

Analisi multivariata (DPRS)

8b-Esempio di fattoriale esplorativa vers. 1.0

Germano Rossi¹

`germano.rossi@unimib.it`

¹Dipartimento di Psicologia, Università di Milano-Bicocca

2010-2011

Esempio di fattoriale esplorativa

- Questo esempio utilizza dei dati fittizi derivati da Catena y Ramos (19??), con alcune modifiche (file `CatenaYRamos.sav` allegato)
- Si tratta di 10 variabili (identificate da X_1 a X_{10}) e 20 soggetti (troppo pochi per un'analisi reale).
- L'obiettivo di questo esempio è solo quello di vedere praticamente una serie di procedure per l'analisi fattoriale esplorativa (AFE).
- Proprio perché i dati sono stati costruiti per dare risultati "puliti", non è neppure necessario conoscere il significato degli item.

Passi necessari

- Ispezione visiva della matrice di correlazione (quando ci sono poche variabili)
- La prima analisi fattoriale serve solo per avere informazioni generiche:
 - Statistiche sulla matrice di correlazione per vedere la possibilità statistica di effettuare un'AFE
 - Numero dei fattori da estrarre
 - Tramite gli autovalori
 - Tramite lo *scree-plot*
- Adesso possiamo decidere il numero dei fattori, metodo di estrazione e di rotazione

■ **Analizza » Correlazione » Bivariata**

```
CORRELATIONS
```

```
/VARIABLES=x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/PRINT=TWOTAIL NOSIG /MISSING=PAIRWISE.
```

■ **Analizza » Riduzione dimensioni » Fattoriale** (in Descrittive check su Determinante, Test KMO; in Estrazione, check su Grafico decrescente autovalori)

```
FACTOR /VARIABLES x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/MISSING LISTWISE  
/ANALYSIS x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/PRINT INITIAL DET KMO EXTRACTION /PLOT EIGEN  
/CRITERIA MINEIGEN(1) ITERATE(25)  
/EXTRACTION PC /ROTATION NOROTATE  
/METHOD=CORRELATION.
```

Ispezione visiva della matrice di correlazione

Dopo aver evidenziato in grassetto le correlazioni elevate (superiori a .40), usiamo uno sfondo diverso per evidenziare i blocchi di correlazioni reciproche.

Correlazioni

	x1	x2	x3	x4	x5	x6	x7	x8	x9	x10
x1	1									
x2	.891	1								
x3	-.300	-.177	1							
x4	.126	.000	-.804	1						
x5	-.221	-.124	.876	-.721	1					
x6	.842	.799	-.127	-.086	-.074	1				
x7	.747	.704	-.161	-.150	-.206	.878	1			
x8	-.258	-.151	.885	-.880	.798	-.084	-.001	1		
x9	-.763	-.787	.280	.106	.273	-.787	-.850	.201	1	
x10	.264	.141	-.884	.928	-.820	.076	.036	-.915	-.109	1

Ci sono due blocchi, quindi due probabili fattori

Statistiche sulla matrice di correlazione

Matrice di correlazione

a. Determinante = ,000

Test KMO e di Bartlett

Misura di adeguatezza campionaria KMO (Keiser Meyer Olkin).		0,703
Test di sfericità di Bartlett	Chi-quadrato appross.	228,74
	df	45
	Sig.	,000

- 1 Il determinante è diverso da 0, quindi ci sono delle righe che possono essere espresse come combinazione lineare delle altre.
- 2 KMO è maggiore di .6, quindi ci sono abbastanza correlazioni elevate.
- 3 Bartlett è significativo, quindi non ci sono correlazioni nulle fuori dalla diagonale.

Numero dei fattori da estrarre

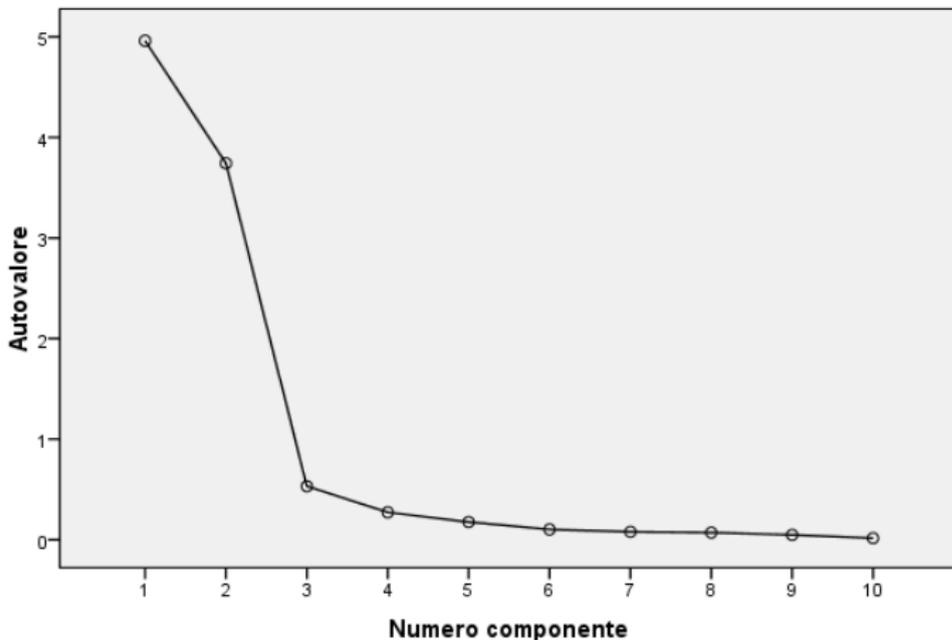
Varianza totale spiegata

Componente	Autovalori iniziali			Pesi dei fattori non ruotati		
	Totale	% varianza	% cumulata	Totale	% varianza	% cumulata
1	4,961	49,611	49,611	4,961	49,611	49,611
2	3,743	37,429	87,041	3,743	37,429	87,041
3	,532	5,319	92,359			
4	,272	2,724	95,084			
5	,177	1,767	96,850			
6	,102	1,020	97,870			
7	,078	,785	98,655			
8	,071	,710	99,365			
9	,048	,480	99,845			
10	,015	,155	100,000			

- Ci sono due componenti che spiegano più varianza di un singolo item (cioè 1).
- Le due componenti spiegherebbero l'87% della varianza totale

Numero dei fattori da estrarre

Grafico decrescente autovalori



Il punto di flesso è al 3° fattore. Si possono estrarre due o tre componenti. Tuttavia la terza spiega meno di 1, quindi 2.

Passi necessari

■ Metodo di estrazione

- Componenti principali: anche se è più diffusa in letteratura, non è una AF perché non prevede i fattori unici
- Assi principali (AP): matematicamente è la stessa cosa delle componenti principali, ma ipotizza i fattori unici
- Massima verosimiglianza (ML): è un approccio statistico che lavora ipotizzando di stimare i dati della popolazione

■ Tipo di rotazione

- Ortogonale: i fattori non correlano fra loro
- Obliqua: i fattori “possono” essere fra loro correlati

Metodo di estrazione

- Tutti i metodi di estrazione dovrebbero portare alla stessa soluzione fattoriale; con campioni molto grandi o con dati “perfetti” succede.
- Componenti principali e assi principali tendono a privilegiare il primo fattore, mentre gli altri metodi (ML, residui minimi) no
- Siccome di solito si fanno più analisi, la prima può essere AP e la successiva ML; in entrambi i casi abbiamo stabilito di chiedere 2 fattori (tramite autovalori > 1 e scree-test).

Metodo di rotazione

- I metodi ortogonali (Varimax, Equamax, Quartimax), sono da preferire quando vogliamo ottenere fattori/componenti il più possibile “puliti” (cioè che misurano un determinato costrutto e solo quello). Ad es. quando stiamo costruendo o lavorando su più scale. Il metodo Varimax è il più famoso e utilizzato in letteratura; privilegia il primo fattore. Equamax e Quartimax cercano invece di sparpagliare le saturazioni delle variabili su tutti i fattori (ma non sempre ci riescono)
- I metodi obliqui (Oblimin, Promax) quando vogliamo ottenere fattori che “possono” misurare una parte diversa di un medesimo costrutto. Solitamente quando uno strumento di misura è costituito da sottoscale che misurano aspetti diversi della stessa cosa. Il metodo Oblimin è il più vecchio, mentre il Promax è più recente ed è quello preferito attualmente.
- Tentiamo una soluzione Varimax e una Promax

Soluzioni ortogonali

■ Assi principali, Varimax

```
FACTOR /VARIABLES x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/MISSING LISTWISE /ANALYSIS x1 TO x10  
/PRINT INITIAL EXTRACTION ROTATION  
/FORMAT SORT BLANK(.25) /PLOT ROTATION  
/CRITERIA FACTORS(2) ITERATE(25) /EXTRACTION PAF  
/ROTATION VARIMAX /METHOD=CORRELATION.
```

■ Massima verosimiglianza, Varimax

```
FACTOR /VARIABLES x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/MISSING LISTWISE /ANALYSIS x1 TO x10  
/PRINT INITIAL EXTRACTION ROTATION  
/FORMAT SORT BLANK(.25) /PLOT ROTATION  
/CRITERIA FACTORS(2) ITERATE(25) /EXTRACTION ML  
/ROTATION VARIMAX /METHOD=CORRELATION.
```

Soluzioni ortogonali

Matrice fattoriale ruotata Assi principali			Matrice fattoriale ruotata Massimaverosimiglianza		
	Fattore			Fattore	
	1	2		1	2
x10	-.970		x10	-.977	
x8	.946		x4	-.946	
x3	.926		x8	.939	
x4	-.922		x3	.901	
x5	.843		x5	.831	
x6		.929	x6		.929
x7		.895	x7		.906
x9		-.891	x9		-.892
x1		.890	x1		.882
x2		.872	x2		.867

Le due soluzioni sono uguali (appartenenza ai fattori, segno; cambia il valore della saturazione).

Soluzioni oblique

■ Assi principali, Promax

FACTOR

```
/VARIABLES x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/MISSING LISTWISE /ANALYSIS x1 TO x10  
/PRINT INITIAL EXTRACTION ROTATION  
/FORMAT SORT BLANK(.25) /PLOT ROTATION  
/CRITERIA FACTORS(2) ITERATE(25) /EXTRACTION PAF  
/ROTATION PROMAX(4) /METHOD=CORRELATION.
```

■ Massima verosimiglianza, Promax

FACTOR

```
/VARIABLES x1 x2 x3 x4 x5 x6 x7 x8 x9 x10  
/MISSING LISTWISE /ANALYSIS x1 TO x10  
/PRINT INITIAL EXTRACTION ROTATION  
/FORMAT SORT BLANK(.25) /PLOT ROTATION  
/CRITERIA FACTORS(2) ITERATE(25) /EXTRACTION ML  
/ROTATION PROMAX(4) /METHOD=CORRELATION.
```

Soluzioni oblique

Matrice dei modelli Assi principali			Matrice dei modelli Massimaverosimiglianza		
	Fattore			Fattore	
	1	2		1	2
x10	-.973		x10	-.980	
x8	.947		x4	-.962	
x4	-.938		x8	.940	
x3	.921		x3	.896	
x5	.840		x5	.827	
x6		.937	x6		.935
x7		.903	x7		.915
x9		-.891	x9		-.893
x1		.883	x1		.874
x2		.873	x2		.868

Anche in queste due soluzioni, gli item vengono assegnati agli stessi fattori della soluzione ortogonale. Le correlazioni fra i due fattori sono basse per entrambe le soluzioni (rispettivamente $-.143$ e $-.140$)

Conclusioni

- Avendo eseguito 4 analisi fattoriali esplorative con due metodi di estrazione e due metodi di rotazione...
- avendo trovato la stessa struttura fattoriale in tutte e tre le soluzioni...
- possiamo concludere che gli item x10, x8, x4, x3 e x5 sono spiegati dal primo fattore, mentre gli item x6, x7, x9, x1, x2 sono spiegati dal secondo fattore
- Considerare che i due fattori (nelle soluzioni oblique) non correlano molto, possiamo considerarli come se fossero indipendenti.