

Tre modi di utilizzare l'analisi fattoriale

- 1) L'analisi fattoriale si propone di trovare delle variabili latenti, o fattori, che spieghino le correlazioni tra le variabili osservate.
- 2) E' usata per ridurre l'insieme dei dati, identificando un numero ridotto di fattori che spiegano la maggior parte della varianza osservata.
- 3) E' usata per generare ipotesi riguardanti meccanismi causali tra variabili o per ridurre le variabili per successive analisi (per esempio, per identificare la collinearità tra variabili in un'analisi di regressione lineare).

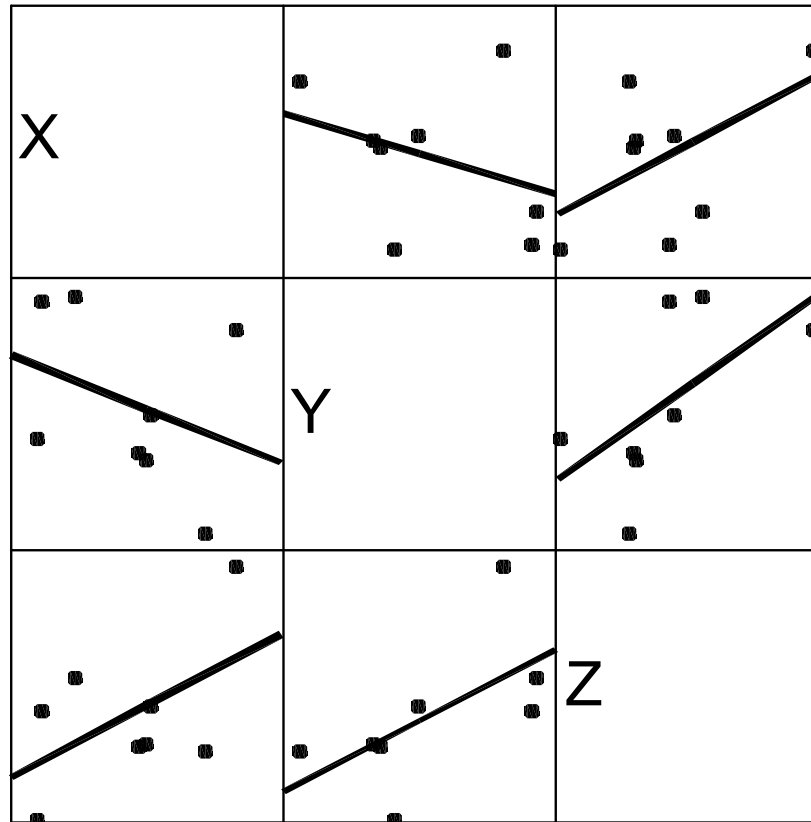
ACP - esempio multicollinearità

Si hanno 3 variabili moderatamente correlate:

X	Y	Z
-3.138	1.286	.169
-.297	.250	.044
-4.582	1.247	.109
.301	.498	.117
2.729	-.280	.035
-4.836	.350	-.094
.065	.208	.047
4.102	1.069	.375

Correlations

	X	Y	Z
X	1.000	-.346	.533
Y	-.346	1.000	.610
Z	.533	.610	1.000



Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	1.659	55.308	55.308
2	1.341	44.692	100.000
3	6.829E-06	2.276E-04	100.000

Extraction Method: Principal Component Analysis.

Il terzo autovalore è nullo poiché una delle tre variabili è combinazione lineare delle altre due e dunque il rango della matrice è pari a 2

NB: le componenti principali sono già standardizzate

[illegible]

ACP - Eurojob

Dati: Percentuali di occupati in diversi settori industriali

Matrice dati: (26 x 9)

L'analisi si effettua sulla **matrice di correlazione** perché si vuole considerare l'**intera struttura occupazionale** e quindi anche di quei settori industriali che sebbene pesino poco all'interno del paese, tuttavia ne caratterizzano la struttura

Osservazione: i dati percentuali introducono una relazione tra le 9 variabili: per ogni riga la somma è pari a 100. I 26 paesi si situano perciò in uno spazio di dimensione 8



Correlations

Pearson Correlation

	% occupati in agricoltura	% occupati industria edile	% occupati industria energetica	% occupati finanza	% occupati industria manifatturiera	% occupati industria mineraria	% occupati industria servizi	% occupati servizi sociali	% occupati industria telecomunicazione e trasporti
% occupati in agricoltura	1.000	-.538**	-.400*	-.220	-.671**	.036	-.737**	-.747**	-.565**
% occupati industria edile	-.538**	1.000	.060	.016	.494*	-.026	.356	.158	.388
% occupati industria energetica	-.400*	.060	1.000	.110	.385	.405*	.202	.132	.375
% occupati finanza	-.220	.016	.110	1.000	-.156	-.443*	.366	.108	-.246
% occupati industria manifatturiera	-.671**	.494*	.385	-.156	1.000	.445*	.204	.154	.351
% occupati industria mineraria	.036	-.026	.405*	-.443*	.445*	1.000	-.397*	-.281	.157
% occupati industria servizi	-.737**	.356	.202	.366	.204	-.397*	1.000	.572**	.188
% occupati servizi sociali	-.747**	.158	.132	.108	.154	-.281	.572**	1.000	.568**
% occupati industria telecomunicazione e trasporti	-.565**	.388	.375	-.246	.351	.157	.188	.568**	1.000

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Total Variance Explained

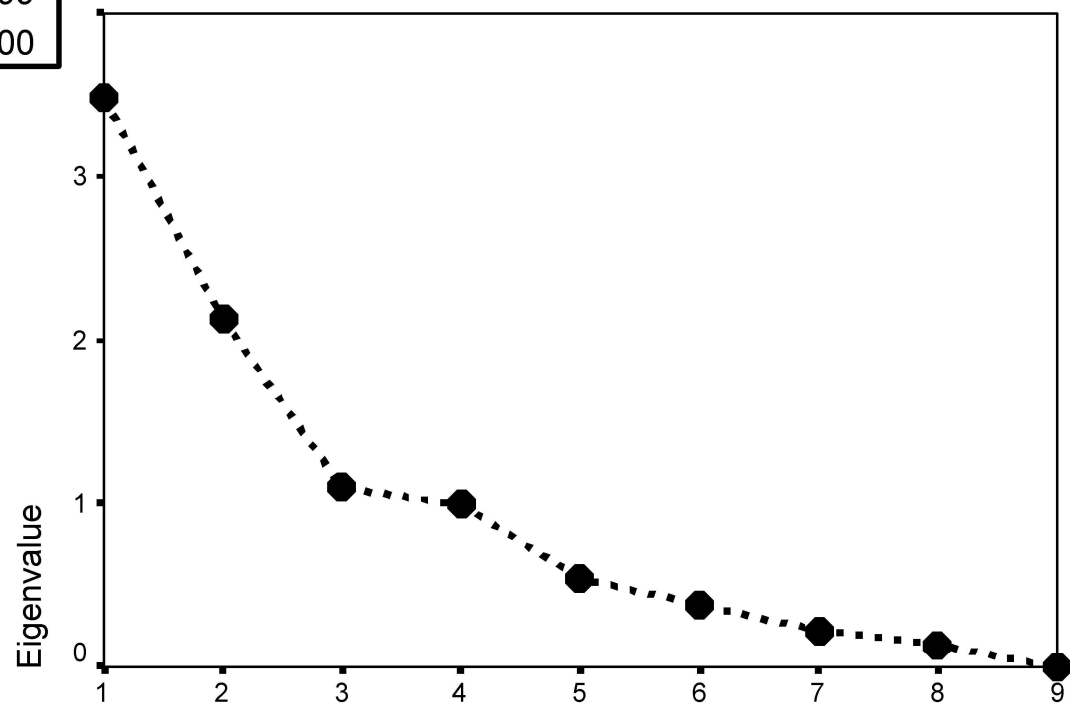
Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	3.487	38.746	38.746
2	2.130	23.669	62.415
3	1.099	12.211	74.625
4	.994	11.050	85.675
5	.543	6.036	91.711
6	.383	4.260	95.971
7	.226	2.508	98.480
8	.137	1.520	99.999
9	4.563E-05	5.069E-04	100.000

Extraction Method: Principal Component Analysis.

Correlation Matrix^a

a. Determinant =
2.382E-06

Scree Plot



Component Number

**Problema degli autovalori
multipli 3 e 4:**

o si includono entrambi
o si escludono entrambi

Correlazione delle variabili con le componenti principali

Correlations

	factor 1	factor 2	factor 3	factor 4	factor 5	factor 6	factor 7	fact or 8	factor 9
agricoltura	-,978	,078	-,051	,029	,157	-,095	-,010	,003	,005
industria edile	,607	,075	-,161	-,666	,348	-,081	,105	-,13	,001
industria energetica	,478	,381	,588	,392	,218	-,221	-,121	-,13	,000
finanza	,139	-,662	,616	-,051	,206	,325	,089	,064	,001
industria manifatturiera	,649	,518	,158	-,345	-,284	,178	-,228	,047	,002
industria mineraria	-,002	,902	,211	,064	-,121	,062	,345	,033	,000
industria servizi	,708	-,511	,121	-,050	-,209	-,381	,109	,143	,002
servizi sociali	,723	-,323	-,327	,411	-,162	,163	,091	-,19	,002
industria telecomunicazione e trasporti	,685	,296	-,393	,314	,378	,077	-,032	,201	,000

L'interpretazione delle componenti fattoriali è un tentativo di cogliere delle dimensioni latenti soggiacenti al fenomeno osservato.

Tuttavia, non sempre è lecita l'interpretazione dei risultati, giacché le componenti principali risultanti sono il frutto sostanzialmente di una procedura di ottimizzazione su una funzione obiettivo.

Possiamo dire che l'interpretazione è lecita soprattutto quando questa va a convalidare delle ipotesi su un modello già precostituito.

Interpretazione delle prime due componenti principali

Fattore 1:

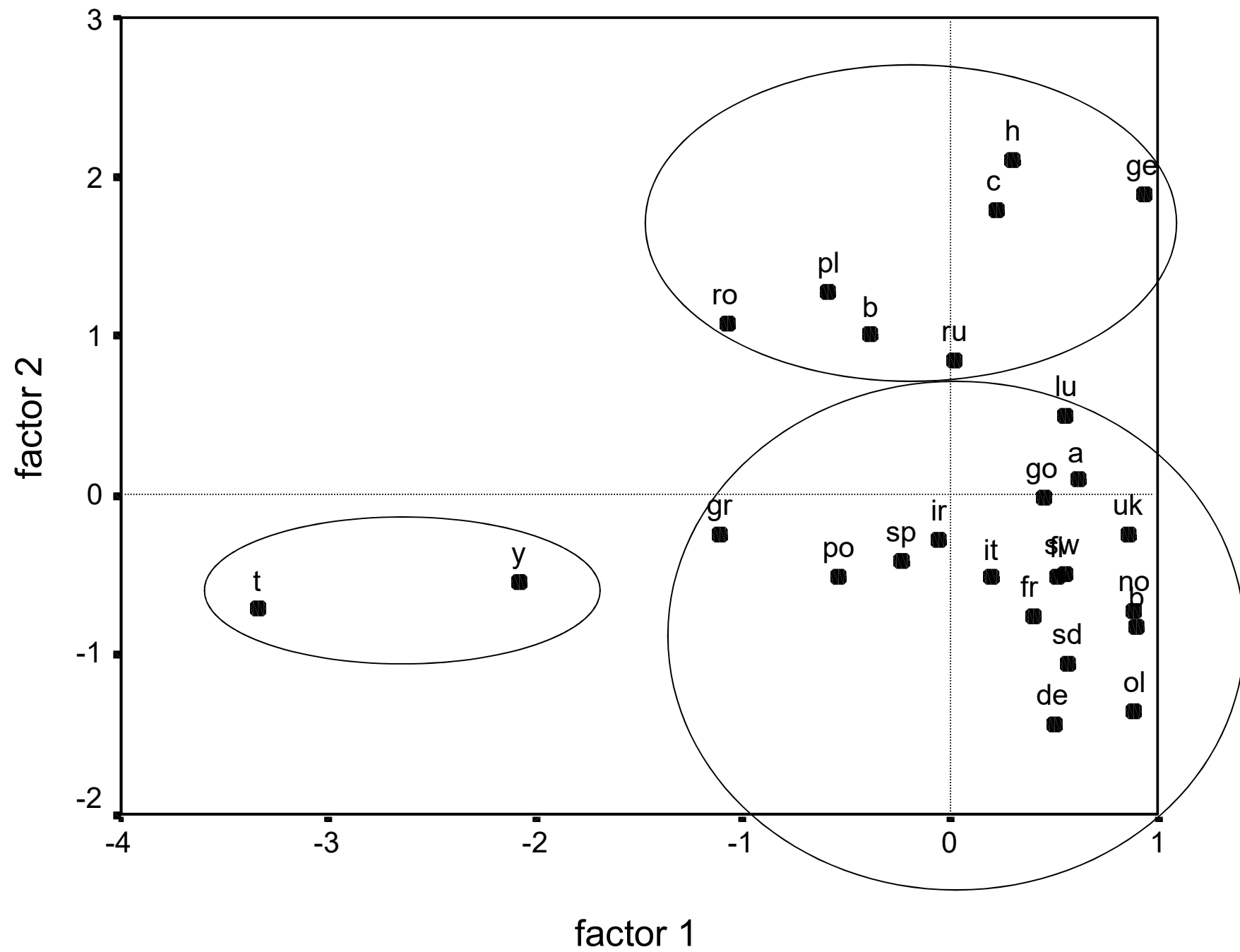
E' fortemente correlato negativamente con l'agricoltura mentre è incorrelato o positivamente correlato con tutte le altre variabili. Questo fattore discrimina paesi con economia agricola da quelli con economia industriale.

Il grafico mostra che la Turchia e la Jugoslavia hanno sul fattore 1 valori più bassi degli altri paesi e ciò indica che le loro economie si basano sull'agricoltura molto più di quelle degli altri paesi.

Fattore 2:

Sono correlate negativamente le variabili relative all'industria del settore Terziario, della Finanza e dei Servizi Sociali, mentre le restanti variabili sono incorrelate o correlate positivamente.

Questo fattore discrimina tra paesi rispetto al loro sviluppo in termini di settori industriali dei Servizi e Finanziari, opponendo i paesi dell'Europa occidentale ai paesi dell'Europa Orientale.



ACP -Città

DATI: 5 variabili Socio-Economiche osservate su 12 quartieri di una città americana

Matrice dati: (12 x 5)

Variabili:

POPOL	= Popolazione (in migliaia)
SCUOLA	= Num. Medio anni studio
OCCUPATI	= Occupati (in migliaia)
SERVIZI	= Num. Servizi pubblici
CASE	= Costo medio dell'abitazione

l'analisi si basa sulla matrice di correlazione

Descriptive Statistics

	Mean	Std. Deviation	N
costo medio abitazione	17000,0000	6367,5313	12
OCCUPATI	2333,3333	1241,2115	12
POPOL	6241,6667	3439,9943	12
num. medio anni studio	11,4417	1,7865	12
num. servizi pubbl.	120,8333	114,9275	12

- 1) Le dev. Stand. sono molto alte;
 - 2) le variabili POPOL e OCCUPATI
 - 3) sono molto correlate;
- tutte le variabili sono correlate positivamente e questo implica che i coefficienti del 1° fattore saranno tutti positivi.

Correlations

		costo medio abitazione	OCCUPATI	POPOL	num. medio anni studio	num. servizi pubbl.
Pearson Correlation	costo medio abitazione	1,000	,122	,022	,863	,778
	OCCUPATI	,122	1,000	,972	,154	,515
	POPOL	,022	,972	1,000	,010	,439
	num. medio anni studio	,863	,154	,010	1,000	,691
	num. servizi pubbl.	,778	,515	,439	,691	1,000

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2,873	57,466	57,466
2	1,797	35,933	93,399
3	,215	4,297	97,696
4	9,993E-02	1,999	99,695
5	1,526E-02	,305	100,000

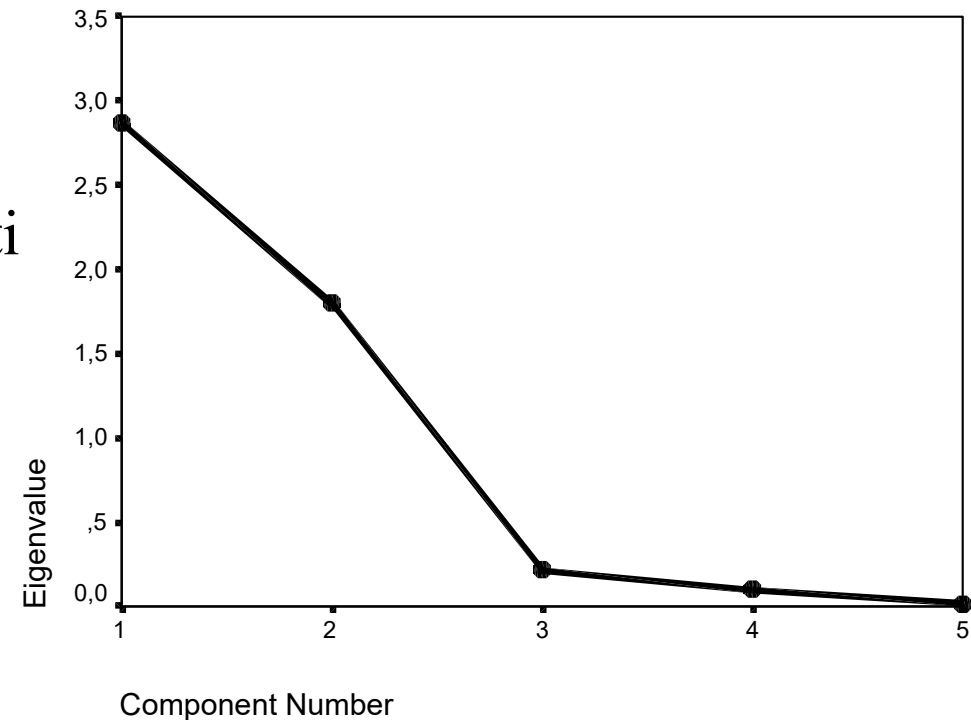
Correlation Matrix^a

a. Determinant =
1,691E-03

Extraction Method: Principal Component Analysis.

I primi due fattori sono sufficienti
l'ultimo autovalore è nullo

Scree Plot



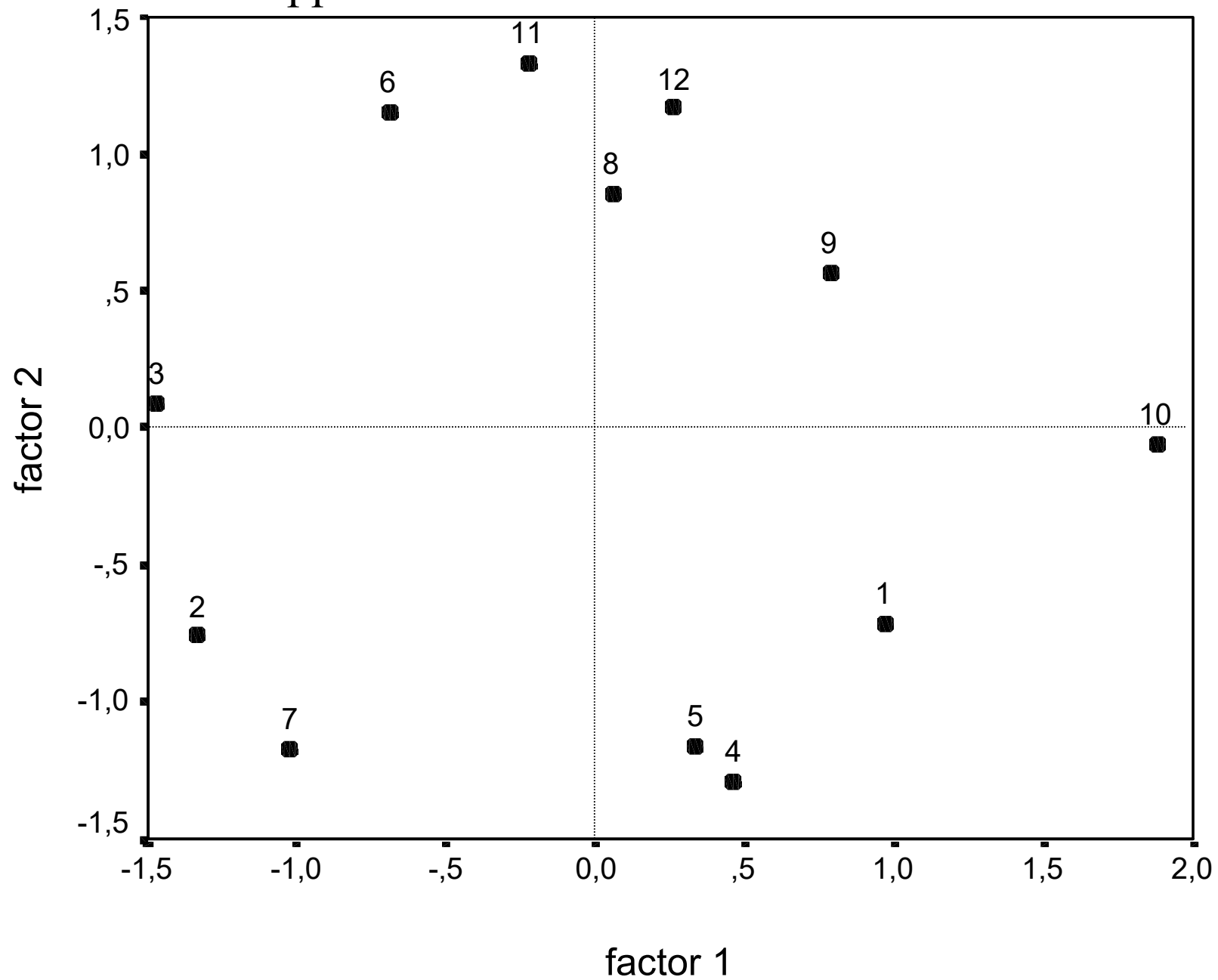
Tutte correlate
positivamente

Component Matrix

	Component				
	1	2	3	4	5
costo medio abitazione	,791	-,558	-6,473E-02	,241	-1,018E-02
OCCUPATI	,672	,726	,115	7,251E-03	-8,621E-02
POPOL	,581	,806	2,759E-02	6,450E-02	8,516E-02
num. medio anni studio	,767	-,545	,319	-,112	2,160E-02
num. servizi pubbl.	,932	-,104	-,308	-,158	-1,526E-05

Extraction Method: Principal Component Analysis.

Rappresenta il 93% della variabilità totale



Interpretazione dei fattori

Fattore 1:

Sintetizza globalmente l'informazione di tutte le variabili originarie, in particolare della variabile **Servizi**.

In questo modo riesce a ordinare i quartieri soprattutto in base alle loro caratteristiche urbanistiche (servizi e infrastrutture)

Fattore 2:

Contrappone le variabili **Popolaz** e **Occupati** alle variabili **Case** e **Scuola**.

Contrappone “quartieri popolari” ad alta densità di popolazione e di occupati a “quartieri borghesi” con abitazioni più lussuose i cui abitanti possiedono un livello di studio più alto (presenza di lib. Professionisti e dirigenti)

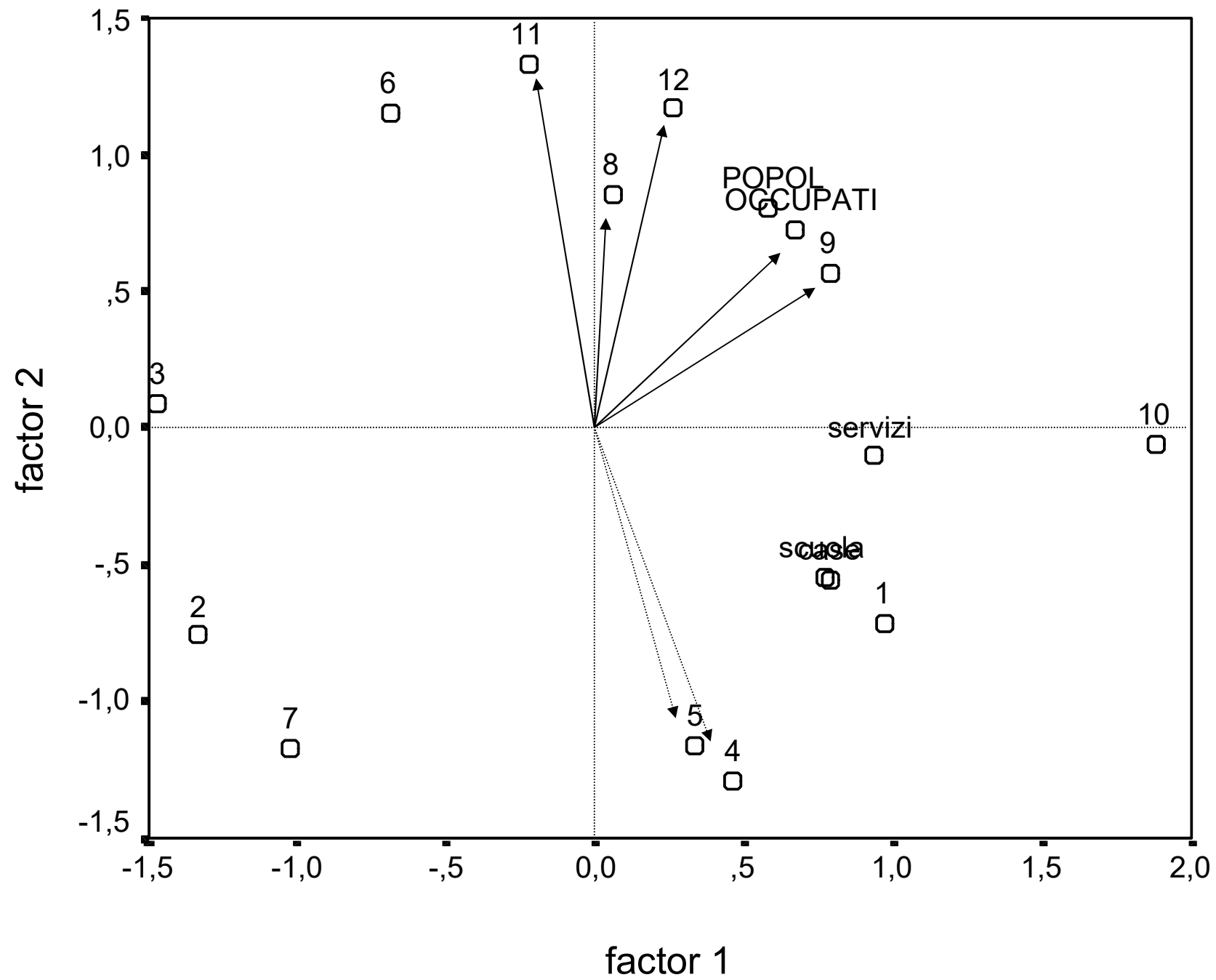
BIPLOT

E' una rappresentazione grafica utile all'interpretazione delle relazioni tra unità e variabili sul piano fattoriale

Il grafico rappresenta sullo stesso piano fattoriale le proiezioni sia delle unità che delle variabili originarie

Consideriamo il coseno dell'angolo tra il vettore che dall'origine va ad una variabile e quello che dall'origine va ad un'unità :

se il coseno è positivo ($<90^\circ$), l'unità presenta un valore superiore al valore medio della variabile; il contrario se il coseno è negativo ($>90^\circ$)



VARIABILI SCARTO

	CASE	OCCUP	POPOL	SCUOL	SERV	FACTOR 1	FACTOR 2
1	8000,00	166,67	-541,67	1,36	149,17	,96967	-,71262
2	-7000,00	-1733,33	-5241,67	-,54	-110,83	-1,33148	-,75880
3	-8000,00	-1333,33	-2841,67	-2,64	-110,83	-1,47203	,08977
4	8000,00	-633,33	-2441,67	2,16	19,17	,45964	-1,29109
5	8000,00	-733,33	-2241,67	1,36	19,17	,33300	-1,16178
6	-5000,00	266,67	1958,33	-3,14	-60,83	-,69203	1,15021
7	-1000,00	-1933,33	-5041,67	-,04	-110,83	-1,02326	-1,17543
8	-3000,00	966,67	2858,33	,06	-60,83	,05749	,85488
9	1000,00	1066,67	3658,33	1,06	59,17	,78458	,56632
10	8000,00	1266,67	3358,33	2,26	269,17	1,87960	-,05898
11	-5000,00	966,67	3358,33	-1,84	-40,83	-,22704	1,33005
12	-4000,00	1666,67	3158,33	-,04	-20,83	,26186	1,16748

Qualità di rappresentazione delle unità

$$\cos^2\theta_{\mathbf{x}_i\mathbf{P}} = \|\mathbf{x}_{i,\mathbf{P}}\|^2 / \|\mathbf{x}_i\|^2 = [(c_i^1)^2 + (c_i^2)^2] / [(c_i^1)^2 + \dots + (c_i^m)^2]$$

Esprime la vicinanza, in termini di angolo, tra \mathbf{x}_i e $\mathbf{x}_{i,\mathbf{P}}$ e dipende dai punteggi assunti dall'unità i -sima nelle prime due componenti principali.

(nell'spss è necessario riportare i fattori all'unità di misura originaria, moltiplicando gli scores per la radice degli autovalori corrispondenti)

Qualità di rappresentazione delle unità statistiche sul piano fattoriale

	1	2	3	4	5	6	7	8	9	10	11	12
COS	,93	,94	,96	,94	,93	,86	,99	,73	,93	,95	,99	,86

Qualità di rappresentazione delle variabili

$$\cos^2 \theta_{x^j p} = \|\mathbf{x}^j_{,p}\|^2 / \|\mathbf{x}^j\|^2$$

	variabili	qualità
	case	,94
	OCCUPATI	,98
	POPOL	,99
	scuola	,89
	servizi	,88

Si fa la somma dei quadrati delle correlazioni delle variabili con i primi due fattori

Le variabili originarie sono misurate in maniera disomogenea. POPOL, OCCUPATI e SERVIZI sono valori assoluti, mentre SCUOLA e CASE sono delle medie.

Al posto delle variabili POPOL e OCCUPATI consideriamo la variabile % di Occupati su Popolazione

Al posto della variabile SERVIZI consideriamo la variabile % di Servizi su Popolazione.

Correlations

	CASE	%OCCUP	%SERV	SCUOLA
CASE	1,000	,113	,948	,863
%OCCUP	,113	1,000	,273	,355
%SERV	,948	,273	1,000	,803
SCUOLA	,863	,355	,803	1,000

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	2,842	71,061	71,061
2	,941	23,524	94,586
3	,197	4,935	99,521
4	1,916E-02	,479	100,000

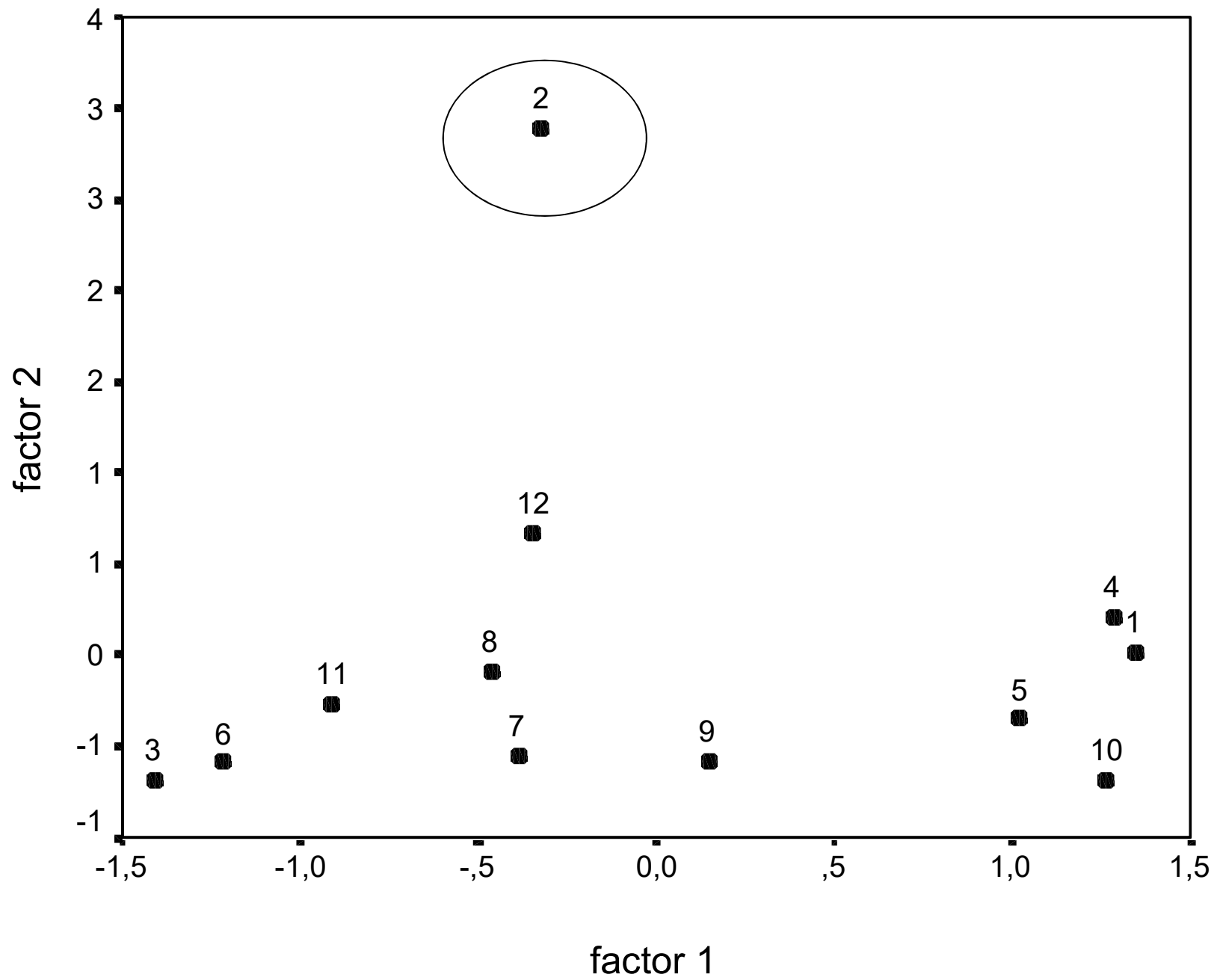
Extraction Method: Principal Component Analysis.

Component Matrix^a

	Component			
	1	2	3	4
POCCUP	,381	,923	5,695E-02	2,314E-02
costo medio abitazione	,953	-,277	5,610E-02	,106
PSERV	,955	-,112	,263	-7,764E-02
num. medio anni studio	,936	2,133E-02	-,349	-3,782E-02

Extraction Method: Principal Component Analysis.

a. 4 components extracted.



	SCUO LA	CASE	%OCCUP	%SERV
1	12,80	25000,00	43,86	4,74
2	10,90	10000,00	60,00	1,00
3	8,80	9000,00	29,41	,29
4	13,60	25000,00	44,74	3,68
5	12,80	25000,00	40,00	3,50
6	8,30	12000,00	31,71	,73
7	11,40	16000,00	33,33	,83
8	11,50	14000,00	36,26	,66
9	12,50	18000,00	34,34	1,82
10	13,70	25000,00	37,50	4,06
11	9,60	12000,00	34,38	,83
12	11,40	13000,00	42,55	1,06

Record - ACP

DATI: 8 variabili corrispondenti ai tempi dei record nazionali prima delle Olimpiadi del 1984 in 8 differenti gare di corsa maschile effettuati da 55 diverse nazioni

Matrice dati: (55 x 8)

a100m: tempo record in secondi nei 100 metri

a200m: tempo record in secondi nei 200 metri

a400m: tempo record in secondi nei 400 metri

a800m: tempo record in secondi negli 800 metri

a1500m: tempo record in minuti nei 1500 metri

a5k: tempo record in minuti nei 5000 metri

a10k: tempo record in minuti nei 10000 metri

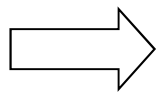
MARAT: tempo record in minuti nella maratona

Obiettivo: Ridurre la dimensione dei dati

Descriptive Statistics

	Mean	Std. Deviation
A100M	10,4711	,3514
A10K	28,9891	1,8077
A1500M	3,6982	,1559
A200M	20,9404	,6446
A400M	46,4387	1,4570
A5K	13,8458	,8012
A800M	1,7933	6,368E-02
MARAT	136,6240	9,2270

Variabilità e scala molto diverse



Analisi sulla matrice di correlazione

Correlations

	A100M	A200M	A400M	A800M	A1500M	A5K	A10K	MARAT
A100M	1,000	,923	,841	,756	,700	,619	,633	,520
A200M	,923	1,000	,851	,807	,775	,695	,697	,596
A400M	,841	,851	1,000	,870	,835	,779	,787	,705
A800M	,756	,807	,870	1,000	,918	,864	,869	,806
A1500M	,700	,775	,835	,918	1,000	,928	,935	,866
A5K	,619	,695	,779	,864	,928	1,000	,975	,932
A10K	,633	,697	,787	,869	,935	,975	1,000	,943
MARAT	,520	,596	,705	,806	,866	,932	,943	1,000

Tutte le correlazioni sono di segno positivo e $>0,5$.
 Inoltre, la correlazione diminuisce all'aumentare della
 “diversità”(data dalla differenza di metri da percorrere)
 tra le corse.

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	6,622	82,777	82,777
2	,878	10,970	93,747
3	,159	1,992	95,739
4	,124	1,551	97,289
5	7,988E-02	,999	98,288
6	6,797E-02	,850	99,137
7	4,642E-02	,580	99,717
8	2,260E-02	,283	100,000

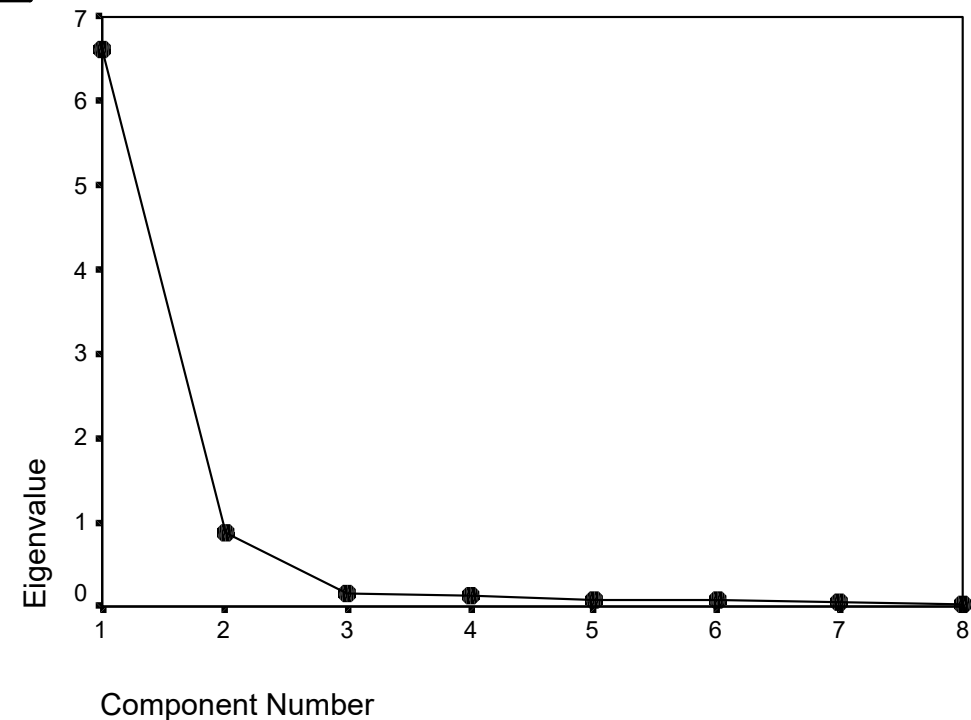
Extraction Method: Principal Component Analysis.

Le prime due componenti spiegano quasi tutta la variabilità.

Correlation Matrix^a

a. Determinant =
6,542E-07

Scree Plot

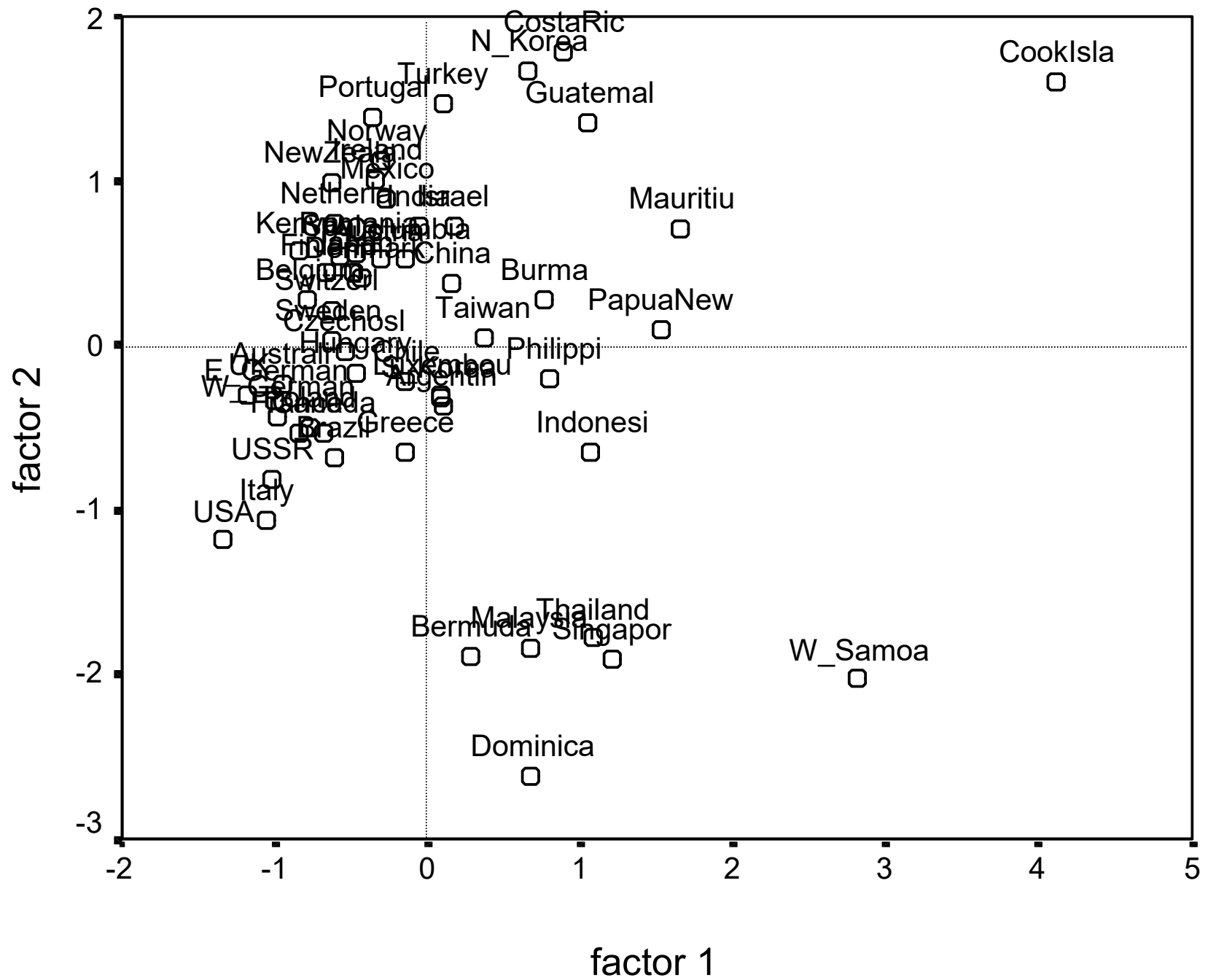


Correlazione tra variabili e fattori

	Component	
	1	2
A100M	,817	,531
A200M	,867	,432
A400M	,915	,233
A800M	,949	1,164E-02
A1500M	,959	-,131
A5K	,938	-,292
A10K	,944	-,287
MARAT	,880	-,411

Il primo fattore è riassuntivo dell'intero insieme delle variabili originarie; si può considerare come una variabile di classificazione delle unità osservate.

Sul secondo fattore si contrappongono le gare dei velocisti (100, 200, 400, 800) con quelle dei fondisti (1500, 5000, 1000 e maratona), mettendo in evidenza la particolare struttura della matrice di correlazione



	PAESE	factor 1
1	USA	-1,33311
2	UK	-1,17521
3	Italy	-1,05969
4	USSR	-1,02079
5	E_German	-1,00650
6	W_German	-,99199
7	Australi	-,95066
8	France	-,84400
9	Kenya	-,84260
10	Belgium	-,79323

50	Thailand	1,07324
51	Singapor	1,21325
52	PapuaNew	1,51910
53	Mauritiu	1,65491
54	W_Samoa	2,81004
55	CookIsle	4,10190

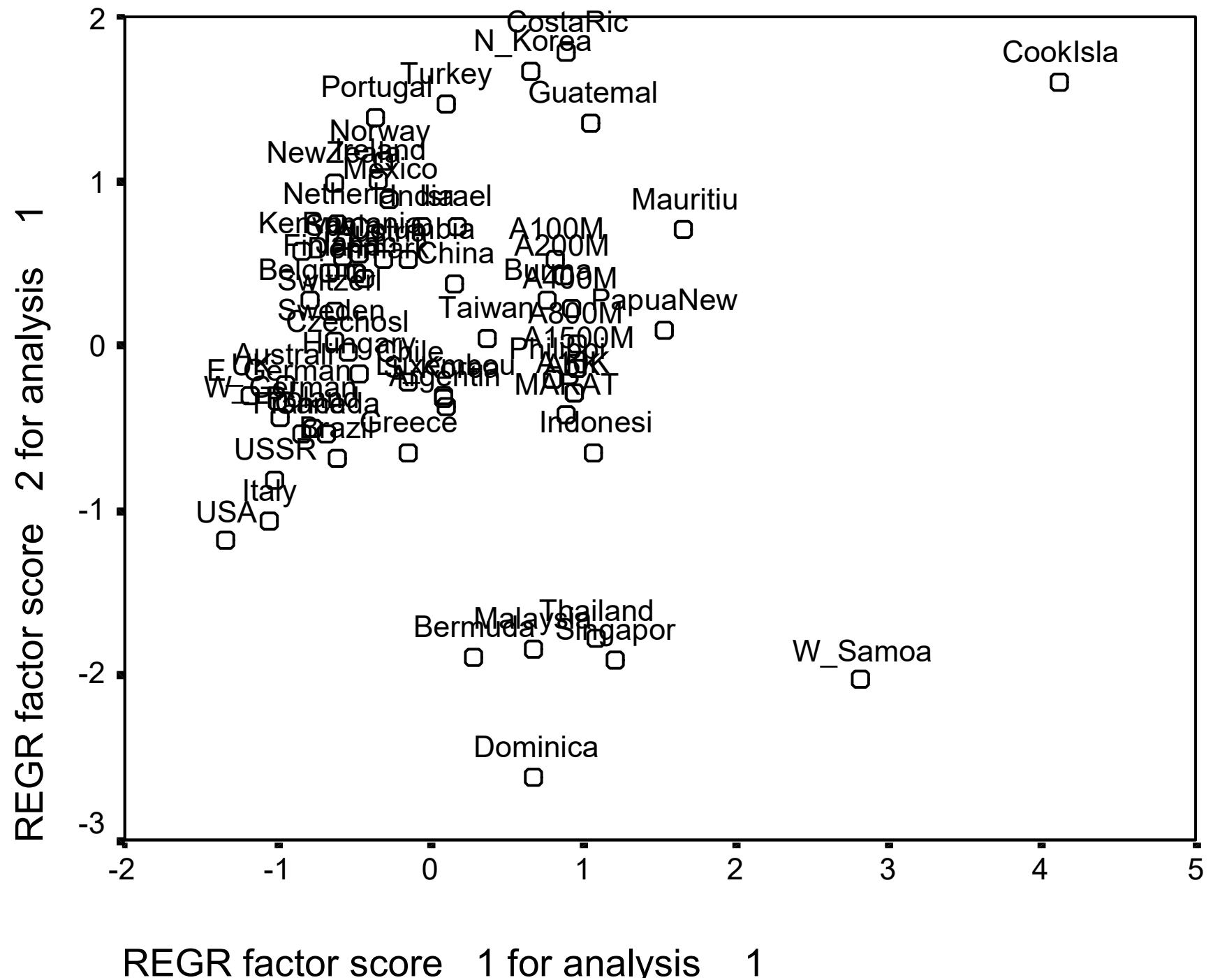
Ecco una graduatoria
complessiva fra le nazioni
rispetto ai record di corsa
maschile

Il secondo fattore misura la differenza nelle performance di una nazione tra le corse brevi e lunghe.

A valori positivi corrispondono migliori performance nelle gare lunghe.

Esempio:

Sebbene le performance sia delle Isole Cook che di W_Samoa sono complessivamente scarse, la prima nazione è migliore nelle gare lunghe mentre la seconda nelle gare brevi.



DATI: 15 variabili corrispondenti ad indicatori demografici
nelle regioni italiane al 1984

Matrice dati: (20 x 15)

NAT	Natalità (nati vivi per 1000 ab.)
MOR	Mortalità (morti per 1000 ab.)
INC	Incremento naturale (natalità -mortalità)
TFT	Tasso di fecondità totale (n. medio figli/donna)
ILL	Illegittimi per 1000 nati
ABP	Aborti procurati per 100 nati
MINF	Mortalità infantile (per 1000 nati vivi)
VMM	Vita media maschile
VMF	Vita media femminile
SMM	Supermortalità maschile (per 1000 nati vivi)
NUZ	Nuzialità (matrimoni per 1000 abitanti)
MATC	Matrimoni civili per 100 abitanti
DIV	Divorzi per 1000 ab.
VEC	Vecchi per 100 giovani
AMF	Ampiezza media delle famiglie

Descriptive Statistics

	Mean	Std. Deviation	N
ABP	41,4800	18,3503	20
AMF	3,0200	,2821	20
DIV	2,1450	1,3260	20
ILL	49,1000	23,6841	20
INC	,2200	4,1390	20
MATC	13,7500	6,3776	20
MINF	10,9550	3,1235	20
MOR	10,0800	1,6873	20
NAT	10,3000	2,5507	20
NUZ	5,2550	,7338	20
SMM	108,9550	1,8549	20
TFT	1,4620	,3342	20
VEC	69,3500	23,0523	20
VMF	77,5250	,7419	20
VMM	71,1850	1,3546	20

Correlations

	ABP	AMF	DIV	ILL	INC	MATC	MINF	MOR	NAT	NUZ	SMM	TFT	VEC	VMF	VMM
ABP	1,000	-,774	,643	,488	-,786	,618	-,532	,735	-,789	-,795	,287	-,692	,779	,295	-,130
AMF	-,774	1,000	-,833	-,606	,922	-,632	,725	-,886	,910	,841	-,457	,782	-,846	-,153	,353
DIV	,643	-,833	1,000	,788	-,637	,734	-,571	,612	-,630	-,599	,464	-,479	,488	-,254	-,552
ILL	,488	-,606	,788	1,000	-,490	,664	-,413	,483	-,475	-,435	,538	-,404	,340	-,227	-,608
INC	-,786	,922	-,637	-,490	1,000	-,651	,728	-,965	,985	,932	-,582	,891	-,948	-,372	,357
MATC	,618	-,632	,734	,664	-,651	1,000	-,458	,667	-,615	-,612	,617	-,499	,567	-,060	-,583
MINF	-,532	,725	-,571	-,413	,728	-,458	1,000	-,610	,778	,731	-,574	,730	-,602	-,272	,390
MOR	,735	-,886	,612	,483	-,965	,667	-,610	1,000	-,904	-,866	,533	-,819	,963	,293	-,349
NAT	-,789	,910	-,630	-,475	,985	-,615	,778	-,904	1,000	,940	-,591	,904	-,901	-,410	,349
NUZ	-,795	,841	-,599	-,435	,932	-,612	,731	-,866	,940	1,000	-,567	,920	-,866	-,477	,295
SMM	,287	-,457	,464	,538	-,582	,617	-,574	,533	-,591	-,567	1,000	-,561	,376	,036	-,882
TFT	-,692	,782	-,479	-,404	,891	-,499	,730	-,819	,904	,920	-,561	1,000	-,811	-,467	,295
VEC	,779	-,846	,488	,340	-,948	,567	-,602	,963	-,901	-,866	,376	-,811	1,000	,456	-,129
VMF	,295	-,153	-,254	-,227	-,372	-,060	-,272	,293	-,410	-,477	,036	-,467	,456	1,000	,439
VMM	-,130	,353	-,552	-,608	,357	-,583	,390	-,349	,349	,295	-,882	,295	-,129	,439	1,000

Faccio l'analisi sulla matrice di correlazione di 12 variabili,
escludendo NAT, MOR, INC.

Total Variance Explained

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	6,913	57,607	57,607
2	2,297	19,144	76,752
3	1,060	8,832	85,584
4	,537	4,479	90,063
5	,419	3,490	93,553
6	,295	2,461	96,015
7	,212	1,766	97,781
8	,131	1,092	98,872
9	,119	,990	99,863
10	1,422E-02	,118	99,981
11	2,183E-03	1,819E-02	99,999
12	9,283E-05	7,736E-04	100,000

Extraction Method: Principal Component Analysis.

Consideriamo i primi due fattori che spiegano circa il 77% della variabilità totale.

Component Matrix

	Component	
	1	2
ABP	,804	-,273
AMF	-,921	9,439E-02
DIV	,808	,364
ILL	,695	,481
MATC	,789	,291
MINF	-,786	,117
NUZ	-,912	,329
SMM	,697	,365
TFT	-,854	,342
VEC	,821	-,426
VMF	,199	-,858
VMM	-,543	-,732

Il primo fattore è fortemente correlato con 8 delle 12 variabili originarie ed ha un significato demografico molto ampio.

Il secondo fattore rappresenta unicamente l'aspetto della durata media di vita della popolazione

Contrapposti sul primo fattore ci sono:

Sul semiasse positivo:

Comportamenti laici (matrimoni civili, illegittimità, divorzialità, abortività elevate);

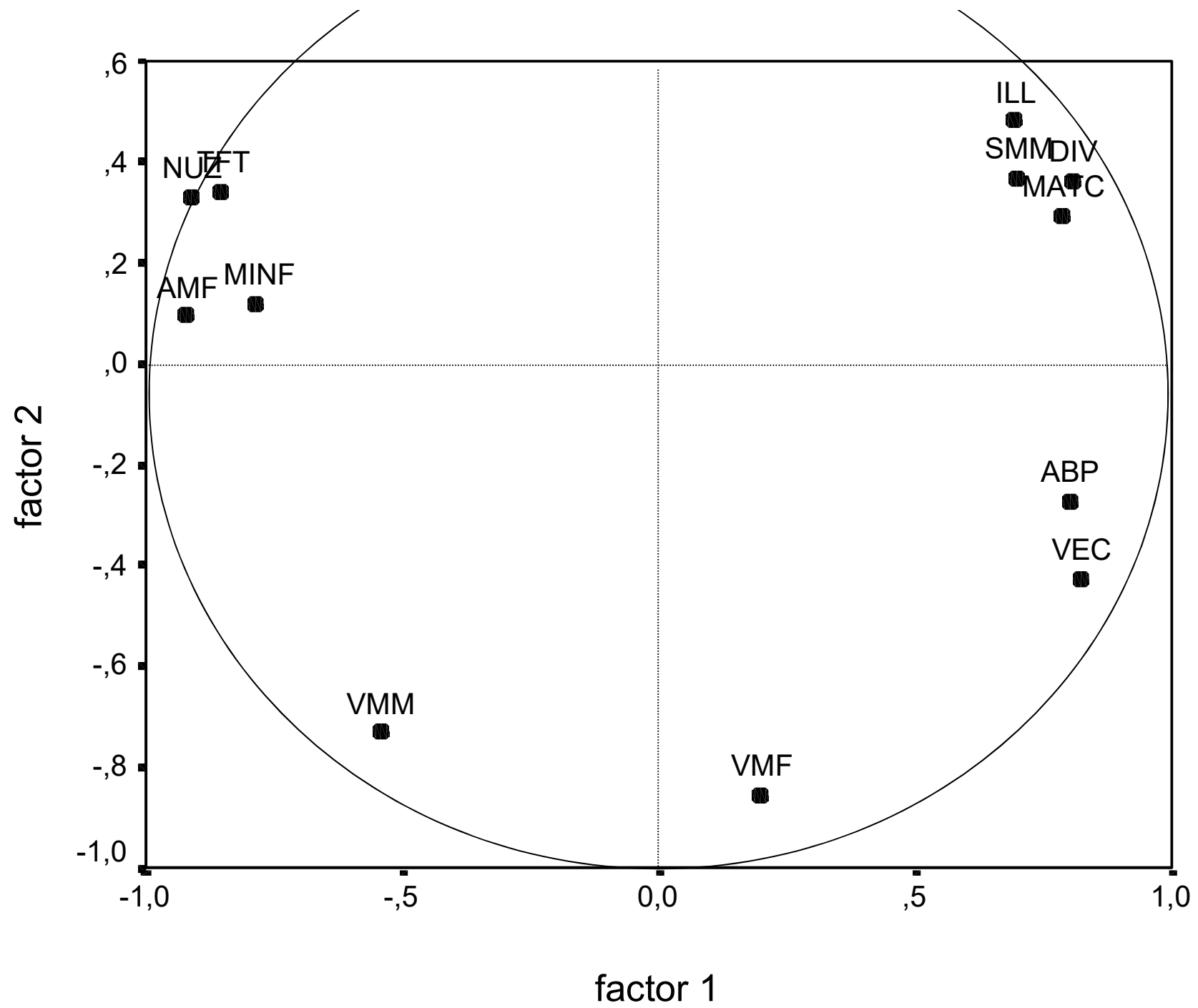
elevato invecchiamento demografico

supermortalità

Sul semiasse negativo:

comportamenti tradizionali (nuzialità e fecondità elevate con ampie dimensioni familiari)

mortalità infantile più elevata



Qualità di rappresentazione unità:

Piem.=0,75

V.Aosta= 0,75

Lomb=0,68

Trent=0,47

Venet=0,06

Friuli=0,91

Lazio=0,10

Abruz=0,86

Molis=0,56

Camp=0,88

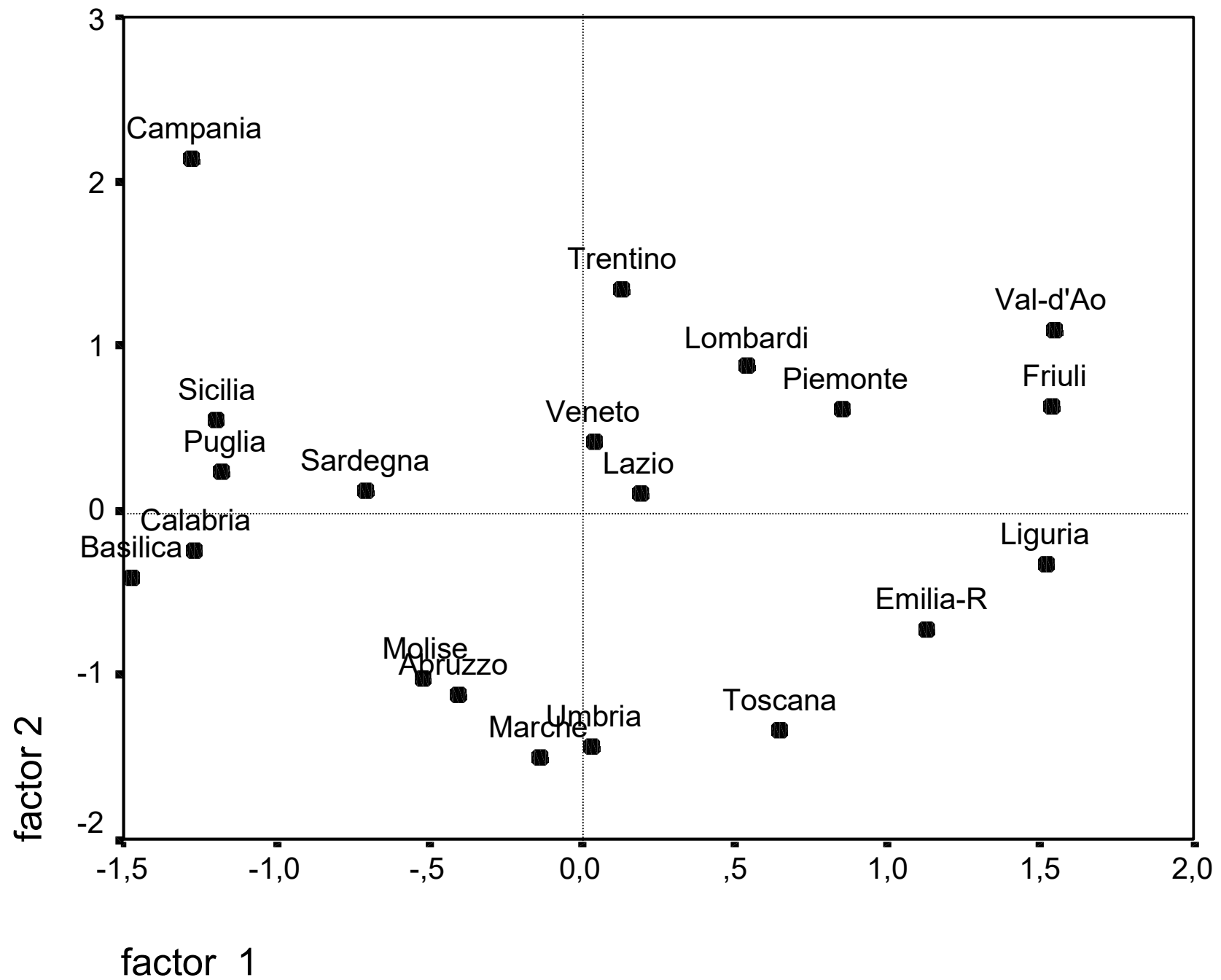
Pugl=0,77

Basil=0,93

Calab=0,94

Sicil=0,79

Sard=0,59



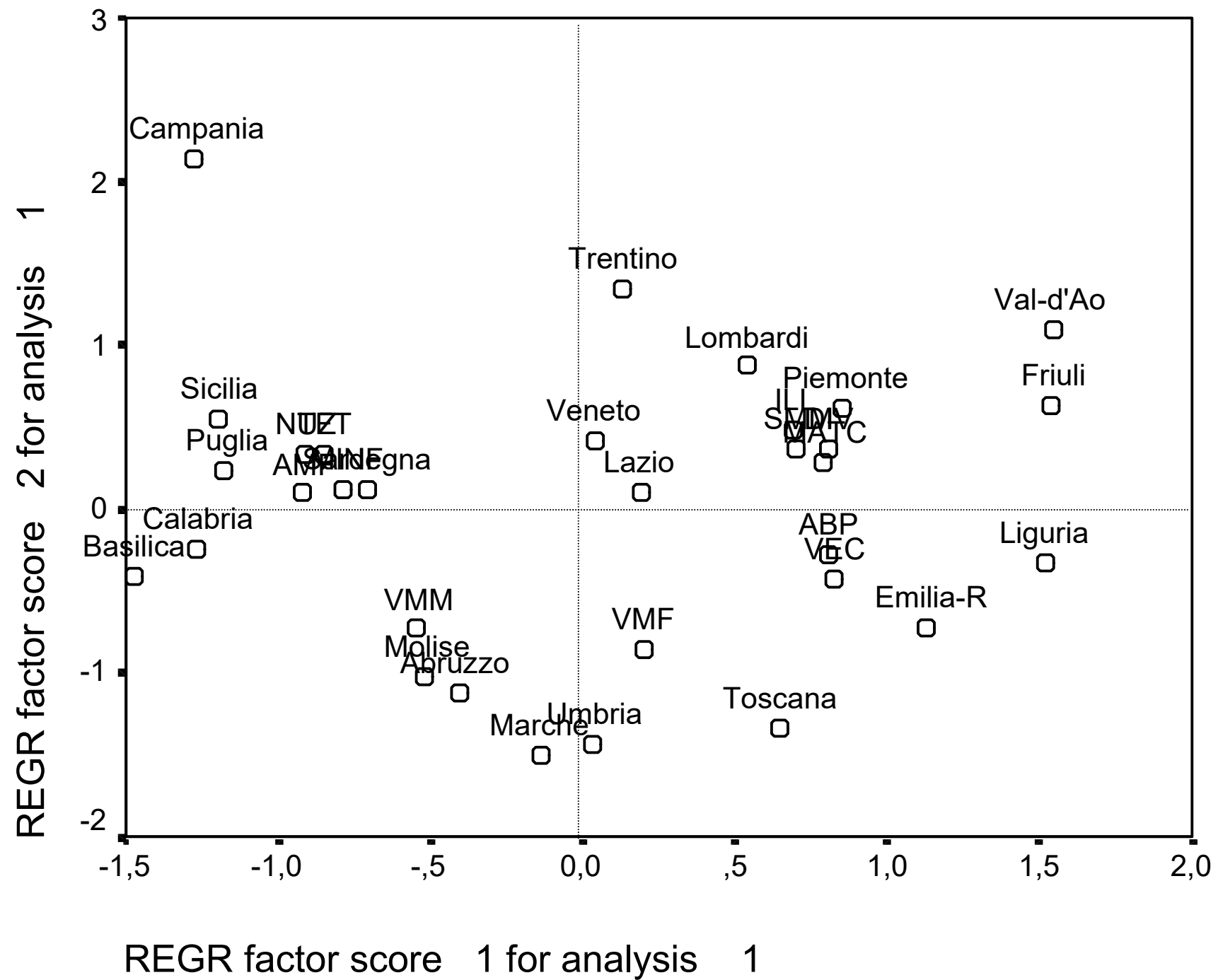
Dal piano fattoriale si osserva che nella zona caratterizzata:

dai **comportamenti tradizionali**, si trovano tutte le regioni meridionali, eccetto Abruzzo e Molise

dai valori più elevati della **vita media maschile e femminile** si trovano alcune regioni del centro e Abruzzo e Molise

dall'**abortività e invecchiamento** si trovano Liguria e Emilia-Romagna

dai **comportamenti laici** e da forte **supermortalità maschile** si trovano tutte le regioni del Nord



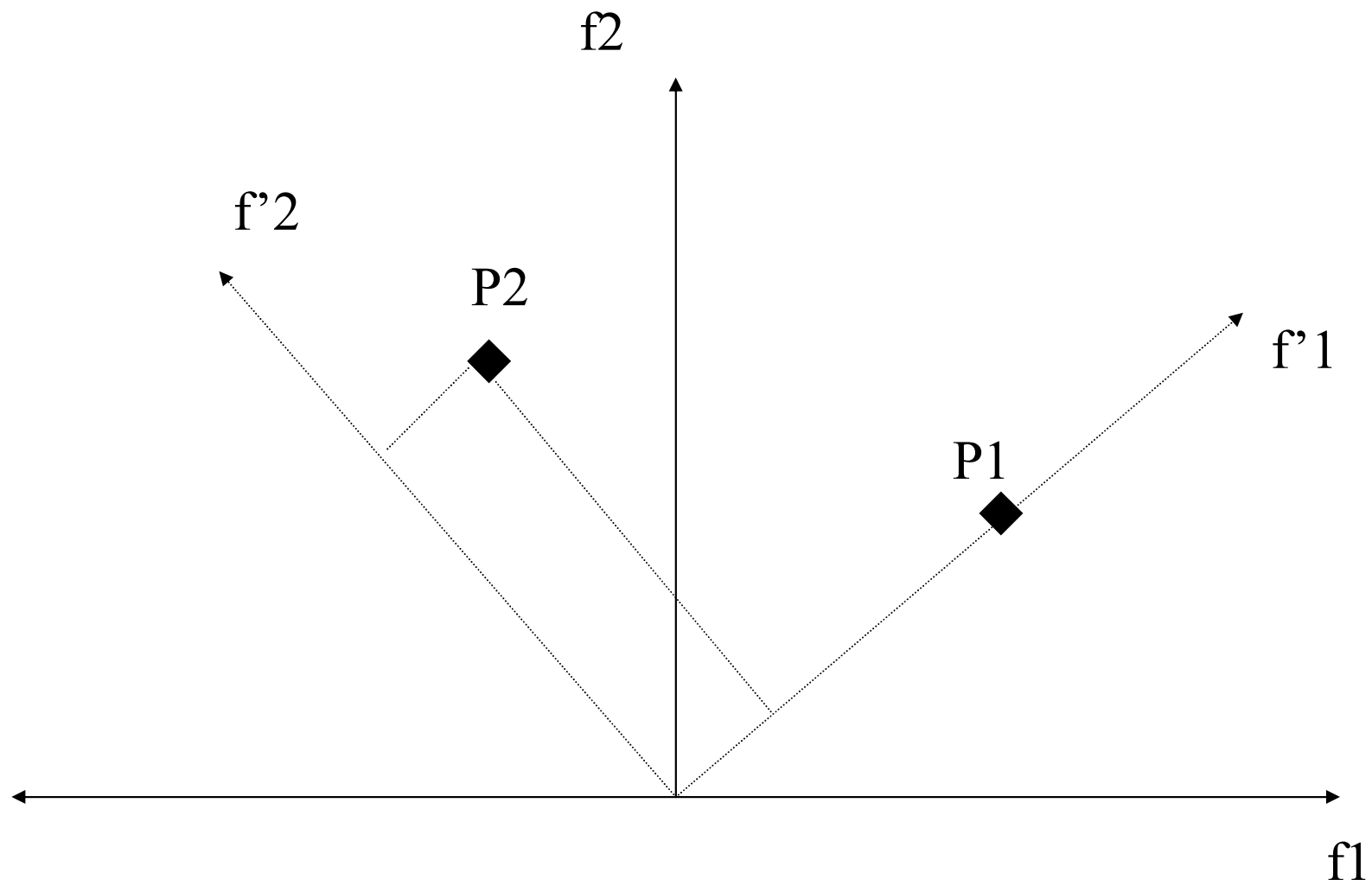
ACP - Rotazioni

I fattori possono essere **ortogonali**, ossia incorrelati oppure **obliqui**, ossia correlati.

L'ACP da fattori ortogonali.

La rotazione degli assi è un cambiamento di posizione dei fattori estratti nella prima fase dell'analisi, mantenendo fissa l'origine.

La soluzione ideale è quella in cui tutti i pesi fattoriali sono prossimi a 0 o a 1.



Esempio rotazione

Rotazioni ortogonali:

Varimax: si cerca di massimizzare la varianza dei pesi fattoriali
E' efficace quando:

- si vuole ottenere una netta separazione tra i fattori
- se la rotazione è effettuata alla “cieca”, senza precisi criteri

Quartimax: tende a far corrispondere una variabile ad uno o pochi fattori.

E' efficace quando:

- si vuole identificare i fattori che governano la variabilità dei caratteri osservati
- da risultati migliori della varimax quando si vuole estrarre un solo fattore generale

Equamax: è un compromesso tra le due precedenti; non è efficace nella ricerca di strutture semplici

Rotazione obliqua

Sono procedimenti iterativi di variazione dell'angolo tra le coppie di assi dopo una rotazione ortogonale

Promax: parte da una rotazione ortogonale Varimax dei pesi originari. Poi cerca una trasformazione dei pesi ruotati che incrementi i pesi già grandi in assoluto e riduca quelli più piccoli, con un procedimento iterativo di aggiustamento.