondono i vettori principali

$$\tilde{\mathbf{c}}_1 = \begin{bmatrix} 0.1155 \\ 0.2309 \\ -0.3464 \\ 0.2309 \end{bmatrix}, \quad \tilde{\mathbf{c}}_2 = \begin{bmatrix} -0.2582 \\ 0.2582 \\ 0.0000 \\ -0.2582 \end{bmatrix}$$

e le componenti principali

$$\tilde{\mathbf{y}}_1 = \begin{bmatrix} 1.1547 \\ -1.7321 \\ -1.7321 \\ 1.1547 \\ 1.1547 \end{bmatrix}, \quad \tilde{\mathbf{y}}_2 = \begin{bmatrix} 1.9365 \\ 0.0000 \\ 0.0000 \\ -1.9365 \\ 0.0000 \end{bmatrix}.$$

Come si può facilmente verificare, risulta

$$-\tilde{r}_1 = 1 - \tilde{\lambda}_1 = 1 - 2 = -1, \quad -\tilde{r}_2 = 1 - \tilde{\lambda}_2 = 1 - 1.5 = -0.5$$

e

$$\begin{bmatrix} \tilde{\mathbf{a}}_{(1)1} \\ \tilde{\mathbf{a}}_{(2)1} \end{bmatrix} = \sqrt{2} \mathbf{Q} \tilde{\mathbf{c}}_1 = \sqrt{2} \begin{bmatrix} 7.5 & 2.5 & 0 & 0 \\ 2.5 & 5.0 & 0 & 0 \\ 0 & 0 & 7.5 & 5.0 \\ 0 & 0 & 5.0 & 7.5 \end{bmatrix} \begin{bmatrix} 0.1155 \\ 0.2309 \\ -0.3464 \\ 0.2309 \end{bmatrix} = \begin{bmatrix} 2.0412 \\ 2.0412 \\ -2.0412 \\ 0.0000 \end{bmatrix},$$

$$\begin{bmatrix} \tilde{\mathbf{a}}_{(1)2} \\ \tilde{\mathbf{a}}_{(2)2} \end{bmatrix} = \sqrt{2} \mathbf{Q} \tilde{\mathbf{c}}_2 = \sqrt{2} \begin{bmatrix} 7.5 & 2.5 & 0 & 0 \\ 2.5 & 5.0 & 0 & 0 \\ 0 & 0 & 7.5 & 5.0 \\ 0 & 0 & 5.0 & 7.5 \end{bmatrix} \begin{bmatrix} -0.2582 \\ 0.2582 \\ 0.0000 \\ -0.2582 \end{bmatrix} = \begin{bmatrix} -1.8257 \\ 0.9129 \\ -1.8257 \\ -2.7386 \end{bmatrix}.$$

Inoltre,

$$\tilde{\mathbf{z}}_{(1)1} = \frac{\sqrt{2}}{\tilde{\lambda}_1} \mathbf{P}_1 \tilde{\mathbf{y}}_1 = \frac{\sqrt{2}}{2} \begin{bmatrix} 0.3 & -0.2 & -0.2 & -0.2 & 0.3 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ 0.3 & -0.2 & -0.2 & -0.2 & 0.3 \end{bmatrix} \begin{bmatrix} 1.1547 \\ -1.7321 \\ -1.7321 \\ 1.1547 \\ 1.1547 \end{bmatrix} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix},$$

$$\tilde{\mathbf{z}}_{(1)2} = \frac{\sqrt{2}}{\tilde{\lambda}_2} \mathbf{P}_1 \tilde{\mathbf{y}}_2 = \frac{\sqrt{2}}{1.5} \begin{bmatrix} 0.3 & -0.2 & -0.2 & -0.2 & 0.3 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.8 & -0.2 \\ 0.3 & -0.2 & -0.2 & -0.2 & 0.3 \end{bmatrix} \begin{bmatrix} 1.9365 \\ 0.0000 \\ 0.0000 \\ -1.9365 \\ 0.0000 \end{bmatrix} = \begin{bmatrix} 0.9129 \\ 0.0000 \\ 0.0000 \\ -1.8257 \\ 0.9129 \end{bmatrix},$$

$$\tilde{\mathbf{z}}_{(2)1} = \frac{\sqrt{2}}{\tilde{\lambda}_1} \mathbf{P}_2 \tilde{\mathbf{y}}_1 = \frac{\sqrt{2}}{2} \begin{bmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.3 & 0.3 \\ -0.2 & -0.2 & -0.2 & 0.3 & 0.3 \end{bmatrix} \begin{bmatrix} 1.1547 \\ -1.7321 \\ -1.7321 \\ 1.1547 \\ 1.1547 \end{bmatrix} = \begin{bmatrix} 0.8165 \\ -1.2247 \\ -1.2247 \\ 0.8165 \\ 0.8165 \end{bmatrix},$$

$$\tilde{\mathbf{z}}_{(2)2} = \frac{\sqrt{2}}{\tilde{\lambda}_2} \mathbf{P}_2 \tilde{\mathbf{y}}_2 = \frac{\sqrt{2}}{1.5} \begin{bmatrix} 0.8 & -0.2 & -0.2 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & 0.3 & 0.3 & -0.2 & -0.2 \\ -0.2 & -0.2 & -0.2 & 0.3 & 0.3 \\ -0.2 & -0.2 & -0.2 & 0.3 & 0.3 \end{bmatrix} \begin{bmatrix} 1.9365 \\ 0.0000 \\ 0.0000 \\ -1.9365 \\ 0.0000 \end{bmatrix} = \begin{bmatrix} 1.8257 \\ 0.0000 \\ 0.0000 \\ -0.9129 \\ -0.9129 \end{bmatrix}.$$

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 ACC.

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 of All Variables, Preliminary Statistics of the First Set of Variables, Preliminary Statistics of the Second Set of Variables,*) è dedicata alla esposizione delle principali caratteristiche dei dati a nostra disposizione.

- La seconda parte (*Canonical Correlation Coefficients, Canonical Factors, Canonical Variables*) mostra alcuni risultati che si ottengono dalla ACC.

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

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

La qualità della rappresentazione delle variabili del primo insieme, sia sul cerchio di correlazione I sia sul cerchio di correlazione II, è ottima; lo stesso può dirsi della qualità della rappresentazione delle variabili del secondo insieme, con esclusione della variabile f .

Gli individui 4, 5, 6, 7, 9 del primo insieme risultano ben rappresentati sul piano I, mentre gli individui 1, 2, 3 del secondo insieme risultano ben rappresentati sul piano II.

RISULTATI NUMERICI E GRAFICI*PRELIMINARY STATISTICS OF ALL VARIABLES*

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

rY: 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		

PRELIMINARY STATISTICS OF THE FIRST SET OF VARIABLES

CHOICE OF p1: NUMBER OF VARIABLES OF THE FIRST SET

p1 = 4

X1: RAW DATA MATRIX OF THE FIRST SET OF VARIABLES

	a	b	c	d
1	2.00	1.00	3.00	0.00
2	3.00	1.00	2.00	0.00
3	1.00	2.00	3.00	0.00
4	0.00	0.00	0.00	2.00
5	0.00	0.00	0.00	6.00
6	0.00	0.00	0.00	4.00
7	0.00	0.00	0.00	0.00
8	0.00	0.00	0.00	0.00
9	0.00	0.00	0.00	0.00

(n,p1): ORDER OF X1

n = 9 , p1 = 4

rX1: RANK OF X1

4

g1: BARYCENTRE

a	0.6667
b	0.4444
c	0.8889
d	1.3333

Y1: MEAN CENTRED DATA MATRIX

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

rY1: RANK OF Y1

4

V11: COVARIANCE MATRIX

	a	b	c	d
a	1.1111	0.4815	1.0741	-0.8889
b	0.4815	0.4691	0.8272	-0.5926
c	1.0741	0.8272	1.6543	-1.1852
d	-0.8889	-0.5926	-1.1852	4.4444

Vars1: VARIANCES

a	1.1111
b	0.4691
c	1.6543
d	4.4444

PRELIMINARY STATISTICS OF THE SECOND SET OF VARIABLES

p2 = 5

X2: RAW DATA MATRIX OF THE SECOND SET OF VARIABLES

	e	f	g	h	i
1	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.00	0.00
4	4.00	6.00	0.00	0.00	0.00
5	4.00	2.00	0.00	0.00	0.00
6	6.00	2.00	0.00	0.00	0.00
7	0.00	0.00	6.00	5.00	7.00
8	0.00	0.00	7.00	5.00	6.00
9	0.00	0.00	5.00	6.00	7.00

(n,p2): ORDER OF X2

n = 9, p2 = 5

rX2: RANK OF X2

5

g2: BARYCENTRE

e 1.5556
f 1.1111
g 2.0000
h 1.7778
i 2.2222

Y2: MEAN CENTRED DATA MATRIX

	e	f	g	h	i
1	-1.5556	-1.1111	-2.0000	-1.7778	-2.2222
2	-1.5556	-1.1111	-2.0000	-1.7778	-2.2222
3	-1.5556	-1.1111	-2.0000	-1.7778	-2.2222
4	2.4444	4.8889	-2.0000	-1.7778	-2.2222
5	2.4444	0.8889	-2.0000	-1.7778	-2.2222
6	4.4444	0.8889	-2.0000	-1.7778	-2.2222
7	-1.5556	-1.1111	4.0000	3.2222	4.7778
8	-1.5556	-1.1111	5.0000	3.2222	3.7778
9	-1.5556	-1.1111	3.0000	4.2222	4.7778

rY2: RANK OF Y2

5

V22: COVARIANCE MATRIX

	e	f	g	h	i
e	5.1358	3.1605	-3.1111	-2.7654	-3.4568
f	3.1605	3.6543	-2.2222	-1.9753	-2.4691
g	-3.1111	-2.2222	8.2222	7.0000	8.7778
h	-2.7654	-1.9753	7.0000	6.3951	7.9383
i	-3.4568	-2.4691	8.7778	7.9383	9.9506

Vars2: VARIANCES

e 5.1358
f 3.6543
g 8.2222
h 6.3951
i 9.9506

V12: CROSS COVARIANCE MATRIX

	e	f	g	h	i
a	-1.0370	-0.7407	-1.3333	-1.1852	-1.4815
b	-0.6914	-0.4938	-0.8889	-0.7901	-0.9877
c	-1.3827	-0.9877	-1.7778	-1.5802	-1.9753
d	4.1481	2.0741	-2.6667	-2.3704	-2.9630

k: RANK OF V12

2

CANONICAL CORRELATION COEFFICIENTS, CANONICAL FACTORS, CANONICAL VARIABLES

R: POSITIVE CANONICAL CORRELATION COEFFICIENT MATRIX

	(1)	(2)
(1)	1.0000	0.0000
(2)	0.0000	0.8858

Rv: POSITIVE CANONICAL CORRELATION COEFFICIENT VECTOR

(1)	1.0000
(2)	0.8858

A1: CANONICAL FACTOR MATRIX OF THE FIRST SET

	(1)	(2)
(1)	0.3939	0.0328
(2)	0.3939	0.0328
(3)	0.3939	0.0328
(4)	0.1970	0.4924

CheckA1: Does transpose(A1)*V11*A1 equal Ik?

ans:

	(1)	(2)
(1)	1.0000	-0.0000
(2)	-0.0000	1.0000

Rsquare: SQUARE POSITIVE CANONICAL CORRELATION COEFFICIENT MATRIX

	(1)	(2)
(1)	1.0000	0.0000
(2)	0.0000	0.7847

CheckA1: Does $\text{transpose}(A1)*V12*\text{inv}(V22)*V21*A1$ equal R_{square} ?

ans:

	(1)	(2)
(1)	1.0000	-0.0000
(2)	-0.0000	0.7847

Z1: CANONICAL VARIABLE MATRIX OF THE FIRST SET

	(1)	(2)
1	1.3131	-0.5252
2	1.3131	-0.5252
3	1.3131	-0.5252
4	-0.6565	0.2626
5	0.1313	2.2322
6	-0.2626	1.2474
7	-1.0505	-0.7222
8	-1.0505	-0.7222
9	-1.0505	-0.7222

A2: CANONICAL FACTOR MATRIX OF THE SECOND SET

	(1)	(2)
(5)	-0.1970	0.4975
(6)	-0.1970	-0.1728
(7)	-0.1313	-0.0194
(8)	-0.1313	-0.0194
(9)	-0.1313	-0.0194

CheckA2: Does $\text{transpose}(A2)*V22*A2$ equal I_k ?

ans:

	(1)	(2)
(1)	1.0000	0.0000
(2)	0.0000	1.0000

CheckA2: Does $\text{transpose}(A2)*V21*\text{inv}(V11)*V12*A2$ equal R_{square} ?

ans:

	(1)	(2)
(1)	1.0000	0.0000
(2)	0.0000	0.7847

Z2: CANONICAL VARIABLE MATRIX OF THE SECOND SET

	(1)	(2)
1	1.3131	-0.4652
2	1.3131	-0.4652
3	1.3131	-0.4652
4	-0.6565	0.4880
5	0.1313	1.1793
6	-0.2626	2.1743
7	-1.0505	-0.8153
8	-1.0505	-0.8153
9	-1.0505	-0.8153

DIAGNOSTICS

CORY1Z1: CORRELATION COEFFICIENT MATRIX BETWEEN Y1 AND Z1

	(1)	(2)
a	0.8305	-0.3322
b	0.8520	-0.3408
c	0.9075	-0.3630
d	-0.0830	0.9965

CORY2Z1: CORRELATION COEFFICIENT MATRIX BETWEEN Y2 AND Z1

	(1)	(2)
e	-0.1803	0.8562
f	-0.2442	0.4961
g	-0.7327	-0.5037
h	-0.7385	-0.5077
i	-0.7400	-0.5088

CORY1Z2: CORRELATION COEFFICIENT MATRIX BETWEEN Y1 AND Z2

	(1)	(2)
a	0.8305	-0.2942
b	0.8520	-0.3019
c	0.9075	-0.3215
d	-0.0830	0.8827

CORY2Z2: CORRELATION COEFFICIENT MATRIX BETWEEN Y2 AND Z2

	(1)	(2)
e	-0.1803	0.9666
f	-0.2442	0.5600
g	-0.7327	-0.5687
h	-0.7385	-0.5731
i	-0.7400	-0.5743

QRj(1)I: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE
FIRST SET ON CORRELATION CIRCLE I

a 0.8000
b 0.8421
c 0.9552
d 1.0000

QRj(2)I: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE
SECOND SET ON CORRELATION CIRCLE I

e 0.7656
f 0.3057
g 0.7905
h 0.8031
i 0.8065

QRj(1)II: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE
FIRST SET ON CORRELATION CIRCLE II

a 0.7762
b 0.8171
c 0.9269
d 0.7861

QRj(2)II: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE
SECOND SET ON CORRELATION CIRCLE II

e 0.9668
f 0.3733
g 0.8602
h 0.8738
i 0.8775

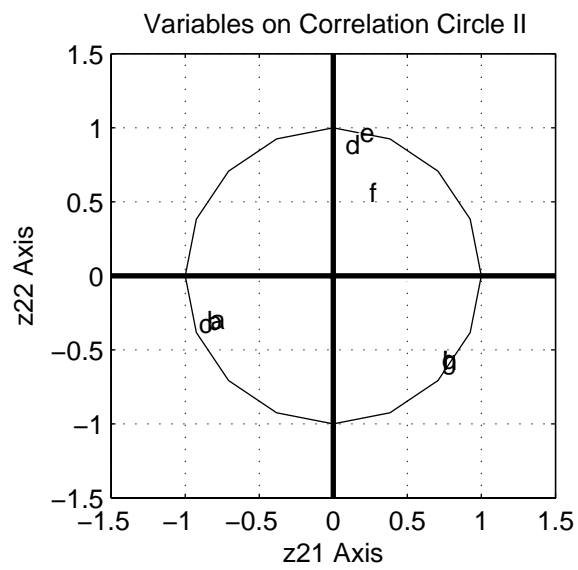
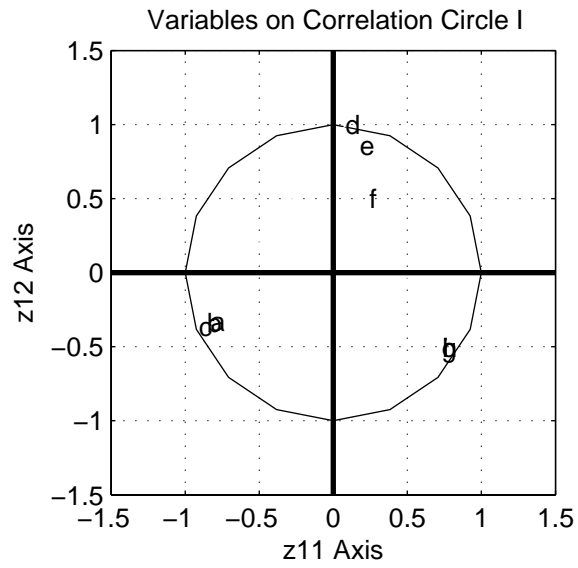
QRi(1)I: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL OF THE
FIRST SET ON PLANE I

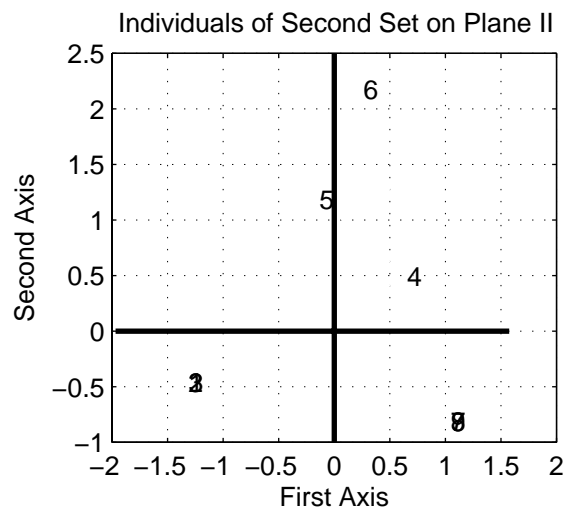
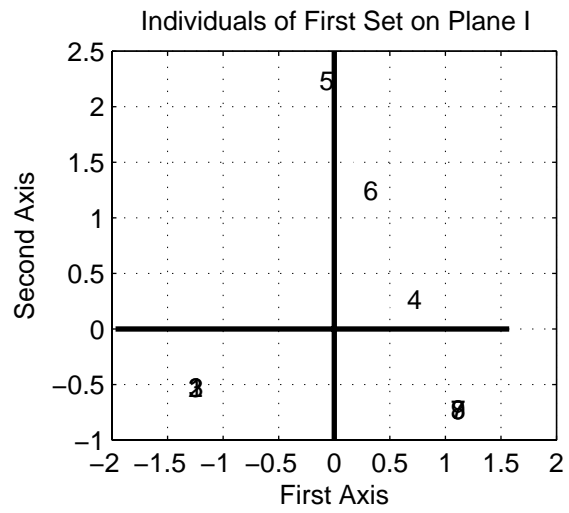
1 0.2500
2 0.2500
3 0.2500
4 1.0000
5 1.0000
6 1.0000
7 1.0000
8 1.0000
9 1.0000

QRi(2)II: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL OF THE
SECOND SET ON PLANE II

1	0.8885
2	0.8885
3	0.8885
4	0.0548
5	0.0119
6	0.0118
7	0.1379
8	0.1379
9	0.1379

GRAPHICS





ESEMPIO 3

Questo esempio analizza i risultati conseguiti, durante i Giochi Olimpici del 1992, dai 28 atleti che hanno partecipato alle seguenti 6 prove di decathlon: 100, 400, 110h, peso, disco, asta.

Le etichette con cui sono identificati gli atleti riflettono anche l'ordinamento basato sul punteggio complessivo da essi ottenuto.

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

COMMENTI

- La prima parte (*Preliminary Statistics of All Variables, Preliminary Statistics of the First Set of Variables, Preliminary Statistics of the Second Set of Variables,*) è dedicata alla esposizione delle principali caratteristiche dei dati a nostra disposizione.
- La seconda parte (*Canonical Correlation Coefficients, Canonical Factors, Canonical Variables*) mostra alcuni risultati che si ottengono dalla ACC.
- La terza parte (*Diagnostics*) è dedicata alla presentazione di alcuni indici che servono a interpretare correttamente i grafici contenuti nella quarta parte (*Graphics*).

La qualità della rappresentazione delle variabili sul cerchio di correlazione I è ottima per le variabili del primo insieme, mediocre per le variabili del secondo insieme; considerazioni di segno opposto si possono fare circa la qualità della rappresentazione delle variabili sul cerchio di correlazione II.

RISULTATI NUMERICI E GRAFICI*PRELIMINARY STATISTICS OF ALL VARIABLES*

X: RAW DATA MATRIX

	100	400	110h	shot	disc	pole
1	10.7800	48.6500	13.9500	14.5300	45.0000	5.1000
2	11.0900	49.6600	14.5800	16.5000	49.6800	4.9000
3	11.1600	49.7600	14.7600	15.2800	49.1200	5.1000
4	11.0900	48.2400	14.8600	13.7300	39.2200	5.3000
5	11.3600	50.0000	14.7500	16.0200	50.7400	4.9000
6	10.7500	48.3300	15.2200	15.3400	42.1400	4.6000
7	11.4200	50.4400	15.0200	15.4400	50.5800	4.7000
8	10.9700	49.3000	14.7800	14.3500	45.0800	4.9000
9	11.1100	48.9100	14.9700	14.3200	45.7600	4.9000
10	11.0900	49.2400	14.8900	15.4800	48.3400	4.8000
11	10.4900	47.9100	14.6400	14.5600	45.0600	4.8000
12	11.3000	50.6200	14.7800	15.2100	44.8400	4.9000
13	10.9500	47.9800	14.5100	13.5500	42.5600	4.6000
14	11.0300	49.0300	14.6000	13.1100	44.5800	4.8000
15	11.3100	49.1900	14.5000	13.5600	37.7800	5.1000
16	10.9500	48.1400	15.1800	14.4100	43.5200	4.5000
17	10.8000	48.3300	14.7600	15.2900	44.7000	4.5000
18	11.1400	47.7300	14.9400	12.5600	37.4200	4.5000
19	10.9700	50.0600	15.2900	13.6700	42.5400	4.5000
20	11.1300	48.6900	14.8100	12.5600	41.2200	4.4000
21	11.5600	50.0800	15.1700	12.9200	43.2600	4.6000
22	11.4100	49.6000	14.8600	12.9900	40.1000	4.5000
23	11.4000	51.5400	15.4200	13.4000	39.8000	4.5000
24	11.4000	50.7600	15.0800	13.2100	39.9000	3.8000
25	11.5500	51.7800	16.6800	13.5700	39.5200	4.2000
26	11.1400	51.4500	15.7000	11.9700	40.6800	4.0000
27	11.6300	51.4000	15.9800	8.6200	24.5800	4.1000
28	11.7500	54.8100	16.2000	9.5400	30.2600	4.0000

(n,p): ORDER OF X

n = 28 , p = 6

rX: RANK OF X

6

CHOICE OF METRIC M

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

c = 1

g: BARYCENTRE

100	11.1689	shot	13.7746
400	49.7011	disc	42.4279
110h	15.0314	pole	4.6250

Y: MEAN CENTRED DATA MATRIX

	100	400	110h	shot	disc	pole
1	-0.3889	-1.0511	-1.0814	0.7554	2.5721	0.4750
2	-0.0789	-0.0411	-0.4514	2.7254	7.2521	0.2750
3	-0.0089	0.0589	-0.2714	1.5054	6.6921	0.4750
4	-0.0789	-1.4611	-0.1714	-0.0446	-3.2079	0.6750
5	0.1911	0.2989	-0.2814	2.2454	8.3121	0.2750
6	-0.4189	-1.3711	0.1886	1.5654	-0.2879	-0.0250
7	0.2511	0.7389	-0.0114	1.6654	8.1521	0.0750
8	-0.1989	-0.4011	-0.2514	0.5754	2.6521	0.2750
9	-0.0589	-0.7911	-0.0614	0.5454	3.3321	0.2750
10	-0.0789	-0.4611	-0.1414	1.7054	5.9121	0.1750
11	-0.6789	-1.7911	-0.3914	0.7854	2.6321	0.1750
12	0.1311	0.9189	-0.2514	1.4354	2.4121	0.2750
13	-0.2189	-1.7211	-0.5214	-0.2246	0.1321	-0.0250
14	-0.1389	-0.6711	-0.4314	-0.6646	2.1521	0.1750
15	0.1411	-0.5111	-0.5314	-0.2146	-4.6479	0.4750
16	-0.2189	-1.5611	0.1486	0.6354	1.0921	-0.1250
17	-0.3689	-1.3711	-0.2714	1.5154	2.2721	-0.1250
18	-0.0289	-1.9711	-0.0914	-1.2146	-5.0079	-0.1250
19	-0.1989	0.3589	0.2586	-0.1046	0.1121	-0.1250
20	-0.0389	-1.0111	-0.2214	-1.2146	-1.2079	-0.2250
21	0.3911	0.3789	0.1386	-0.8546	0.8321	-0.0250
22	0.2411	-0.1011	-0.1714	-0.7846	-2.3279	-0.1250
23	0.2311	1.8389	0.3886	-0.3746	-2.6279	-0.1250
24	0.2311	1.0589	0.0486	-0.5646	-2.5279	-0.8250
25	0.3811	2.0789	1.6486	-0.2046	-2.9079	-0.4250
26	-0.0289	1.7489	0.6686	-1.8046	-1.7479	-0.6250
27	0.4611	1.6989	0.9486	-5.1546	-17.8479	-0.5250
28	0.5811	5.1089	1.1686	-4.2346	-12.1679	-0.6250

rY: RANK OF Y

6

V: COVARIANCE MATRIX

	100	400	110h	shot	disc	pole
100	0.0819	0.3357	0.0929	-0.2515	-0.7129	-0.0428
400	0.3357	2.3165	0.6012	-1.3064	-3.5526	-0.3018
110h	0.0929	0.6012	0.3031	-0.5313	-1.7543	-0.1394
shot	-0.2515	-1.3064	-0.5313	2.9483	8.7387	0.3854
disc	-0.7129	-3.5526	-1.7543	8.7387	31.1583	1.1269
pole	-0.0428	-0.3018	-0.1394	0.3854	1.1269	0.1312

Vars: VARIANCES

100	0.0819	shot	2.9483
400	2.3165	disc	31.1583
110h	0.3031	pole	0.1312

PRELIMINARY STATISTICS OF THE FIRST SET OF VARIABLES

CHOICE OF p1: NUMBER OF VARIABLES OF THE FIRST SET

p1 = 3

X1: RAW DATA MATRIX OF THE FIRST SET OF VARIABLES

	100	400	110h
1	10.78	48.65	13.95
2	11.09	49.66	14.58
3	11.16	49.76	14.76
4	11.09	48.24	14.86
5	11.36	50.00	14.75
6	10.75	48.33	15.22
7	11.42	50.44	15.02
8	10.97	49.30	14.78
9	11.11	48.91	14.97
10	11.09	49.24	14.89
11	10.49	47.91	14.64
12	11.30	50.62	14.78
13	10.95	47.98	14.51
14	11.03	49.03	14.60
15	11.31	49.19	14.50
16	10.95	48.14	15.18
17	10.80	48.33	14.76
18	11.14	47.73	14.94
19	10.97	50.06	15.29
20	11.13	48.69	14.81
21	11.56	50.08	15.17
22	11.41	49.60	14.86
23	11.40	51.54	15.42
24	11.40	50.76	15.08
25	11.55	51.78	16.68
26	11.14	51.45	15.70
27	11.63	51.40	15.98
28	11.75	54.81	16.20

(n,p1): ORDER OF X1

n = 28 , p1 = 3

rX1: RANK OF X1

3

g1: BARYCENTRE

```
100  11.1689
400  49.7011
110h 15.0314
```

Y1: MEAN CENTRED DATA MATRIX

```
      100      400      110h
1  -0.3889 -1.0511 -1.0814
2  -0.0789 -0.0411 -0.4514
3  -0.0089  0.0589 -0.2714
4  -0.0789 -1.4611 -0.1714
5   0.1911  0.2989 -0.2814
6  -0.4189 -1.3711  0.1886
7   0.2511  0.7389 -0.0114
8  -0.1989 -0.4011 -0.2514
9  -0.0589 -0.7911 -0.0614
10 -0.0789 -0.4611 -0.1414
11 -0.6789 -1.7911 -0.3914
12  0.1311  0.9189 -0.2514
13 -0.2189 -1.7211 -0.5214
14 -0.1389 -0.6711 -0.4314
15  0.1411 -0.5111 -0.5314
16 -0.2189 -1.5611  0.1486
17 -0.3689 -1.3711 -0.2714
18 -0.0289 -1.9711 -0.0914
19 -0.1989  0.3589  0.2586
20 -0.0389 -1.0111 -0.2214
21  0.3911  0.3789  0.1386
22  0.2411 -0.1011 -0.1714
23  0.2311  1.8389  0.3886
24  0.2311  1.0589  0.0486
25  0.3811  2.0789  1.6486
26 -0.0289  1.7489  0.6686
27  0.4611  1.6989  0.9486
28  0.5811  5.1089  1.1686
```

rY1: RANK OF Y1

3

V11: COVARIANCE MATRIX

```
      100      400      110h
100  0.0819  0.3357  0.0929
400  0.3357  2.3165  0.6012
110h 0.0929  0.6012  0.3031
```

Vars1: VARIANCES

100	0.0819
400	2.3165
110h	0.3031

PRELIMINARY STATISTICS OF THE SECOND SET OF VARIABLES

p2 = 3

X2: RAW DATA MATRIX OF THE SECOND SET OF VARIABLES

	shot	disc	pole
1	14.53	45.00	5.10
2	16.50	49.68	4.90
3	15.28	49.12	5.10
4	13.73	39.22	5.30
5	16.02	50.74	4.90
6	15.34	42.14	4.60
7	15.44	50.58	4.70
8	14.35	45.08	4.90
9	14.32	45.76	4.90
10	15.48	48.34	4.80
11	14.56	45.06	4.80
12	15.21	44.84	4.90
13	13.55	42.56	4.60
14	13.11	44.58	4.80
15	13.56	37.78	5.10
16	14.41	43.52	4.50
17	15.29	44.70	4.50
18	12.56	37.42	4.50
19	13.67	42.54	4.50
20	12.56	41.22	4.40
21	12.92	43.26	4.60
22	12.99	40.10	4.50
23	13.40	39.80	4.50
24	13.21	39.90	3.80
25	13.57	39.52	4.20
26	11.97	40.68	4.00
27	8.62	24.58	4.10
28	9.54	30.26	4.00

(n,p2): ORDER OF X2

n = 28 , p2 = 3

rX2: RANK OF X2

3

g2: BARYCENTRE

shot 13.7746
disc 42.4279
pole 4.6250

Y2: MEAN CENTRED DATA MATRIX

	shot	disc	pole
1	0.7554	2.5721	0.4750
2	2.7254	7.2521	0.2750
3	1.5054	6.6921	0.4750
4	-0.0446	-3.2079	0.6750
5	2.2454	8.3121	0.2750
6	1.5654	-0.2879	-0.0250
7	1.6654	8.1521	0.0750
8	0.5754	2.6521	0.2750
9	0.5454	3.3321	0.2750
10	1.7054	5.9121	0.1750
11	0.7854	2.6321	0.1750
12	1.4354	2.4121	0.2750
13	-0.2246	0.1321	-0.0250
14	-0.6646	2.1521	0.1750
15	-0.2146	-4.6479	0.4750
16	0.6354	1.0921	-0.1250
17	1.5154	2.2721	-0.1250
18	-1.2146	-5.0079	-0.1250
19	-0.1046	0.1121	-0.1250
20	-1.2146	-1.2079	-0.2250
21	-0.8546	0.8321	-0.0250
22	-0.7846	-2.3279	-0.1250
23	-0.3746	-2.6279	-0.1250
24	-0.5646	-2.5279	-0.8250
25	-0.2046	-2.9079	-0.4250
26	-1.8046	-1.7479	-0.6250
27	-5.1546	-17.8479	-0.5250
28	-4.2346	-12.1679	-0.6250

rY2: RANK OF Y2

3

V22: COVARIANCE MATRIX

	shot	disc	pole
shot	2.9483	8.7387	0.3854
disc	8.7387	31.1583	1.1269
pole	0.3854	1.1269	0.1312

Vars2: VARIANCES

shot 2.9483
disc 31.1583
pole 0.1312

V12: CROSS COVARIANCE MATRIX

	shot	disc	pole
100	-0.2515	-0.7129	-0.0428
400	-1.3064	-3.5526	-0.3018
110h	-0.5313	-1.7543	-0.1394

k: RANK OF V12

3

CANONICAL CORRELATION COEFFICIENTS, CANONICAL FACTORS, CANONICAL VARIABLES

R: POSITIVE CANONICAL CORRELATION COEFFICIENT MATRIX

	(1)	(2)	(3)
(1)	0.7351	0.0000	0.0000
(2)	0.0000	0.3379	0.0000
(3)	0.0000	0.0000	0.2308

Rv: POSITIVE CANONICAL CORRELATION COEFFICIENT VECTOR

(1)	0.7351
(2)	0.3379
(3)	0.2308

A1: CANONICAL FACTOR MATRIX OF THE FIRST SET

	(1)	(2)	(3)
(1)	0.1224	-2.9833	4.6219
(2)	0.0360	-0.3883	-1.1349
(3)	1.7042	1.7908	0.8560

Rsquare: SQUARE POSITIVE CANONICAL CORRELATION COEFFICIENT MATRIX

	(1)	(2)	(3)
(1)	0.5403	0.0000	0.0000
(2)	0.0000	0.1142	0.0000
(3)	0.0000	0.0000	0.0532

Z1: CANONICAL VARIABLE MATRIX OF THE FIRST SET

	(1)	(2)	(3)
1	-1.9284	-0.3682	-1.5304
2	-0.7805	-0.5570	-0.7046
3	-0.4615	-0.4823	-0.3405
4	-0.3544	0.4958	1.1466
5	-0.4455	-1.1901	0.3030
6	0.2208	2.1199	-0.2188
7	0.0378	-1.0564	0.3121
8	-0.4673	0.2989	-0.6795
9	-0.1404	0.3730	0.5728
10	-0.2673	0.1612	0.0374
11	-0.8146	2.0199	-1.4404
12	-0.3794	-1.1981	-0.6523
13	-0.9774	0.3876	0.4950
14	-0.7764	-0.0976	-0.2498
15	-0.9068	-1.1741	0.7771
16	0.1702	1.5253	0.8869
17	-0.5571	1.1469	-0.3815
18	-0.2303	0.6879	2.0249
19	0.4292	0.9172	-1.1054
20	-0.4185	0.1122	0.7780
21	0.2976	-1.0657	1.4961
22	-0.2663	-0.9869	1.0822
23	0.7567	-0.7075	-0.6863
24	0.1492	-1.0136	-0.0922
25	2.9310	1.0082	0.8131
26	1.1988	0.6045	-1.5462
27	1.7341	-0.3365	1.0149
28	2.2465	-1.6246	-2.1120

A2: CANONICAL FACTOR MATRIX OF THE SECOND SET

	(1)	(2)	(3)
(4)	0.0650	1.4485	0.3859
(5)	-0.0814	-0.3104	-0.2956
(6)	-2.1218	-1.6937	2.2397

Z2: CANONICAL VARIABLE MATRIX OF THE SECOND SET

	(1)	(2)	(3)
1	-1.1682	-0.5087	0.5950
2	-0.9968	1.2311	-0.4762
3	-1.4548	-0.7011	-0.3335
4	-1.1739	-0.2123	2.4428
5	-1.1143	0.2068	-0.9748
6	0.1783	2.3991	0.6331
7	-0.7146	-0.2450	-1.5992
8	-0.7620	-0.4555	0.0539
9	-0.8193	-0.7100	-0.1586
10	-0.7418	0.3388	-0.6977
11	-0.5346	0.0242	-0.0831
12	-0.6866	0.8647	0.4567
13	0.0277	-0.3241	-0.1817
14	-0.5898	-1.9271	-0.5007
15	-0.6434	0.3272	2.3550
16	0.2176	0.7931	-0.3576
17	0.1788	1.7015	-0.3669
18	0.5940	0.0066	0.7317
19	0.2493	0.0253	-0.3535
20	0.4968	-1.0035	-0.6156
21	-0.0703	-1.4539	-0.6317
22	0.4037	-0.2023	0.1054
23	0.4548	0.4847	0.3523
24	1.9196	1.3640	-1.3184
25	1.1252	1.3259	-0.1713
26	1.3511	-1.0130	-1.5795
27	2.2320	-1.0378	2.1111
28	2.0415	-1.2988	0.5631

DIAGNOSTICS

CORY1Z1: CORRELATION COEFFICIENT MATRIX BETWEEN Y1 AND Z1

	(1)	(2)	(3)
100	0.6301	-0.7283	0.2692
400	0.7549	-0.5417	-0.3697
110h	0.9982	0.0588	0.0115

CORY2Z1: CORRELATION COEFFICIENT MATRIX BETWEEN Y2 AND Z1

	(1)	(2)	(3)
shot	-0.5726	0.1782	-0.0783
disc	-0.5741	0.0653	-0.1370
pole	-0.7002	-0.0127	0.0697

CORY1Z2: CORRELATION COEFFICIENT MATRIX BETWEEN Y1 AND Z2

	(1)	(2)	(3)
100	0.4632	-0.2461	0.0621
400	0.5549	-0.1831	-0.0853
110h	0.7338	0.0199	0.0027

CORY2Z2: CORRELATION COEFFICIENT MATRIX BETWEEN Y2 AND Z2

	(1)	(2)	(3)
shot	-0.7790	0.5274	-0.3392
disc	-0.7811	0.1932	-0.5938
pole	-0.9526	-0.0377	0.3019

QRj(1)I: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE FIRST SET ON CORRELATION CIRCLE I

100	0.9275
400	0.8634
110h	0.9999

QRj(2)I: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE SECOND SET ON CORRELATION CIRCLE I

shot	0.3597
disc	0.3339
pole	0.4905

QRj(1)II: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE FIRST SET ON CORRELATION CIRCLE II

100	0.2751
400	0.3415
110h	0.5388

QRj(2)II: QUALITY OF REPRESENTATION OF EACH VARIABLE OF THE SECOND SET ON CORRELATION CIRCLE II

shot	0.8850
disc	0.6474
pole	0.9088

QRi(1)I: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL OF THE
FIRST SET ON PLANE I

1	0.6220	15	0.7847
2	0.6494	16	0.7497
3	0.7936	17	0.9178
4	0.2203	18	0.1137
5	0.9462	19	0.4563
6	0.9896	20	0.2368
7	0.9198	21	0.3536
8	0.3999	22	0.4715
9	0.3261	23	0.6950
10	0.9859	24	0.9920
11	0.6957	25	0.9356
12	0.7878	26	0.4299
13	0.8186	27	0.7518
14	0.9075	28	0.6328

QRi(2)II: QUALITY OF REPRESENTATION OF EACH INDIVIDUAL OF THE
SECOND SET ON PLANE II

1	0.6901	15	0.0682
2	0.3632	16	0.0589
3	0.7784	17	0.0104
4	0.1865	18	0.3972
5	0.5556	19	0.3310
6	0.0051	20	0.1512
7	0.1632	21	0.0020
8	0.7340	22	0.7580
9	0.5591	23	0.3656
10	0.4777	24	0.5059
11	0.9745	25	0.4146
12	0.3302	26	0.3414
13	0.0055	27	0.4737
14	0.0807	28	0.6753

GRAPHICS

