L'ANALISI FATTORIALE ESPLORATIVA (EFA)

Sommario

* Concetti fondamentali
* Equazioni fondamentali
* Metodi di estrazione dei fattori
* Metodi per stabilire il numero di fattori
* Metodi di rotazione dei fattori
* Assunzioni statistiche e prerequisiti

Analisi Fattoriale Esplorativa

Scopo dell'Analisi Fattoriale è quello di studiare le relazioni in un insieme di variabili per:

a) individuare "*dimensioni latenti*" che spieghino le relazioni tra le variabili

questo solitamente porta a...

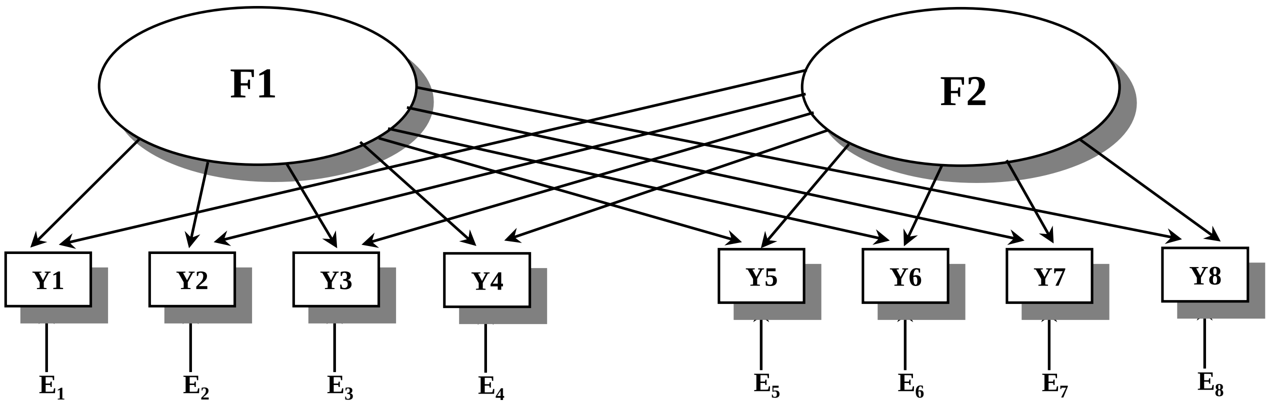
b) ridurre l'informazione in un insieme di dati

Da dove si parte.... Dati non strutturati

1.00							
.36	1.00						
.25	.37	1.00					
.33	.43	.41	1.00				
.05	.16	.12	.12	1.00			
.04	.05	.16	.06	.31	1.00		
.08	.06	.12	.14	.31	.24	1.00	
.02	.10	.17	.04	.29	.34	.29	1.00

Non viene formulata nessuna ipotesi su cosa genera le correlazioni tra le variabili. Si osserva semplicemente che alcune variabili sono più correlate tra loro di altre.

Dove si arriva.... Dati strutturati



Le relazioni tra le variabili osservate sono ricondotte alla presenza di fattori latenti.

E' un'ipotesi teorica sottoponibile a verifica empirica

Dove si arriva.... Dati strutturati

	F1	F2
Y1	0.516	-0.061
Y2	0.659	-0.010
Y3	0.539	0.119
Y4	0.685	-0.027
Y5	0.047	0.531
Y6	-0.039	0.570
Y7	0.033	0.481
Y8	-0.034	0.594

Le relazioni tra variabili osservate e fattori sono le saturazioni fattoriali.

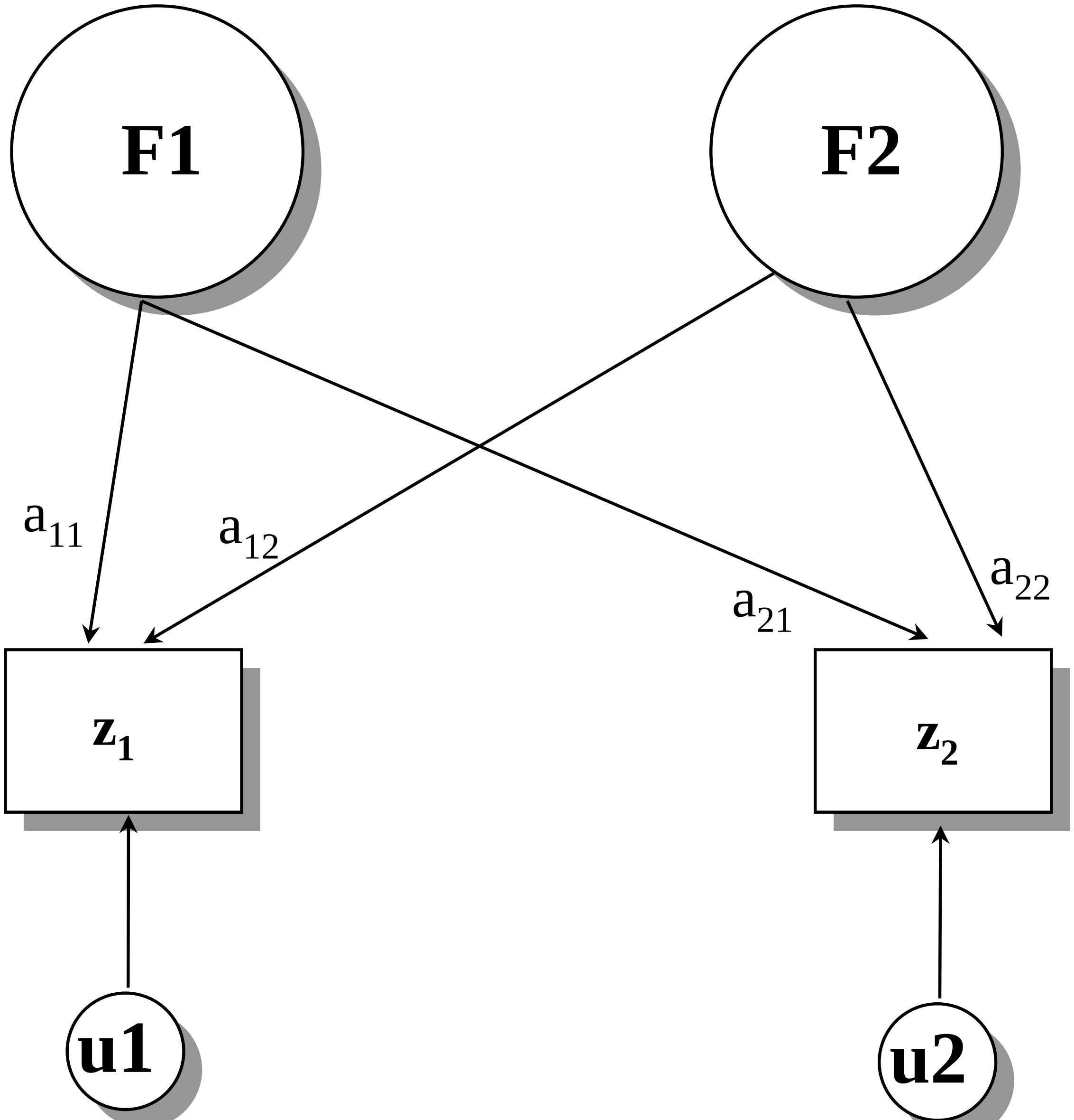
Modello teorico dell’analisi fattoriale

Esame della varianza che le variabili hanno in comune, ovvero della varianza comune.

Ipotesi di base:

La correlazione tra le variabili è determinata da dimensioni non osservabili (i fattori) che “causano” le variabili osservate.

Modello dell’analisi Fattoriale – Rappresentazione Grafica



Modello dell’analisi Fattoriale

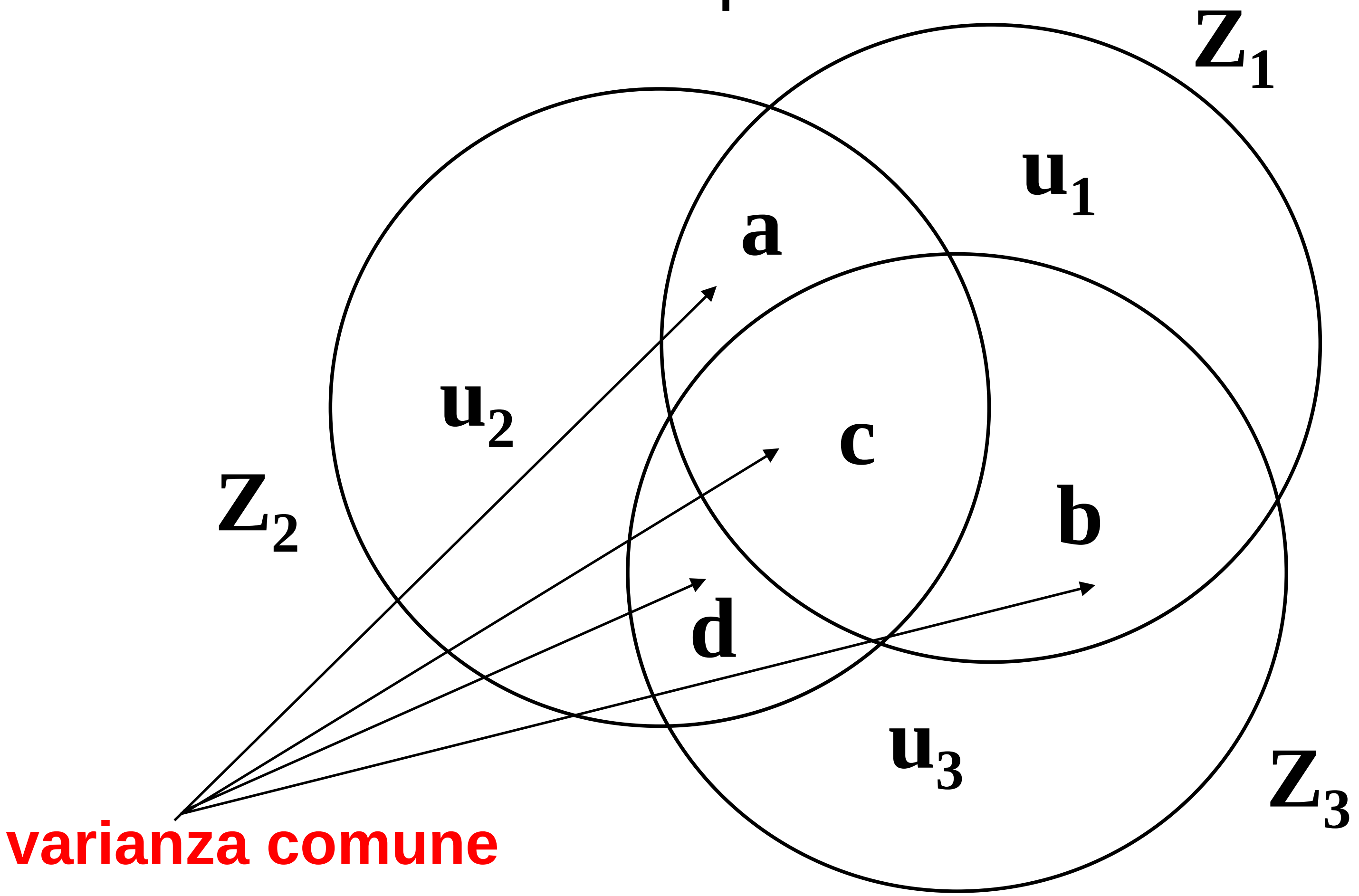
- F = fattori comuni; rappresentano la variabilità condivisa tra le variabili in analisi. Possono influenzare più di una variabile osservata.

- a = saturazioni; relazioni tra variabili e fattori.

- u = termine unico o “unicità della variabile”. Parte di varianza <u>non</u> condivisa. Dovuta a cause sistematiche specifiche, o all’errore casuale di misurazione.

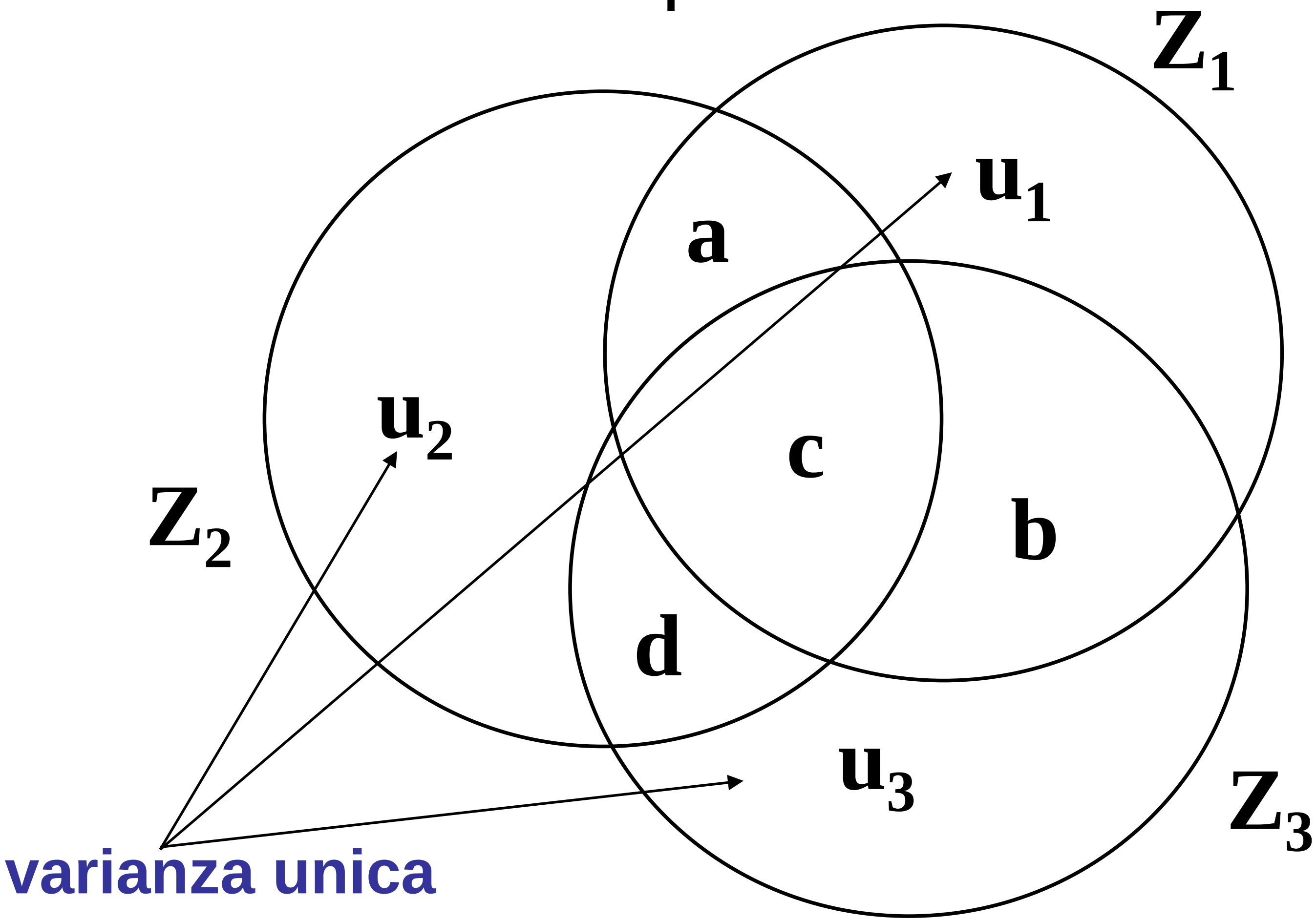
Rappresentazione grafica della varianza comune

Parte di varianza comune delle 3 var: area (a+b+c+d).
Varianza unica: aree u_i.



Rappresentazione grafica della varianza comune

Parte di varianza comune delle 3 var: area (a+b+c+d).
Varianza unica: aree u_i.



Modello dell’analisi fattoriale

Punteggio (standardizzato) di un soggetto in una variabile = somma "ponderata" del punteggio ottenuto dallo stesso soggetto:

a) nei fattori comuni;

b) in una componente unica.

a) e b) sono individuati tramite l'analisi fattoriale.

Equazione del modello teorico dell'analisi fattoriale

$$z_{ik} = F_{i1}\, a_{k1} + F_{i2}\, a_{k2} + \ldots + F_{im} a_{km} + u_{ik} \qquad (1)$$

z_{ik} = punteggio standardizzato per la persona i nella variabile k

F_{ij} = punteggi standardizzati per la persona i nei fattori comuni j

a_{kj} = saturazioni fattoriali della variabile k nei fattori comuni j

u_{ik} = punteggio standard. per la persona i nella componente unica associata alla variabile k

Espressione matriciale dell'equazione (1):

$$Z = FA' + U$$

Z: matrice dei punteggi standardizzati nelle variabili,

F: matrice dei punteggi nei fattori comuni,

A: matrice delle saturazioni delle variabili nei fattori,

U: matrice delle componenti uniche delle variabili.

Scomposizione della varianza di ogni variabile:

Varianza totale= $1 = h^2 + u^2$

Comunalità = h^2.
Parte di varianza totale spiegata dai fattori comuni

Unicità o **varianza unica** = $u^2 = 1 - h^2$.
Parte di varianza totale non spiegata dai fattore comuni

Assunzioni:

$Cov(u_i, F_j)=0$, per ogni i e per ogni j
$Cov(u_i, u_j)=0$ per ogni i diversa da j
$Cov(F_i, F_j)$ diversa da 0 solo nelle soluzioni "oblique"

Espressioni matriciali

In base alle assunzioni e considerando che:

$R = Z'Zn^{-1}$ **e che** $Z = FA' + U$

si ha che:

$$R = AA' + U^2$$

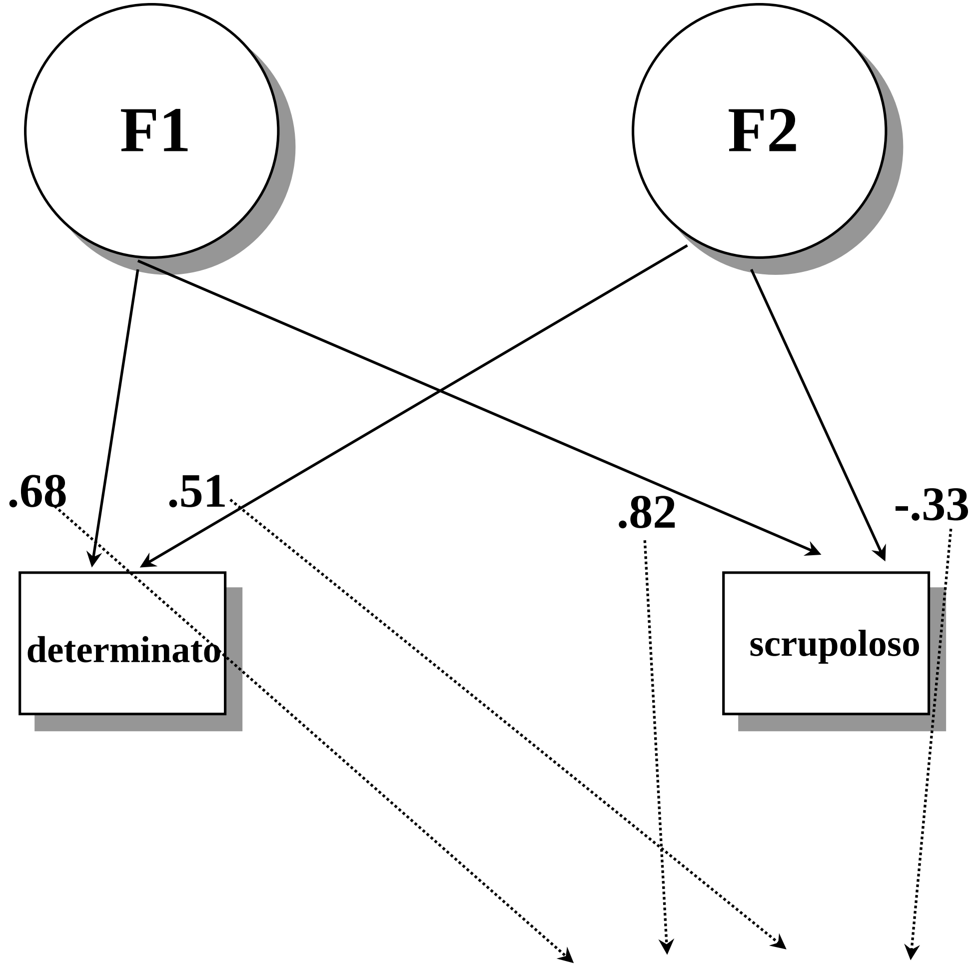
A = matrice delle saturazioni nei fattori comuni
U^2 = matrice diagonale delle varianze uniche.
AA': rende conto degli elementi fuori della diagonale principale, e della comunalità di ogni variabile.
U^2: contribuisce a rendere conto degli elementi sulla diagonale principale di R.

La correlazione tra due variabili i e j può essere riprodotta dalla somma dei prodotti delle loro saturazioni in ciascuno dei fattori comuni:

$$r_{ij} = a_{i1}a_{j1} + a_{i2}a_{j2} + .. + a_{im}a_{jm} = \Sigma\, a_{ir}a_{jr}, \text{ se } i \neq j$$

$$r_{ii} = a_{i1}a_{i1} + a_{i2}a_{i2} + .. + a_{im}a_{im} + u_i^2 = \Sigma\, a_{ir}a_{ir} + u_i^2, \text{ se } i = j$$

Rappresentiamolo con un diagramma….



r^ (determinato, scrupoloso) = (.68*.82)+(.51*-.33)= .56 - .17 = .39

Correlazione residua

Differenza tra la correlazione osservata e la correlazione riprodotta tramite le saturazioni.

r(determinato, scrupoloso) = .40

r^(determinato, scrupoloso) = .39

r residua (e) = .40-.39 = .01

Interpretazione dei fattori e grandezza delle saturazioni

I fattori si interpretano in base alle variabili con le quali presentano correlazioni (saturazioni) più elevate.
Regola pratica: livello soglia di circa |.30| (circa 9% di varianza in comune tra fattore e variabile, Comrey e Lee, 1992; Tabachnick e Fidell, 2007).

a) |.71| (50% varianza comune): eccellente;
b) |.63| (40% varianza comune): molto buona;
c) |.55| (30% varianza comune): buona;
d) |.45| (20% varianza comune): sufficiente;
e) |.32| (10% varianza comune): scarsa.

Saturazioni sotto |.30| inadeguate.

$$R=AA'+U^2$$

"Equazione fondamentale dell'analisi fattoriale" (Thurstone, 1947).

Mette in relazione il punto di partenza dell'Analisi fattoriale con il suo punto di arrivo.

- Per riprodurre le correlazioni tra le variabili che stanno fuori la diagonale principale sono necessari solo i fattori comuni.

- Per "riprodurre" anche gli elementi sulla diagonale principale (varianza totale delle variabili) sono necessarie anche le unicità.

"<u>Equazione fondamentale dell'analisi fattoriale</u>" Equazione che definisce la Struttura di R (Thurstone, 1947).

$$R^* = AA' \qquad (1)$$

$$R = AA' + U^2 \qquad (2)$$

R*: matrice delle correlazioni che contiene le comunalità sulla diagonale principale.

Come ricavare A, la matrice delle saturazioni nei fattori comuni, in maniera tale che il numero di fattori comuni sia strettamente minore del numero di variabili osservate.

Una soluzione di questo problema è rappresentata dal calcolo delle componenti principali (vedi oltre).

Calcolo di alcuni <u>elementi</u> che caratterizzano la matrice di correlazione

- radici caratteristiche di R (autovalori, L)
- vettori ad essi associati (autovettori, V)

Autovalori e autovettori di una matrice

Gli autovalori sono scalari di enorme importanza nell’analisi multivariata (es., nell’analisi fattoriale).

Per identificare gli autovalori di A è necessario effettuare alcuni calcoli sulla matrice, per i quali si rimanda al libro di testo.

In una matrice quadrata ci sono tanti autovalori quante sono le righe (ovvero le colonne) della matrice.

Ogni autovettore relativo ad un autovalore è un vettore che ha una colonna e tante righe quante quelle della matrice

Autovalori e autovettori di una matrice

Esempio: data la matrice seguente:

$$\mathbf{A} = \begin{bmatrix} 1 & .50 \\ .50 & 1 \end{bmatrix}$$

Gli autovalori sono: λ_1 = 1.5, e λ_2 =.5.

Gli autovettori x_1 relativo a λ_1 e x_2 relativo a λ_2 sono:

$$\mathbf{x}_1 = \begin{bmatrix} .707 \\ .707 \end{bmatrix}; \mathbf{x}_2 = \begin{bmatrix} .707 \\ -.707 \end{bmatrix}$$

Autovalori e autovettori di R

Elementi che sintetizzano l'informazione relativa alla varianza delle variabili, e alla correlazione tra le variabili.

Il calcolo di questi elementi è un passo preliminare per il calcolo delle soluzioni di analisi fattoriale.

Ogni autovalore è associato ad un autovettore.

Scomposizione della matrice di correlazione R
Se si considerano V la matrice degli autovettori e L la matrice degli autovalori, allora è possibile dimostrare che $R = VLV'$.

Una volta calcolate le matrici V e L, è possibile ricavare da V e da L la matrice A. In particolare:

$$A = V\sqrt{L}$$

E' possibile dimostrare che: $R = V\sqrt{L}(V\sqrt{L})' = AA'$.

Scomposizione della matrice di correlazione R

R= VLV'

$$
\begin{vmatrix} 1 & .45 & .30 \\ .45 & 1 & .50 \\ .30 & .50 & 1 \end{vmatrix}
=
\begin{vmatrix} -.54 & .76 & .36 \\ -.63 & -.08 & -.78 \\ -.57 & -.64 & .52 \end{vmatrix}
*
\begin{vmatrix} 1.84 & 0 & 0 \\ 0 & .70 & 0 \\ 0 & 0 & .46 \end{vmatrix}
*
\begin{vmatrix} -.54 & -.63 & -.57 \\ .76 & -.08 & -.64 \\ .30 & -.78 & .52 \end{vmatrix}
$$

R = V L V'

Primo autovettore e primo autovalore di R

Secondo autovettore e secondo autovalore di R

Secondo autovettore e secondo autovalore di R

Una volta calcolate le matrici V e L, è possibile ricavare da V e da L la matrice A delle saturazioni fattoriali.
In particolare, A = V√L

Nell'esempio precedente la matrice A è data da:

$$\begin{bmatrix} -.73 & .64 & .24 \\ -.85 & -.07 & -.53 \\ -.77 & -.54 & .35 \end{bmatrix} = \begin{bmatrix} -.54 & .76 & .36 \\ -.63 & -.08 & -.78 \\ -.57 & -.64 & .52 \end{bmatrix} * \begin{bmatrix} 1.36 & 0 & 0 \\ 0 & .84 & 0 \\ 0 & 0 & .68 \end{bmatrix}$$

A **V** **√L**

Primo autovalore di R: quello più elevato di tutti, associato al primo fattore che spiega una proporzione di varianza maggiore degli altri.

Secondo autovalore: quello più elevato dopo il primo, associato al secondo fattore.

La grandezza degli autovalori rappresenta una progressione decrescente che corrisponde alla progressione della varianza spiegata dai fattori associati ad essi.

Nella matrice precedente:
1.84, .70, .46

Autovalori e varianza spiegata

- **Somma delle saturazioni elevate al quadrato per ogni fattore (colonna) = autovalore associato al fattore;**

- **Autovalore diviso per il numero di variabili osservate in analisi = proporzione di varianza spiegata dal fattore;**

- **Somma delle saturazioni al quadrato per ogni variabile (riga) = comunalità delle variabili.**

Autovalori e varianza spiegata

	F1	F2	h^2
Determinato	.68	.51	.72
Dinamico	.74	.48	.78
Energico	.78	.33	.72
Affidabile	.80	-.41	.81
Responsabile	.84	-.43	.89
Scrupoloso	.82	-.33	.78
Autovalori	3.66	1.08	
Proporzione di Varianza Spiegata	.61	.18	

Metodi di Estrazione dei Fattori

Metodi che cercano di rendere conto di R tramite fattori che ne spiegano il massimo di varianza.

Metodi che cercano invece di rendere conto di R massimizzandone la "riproduzione".

Metodi che richiedono una stima iniziale delle comunalità.

Metodi che utilizzano solo gli elementi al di fuori della diagonale principale e richiedono una stima del numero di fattori da estrarre.

Metodi di Estrazione dei Fattori

Nelle analisi che utilizzano stime delle comunalità viene analizzata la matrice di correlazione R*, con la stima delle comunalità ($\hat{h}_j^2$) sulla diagonale principale.

Nelle analisi che non utilizzano stime delle comunalità viene analizzata la matrice R_1 nella quale non si considerano gli elementi fuori della diagonale principale.

Analisi delle Componenti Principali (ACP)

- Identifica una serie di combinazioni lineari ortogonali delle variabili originali X_i ($c_i = X_i V$, con V = autovettori di R) tali che spieghino più varianza possibile delle variabili originali X_i, e che riducano la complessità dei dati iniziali.

- L'ACP analizza la varianza totale delle variabili (analizza R con valori 1 sulla diagonale principale). La varianza unica è assorbita dai fattori comuni. Nella soluzione ci sono solo "fattori comuni" (le componenti principali).

- Le saturazioni si basano sul calcolo diretto degli autovalori e degli autovettori di R: $A = V\sqrt{L}$

Analisi delle Componenti Principali (ACP)

- Le saturazioni fattoriali risultano gonfiate dalla presenza di varianza comune e varianza unica.

- L'ACP estrae il massimo della varianza per ogni componente, cioè massimizza la varianza spiegata ad ogni estrazione.

- La prima componente è la combinazione lineare dei dati originali che spiega più varianza, la seconda è quella che spiega più varianza dopo la prima, ecc.

- Le componenti principali sono semplici trasformazioni lineari delle variabili originali che forniscono un sommario empirico dei dati. La matrice R è perfettamente replicata se vengono estratte tante componenti quante sono le variabili.

Analisi dei Fattori Principali (AFP o PAF)

- Massimizza lo stesso criterio della ACP, ma con stime della comunalità inserite nella diagonale principale.

- Analizza solo la varianza attribuibile ai fattori “comuni” (ovvero la comunalità) per ottenere una soluzione non contaminata dalla varianza unica.

- Estrae il massimo di varianza per ogni fattore, ma considera solo la varianza dovuta ai fattori comuni, quindi spiega meno varianza della ACP.

Analisi dei Fattori Principali (AFP o PAF)

- Primo passo: rimuovere dalla diagonale principale di R la varianza unica (cioè, $u^2 = 1-h^2$).

- Stima iniziale delle comunalità delle variabili:
* coefficiente di correlazione multipla al quadrato (SMC)
* correlazione più elevata
* media delle correlazioni

- Le saturazioni si basano sul calcolo diretto degli autovalori e degli autovettori di R_1: $A = V\sqrt{L}$, dove R_1 è la matrice delle correlazioni con le stime delle comunalità sulla diagonale principale

Analisi dei Fattori Principali (AFP o PAF)

	fp10	fp15	fp26	fp30
fp10	1,000	,368	,256	,344
fp15	,368	1,000	,390	,444
fp26	,256	,390	1,000	,418
fp30	,344	,444	,418	1,000

Matrice di correlazione originale

fp10	,180
fp15	,288
fp26	,231
fp30	,294

Stima delle Comunalità

	fp10	fp15	fp26	fp30
fp10	**,180**	,368	,256	,344
fp15	,368	**,288**	,390	,444
fp26	,256	,390	**,231**	,418
fp30	,344	,444	,418	**,294**

Matrice di correlazione analizzata da PAF

Analisi dei Fattori Principali (AFP o PAF)
Calcolo iterativo delle comunalità

Nell'AFP le stime delle comunalità rappresentano soltanto un *valore iniziale* che viene cambiato e ricalcolato nel corso dell'estrazione dei fattori.

Le stime iniziali servono per estrarre gli autovalori e gli autovettori di R_1 e quindi per individuare la matrice A delle saturazioni.

Questa soluzione iniziale consente di calcolare empiricamente le comunalità: le comunalità empiriche vengono quindi sostituite alle stime iniziali e il processo di estrazione dei fattori viene ripetuto, dando origine a nuove saturazioni e quindi a nuove comunalità. Il processo si interrompe quando i valori empirici delle comunalità diventano stabili.

Analisi dei Fattori Principali (AFP o PAF)
Calcolo iterativo delle comunalità

Durante il processo di iterazione delle comunalità, il numero di fattori deve rimanere costante.

Una volta che la soluzione fattoriale si è stabilizzata, i valori finali delle comunalità possono essere ricavati dalla soluzione stessa, elevando al quadrato le saturazioni di ogni variabile in ogni fattore comune e sommando tali quadrati.

Ci sono dei problemi che possono presentarsi nel processo iterativo di calcolo delle comunalità. In alcuni casi si può assistere a valori di comunalità che eccedono 1 (tale problema viene definita spesso "*Heywood case*").

Minimi Quadrati (ULS e GLS) (Minimum residuals /Minres)

Minimizza le differenze al quadrato tra gli elementi della matrice di correlazione osservata (R), e quella riprodotta (R^) utilizzando i fattori estratti.

Minimizza le correlazioni residue (R- R^) cioè la parte di correlazione tra le variabili non spiegata dai fattori.

Funzione dei minimi quadrati ordinari (ULS) minimizzata nel processo di estrazione dei fattori:

$$\Sigma_j \Sigma_k (r_{jk} - r_{jk}\hat{})^2$$

Minimi Quadrati (ULS e GLS) (Minimum residuals /Minres)

Massimizza la riproduzione dei coefficienti <u>fuori</u> della diagonale principale di R.

Si inizia il processo stabilendo il numero di fattori.

Si stimano le saturazioni iniziali con l'ACP.

Le saturazioni vengono modificate iterativamente finché lo scarto tra R e R^ non è molto piccolo.

Minimi quadrati generalizzati (GLS): si introduce un fattore di ponderazione, per cui le variabili con fattore unico più elevato hanno peso minore.

Maximum Likelihood (Massima verosimiglianza)

Calcola le saturazioni che rendono massima la probabilità di riprodurre la matrice R, ovvero identifica la soluzione che meglio riproduce R.

Stima le saturazioni della popolazione che hanno la massima verosimiglianza (ovvero la massima probabilità) nel riprodurre R, quindi che rendono minima la differenza tra matrice osservata e riprodotta.

Si considerano gli elementi fuori della diagonale principale, si fornisce il numero di fattori da estrarre.

Maximum Likelihood (Massima verosimiglianza)

La stima delle saturazioni avviene attraverso la minimizzazione di una funzione (F_{ML}) delle matrici delle correlazioni osservate R e riprodotte R^.

$$F_{ML} = tr(RC^{-1}) + \ln|C| - \ln|R| - n$$

dove C =AA'+U², è la matrice riprodotta dalla soluzione (A matrice delle saturazioni, U^2 stima della varianza unica), *n* è il numero di variabili, | | indica il determinante e tr() la traccia della matrice.

Per calcolare la funzione è necessario che:

$$(n-k)^2 > (n + k)$$

dove n = numero di variabili, k = numero di fattori.

NUMERO MASSIMO DI FATTORI

Il numero massimo di fattori che è possibile estrarre dipende dai gradi di libertà che sono determinati dal numero di parametri da stimare e dal numero di correlazioni non ridondanti.

Gradi di libertà = $[(n-k)^2 - (n+k)]/2$,

Deve valere sempre: $(n-k)^2 > (n+k)$

n = numero di variabili; k = il numero di fattori.

Esempio: n = 8

Se k = 1, $(8-1)^2-(8+1)= 49-9=40$, gdl=20

Se k = 2, $(8-2)^2-(8+2)= 36-10=26$, gdl=13

Se k = 3, $(8-3)^2-(8+3)= 25-11=14$, gdl=7

Se k = 4, $(8-4)^2-(8+4)= 16-12=4$, gdl=2

Se k = 5, $(8-5)^2-(8+5)= 9-13=-4$, gdl=-2

Con 8 variabili osservate si possono estrarre al massimo 4 fattori.

Test di bontà dell'adattamento (goodness of fit)

Si ottiene dalle funzioni ML e GLS che vengono minimizzate se le variabili seguono la distribuzione normale multivariata.

Ipotesi nulla: R = R^

Il test segue la distribuzione del χ^2

Gradi di libertà del test: $df=[(n-k)^2 - (n+k)]/2$

χ^2 non significativo: il modello che ipotizza k fattori è consistente con i dati empirici, non si può rifiutare l'ipotesi nulla H_0:R=R^, quindi non vi sono più fattori da estrarre

χ^2 significativo: il modello che ipotizza k fattori è consistente con i dati empirici, quindi è necessario procedere all'estrazione di fattori ulteriori.

Test fortemente dipendente dal numero di casi.

Stabilire il numero dei fattori da estrarre

Decisione che ha conseguenze cruciali sulla soluzione fattoriale.

Salvaguardare la parsimonia della soluzione, e la sua adeguatezza (capacità di riprodurre R).

Metodi per stabilire il numero di fattori

- Mineigen (Kaiser-Guttman rule)
- Scree test degli autovalori (Cattell e Vogelman)
- Test statistico, indici di bontà dell’adattamento
- Percentuale di varianza spiegata
- Massima correlazione residua

Mineigen (Kaiser-Guttman rule)

Estrae tutti quei fattori che hanno un autovalore maggiore di 1 quando viene analizzata la matrice R completa (con 1 sulla diagonale principale).

I fattori devono spiegare almeno la stessa varianza spiegata dalle variabili osservate.

Il numero di autovalori maggiori di 1 è uguale approssimativamente ad un numero compreso tra 1/3 e 1/5 del numero delle variabili.

Criterio inappropriato per soluzioni diverse dall'ACP

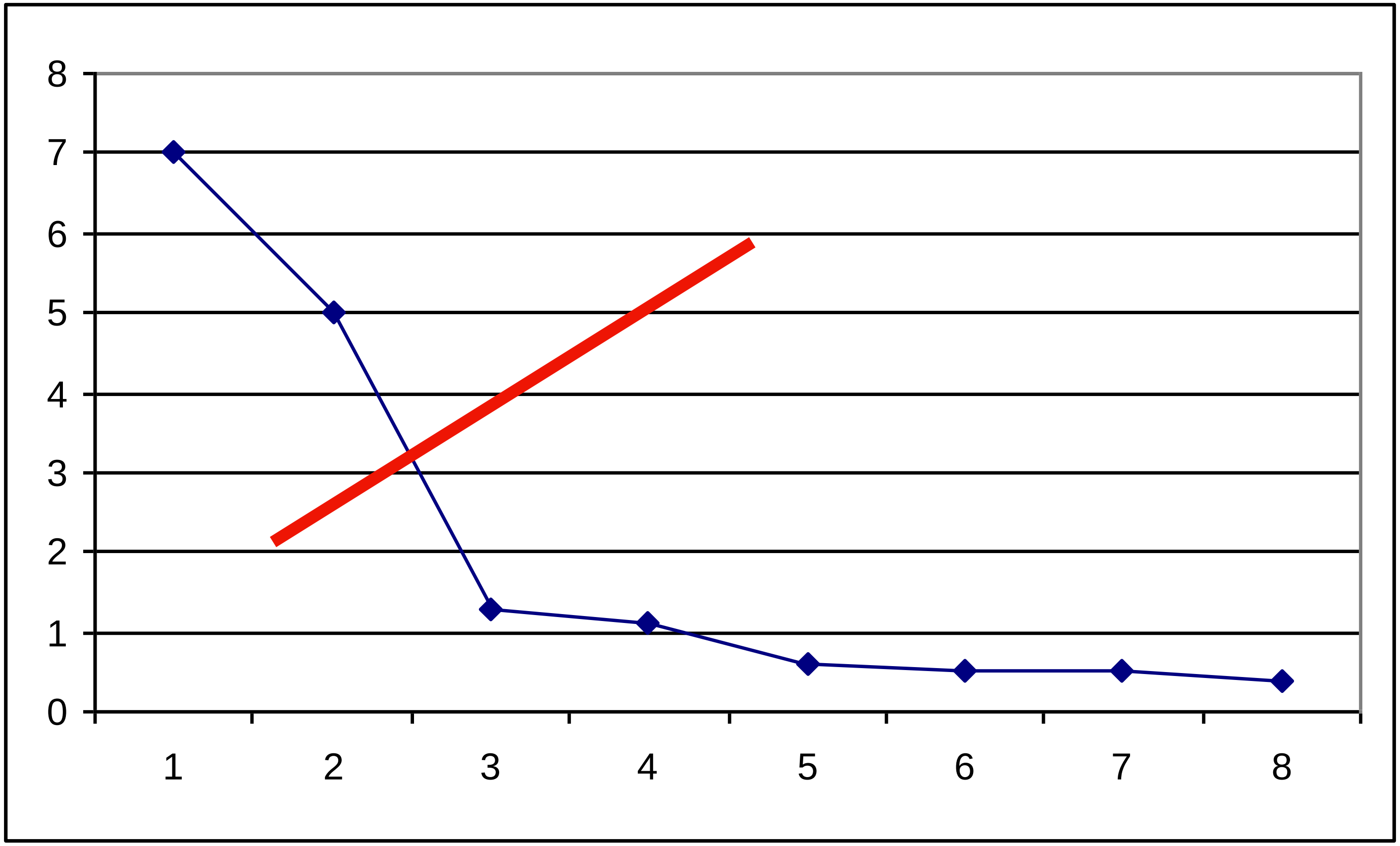
Scree test degli autovalori (Cattell e Vogelman)

I primi fattori sono i più attendibili e i più validi, poiché spiegano una percentuale di varianza maggiore rispetto ai rimanenti, e avranno autovalori più grandi degli altri.

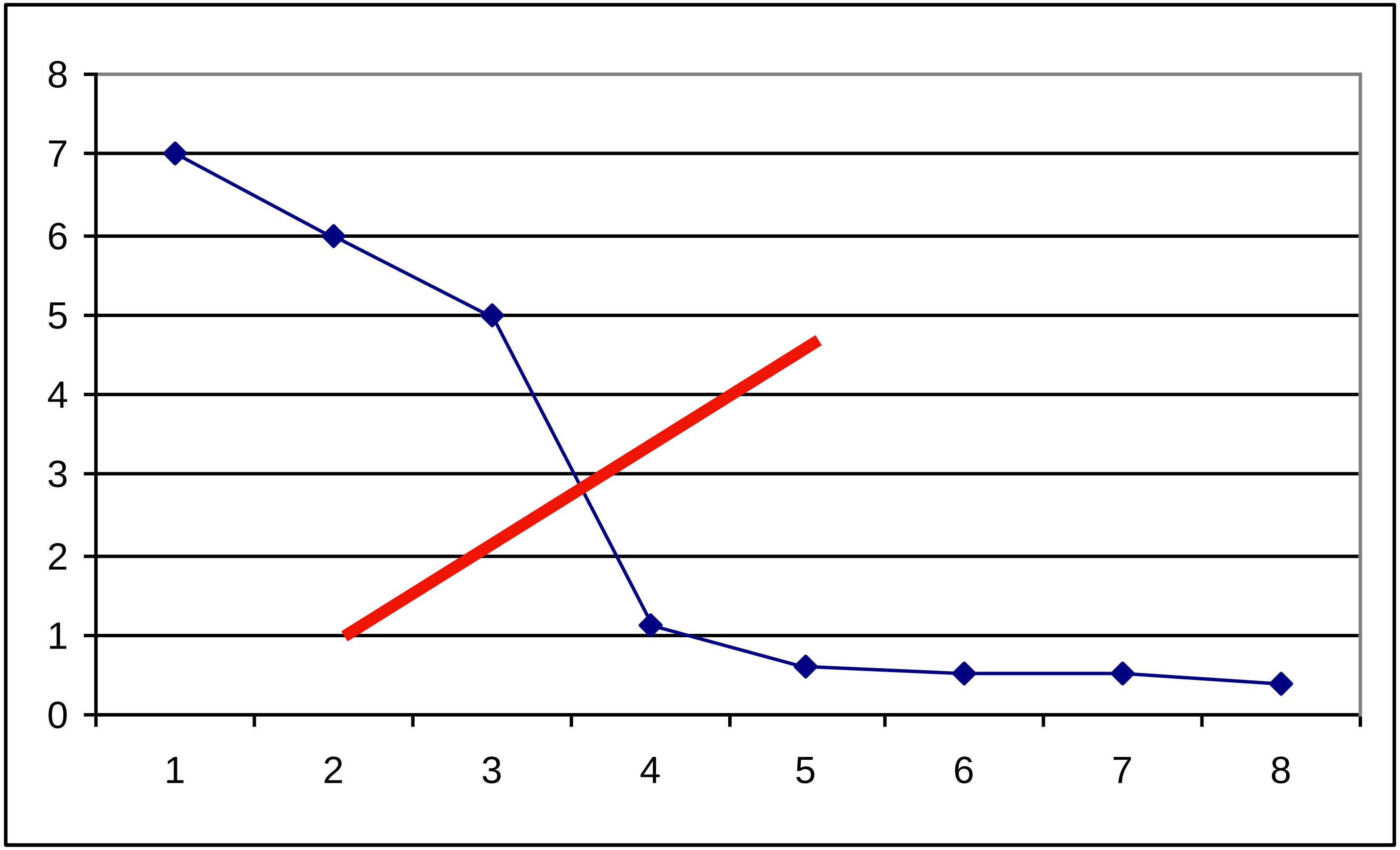
Progressione decrescente degli autovalori: grafico in cui ogni autovalore è in ordinata, e il numero del fattore ad esso relativo in ascissa.

Il processo di estrazione si interrompe nel punto in cui la curva degli autovalori decrescenti cambia pendenza e diventa sostanzialmente piatta. Vanno presi quei fattori i cui autovalori sono al di sopra della linea piatta formata dagli autovalori dei fattori più piccoli.

Applicazione più attendibile quando il campione è grande, le comunalità elevate, e ogni fattore satura diverse variabili.



2 fattori



3 fattori

Analisi parallela

E' un procedimento che si basa sul confronto tra gli autovalori della matrice di correlazione campionaria e gli autovalori ottenuti da una matrice calcolata su un set di dati casuali generati artificialmente.

I risultati di questi studi suggeriscono di mantenere nella soluzione quei fattori associati ad un autovalore superiore a quello associato ad un fattore omologo estratto nei dati artificiali.

Analisi parallela

Supponiamo che i primi 5 autovalori dei dati reali siano: 5.72 1.51 1.03 0.50 0.40, e quelli ricavati dai dati "artificiali" siano: 1.64 1.45 1.31 1.19 1.09.

Verranno mantenuti quei fattori reali che presentano un autovalore maggiore di quello del corrispondente fattore dei dati artificiali, quindi nel nostro caso verranno mantenuti 2 fattori perché solo per i primi due fattori gli autovalori nei dati reali sono maggiori degli autovalori associati ai corrispondenti fattori nei dati artificiali [5.72>1.64;1.51>1.45;1.03<1.31;0.50< 1.19; 0.40< 1.09]

Sintassi SPSS per l'analisi parallela:
https://people.ok.ubc.ca/brioconn/nfactors/nfactors.html

Test statistico e indici di bontà dell'adattamento

Il test statistico associato ai metodi di estrazione ML e GLS (chi-quadrato) da un punto di vista puramente statistico, è il migliore. Da un punto di vista pratico, però, questo test tende ad essere fortemente dipendente dall'ampiezza del campione.

Gli indici alternativi di bontà dell'adattamento (che introdurremo quando affronteremo i modelli confermativi) possono spesso dare risultati più verosimili: tra questi indici l'SRMR e l'RMSEA sembrano i più affidabili.
Come regola pratica il ricercatore dovrebbe considerare più indici alternativi per ciascuna soluzione, e privilegiare le soluzioni nelle quali i diversi indici mostrano maggiore convergenza.

Percentuale di varianza spiegata

Contributo minimo di un fattore alla spiegazione della varianza, oppure proporzione di varianza spiegata dall'ultimo fattore. Metodo troppo soggettivo.

Replicabilità della soluzione

I fattori "validi" sono quelli che risultano più facilmente replicabili su campioni diversi da quelli nei quali sono stati individuati.

I fattori "spuri" risultano poco generalizzabili e sono determinati sostanzialmente dall'errore campionario.

Massima correlazione residua

Per ogni elemento di R fuori della diagonale principale si può definire un residuo che è uguale a (r- r^), ovvero correlazione osservata meno correlazione riprodotta. La matrice dei residui quindi si ottiene nel modo seguente: E = (R-R^).

Se dopo aver effettuato l'estrazione di un certo numero di fattori tutti i residui sono minori di |.10|, non è necessario continuare il processo di estrazione: il nuovo fattore estratto avrebbe saturazioni molto basse.

Rotazione dei fattori

E' un'operazione che rende la soluzione fattoriale più interpretabile senza cambiarne le fondamentali proprietà matematiche (capacità di riprodurre R, % var. spiegata).

Esistono infinite matrici T che trasformano una matrice di saturazioni non ruotata A in modo che:

$$AT = B, \text{ e } R = BB'$$

T è la matrice di trasformazione (gli elementi sono seni e coseni di un generico angolo di rotazione "ϕ"), B è la matrice ruotata

Rotazioni ortogonali: i fattori ruotati non sono correlati. Rotazioni oblique: i fattori ruotati possono essere correlati tra loro.

Rotazione dei fattori

T è la matrice di <u>trasformazione</u> (gli elementi sono seni e coseni di un generico angolo di rotazione "ϕ").
Nel caso di due fattori T è come la matrice seguente:

$$\mathbf{T} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$$

Nella rotazione dei fattori, AT=B.
R=BB', ma B=AT, quindi R=(AT)(AT)'=ATT'A'
E' possibile dimostrare che TT'=T'T=I, quindi:

R=BB'=AA'

Questo fenomeno viene definito indeterminatezza della soluzione fattoriale: esistono infinite matrici A tali che R=AA', cioè che riproducono una data matrice R altrettanto bene.

La rotazione dei fattori – La struttura semplice (Thurstone, 1947)

Guida il processo di rotazione dei fattori. Si pone l'obiettivo di rendere più interpretabile la soluzione massimizzando il numero di zeri nelle righe e nelle colonne della matrice delle saturazioni.

Ogni fattore deve saturare una minoranza di variabili; ogni variabile deve essere spiegata da pochi fattori (possibilmente uno solo).

Soluzioni fattoriali

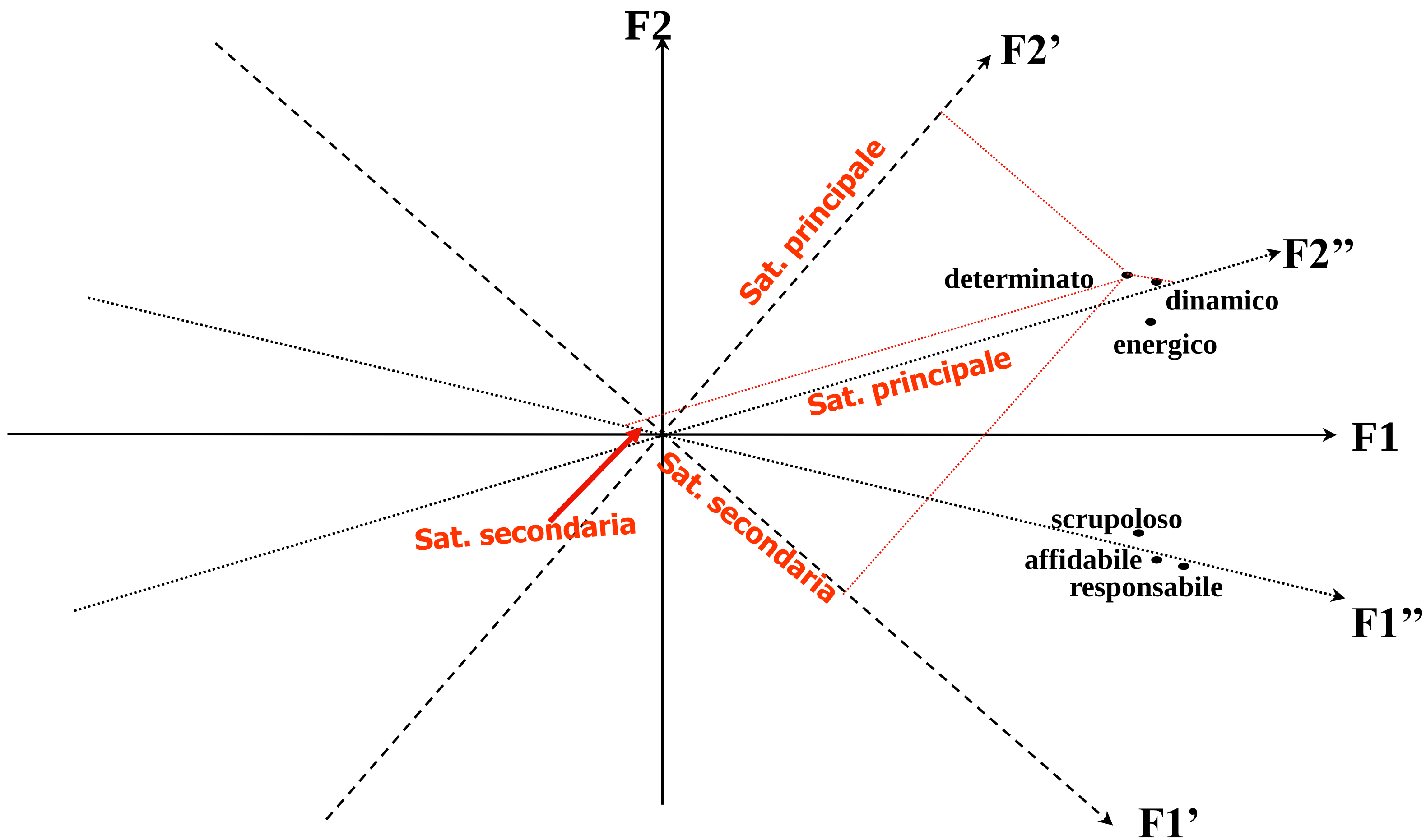
	Soluzione iniziale		
	F1	F2	h^2
Determinato	.68	.51	.72
Dinamico	.74	.48	.78
Energico	.78	.33	.72
Affidabile	.80	-.41	.81
Responsabile	.84	-.43	.89
Scrupoloso	.82	-.33	.78
Varianza Spiegata	61%	18%	

La rotazione dei fattori – La struttura semplice (Thurstone, 1947)

Fa in modo che le variabili cadano il più vicino possibile agli assi fattoriali. Gli spazi "interstiziali" tendono ad essere più vuoti degli spazi vicini agli assi.

Numero di saturazioni prossime a 0 in un fattore: indice della semplicità del fattore.

La struttura fattoriale più semplice possibile è quella in cui le variabili hanno saturazioni uguali a 0 in tutti i fattori tranne che in un unico fattore comune.



Assi fattoriali prima della rotazione
Assi fattoriali dopo la rotazione ortogonale
Assi fattoriali dopo la rotazione obliqua

Rotazioni ortogonali - Varimax

Aumenta la semplicità dei fattori.

Massimizza la varianza delle saturazioni delle variabili all'interno di ogni fattore (nelle colonne di A).

Per ogni fattore, tende a far diventare le saturazioni elevate più elevate e quelle più basse ancora più basse.

La variabilità delle saturazioni è massimizzata, e la varianza redistribuita.

Varimax tende a produrre fattori che presentano alcune saturazioni elevate, poche intermedie e molte basse. Risultati più chiari e più generalizzabili, e fattori diversi separati meglio.

Soluzioni fattoriali

	Soluzione ortogonale		
	F1	F2	h^2
Determinato	.17	.83	.72
Dinamico	.24	.85	.78
Energico	.36	.77	.72
Affidabile	.87	.23	.81
Responsabile	.91	.24	.89
Scrupoloso	.83	.30	.78
Varianza Spiegata	42%	37%	

La rotazione ridistribuisce la varianza spiegata dai singoli fattori, ma la varianza totale rimane identica (79%)

Rotazioni ortogonali - Quartimax

Massimizza la semplicità delle variabili a scapito dei fattori.

Massimizza la varianza delle saturazioni di ogni variabile per riga.

Concentra più varianza possibile per ogni variabile su un solo fattore, creando fattori generale.

Rotazioni oblique

Oblimin: Fa in modo che le variabili abbiano saturazioni il più possibile vicine a 0 in tutti i fattori tranne uno. Massimizza una funzione che comprende anche le covarianze tra i fattori.

Promax: Parte da una rotazione ortogonale, e la modifica per renderla più semplice, consentendo che i fattori siano correlati.

Rotazioni di Procuste: La matrice originale viene ruotata verso una matrice "bersaglio" che ha certe caratteristiche ipotizzate dal ricercatore. La soluzione iniziale viene ruotata in modo da renderla più simile possibile alla matrice bersaglio.

Nelle soluzioni ortogonali:
l'impatto del fattore sulla variabile
è uguale alla correlazione tra variabile e fattore
(saturazione fattoriale).

Nelle soluzioni oblique
è possibile distinguere tra:

- **correlazione tra variabile e fattore**
- **impatto del fattore sulla variabile (contributo unico del fattore al netto degli altri fattori)**

Nelle soluzioni oblique

La variabile osservata può condividere una parte di varianza simultaneamente con più fattori.

La correlazione tra variabile e fattore comprende sia il contributo unico del fattore sia il contributo condiviso con gli altri fattori.

Per questo ci sono due diverse matrici che riassumono le relazioni tra variabili e fattori

Matrice Pattern (P)

Impatto diretto di ciascun fattore sulle variabili, al netto dell'impatto degli altri fattori.

Influenza *unica* di ciascun fattore sulle variabili (pesi beta, β).

Matrice Struttura (S)

Correlazioni tra le variabili e i fattori.

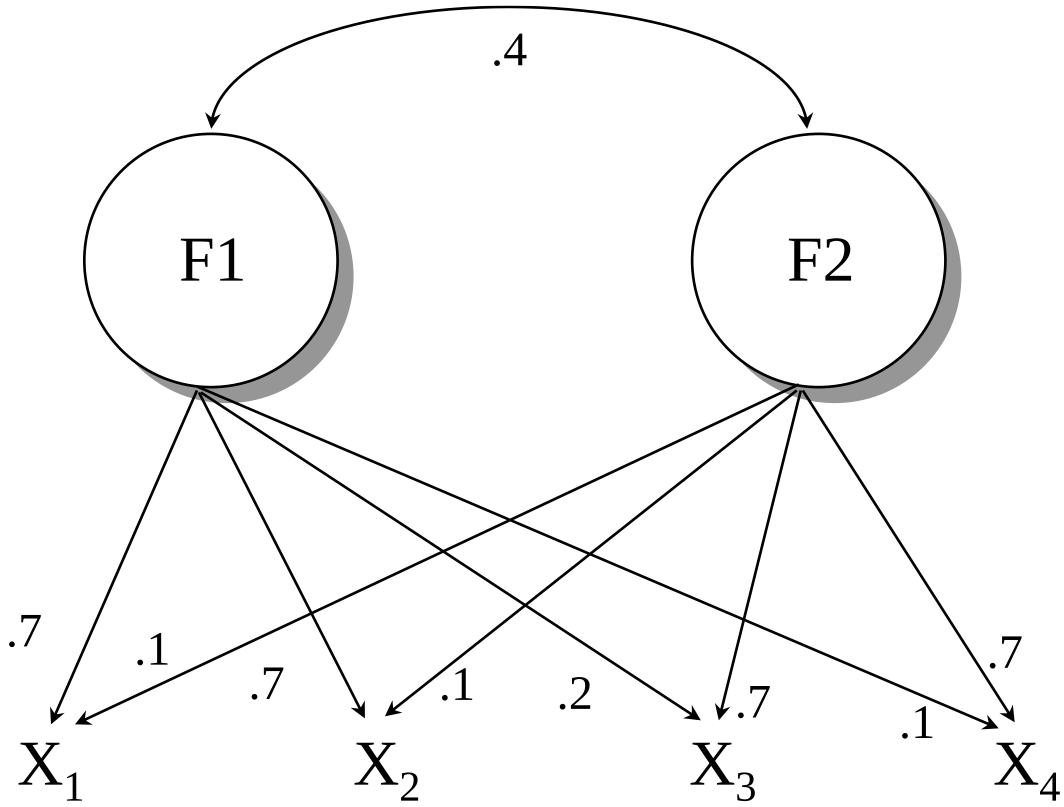
Risultano tanto più "gonfiate" quanto più è elevata la correlazione tra i fattori.

Per interpretare i fattori nelle rotazioni oblique si esamina la matrice pattern.

Soluzioni fattoriali

	Soluzione obliqua		
	F1	F2	h^2
Determinato	-.08	.88	.72
Dinamico	-.01	.89	.78
Energico	.16	.76	.72
Affidabile	.91	-.02	.81
Responsabile	.95	-.02	.89
Scrupoloso	.85	.07	.78
Varianza Spiegata	42%	37%	

La rotazione ridistribuisce la varianza spiegata dai singoli fattori, ma la varianza totale rimane identica (79%)



Effetto diretto di F1 su X1 = .7 (P)

Effetto di F1 su X1 dovuto alla correlazione tra F1 e F2= .4*.1=.04

Effetto totale di F1 su X1 = .7 + .04 = .74 (S)

Riprodurre R dopo una rotazione obliqua:

Matrici Struttura (S) e Pattern (P)

$$S = P\Phi.$$

Φ= Matrice delle correlazioni tra i fattori

$$R^* = SP' \quad \text{e} \quad R = SP' + U^2$$

$$R^* = P\Phi P' \quad \text{e} \quad R = P\Phi P' + U^2$$

Nelle rotazioni ortogonali invece:

$S = P = A$, $\Phi = I$, quindi

$$R^* = P\Phi P' = AA'$$

Varianza spiegata dopo la rotazione obliqua

- Moltiplicare P e S elemento per elemento
- Sommare i prodotti per colonna
- Dividere i totali di colonna per il numero di variabili e moltiplicare per 100.

Variabili	PATTERN (P)		STRUCTURE (S)		PRODOTTO (P * S)		
	F1	F2	F1	F2	F1	F2	h^2
X1	0,70	0,10	0,74	0,38	0,52	0,04	0,56
X2	0,70	0,10	0,74	0,38	0,52	0,04	0,56
X3	0,20	0,70	0,48	0,78	0,10	0,55	0,65
X4	0,10	0,70	0,38	0,74	0,04	0,52	0,56
% Var.					29	29	

Somma per riga dei prodotti: comunalità.

Interpretazione dei fattori tramite le saturazioni

Sono le saturazioni degli item (i cosiddetti *factor loadings*) a rappresentare il principale strumento per l'interpretazione del fattore, e per l'eventuale selezione delle variabili *migliori*.
E' prassi interpretare i fattori considerando gli item che presentano saturazioni elevate nel fattore e contemporaneamente basse negli altri fattori. Questi item vengono definiti *markers*.
Sotto il livello soglia di circa |.30| le saturazioni possono essere considerate inadeguate. E' preferibile che le saturazioni principali siano decisamente superiori a .30. E' augurabile che le saturazioni secondarie siano il più possibile vicine a 0, per raggiungere la *struttura semplice*

Interpretazione dei fattori tramite le saturazioni

Matrice "ruotata"

	F1	F2	h^2
Determinato	.17	.83	.72
Dinamico	.24	.85	.78
Energico	.36	.77	.72
Affidabile	.87	.23	.81
Responsabile	.91	.24	.89
Scrupoloso	.83	.30	.78
Proporzione di Varianza Spiegata	.42	.37	

Saturazioni secondarie (sui fattori non pertinenti)

Saturazioni principali (sul fattore di pertinenza): SU DI ESSE SI BASA L'INTERPRETARE DELLA SOLUZIONE

76

Interpretazione dei fattori tramite le saturazioni

La struttura semplice rende più facile l'interpretazione della soluzione fattoriale ed è per questo che risulta particolarmente augurabile.

In generale una struttura è tanto più semplice quanto maggiore è la differenza tra la saturazione principale e la saturazione secondaria più elevata

Si suggeriscono i seguenti criteri di semplicità fattoriale:

- La saturazione secondaria inferiore a |.32| (Tabachnick e Fidell, 2007)
- Differenza tra loading principale e secondo loading più elevato maggiore di |.30| (Matsunaga, 2011)
- Rapporto tra loading principale e secondo loading più elevato maggiore di 2 (Ercolani e Perugini, 1997)

Interpretazione dei fattori tramite le saturazioni

Esempio empirico per comprendere l'interpretazione della struttura fattoriale tramite saturazioni

Item	1	2	Diff (S_Pri,S_Sec)	S_Pri/S_Sec
IT7	**0,876**	-0,123	0,753	7.14
IT8	**0,845**	-0,116	0,729	7.25
IT9	**0,837**	-0,101	0,736	8.28
IT6	**0,545**	0,285	0,26	1.91
IT10	**0,300**	-0,051	0,249	5.84
IT2	-0,120	**0,872**	0,752	7.27
IT1	-0,077	**0,828**	0,751	10.74
IT3	-0,046	**0,812**	0,766	17.73
IT5	**0,365**	**0,487**	0,122	1.33
IT4	-0,043	0,189	0,146	4.41

Diff(S_Pri,S_Sec) = differenza (in valore assoluto) tra saturazione principale (S_Pri) e secondaria (S_Sec)
S_Pri/S_Sec = rapporto tra la saturazione principale (S_Pri) e quella secondaria (S_Sec)

Interpretazione dei fattori tramite le saturazioni

Consideriamo il *primo fattore*:

- 4 item presentano saturazioni principali superiori a .30, di cui tre (il 7, l'8 e il 9) sono anche fattorialmente semplici, mentre uno (il 6) ha una differenza tra le 2 saturazioni inferiore a |.3| e il rapporto leggermente inferiore a 2, quindi è meno semplice degli altri

- un item presenta una saturazione principale esattamente uguale a .30

I primi tre item (7, 8, 9) possono essere considerati *marker* della dimensione misurata dal primo fattore.

Interpretazione dei fattori tramite le saturazioni

Consideriamo il *secondo fattore*:

- Quattro item presentano saturazioni principali superiori a .30, ma tre (il 2, l'1, e il 3) sono fattorialmente semplici, mentre l'item 5 ha un indice di semplicità decisamente inferiore a |.3| così come il rapporto tra saturazioni decisamente più basso di 2
- un item presenta una saturazione principale inferiore a .30.

Anche in questo caso i primi tre item possono essere considerati *marker* della dimensione misurata dal secondo fattore.

Interpretazione dei fattori tramite le saturazioni

Marker Index (Gallucci e Perugini, 2007)

$$M_{ik} = 1 - \sqrt{1 - 2a_{ik} + h_i^2}$$

a_{ik} è il factor loading dell'item i nel fattore k
h_i^2 è la comunalità dell'item

Gli autori consigliano come adeguati valori maggiori di 0.40.

Interpretazione dei fattori tramite le saturazioni

Esempio empirico per comprendere l'interpretazione della struttura fattoriale tramite saturazioni

Item	1	2	Marker Index
IT7	**0,876**	-0,123	0,825
IT8	**0,845**	-0,116	0,806
IT9	**0,837**	-0,101	0,808
IT6	***0,545***	*0,285*	*0,463*
IT10	**0,300**	-0,051	0,298
IT2	-0,120	**0,872**	0,825
IT1	-0,077	**0,828**	0,812
IT3	-0,046	**0,812**	0,806
IT5	**0,365**	**0,487**	0,370
IT4	-0,043	0,189	0,188

Interpretazione dei fattori tramite le saturazioni

In generale, tutte le variabili che presentano una saturazione principale maggiore di |.30|, nessuna saturazione secondaria superiore a |.30|, e soprattutto una differenza o un rapporto tra saturazione principale e saturazione secondaria molto elevato possono essere considerate *marker puri* della dimensione misurata dal fattore.
Gli item con saturazione principale elevata ma con una minore differenza o un minor rapporto tra saturazione principale e secondaria (es., rapporto tra 1.5 e 2) possono essere presi in considerazione nell'interpretazione del fattore, ma non possono essere considerati come *marcatori* (*marker*) *puri* dello stesso.
Gli item che presentano una saturazione secondaria sostanzialmente identica a quella principale sono decisamente più ambigui, e non vanno presi in considerazione nell'interpretazione del fattore, a meno che per farlo non ci siano ragioni teoriche (ad esempio, item che definiscono una dimensione interstiziale, ovvero che si colloca a metà strada tra 2 fattori, come avviene nei modelli *circomplessi*).

Assunzioni e prerequisiti - Fattorializzabilità di R

- Test di sfericità di Bartlett:
H_0: R = I (I = matrice identità).
Se significativo, e il campione è sufficientemente ampio, la matrice è fattorializzabile.

- Indice di adeguatezza campionaria KMO:

$$KMO = \Sigma\Sigma r^2 / (\Sigma\Sigma r^2 + \Sigma\Sigma p^2)$$

r = correlazioni tra ogni coppia di variabili
p = correlazioni tra ogni coppia di variabili, parzializzate rispetto a tutte le altre variabili

Assunzioni e prerequisiti - Fattorializzabilità di R

- Test di adeguatezza campionaria di Kaiser (KMO):

Interpretazione dei valori del KMO:

>0.90: eccellenti;
0.80-0.90: buoni;
0.70-0.80: accettabili;
0.60-0.70: mediocri;
<0.60: scarsi/non accettabili (l'analisi è sconsigliata)

Assunzioni e prerequisiti

Livelli di misura e distribuzione delle variabili: Almeno intervalli equivalenti. Anche ordinali se il numero di categorie ordinabili di una variabile è sufficiente (es., da 5 in su), e se la distribuzione delle variabili è normale.

Coefficienti di correlazione: Coefficiente di correlazione di Pearson (dà stime più stabili). Variabili dicotomiche o ordinali: coefficienti di correlazione "speciali" (tetracorici e policorici: sono ottenibili in Prelis, non in SPSS).

Assunzioni e prerequisiti

Normalità multivariata: Se le distribuzioni sono normali la soluzione è migliore. Richiesta Con il metodo di estrazione della Maximum Likelihood.

Linearità: Necessaria perché l'analisi si basa sui coefficienti di Pearson. Metodi di AF non lineari: basati su coefficienti speciali (utili per dati non normali).

Outliers tra i casi e tra le variabili:
Casi estremi univariati e multivariati possono distorcere i risultati. Variabili "outlier": non correlano con le altre variabili in analisi, e vanno a definire fattori "residuali" e poco attendibili (saturati soltanto da quella variabile).

Assunzioni e prerequisiti

Numero di variabili:

- Numero di variabili 3 o 4 volte superiore al numero dei fattori

- Non meno di 3 variabili "marker" per ogni fattore che si vuole identificare (fattori "sovradeterminati").

Ampiezza e qualità del campione:

- Campioni piccoli producono stime poco stabili di r

- Consigliabile non scendere mai sotto i <u>100</u> soggetti e non avere mai meno di cinque casi per ogni variabile.

- Variabilità delle variabili e/o dei fattori: sufficientemente ampia. Campione molto selezionato ed omogeneo: riduzione della variabilità e quindi delle correlazioni; mancata individuazione di fattori, minore percentuale di var. spiegata, saturazioni più basse.

Ambiti di applicazione dell'Analisi Fattoriale

- Costruzione di test psicologici
- Costruzione di scale e questionari

Esame della qualità di strumenti di misura per:

* identificare indicatori adeguati e non adeguati
* identificare fattori misurati in maniera non adeguata (es., da una sola variabile)

Esplorazione di dati ☹

Problemi

Applicazione poco attenta delle opzioni di *default*

Fiducia in una *erronea* tradizione consolidata

Scarsa conoscenza del modello statistico di base

Decisioni da prendere

Adeguatezza delle variabili

Fattorializzabilità di R

Tecnica per l'estrazione

Numero di fattori

Tecnica per la rotazione

Interpretazione dei fattori

Adeguatezza della soluzione

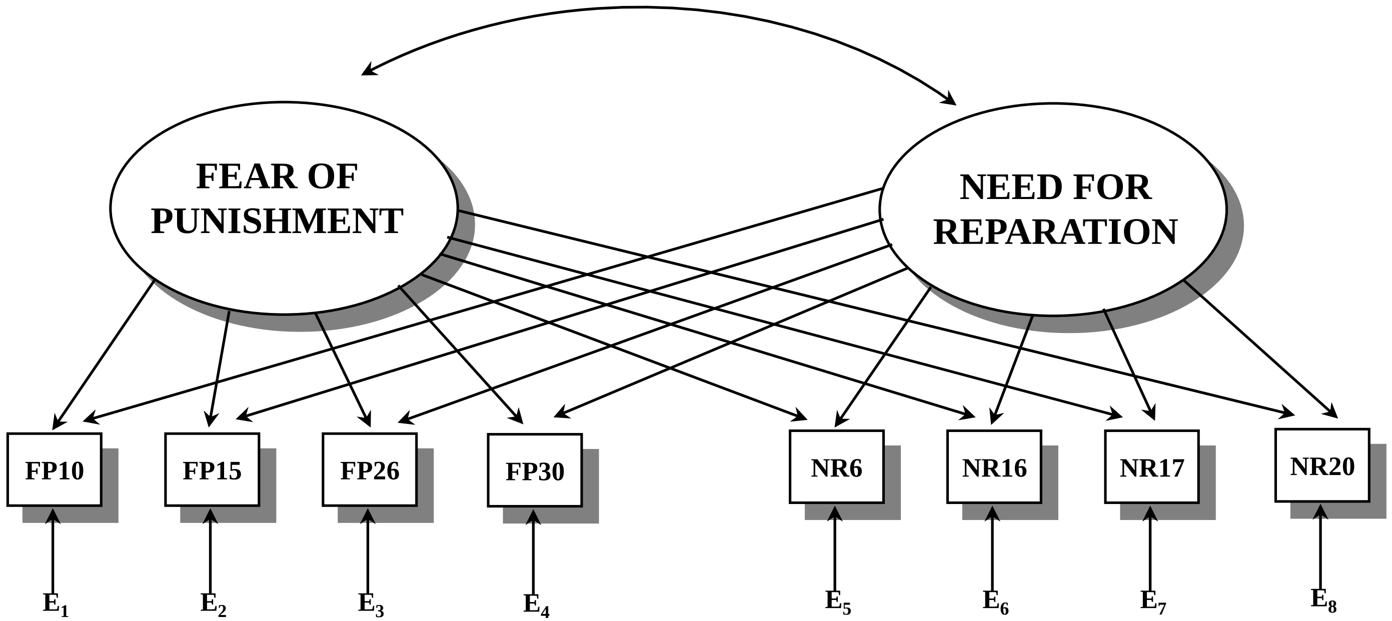
Come ottenere buone soluzioni

- Numero di indicatori per ogni fattore (> 3)
- Almeno 100 soggetti
- Campione *non* selezionato
- Non utilizzare l'analisi delle componenti principali ma un metodo fattoriale vero
- Rotazione obliqua (ortogonale solo se i fattori non correlano)
- Più metodi per scegliere il numero di fattori: non utilizzare il criterio dell'autovalore >1

ANALISI FATTORIALE ESPLORATIVA

(EFA) CON SPSS

MODELLO CONCETTUALE



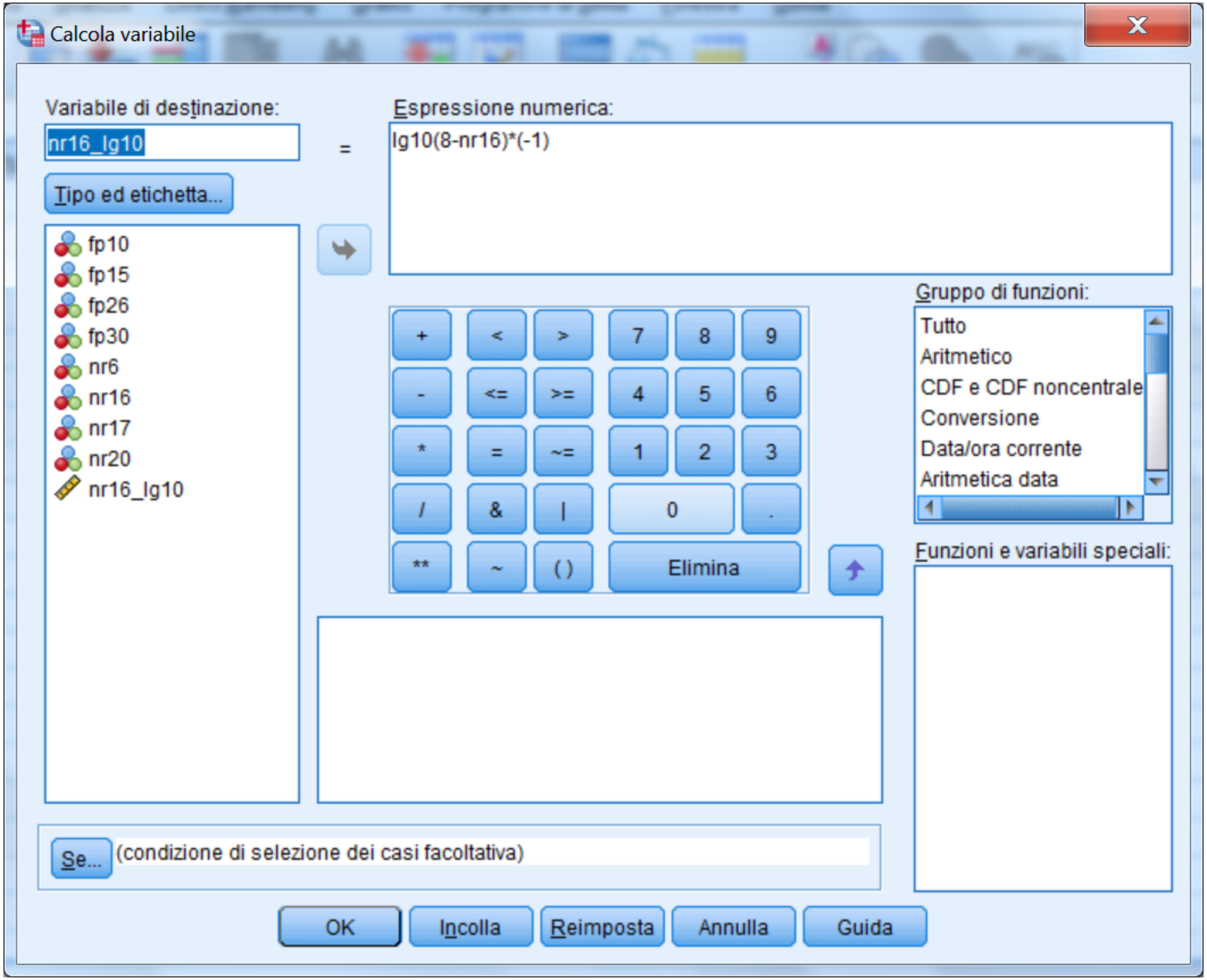
efa_dati.sav

EFA IN SPSS – analisi preliminari

Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	Asimmetria		Curtosi	
	Statistica	Statistica	Statistica	Statistica	Statistica	Statistica	Errore std.	Statistica	Errore std.
fp10	819	1	6	2,66	1,582	,481	,085	-,942	,171
fp15	819	1	6	3,17	1,492	,066	,085	-,945	,171
fp26	819	1	6	3,68	1,393	-,310	,085	-,599	,171
fp30	819	1	6	3,15	1,553	,059	,085	-1,103	,171
nr6	819	1	6	4,37	1,379	-,852	,085	,186	,171
nr16	819	1	6	4,67	1,261	-1,097	,085	1,061	,171
nr17	819	1	6	4,70	1,272	-1,100	,085	,951	,171
nr20	819	1	6	5,04	1,163	-1,496	,085	2,321	,171
Numero di casi validi (listwise)	819								

EFA IN SPSS – analisi preliminari

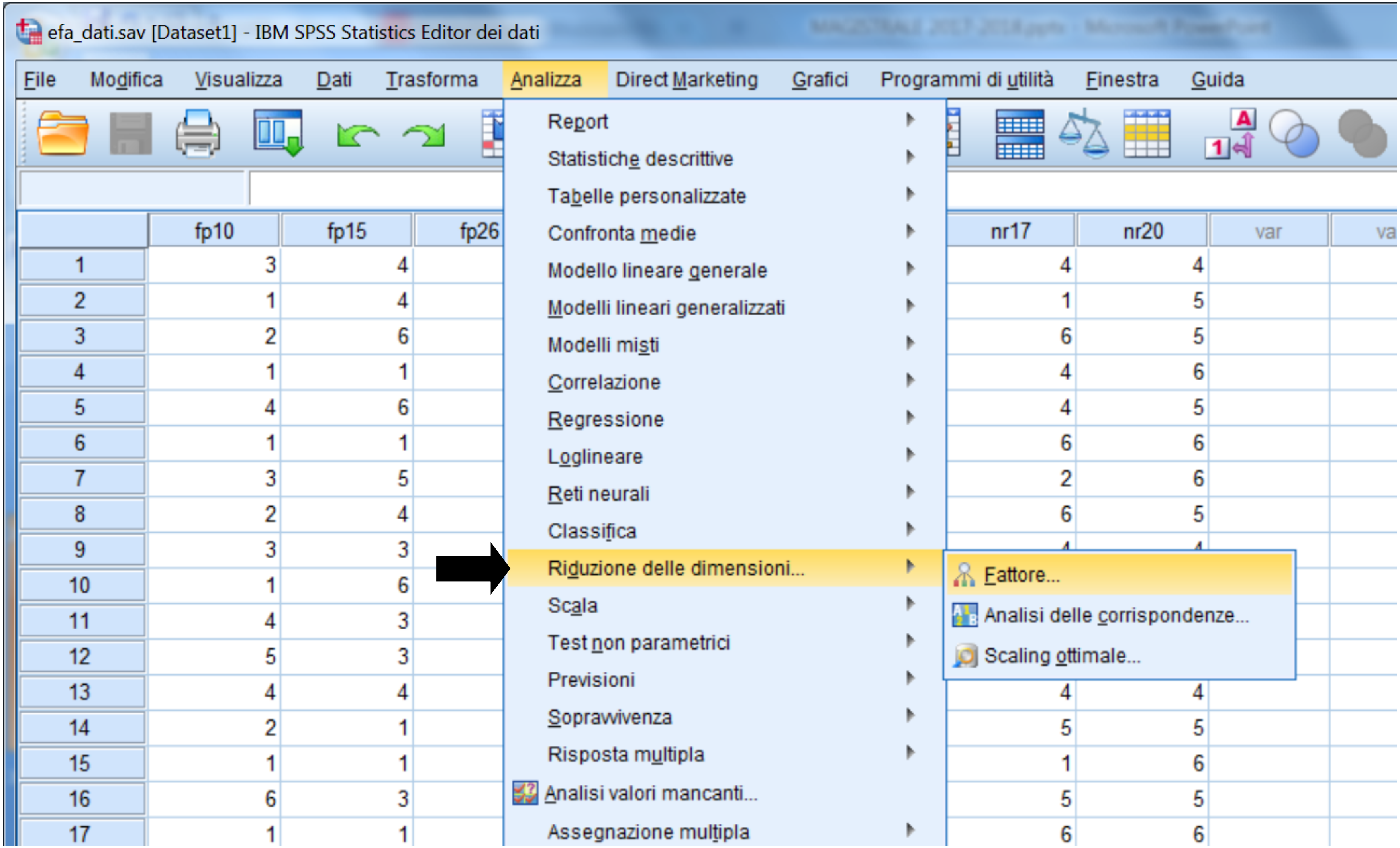


EFA IN SPSS – analisi preliminari

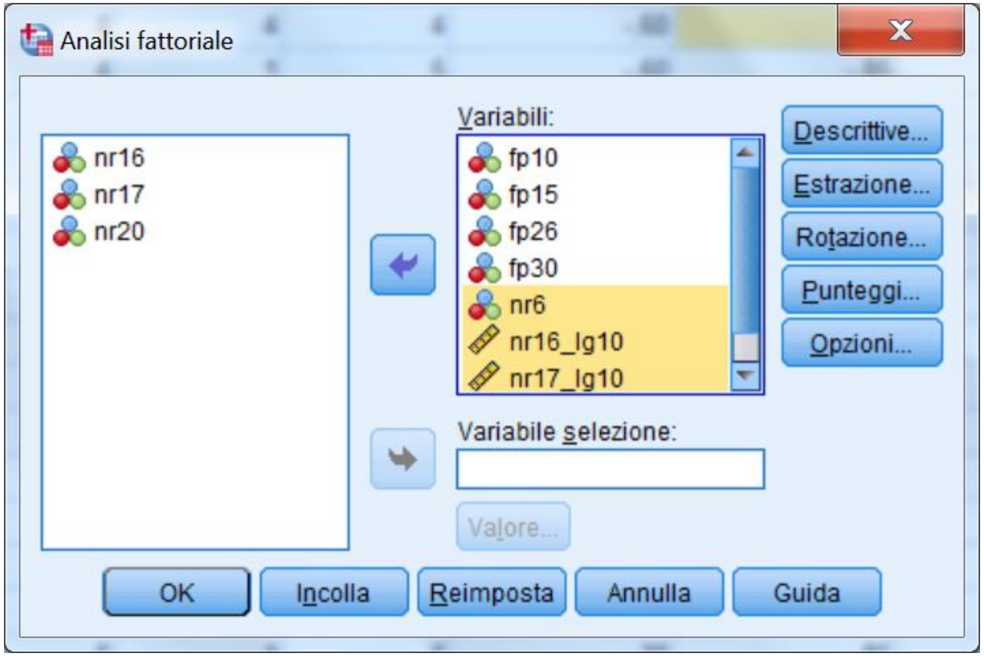
Statistiche descrittive

	N	Minimo	Massimo	Media	Deviazione std.	Asimmetria		Curtosi	
	Statistica	Statistica	Statistica	Statistica	Statistica	Statistica	Errore std.	Statistica	Errore std.
nr16	819	1	6	4,67	1,261	-1,097	,085	1,061	,171
nr17	819	1	6	4,70	1,272	-1,100	,085	,951	,171
nr20	819	1	6	5,04	1,163	-1,496	,085	2,321	,171
nr16_lg10	819	-,85	-,30	-,4936	,15479	-,306	,085	-,641	,171
nr17_lg10	819	-,85	-,30	-,4891	,15665	-,350	,085	-,686	,171
nr20_lg10	819	-,85	-,30	-,4435	,14996	-,721	,085	-,294	,171
Numero di casi validi (listwise)	819								

EFA IN SPSS

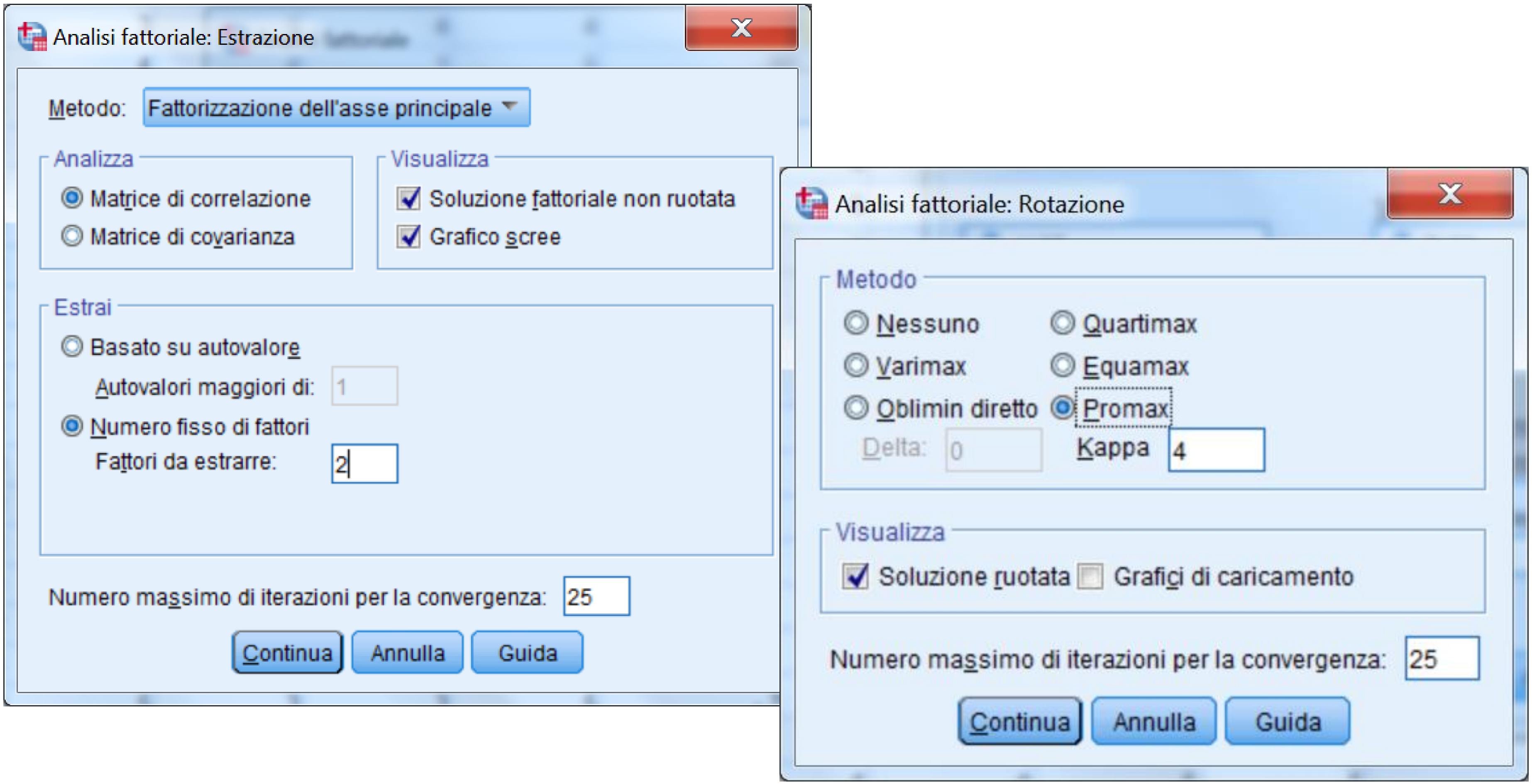

efa_dati.sav [Dataset1] - IBM SPSS Statistics Editor dei dati
File
Modifica
Visualizza
Dati
Trasforma
Analizza
Direct Marketing
Grafici
Programmi di utilità
Finestra
Guida
fp10
fp15
fp26
nr17
nr20
var
Report
Statistiche descrittive
Tabelle personalizzate
Confronta medie
Modello lineare generale
Modelli lineari generalizzati
Modelli misti
Correlazione
Regressione
Loglineare
Reti neurali
Classifica
Riduzione delle dimensioni...
Scala
Test non parametrici
Previsioni
Sopravvivenza
Risposta multipla
Analisi valori mancanti...
Assegnazione multipla
Fattore...
Analisi delle corrispondenze...
Scaling ottimale...

EFA IN SPSS

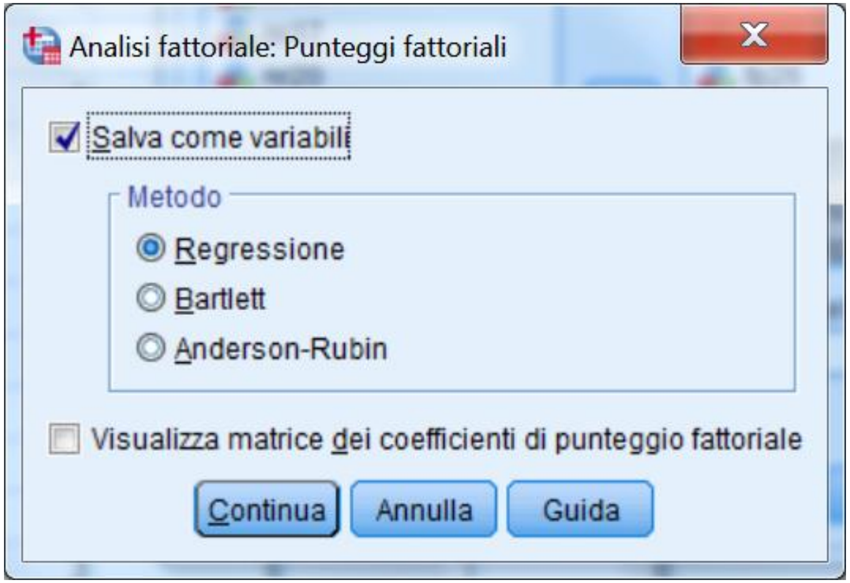

Analisi fattoriale
nr16
nr17
nr20
Variabili:
fp10
fp15
fp26
fp30
nr6
nr16_lg10
nr17_lg10
Variabile selezione:
Valore...
Descrittive...
Estrazione...
Rotazione...
Punteggi...
Opzioni...
OK
Incolla
Reimposta
Annulla
Guida

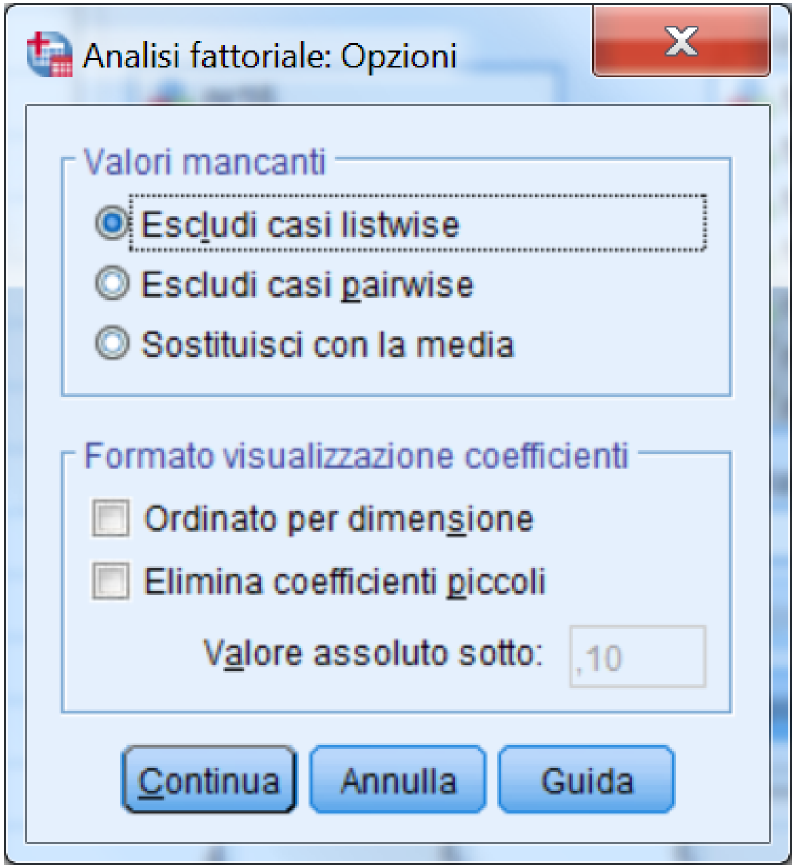

Analisi fattoriale: Descrittive
Statistiche
Descrittive univariate
Soluzione iniziale
Matrice di correlazione
Coefficienti
Inverso
Livelli di significatività
Riprodotto
Determinante
Anti-immagine
Test KMO e test di sfericità di Bartlett
Continua
Annulla
Guida

EFA IN SPSS


Analisi fattoriale: Estrazione
Metodo: Fattorizzazione dell'asse principale
Analizza
Matrice di correlazione
Matrice di covarianza
Visualizza
Soluzione fattoriale non ruotata
Grafico scree
Estrai
Basato su autovalore
Autovalori maggiori di: 1
Numero fisso di fattori
Fattori da estrarre: 2
Numero massimo di iterazioni per la convergenza: 25
Continua
Annulla
Guida
Analisi fattoriale: Rotazione
Metodo
Nessuno
Quartimax
Varimax
Equamax
Oblimin diretto
Promax
Delta: 0
Kappa 4
Visualizza
Soluzione ruotata
Grafici di caricamento
Numero massimo di iterazioni per la convergenza: 25
Continua
Annulla
Guida

EFA IN SPSS


Analisi fattoriale: Punteggi fattoriali
Salva come variabili
Metodo
Regressione
Bartlett
Anderson-Rubin
Visualizza matrice dei coefficienti di punteggio fattoriale
Continua
Annulla
Guida


Analisi fattoriale: Opzioni
Valori mancanti
Escludi casi listwise
Escludi casi pairwise
Sostituisci con la media
Formato visualizzazione coefficienti
Ordinato per dimensione
Elimina coefficienti piccoli
Valore assoluto sotto:
,10
Continua
Annulla
Guida

EFA IN SPSS

Statistiche descrittive

	Media	Deviazione std.	N analisi
fp10	2,66	1,582	819
fp15	3,17	1,492	819
fp26	3,68	1,393	819
fp30	3,15	1,553	819
nr6	4,37	1,379	819
nr16_lg10	-,4936	,15479	819
nr17_lg10	-,4891	,15665	819
nr20_lg10	-,4435	,14996	819

Matrice di correlazione[a]

		fp10	fp15	fp26	fp30	nr6	nr16_lg10	nr17_lg10	nr20_lg10
Correlazione	fp10	1,000	,368	,256	,344	,050	,039	,080	-,010
	fp15	,368	1,000	,390	,444	,155	,060	,074	,074
	fp26	,256	,390	1,000	,418	,122	,163	,130	,141
	fp30	,344	,444	,418	1,000	,120	,068	,154	,036
	nr6	,050	,155	,122	,120	1,000	,301	,310	,307
	nr16_lg10	,039	,060	,163	,068	,301	1,000	,256	,354
	nr17_lg10	,080	,074	,130	,154	,310	,256	1,000	,302
	nr20_lg10	-,010	,074	,141	,036	,307	,354	,302	1,000

a. Determinante = ,299

EFA IN SPSS

Test di KMO e Bartlett

Misura di Kaiser-Meyer-Olkin di adeguatezza del campionamento.		,737
Test della sfericità di Bartlett	Appross. Chi-quadrato	982,770
	gl	28
	Sign.	,000

Comunalità

	Iniziale	Estrazione
fp10	,184	,261
fp15	,299	,456
fp26	,253	,341
fp30	,304	,473
nr6	,189	,296
nr16_lg10	,188	,311
nr17_lg10	,172	,256
nr20_lg10	,208	,375

Metodo di estrazione: Fattorizzazione dell'asse principale.

Varianza totale spiegata

Fattore	Autovalori iniziali			Caricamenti somme dei quadrati di estrazione			Caricamenti somme dei quadrati di rotazione[a]
	Totale	% di varianza	% cumulativa	Totale	% di varianza	% cumulativa	Totale
1	2,402	30,021	30,021	1,769	22,118	22,118	1,605
2	1,655	20,688	50,709	1,001	12,508	34,625	1,354
3	,792	9,895	60,603				
4	,724	9,049	69,652				
5	,714	8,920	78,572				
6	,647	8,086	86,658				
7	,549	6,858	93,516				
8	,519	6,484	100,000				

Metodo di estrazione: Fattorizzazione dell'asse principale.

a. Quando i fattori sono correlati, i caricamenti delle somme dei quadrati non possono essere aggiunti per ottenere una varianza totale.

EFA IN SPSS

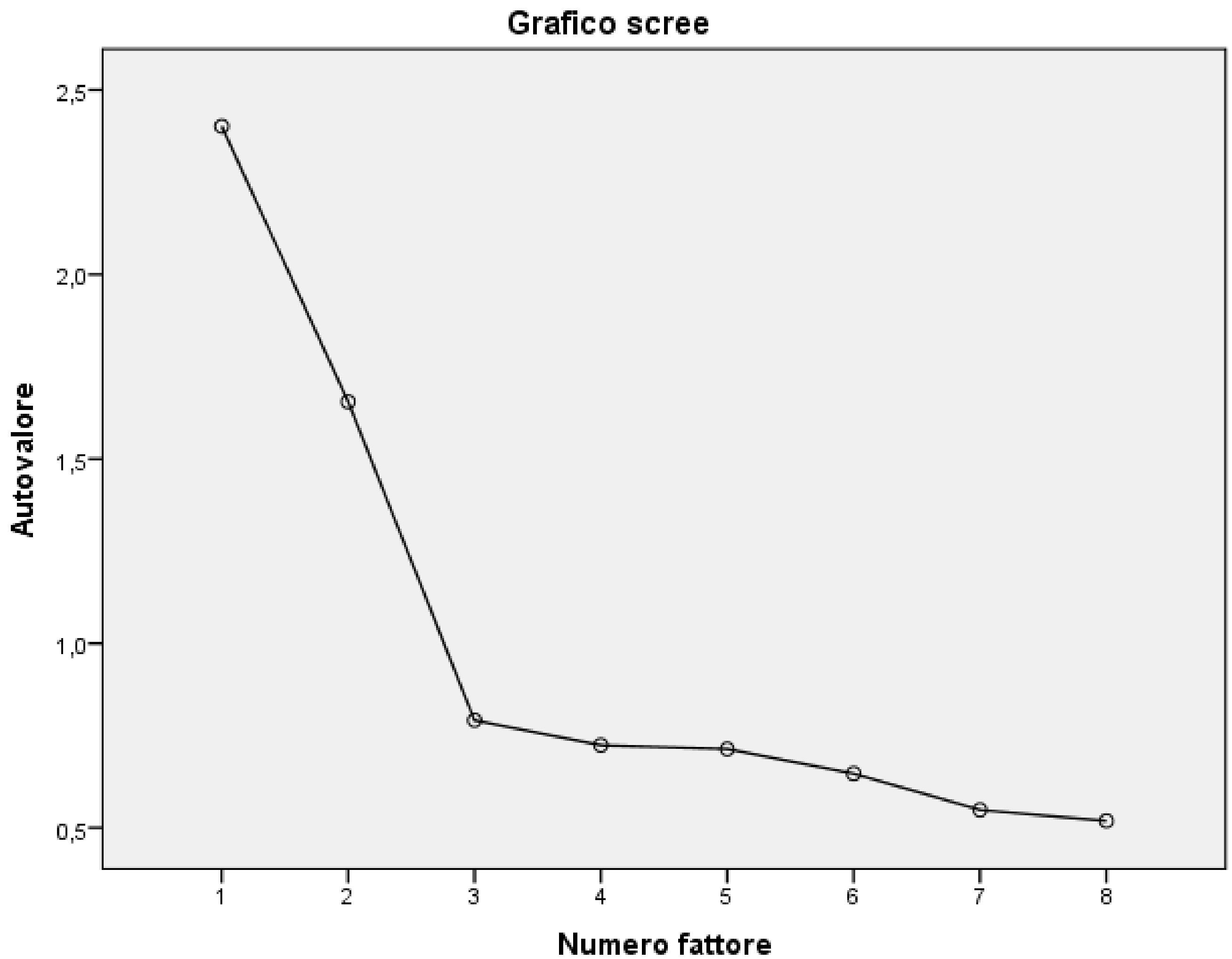


Grafico scree
2,5
2,0
1,5
1,0
0,5
Autovalore
1
2
3
4
5
6
7
8
Numero fattore

EFA IN SPSS

Correlazioni riprodotte

		fp10	fp15	fp26	fp30	nr6	nr16_lg10	nr17_lg10	nr20_lg10
Correlazione riprodotta	fp10	,261[a]	,343	,283	,349	,059	,027	,057	,006
	fp15	,343	,456[a]	,386	,464	,118	,078	,113	,055
	fp26	,283	,386	,341[a]	,394	,162	,131	,153	,119
	fp30	,349	,464	,394	,473[a]	,122	,081	,117	,057
	nr6	,059	,118	,162	,122	,296[a]	,301	,276	,327
	nr16_lg10	,027	,078	,131	,081	,301	,311[a]	,280	,340
	nr17_lg10	,057	,113	,153	,117	,276	,280	,256[a]	,303
	nr20_lg10	,006	,055	,119	,057	,327	,340	,303	,375[a]
Residuo[b]	fp10		,025	-,027	-,006	-,009	,012	,023	-,016
	fp15	,025		,004	-,020	,036	-,018	-,040	,019
	fp26	-,027	,004		,025	-,040	,031	-,023	,022
	fp30	-,006	-,020	,025		-,002	-,013	,037	-,022
	nr6	-,009	,036	-,040	-,002		,000	,034	-,020
	nr16_lg10	,012	-,018	,031	-,013	,000		-,023	,014
	nr17_lg10	,023	-,040	-,023	,037	,034	-,023		-,001
	nr20_lg10	-,016	,019	,022	-,022	-,020	,014	-,001	

Metodo di estrazione: Fattorizzazione dell'asse principale.

a. Comunalità riprodotte

b. I residui vengono calcolati tra le correlazioni osservate e riprodotte. Ci sono 0 (0,0%) residui non ridondanti con valori assoluti maggiori di 0,05.

EFA IN SPSS

Matrice dei fattori[a]

	Fattore 1	Fattore 2
fp10	,418	-,295
fp15	,593	-,323
fp26	,559	-,169
fp30	,605	-,328
nr6	,401	,368
nr16_lg10	,363	,423
nr17_lg10	,376	,339
nr20_lg10	,362	,493

Metodo di estrazione: Fattorizzazione dell'asse principale.

a. 2 fattori estratti. 8 iterazioni richieste.

Matrice del modello[a]

	Fattore 1	Fattore 2
fp10	,525	-,070
fp15	,678	-,013
fp26	,544	,113
fp30	,691	-,011
nr6	,046	,531
nr16_lg10	-,022	,563
nr17_lg10	,048	,492
nr20_lg10	-,071	,627

Metodo di estrazione: Fattorizzazione dell'asse principale.
Metodo di rotazione: Promax con normalizzazione Kaiser.

a. Convergenza per la rotazione eseguita in 3 iterazioni.

Matrice di struttura

	Fattore 1	Fattore 2
fp10	,507	,067
fp15	,675	,164
fp26	,574	,256
fp30	,688	,170
nr6	,185	,543
nr16_lg10	,126	,557
nr17_lg10	,177	,504
nr20_lg10	,093	,608

Metodo di estrazione: Fattorizzazione dell'asse principale.
Metodo di rotazione: Promax con normalizzazione Kaiser.

Matrice di correlazione dei fattori

Fattore	1	2
1	1,000	,261
2	,261	1,000

Metodo di estrazione: Fattorizzazione dell'asse principale.
Metodo di rotazione: Promax con normalizzazione Kaiser.

ESERCIZIO 5: REALIZZAZIONE DI UN MODELLO DI ANALISI FATTORIALE ESPLORATIVA

Effettuare un modello di analisi fattoriale esplorativa.

I dati sono nel file spss ESE_EFA.SAV

items:
Workload: QWI_1 QWI_2 QWI_3 QWI_4 QWI_5
Organizational Constraints: OCS_1 OCS_2 OCS_3 OCS_4 OCS_5 OCS_6 OCS_7 OCS_8 OCS_9 OCS_10 OCS_11

Effettuare l'analisi con SPSS scegliendo il metodo più adeguato per l'estrazione e la rotazione dei fattori, dopo aver esaminato le proprietà distributive degli item