

RENATO LEONI

**Esempi numerici riguardanti l'analisi
semplice delle corrispondenze**
(Parte prima)

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

In questo lavoro sono presentati alcuni esempi numerici riguardanti l'analisi semplice delle corrispondenze (ASC), vista nell'ottica dell'analisi delle componenti principali (ACP) applicata alle matrici dei profili riga e colonna.

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

Leoni, R., *Analisi semplice delle corrispondenze (Parte prima)*, Dipartimento di Statistica "G. Parenti", Firenze, 2007.

ESEMPIO 1

Questo esempio, basato su dati fittizi, ha lo scopo di illustrare i principali concetti teorici della ASC esposti nel lavoro citato nella Avvertenza.

I DATI DI BASE E LA LORO STRUTTURA ALGEBRICA

Si consideri la seguente tavola di contingenza

Modalità di A	Modalità di B		Totale
	b ₁	b ₂	
a ₁	45	15	60
a ₂	5	15	20
Totale	50	30	80

riferita a due caratteri A e B, ciascuno presente con due modalità.

Da questa si deduce che

$$\mathbf{X}_{(1)} = \begin{bmatrix} 0.75 & 0.25 \\ 0.25 & 0.75 \end{bmatrix}, \quad \mathbf{x}_{(1)1} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}, \quad \mathbf{x}_{(1)2} = \begin{bmatrix} 0.25 \\ 0.75 \end{bmatrix}, \quad \mathbf{g}_{(1)} = \begin{bmatrix} 0.625 \\ 0.375 \end{bmatrix}$$

$$\mathbf{M}_{(1)} = \text{diag}(0.75, 0.25), \quad \mathbf{Q}_{(1)} = \text{diag}(1.600, 2.666)$$

e

$$\mathbf{X}_{(2)} = \begin{bmatrix} 0.9 & 0.1 \\ 0.5 & 0.5 \end{bmatrix}, \quad \mathbf{x}_{(2)1} = \begin{bmatrix} 0.9 \\ 0.1 \end{bmatrix}, \quad \mathbf{x}_{(2)2} = \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix}, \quad \mathbf{g}_{(2)} = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$$

$$\mathbf{M}_{(2)} = \text{diag}(0.625, 0.375), \quad \mathbf{Q}_{(2)} = \text{diag}(1.333, 4.000).$$

LA ACP DELLA MATRICE DEI PROFILI RIGA

I PRINCIPALI RISULTATI ALGEBRICI

Considerato che ⁽¹⁾

$$\mathbf{Y}_{(1)} = \begin{bmatrix} 0.125 & -0.125 \\ -0.375 & 0.375 \end{bmatrix}, \quad \mathbf{y}_{(1)1} = \begin{bmatrix} 0.125 \\ -0.125 \end{bmatrix}, \quad \mathbf{y}_{(1)2} = \begin{bmatrix} -0.375 \\ 0.375 \end{bmatrix}$$

(1) Lasciamo al lettore la verifica dei risultati che si ottengono a partire dalle matrici non centrate dei profili riga e colonna.

$$\mathbf{V}_{(1)} = \begin{bmatrix} 0.0469 & -0.0469 \\ -0.0469 & 0.0469 \end{bmatrix}$$

l'equazione

$$\mathbf{V}_{(1)} \mathbf{Q}_{(1)} \mathbf{c}_{(1)} = \begin{bmatrix} 0.075 & -0.125 \\ -0.075 & 0.125 \end{bmatrix} \mathbf{c}_{(1)} = \lambda_{(1)} \mathbf{c}_{(1)}$$

ammette una matrice degli autovalori

$$\tilde{\mathbf{D}}_{(1)} = \text{diag}(\tilde{\lambda}_{(1)1}, \tilde{\lambda}_{(1)2}) = \text{diag}(0.2, 0.0),$$

una matrice dei vettori principali

$$\tilde{\mathbf{C}}_{(1)} = [\tilde{\mathbf{c}}_{(1)1} \quad \tilde{\mathbf{g}}_{(1)}] = \begin{bmatrix} 0.4841 & 0.6250 \\ -0.4841 & 0.3750 \end{bmatrix},$$

una matrice delle componenti principali

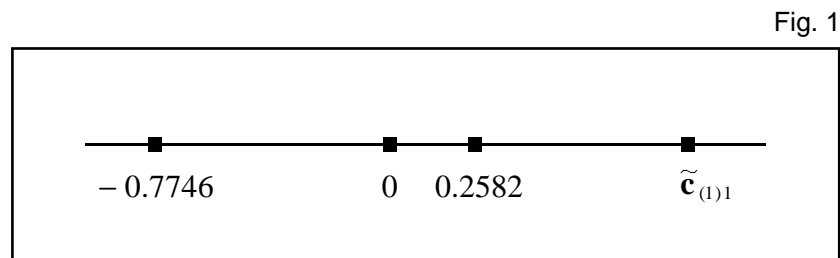
$$\tilde{\mathbf{Y}}_{(1)} = [\tilde{\mathbf{y}}_{(1)1} \quad \tilde{\mathbf{y}}_{(1)2}] = \begin{bmatrix} 0.2582 & 0.0000 \\ -0.7746 & 0.0000 \end{bmatrix}.$$

LA RAPPRESENTAZIONE GRAFICA DEI PROFILI RIGA

Poichè le coordinate, rispetto a $\tilde{\mathbf{c}}_{(1)1}$, delle proiezioni ortogonali di $\mathbf{y}_{(1)1}, \mathbf{y}_{(1)2}$ nel sottospazio generato da $\tilde{\mathbf{c}}_{(1)1}$ sono, rispettivamente,

$$\tilde{y}_{(1)1,1} = 0.2582 \quad , \quad \tilde{y}_{(1)1,2} = -0.7746$$

dove $\tilde{y}_{(1)1,i}$ denota l'elemento i -esimo della componente principale $\tilde{\mathbf{y}}_{(1)1}$, si ottiene il grafico della Fig. 1.



La qualità globale della rappresentazione di $\mathbf{y}_{(1)1}, \mathbf{y}_{(1)2}$ si può misurare

mediante l'indice GQRI1 che assume il valore

$$\text{GQRI1} = \frac{\tilde{\lambda}_{(1)1}}{\tilde{\lambda}_{(1)1} + \tilde{\lambda}_{(1)2}} = \frac{0.2}{0.2} = 1.$$

A sua volta, la qualità della rappresentazione di ciascun $\mathbf{y}_{(1)1}, \mathbf{y}_{(1)2}$ si può misurare attraverso l'indice $\text{QR}(i; \tilde{\mathbf{c}}_{(1)1})$ che assume i valori

$$\begin{aligned} \text{QR}(1; \tilde{\mathbf{c}}_{(1)1}) &= \frac{\tilde{y}_{(1)1,1}^2}{\tilde{y}_{(1)1,1}^2 + \tilde{y}_{(1)2,1}^2} = \frac{(0.2582)^2}{(0.2582)^2 + (0.0000)^2} = 1 \\ \text{QR}(2; \tilde{\mathbf{c}}_{(1)1}) &= \frac{\tilde{y}_{(1)1,2}^2}{\tilde{y}_{(1)1,2}^2 + \tilde{y}_{(1)2,2}^2} = \frac{(-0.7746)^2}{(-0.7746)^2 + (0.0000)^2} = 1. \end{aligned}$$

Come si può vedere, $\mathbf{y}_{(1)1}, \mathbf{y}_{(1)2}$ sono ben rappresentati sia globalmente sia singolarmente ⁽²⁾.

Infine, nell'interpretare i risultati dell'analisi è importante esaminare il contributo di ciascun $\mathbf{y}_{(1)1}, \mathbf{y}_{(1)2}$ all'inerzia spiegata da $\tilde{\mathbf{c}}_{(1)1}$, misurato dall'indice $\text{C}(i; \tilde{\mathbf{c}}_{(1)1})$ il quale assume i valori

$$\begin{aligned} \text{C}(1; \tilde{\mathbf{c}}_{(1)1}) &= \frac{\frac{\mathbf{n}_{1+}}{\mathbf{n}} \tilde{y}_{(1)1,1}^2}{\tilde{\lambda}_{(1)1}} = \frac{0.75 \times (0.2582)^2}{0.2} = 0.25 \\ \text{C}(2; \tilde{\mathbf{c}}_{(1)1}) &= \frac{\frac{\mathbf{n}_{2+}}{\mathbf{n}} \tilde{y}_{(1)1,2}^2}{\tilde{\lambda}_{(1)1}} = \frac{0.25 \times (-0.7746)^2}{0.2} = 0.75. \end{aligned}$$

Il contributo maggiore è fornito dal secondo profilo riga.

LA ACP DELLA MATRICE DEI PROFILI COLONNA

I PRINCIPALI RISULTATI ALGEBRICI

Considerato che

$$\mathbf{Y}_{(2)} = \begin{bmatrix} 0.15 & -0.15 \\ -0.25 & 0.25 \end{bmatrix}, \quad \mathbf{y}_{(2)1} = \begin{bmatrix} 0.15 \\ -0.15 \end{bmatrix}, \quad \mathbf{y}_{(2)2} = \begin{bmatrix} -0.25 \\ 0.25 \end{bmatrix}$$

(2) Data la struttura della tavola di contingenza esaminata, tale risultato è del tutto scontato.

$$\mathbf{V}_{(2)} = \begin{bmatrix} 0.0375 & -0.0375 \\ -0.0375 & 0.0375 \end{bmatrix}$$

l'equazione

$$\mathbf{V}_{(2)} \mathbf{Q}_{(2)} \mathbf{c}_{(2)} = \begin{bmatrix} 0.05 & -0.15 \\ -0.05 & 0.15 \end{bmatrix} \mathbf{c}_{(2)} = \lambda_{(2)} \mathbf{c}_{(2)}$$

ammette una matrice degli autovalori

$$\tilde{\mathbf{D}}_{(2)} = \text{diag}(\tilde{\lambda}_{(2)1}, \tilde{\lambda}_{(2)2}) = \text{diag}(0.2, 0.0),$$

una matrice dei vettori principali

$$\tilde{\mathbf{C}}_{(2)} = [\tilde{\mathbf{c}}_{(2)1} \quad \tilde{\mathbf{g}}_{(2)}] = \begin{bmatrix} 0.433 & 0.750 \\ -0.433 & 0.250 \end{bmatrix},$$

una matrice delle componenti principali

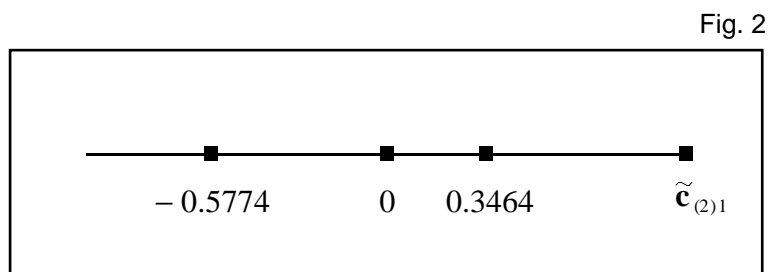
$$\tilde{\mathbf{Y}}_{(2)} = [\tilde{\mathbf{y}}_{(2)1} \quad \tilde{\mathbf{y}}_{(2)2}] = \begin{bmatrix} 0.3464 & 0.0000 \\ -0.5774 & 0.0000 \end{bmatrix}.$$

LA RAPPRESENTAZIONE GRAFICA DEI PROFILI COLONNA

Poichè le coordinate, rispetto a $\tilde{\mathbf{c}}_{(2)1}$, delle proiezioni ortogonali di $\mathbf{y}_{(2)1}, \mathbf{y}_{(2)2}$ nel sottospazio generato da $\tilde{\mathbf{c}}_{(2)1}$ sono, rispettivamente,

$$\tilde{y}_{(2)1,1} = 0.3464 \quad , \quad \tilde{y}_{(2)1,2} = -0.5774$$

dove $\tilde{y}_{(2)1,j}$ denota l'elemento j -esimo della componente principale $\tilde{\mathbf{y}}_{(2)1}$, si ottiene il grafico della Fig. 2.



La qualità globale della rappresentazione di $\mathbf{y}_{(2)1}, \mathbf{y}_{(2)2}$ si può misurare

mediante l'indice GQRI2 che assume il valore

$$\text{GQRI2} = \frac{\tilde{\lambda}_{(2)1}}{\tilde{\lambda}_{(2)1} + \tilde{\lambda}_{(2)2}} = \frac{0.2}{0.2} = 1 .$$

A sua volta, la qualità della rappresentazione di ciascun $\mathbf{y}_{(2)1}, \mathbf{y}_{(2)2}$ si può misurare attraverso l'indice $\text{QR}(j; \tilde{\mathbf{c}}_{(2)1})$ che assume i valori

$$\begin{aligned} \text{QR}(1; \tilde{\mathbf{c}}_{(2)1}) &= \frac{\tilde{y}_{(2)1,1}^2}{\tilde{y}_{(2)1,1}^2 + \tilde{y}_{(2)2,2}^2} = \frac{(0.3464)^2}{(0.3464)^2 + (0.0000)^2} = 1 \\ \text{QR}(2; \tilde{\mathbf{c}}_{(2)1}) &= \frac{\tilde{y}_{(2)1,2}^2}{\tilde{y}_{(2)1,2}^2 + \tilde{y}_{(2)2,2}^2} = \frac{(-0.5774)^2}{(-0.5774)^2 + (0.0000)^2} = 1 . \end{aligned}$$

Come si può vedere, $\mathbf{y}_{(2)1}, \mathbf{y}_{(2)2}$ sono ben rappresentati sia globalmente sia singolarmente ⁽³⁾.

Infine, nell'interpretare i risultati dell'analisi è importante esaminare il contributo di ciascun $\mathbf{y}_{(2)1}, \mathbf{y}_{(2)2}$ all'inerzia spiegata da $\tilde{\mathbf{c}}_{(2)1}$, misurato dall'indice $\text{C}(j; \tilde{\mathbf{c}}_{(2)1})$ il quale assume i valori

$$\begin{aligned} \text{C}(1; \tilde{\mathbf{c}}_{(2)1}) &= \frac{\frac{n_{+1}}{n} \tilde{y}_{(2)1,1}^2}{\tilde{\lambda}_{(2)1}} = \frac{0.625 \times (0.3464)^2}{0.2} = 0.375 \\ \text{C}(2; \tilde{\mathbf{c}}_{(2)1}) &= \frac{\frac{n_{+2}}{n} \tilde{y}_{(2)1,2}^2}{\tilde{\lambda}_{(2)1}} = \frac{0.375 \times (-0.5774)^2}{0.2} = 0.625 . \end{aligned}$$

Il contributo maggiore è fornito dal secondo profilo colonna.

OSSERVAZIONE Come è noto, tra i risultati della ACP dei profili riga e della ACP dei profili colonna intercorrono numerose relazioni e, in particolare, le cosiddette relazioni quasi baricentriche.

Per esempio, si verifica subito che

$$\tilde{y}_{(2)1,1} = \frac{1}{\sqrt{\tilde{\lambda}_1}} \sum_i \tilde{y}_{(1)1,i} \frac{n_{i1}}{n_{+1}} = \frac{1}{\sqrt{0.2}} \left(0.2582 \frac{45}{50} - 0.7746 \frac{5}{50} \right) = 0.3464 .$$

(3) Data la struttura della tavola di contingenza esaminata, tale risultato è del tutto scontato.

ESEMPIO 2

Questo esempio, basato su dati fittizi, 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 ASC vista nell'ottica della ACP applicata alle matrici dei profili riga e colonna.

Le modalità del primo carattere sono identificate con le etichette a, b, c, d; quelle del secondo carattere attraverso le etichette x, y, z.

COMMENTI

- La prima parte (*Preliminary Statistics*) è dedicata alla esposizione di alcune caratteristiche dei dati a nostra disposizione. In particolare, sono indicate sia la matrice dei profili riga (*Row Profile Matrix*), sia la matrice dei profili colonna (*Column Profile Matrix*).
- La seconda parte (*Eigenvalues, Principal Vectors, Principal Components*) contiene i principali risultati che si ottengono dalla ACP delle matrici dei profili riga e colonna.
- La terza parte (*Diagnostics*) è dedicata alla presentazione di alcuni indici che servono a interpretare correttamente i grafici contenuti nella quarta parte (*Graphics*).

Ovviamente, avendo scelto di rappresentare i profili riga e colonna mediante le loro proiezioni ortogonali nei rispettivi piani principali, sia la qualità della rappresentazione globale sia la qualità della rappresentazione dei singoli profili risulta perfetta.

Si noti, inoltre, che:

- I profili riga a, d contribuiscono in misura comparativamente maggiore di b, c all'inerzia spiegata da ciascun asse.
- I profili colonna x, z contribuiscono in egual misura all'inerzia spiegata dal primo asse, mentre il profilo colonna y fornisce un contributo nullo all'inerzia spiegata dal primo asse; i profili colonna x, y, z contribuiscono in egual misura all'inerzia spiegata dal secondo asse.

ANALISI BASATA SULLE MATRICI CENTRATE DEI PROFILI RIGA
E COLONNA

PRELIMINARY STATISTICS

X: FREQUENCY DATA MATRIX

	x	y	z
a	1	0	0
b	1	1	0
c	0	1	1
d	0	0	1

(p,q): ORDER OF X

p =
4

q =
3

n: NUMBER OF INDIVIDUALS

n =
6

X1: ROW PROFILE MATRIX

	x	y	z
a	1.0000	0.0000	0.0000
b	0.5000	0.5000	0.0000
c	0.0000	0.5000	0.5000
d	0.0000	0.0000	1.0000

X2: COLUMN PROFILE MATRIX

	a	b	c	d
x	0.5000	0.5000	0.0000	0.0000
y	0.0000	0.5000	0.5000	0.0000
z	0.0000	0.0000	0.5000	0.5000

g1: ROW MARGINAL PROFILE

x 0.3333

y 0.3333

z 0.3333

g2: COLUMN MARGINAL PROFILE

a 0.1667

b 0.3333

c 0.3333

d 0.1667

Y1: MEAN CENTRED ROW PROFILE MATRIX

	x	y	z
a	0.6667	-0.3333	-0.3333
b	0.1667	0.1667	-0.3333
c	-0.3333	0.1667	0.1667
d	-0.3333	-0.3333	0.6667

r1: RANK OF Y1

r1 =
2

Y2: MEAN CENTRED COLUMN PROFILE MATRIX

	a	b	c	d
x	0.3333	0.1667	-0.3333	-0.1667
y	-0.1667	0.1667	0.1667	-0.1667
z	-0.1667	-0.3333	0.1667	0.3333

r2: RANK OF Y2

r2 =
2

EIGENVALUES, PRINCIPAL VECTORS, PRINCIPAL COMPONENTS

d1: EIGENVALUES OF V1Q1

0.7500
0.2500
0.0000

d2: EIGENVALUES OF V2Q2

0.7500
0.2500
0.0000
0.0000

PV1: PRINCIPAL VECTORS OF V1Q1

-0.4082 0.2357 0.3333
-0.0000 -0.4714 0.3333
0.4082 0.2357 0.3333

Check PV1: Does transpose(PV1)*Q1*PV1 equal Iq?

ans:

1.0000 0.0000 0.0000
0.0000 1.0000 0.0000
0.0000 -0.0000 1.0000

PC1: PRINCIPAL COMPONENTS OF V1Q1

a	-1.2247	0.7071	0.0000
b	-0.6124	-0.3536	0.0000
c	0.6124	-0.3536	0.0000
d	1.2247	0.7071	0.0000

Reconstitution formula: Does PC1*transpose(PV) equal Y1?

ans:

	x	y	z
a	0.6667	-0.3333	-0.3333
b	0.1667	0.1667	-0.3333
c	-0.3333	0.1667	0.1667
d	-0.3333	-0.3333	0.6667

PV2: PRINCIPAL VECTORS OF V2Q2

-0.2357	0.2357	0.1667	-0.1667
-0.2357	-0.2357	0.3333	0.3333
0.2357	-0.2357	0.3333	-0.3333
0.2357	0.2357	0.1667	0.1667

Check PV2: Does transpose(PV2)*Q2*PV2 equal Ip?

ans:

1.0000	-0.0000	0.0000	-0.0000
-0.0000	1.0000	-0.0000	0.0000
0.0000	-0.0000	1.0000	-0.0000
-0.0000	0.0000	-0.0000	1.0000

PC2: PRINCIPAL COMPONENTS OF V2Q2

x	-1.0607	0.3536	0.0000	0.0000
y	0.0000	-0.7071	0.0000	-0.0000
z	1.0607	0.3536	0.0000	0.0000

Reconstitution formula: Does PC2*transpose(PV2) equal Y2?

ans:

	a	b	c	d
x	0.3333	0.1667	-0.3333	-0.1667
y	-0.1667	0.1667	0.1667	-0.1667
z	-0.1667	-0.3333	0.1667	0.3333

Phi_square = 1

DIAGNOSTICS

GQRI1: GLOBAL QUALITY OF REPRESENTATION OF ROW PROFILES ON PRINCIPAL PLANE

1.0000

QRi1: QUALITY OF REPRESENTATION OF EACH ROW PROFILE ON
PRINCIPAL PLANE

a	1.0000
b	1.0000
c	1.0000
d	1.0000

Ci1: CONTRIBUTION OF EACH ROW PROFILE TO INERTIA EXPLAINED BY
AXES

	Ci1;1	Ci1;2
a	0.3333	0.3333
b	0.1667	0.1667
c	0.1667	0.1667
d	0.3333	0.3333

QORI2: GLOBAL QUALITY OF REPRESENTATION OF COLUMN PROFILES ON
PRINCIPAL PLANE

1.0000

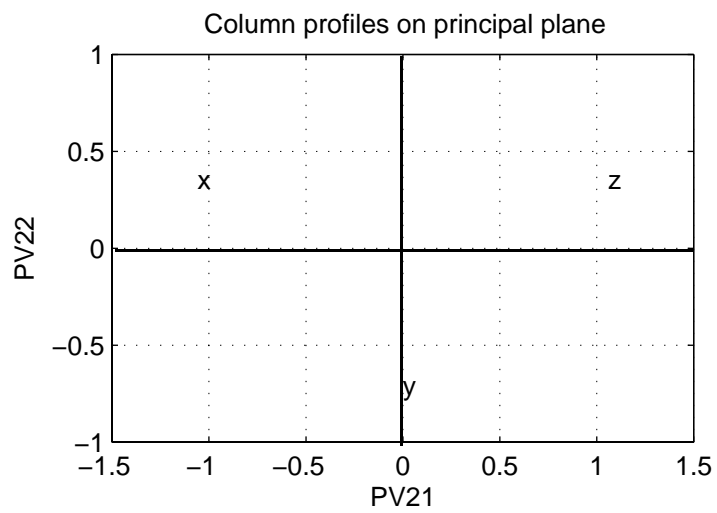
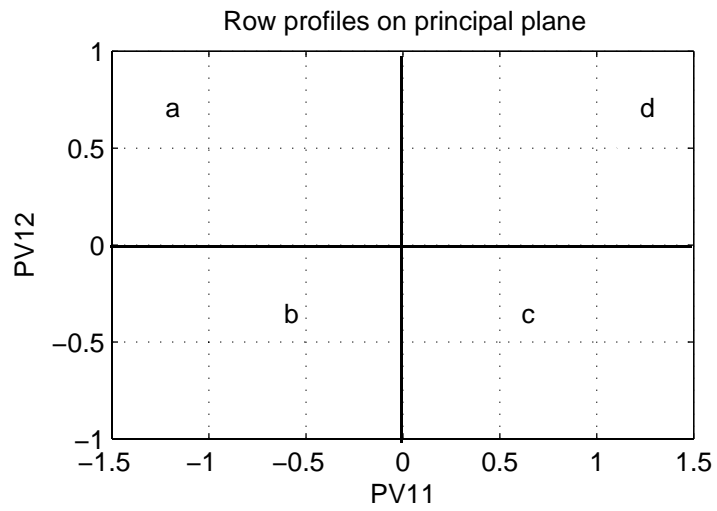
QRi2: QUALITY OF REPRESENTATION OF EACH COLUMN PROFILE ON
PRINCIPAL PLANE

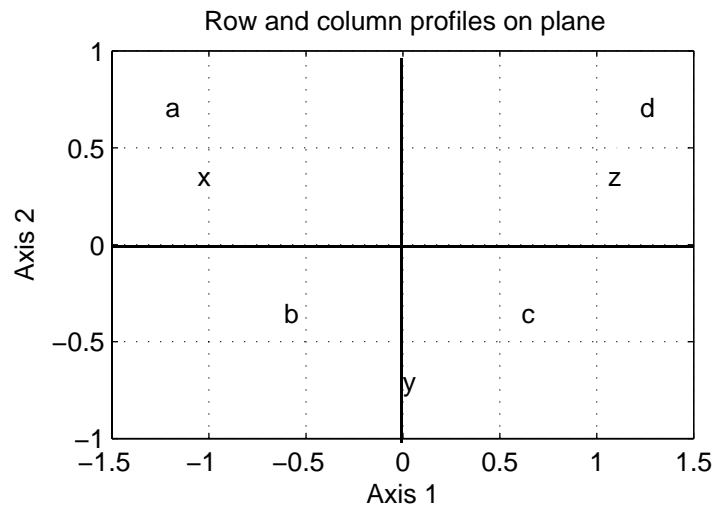
x	1.0000
y	1.0000
z	1.0000

Ci2: CONTRIBUTION OF EACH COLUMN PROFILE TO INERTIA EXPLAINED
BY AXES

	Ci2;1	Ci2;2
x	0.5000	0.1667
y	0.0000	0.6667
z	0.5000	0.1667

GRAPHICS





ANALISI BASATA SULLE MATRICI NON CENTRATE DEI PROFILI RIGA
E COLONNA

PRELIMINARY STATISTICS

X: FREQUENCY DATA MATRIX

	x	y	z
a	1	0	0
b	1	1	0
c	0	1	1
d	0	0	1

(p,q): ORDER OF X

p =
4

q =
3

n: NUMBER OF INDIVIDUALS

n =
6

X1: ROW PROFILE MATRIX

	x	y	z
a	1.0000	0.0000	0.0000
b	0.5000	0.5000	0.0000
c	0.0000	0.5000	0.5000
d	0.0000	0.0000	1.0000

r1: RANK OF X1

r1 =
3

X2: COLUMN PROFILE MATRIX

	a	b	c	d
x	0.5000	0.5000	0.0000	0.0000
y	0.0000	0.5000	0.5000	0.0000
z	0.0000	0.0000	0.5000	0.5000

r2: RANK OF X2

r2 =
3

g1: ROW MARGINAL PROFILE

x 0.3333
y 0.3333
z 0.3333

g2: COLUMN MARGINAL PROFILE

a 0.1667
b 0.3333
c 0.3333
d 0.1667

EIGENVALUES, PRINCIPAL VECTORS, PRINCIPAL COMPONENTS

d1: EIGENVALUES OF S1Q1

0.7500
0.2500
1.0000

d2: EIGENVALUES OF S2Q2

0.7500
0.2500
1.0000
0.0000

PV1: PRINCIPAL VECTORS OF S1Q1

0.4082 0.2357 0.3333
-0.0000 -0.4714 0.3333
-0.4082 0.2357 0.3333

Check PV1: Does $\text{transpose}(\text{PV1}) * \text{Q1} * \text{PV1}$ equal I_q ?

ans:

1.0000 -0.0000 -0.0000
-0.0000 1.0000 -0.0000
-0.0000 -0.0000 1.0000

PC1: PRINCIPAL COMPONENTS OF S1Q1

a 1.2247 0.7071 1.0000
b 0.6124 -0.3536 1.0000
c -0.6124 -0.3536 1.0000
d -1.2247 0.7071 1.0000

Reconstitution formula: Does $PC1 * \text{transpose}(PV)$ equal $X1$?

ans:

	x	y	z
a	1.0000	0.0000	0.0000
b	0.5000	0.5000	0.0000
c	0.0000	0.5000	0.5000
d	0.0000	0.0000	1.0000

PV2: PRINCIPAL VECTORS OF S2Q2

0.2357	-0.2357	0.1667	0.1667
0.2357	0.2357	0.3333	-0.3333
-0.2357	0.2357	0.3333	0.3333
-0.2357	-0.2357	0.1667	-0.1667

Check PV2: Does $\text{transpose}(PV2) * Q2 * PV2$ equal I_p ?

ans:

1.0000	0.0000	-0.0000	0.0000
0.0000	1.0000	-0.0000	-0.0000
-0.0000	-0.0000	1.0000	0.0000
0.0000	0.0000	0.0000	1.0000

PC2: PRINCIPAL COMPONENTS OF S2Q2

x	1.0607	-0.3536	1.0000	0.0000
y	-0.0000	0.7071	1.0000	-0.0000
z	-1.0607	-0.3536	1.0000	0.0000

Reconstitution formula: Does $PC2 * \text{transpose}(PV2)$ equal $X2$?

ans:

	a	b	c	d
x	0.5000	0.5000	-0.0000	0.0000
y	0.0000	0.5000	0.5000	0.0000
z	0.0000	-0.0000	0.5000	0.5000

Phi_square = 1

DIAGNOSTICS

GQR11: GLOBAL QUALITY OF REPRESENTATION OF ROW PROFILES ON PRINCIPAL PLANE

1.0000

QRi1: QUALITY OF REPRESENTATION OF EACH ROW PROFILE ON
PRINCIPAL PLANE

a	1.0000
b	1.0000
c	1.0000
d	1.0000

Ci1: CONTRIBUTION OF EACH ROW PROFILE TO INERTIA EXPLAINED BY
AXES

	Ci1;1	Ci1;2
a	0.3333	0.3333
b	0.1667	0.1667
c	0.1667	0.1667
d	0.3333	0.3333

QORI2: GLOBAL QUALITY OF REPRESENTATION OF COLUMN PROFILES ON
PRINCIPAL PLANE

1.0000

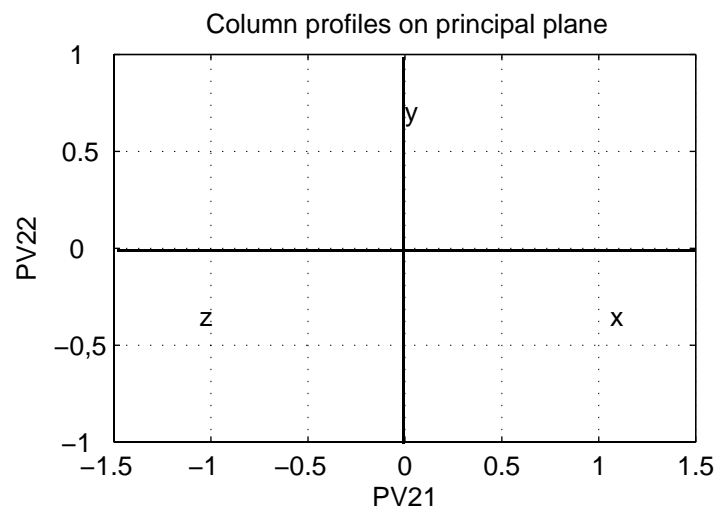
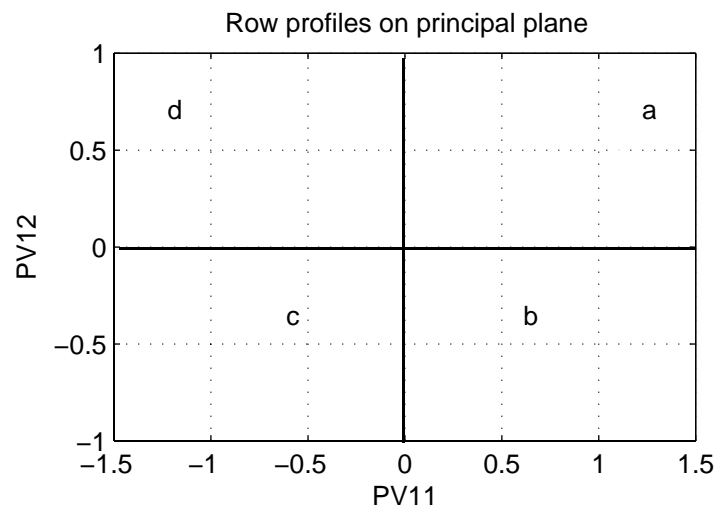
QRi2: QUALITY OF REPRESENTATION OF EACH COLUMN PROFILE ON
PRINCIPAL PLANE

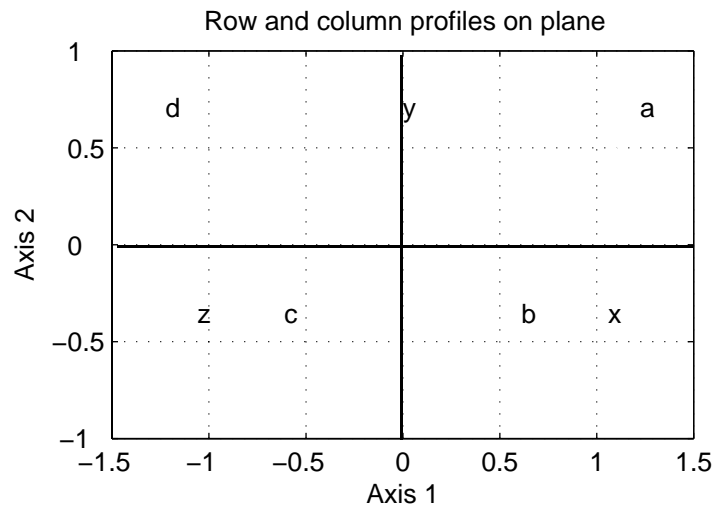
x	1.0000
y	1.0000
z	1.0000

Ci2: CONTRIBUTION OF EACH COLUMN PROFILE TO INERTIA EXPLAINED
BY AXES

	Ci2;1	Ci2;2
x	0.5000	0.1667
y	0.0000	0.6667
z	0.5000	0.1667

GRAPHICS





ESEMPIO 3

Questo esempio analizza una tavola di contingenza riferita ai caratteri «macroregione» e «causa di decesso».

Le modalità assunte dal carattere «macroregione» e le rispettive etichette sono: Nord, Centro, Sud, Isole.

Le modalità assunte dal carattere «causa di decesso» e le rispettive etichette sono: Malattie infettive (inf), Tumori (tum), Malattie nervose (ner), Malattie cardio circolatorie (cir), Malattie apparato respiratorio (res), Malattie apparato digerente (dig).

COMMENTI

- La prima parte (*Preliminary Statistics*) è dedicata alla esposizione di alcune caratteristiche dei dati a nostra disposizione. In particolare, sono indicate sia la matrice dei profili riga (*Row Profile Matrix*), sia la matrice dei profili colonna (*Column Profile Matrix*).
- La seconda parte (*Eigenvalues, Principal Vectors, Principal Components*) contiene i principali risultati che si ottengono dalla ACP delle matrici dei profili riga e colonna.
- La terza parte (*Diagnostics*) è dedicata alla presentazione di alcuni indici che servono a interpretare correttamente i grafici contenuti nella quarta parte (*Graphics*).

Al riguardo, si noti che:

- La qualità della rappresentazione globale dei profili riga e colonna nei rispettivi piani principali è molto buona.
- La qualità della rappresentazione dei singoli profili riga e colonna nei rispettivi piani principali è molto buona.
- I profili riga contribuiscono in modesta misura all'inerzia spiegata sia dal primo asse sia dal secondo asse.
- Il profilo colonna tum contribuisce in misura comparativamente maggiore degli altri profili all'inerzia spiegata dal primo asse; il profilo colonna dig

contribuisce in misura comparativamente maggiore degli altri profili alla inerzia spiegata dal secondo asse.

ANALISI BASATA SULLE MATRICI CENTRATE DEI PROFILI RIGA
E COLONNA

PRELIMINARY STATISTICS

X: FREQUENCY DATA MATRIX

	inf	tum	ner	cir	res	dig
Nord	1268	80665	7382	114942	15513	14367
Centro	614	33206	3074	53052	7817	6180
Sud	346	20058	1739	42597	7160	6323
Isole	186	11719	1127	25603	4055	2991

(p,q): ORDER OF X

p =
4
q =
6

n: NUMBER OF INDIVIDUALS

n =
461984

X1: ROW PROFILE MATRIX

	inf	tum	ner	cir	res	dig
Nord	0.0054	0.3445	0.0315	0.4909	0.0663	0.0614
Centro	0.0059	0.3195	0.0296	0.5104	0.0752	0.0595
Sud	0.0044	0.2564	0.0222	0.5446	0.0915	0.0808
Isole	0.0041	0.2565	0.0247	0.5605	0.0888	0.0655

X2: COLUMN PROFILE MATRIX

	Nord	Centro	Sud	Isole
inf	0.5253	0.2543	0.1433	0.0771
tum	0.5538	0.2280	0.1377	0.0805
ner	0.5541	0.2307	0.1305	0.0846
cir	0.4866	0.2246	0.1803	0.1084
res	0.4491	0.2263	0.2073	0.1174
dig	0.4811	0.2070	0.2117	0.1002

g1: ROW MARGINAL PROFILE

inf 0.0052
tum 0.3153
ner 0.0288
cir 0.5113
res 0.0748
dig 0.0646

g2: COLUMN MARGINAL PROFILE

Nord	0.5068
Centro	0.2250
Sud	0.1693
Isole	0.0989

Y1: MEAN CENTRED ROW PROFILE MATRIX

	inf	tum	ner	cir	res	dig
Nord	0.0002	0.0293	0.0027	-0.0203	-0.0085	-0.0033
Centro	0.0007	0.0042	0.0007	-0.0009	0.0004	-0.0052
Sud	-0.0008	-0.0588	-0.0066	0.0333	0.0168	0.0162
Isole	-0.0012	-0.0587	-0.0042	0.0492	0.0140	0.0008

r1: RANK OF Y1

r1 =
3

Y2: MEAN CENTRED COLUMN PROFILE MATRIX

	Nord	Centro	Sud	Isole
inf	0.0185	0.0294	-0.0260	-0.0218
tum	0.0470	0.0030	-0.0316	-0.0184
ner	0.0473	0.0058	-0.0388	-0.0143
cir	-0.0202	-0.0004	0.0110	0.0095
res	-0.0577	0.0013	0.0379	0.0185
dig	-0.0257	-0.0180	0.0424	0.0013

r2: RANK OF Y2

r2 =
3

EIGENVALUES, PRINCIPAL VECTORS, PRINCIPAL COMPONENTS

d1: EIGENVALUES OF V1Q1

0.0080
0.0004
0.0000
0.0000
0.0000
0.0000

d2: EIGENVALUES OF V2Q2

0.0080
0.0004
0.0000
0.0000

PV1: PRINCIPAL VECTORS OF V1Q1

```

-0.0055  0.0043  0.0534  0.0052 -0.0477 -0.0041
-0.4138 -0.0851  0.0523  0.3153  0.1210  0.1414
-0.0399  0.0245 -0.0206  0.0288  0.0007 -0.1593
 0.2769  0.2688 -0.2217  0.5113 -0.2081  0.0923
 0.1136  0.0160  0.1566  0.0748  0.1746 -0.0319
 0.0688 -0.2285 -0.0201  0.0646 -0.0405 -0.0383

```

Check PV1: Does $\text{transpose}(\text{PV1}) * Q1 * \text{PV1}$ equal I_q ?

ans:

```

1.0000  0.0000  0.0000  0.0000 -0.0000  0.0000
0.0000  1.0000 -0.0000 -0.0000  0.0000 -0.0000
-0.0000 -0.0000  1.0000 -0.0000  0.0000 -0.0000
 0.0000 -0.0000 -0.0000  1.0000 -0.0000  0.0000
-0.0000  0.0000  0.0000 -0.0000  1.0000  0.0000
 0.0000 -0.0000 -0.0000  0.0000  0.0000  1.0000

```

PC1: PRINCIPAL COMPONENTS OF V1Q1

```

Nord      -0.0698 -0.0064 -0.0031  0.0000 -0.0000  0.0000
Centro    -0.0126  0.0180  0.0100 -0.0000 -0.0000  0.0000
Sud       0.1480 -0.0265  0.0024 -0.0000 -0.0000  0.0000
Isole     0.1329  0.0373 -0.0108 -0.0000  0.0000 -0.0000

```

Reconstitution formula: Does $\text{PC1} * \text{transpose}(\text{PV})$ equal $Y1$?

ans:

```

          inf      tum      ner      cir      res      dig
Nord      0.0002  0.0293  0.0027 -0.0203 -0.0085 -0.0033
Centro    0.0007  0.0042  0.0007 -0.0009  0.0004 -0.0052
Sud      -0.0008 -0.0588 -0.0066  0.0333  0.0168  0.0162
Isole    -0.0012 -0.0587 -0.0042  0.0492  0.0140  0.0008

```

PV2: PRINCIPAL VECTORS OF V2Q2

```

-0.3965  0.1732  0.2506  0.5068
-0.0317 -0.2164 -0.3557  0.2250
 0.2809  0.2401 -0.0640  0.1693
 0.1473 -0.1969  0.1692  0.0989

```

Check PV2: Does $\text{transpose}(\text{PV2}) * Q2 * \text{PV2}$ equal I_p ?

ans:

```

1.0000  0.0000  0.0000  0.0000
0.0000  1.0000  0.0000  0.0000
0.0000  0.0000  1.0000 -0.0000
0.0000  0.0000 -0.0000  1.0000

```

PC2: PRINCIPAL COMPONENTS OF V2Q2

inf	-0.0942	-0.0153	-0.0648	-0.0000
tum	-0.1171	0.0051	-0.0011	0.0000
ner	-0.1234	-0.0159	0.0045	0.0000
cir	0.0483	-0.0098	0.0027	0.0000
res	0.1355	-0.0040	-0.0133	-0.0000
dig	0.0949	0.0662	0.0020	0.0000

Reconstitution formula: Does PC2*transpose(PV2) equal Y2?

ans:

	Nord	Centro	Sud	Isole
inf	0.0185	0.0294	-0.0260	-0.0218
tum	0.0470	0.0030	-0.0316	-0.0184
ner	0.0473	0.0058	-0.0388	-0.0143
cir	-0.0202	-0.0004	0.0110	0.0095
res	-0.0577	0.0013	0.0379	0.0185
dig	-0.0257	-0.0180	0.0424	0.0013

Phi_square = 0.0083

DIAGNOSTICS

GORI1: GLOBAL QUALITY OF REPRESENTATION OF ROW PROFILES ON PRINCIPAL PLANE

0.9952

QRi1: QUALITY OF REPRESENTATION OF EACH ROW PROFILE ON PRINCIPAL PLANE

Nord	0.9980
Centro	0.8277
Sud	0.9997
Isole	0.9939

Ci1: CONTRIBUTION OF EACH ROW PROFILE TO INERTIA EXPLAINED BY AXES

	Ci1;1	Ci1;2
Nord	0.3101	0.0592
Centro	0.0045	0.2081
Sud	0.4660	0.3405
Isole	0.2194	0.3921

GORI2: GLOBAL QUALITY OF REPRESENTATION OF COLUMN PROFILES ON PRINCIPAL PLANE

0.9952

**QRi2: QUALITY OF REPRESENTATION OF EACH COLUMN PROFILE ON
PRINCIPAL PLANE**

inf	0.6844
tum	0.9999
ner	0.9987
cir	0.9969
res	0.9905
dig	0.9997

**Ci2: CONTRIBUTION OF EACH COLUMN PROFILE TO INERTIA EXPLAINED
BY AXES**

	Ci2;1	Ci2;2
inf	0.0058	0.0035
tum	0.5432	0.0230
ner	0.0552	0.0209
cir	0.1499	0.1413
res	0.1726	0.0034
dig	0.0732	0.8079

GRAPHICS

