Analisi fattoriale

esplorativa
vers. 1.0

Germano Rossi[1]
germano.rossi@unimib.it

[1] Dipartimento di Psicologia, Università di Milano-Bicocca

2009

Prima di un AFE

- Identificare un dominio di ricerca
- selezionare un certo numero di variabili osservabili che verranno misurate su un buon numero di unità statistiche
- le osservate che correlano molto fra loro possono sottintende un fattore
- le variabili che non correlano con nessun altra, si scartano
  - analizzando la matrice di correlazione
  - usando la correlazione item-totale
  - usando la stessa AFE (comunalità)

Passaggi per un'AFE

- Verificare che l'AFE si possa fare (livelli di misura, normalità e linearità, outliers, n. variabili e n. fattori attesi e n. soggetti)
- Decidere
  1. Quale matrice di coefficienti di associazione da analizzare e quindi verificare la sua adeguatezza
  2. Quanti fattori estrarre
  3. Quale metodo usare per l'estrazione
  4. Quale rotazione usare e se i fattori sono correlati o no
  5. Se servono i punteggi fattoriale e come calcolarli
- Interpretazione dei fattori

Verificare la fattibilità

- Dati quantitativi ”veri” (intervallo / rapporto)
  - è possibile fare un’analisi fattoriale esplorativa con dati ordinali (Lisrel)
  - ma anche nominali (MPlus)
- Variabili che si distribuiscano normalmente
- Esclusione degli *outliers*
- Considerare l’ampiezza del campione.

Ampiezza del campione

- Ci sono orientamenti diversi che possono essere riassunti in:
  - Più soggetti ci sono meglio è
  - Più casi che variabili (si capirà con le componenti principali)
  - Da 10 a 20 casi per ogni variabile
  - Come minimo 5 casi per variabile ma non meno di 100 casi in totale (Gorsuch, 1983)
  - ogni fattore deve avere almeno 4 saturazioni superiori a .60, non importa l'ampiezza del campione (Guadagnoli e Velicer, 1988)
  - ogni fattore deve avere almeno 10 saturazioni superiori a .40 e il campione almeno di 150 casi (idem)
  - un campione di almeno 300 casi (idem)
  - bastano anche 60 casi se tutte le comunalità sono maggiori di .60 (MacCallum, Widaman, Zhang e Hong, 1999)
  - con comunalità attorno a .50 , il campione dev'essere tra 100 e 200 casi (idem)

Quale matrice di coefficienti di associazione usare

L'analisi fattoriale esplorativa si applica ad una matrice di correlazione (o meglio ad una matrice di associazione fra variabili). Gli indici che si possono usare sono:

- **correlazione di Pearson**: è la correlazione più usata in assoluto (il default in quasi tutti i software statistici). Implica variabili misurate a livello intervallo/rapporto
- **correlazione di Spearman**: è una correlazione per variabili di tipo ordinale oppure quando si presume che le variabili intervallo non siano "normali"
- **varianze/covarianze**: è la scelta preferita nelle analisi fattoriali confermative ed è utilizzabile quando le variabili hanno varianza simile
- altri indici di associazione: correlazioni policoriche, poliseriali, punto-biseriale....

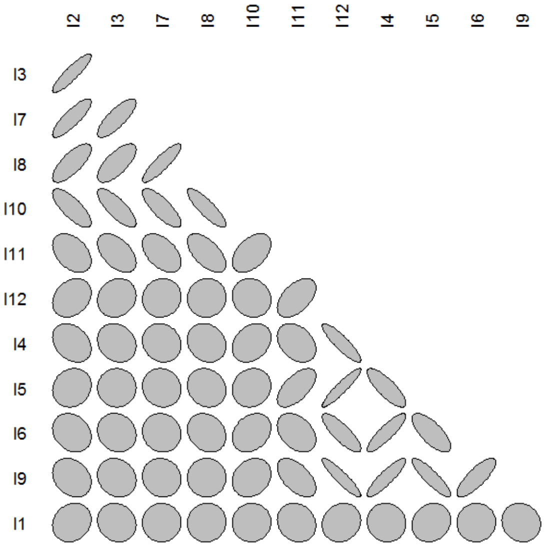
Verificare l'adeguatezza della matrice di correlazione

- Non tutte le correlazioni sono inferiori a |.30|
- Ispezione visiva della matrice di correlazione per vedere che abbia blocchi di correlazioni alte fra loro e basse con le alte
- Determinante: se è nullo non si può fare l'analisi;
- KMO (Test di adeguatezza campionaria di Kaiser-Meyer-Olkin)
- Test della Sfericità di Bartlett [tende a sovrastimare]
- Ispezione della matrice delle correlazioni parzializzate (deve contenere valori bassi) e anti-immagine (deve contenere valori alti)
- Comunalità

Ispezione visiva della matrice di correlazione

I2	I3	I7	I8	I10	I11	I12	I4	I5	I6	I9
1										
0,89	1									
0,84	0,80	1								
0,75	0,70	0,88	1							
-0,76	-0,79	-0,79	-0,85	1						
-0,41	-0,50	-0,48	-0,60	0,47	1					
0,26	0,14	0,08	0,04	-0,11	0,51	1				
-0,30	-0,18	-0,13	-0,16	0,28	-0,34	-0,88	1			
0,13	0,00	-0,09	-0,15	0,11	0,61	0,93	-0,80	1		
-0,22	-0,12	-0,07	-0,21	0,27	-0,38	-0,82	0,88	-0,72	1	
-0,26	-0,15	-0,08	0,00	0,20	-0,55	-0,91	0,89	-0,88	0,80	1
0,16	-0,12	0,00	0,05	0,02	-0,15	0,12	-0,05	0,14	0,02	-0,10

Ispezione visiva della matrice di correlazione



KMO

- KMO (Test di adeguatezza campionaria di Kaiser-Meyer-Olkin)

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j p_{ij}^2}$$

  dove r_{ij} sono le correlazioni e p_{ij} sono le correlazioni parzializzate su tutte le altre
- se le correlazioni parzializzate sono piccole tende a 1, quindi (secondo Kaiser, 197)
- se > 0.90 è eccellente
- fra .80 e .90 buono;
- fra .70 e .80 accettabile
- fra .60 e .70 mediocre
- inferiore a .60, meglio non fare l'analisi
- N.B. la dicitura "campionaria" non si riferisce al campione

Sfericità di Bartlett

- Il test della sfericità di Bartlett verifica l'ipotesi $H_0 : \mathbf{R} = \mathbf{I}$ tramite la formula:

$$\chi^2 = -\left[n - 1 - \frac{1}{6}(2p+5)\right] ln \mid \mathbf{R} \mid$$

  in cui n è il numero dei soggetti, p il numero delle variabili e $\mid \mathbf{R} \mid$ il determinante della matrice di correlazione.
- Si distribuisce con $(p^2 - p)/2$ gradi di libertà
- Se è significativo, significa che $\mathbf{R}$ ha correlazioni sufficientemente elevate da non essere paragonabili a 0; se è non significativo le correlazioni sono basse e non si distinguono da 0
- ma questo test dipende dal numero delle variabile e dalla numerosità del campione, quindi tende ad essere significativo all'aumentare del campione e del numero delle variabili anche se ci sono correlazioni basse.

Adeguatezza

Matrice di correlazione[a]

a. Determinante = ,000

Test KMO e di Bartlett

Misura di adeguatezza campionaria KMO (Keiser Meyer Olkin).		,703
Test di sfericità di Bartlett	Chi-quadrato appross.	228,740
	df	45
	Sig.	,000

Matrice anti-immagine

- La matrice delle correlazioni parzializzate indica il valore di ogni correlazione dopo aver eliminato il contributo di tutte le altre variabili non implicate.
- una parzializzata alta significa che le due variabili sono molto correlate fra loro, ma non hanno legami con nessun altra.
- La logica invece è che ci siano più di due variabili per fattore
- La matrice anti-immagine contiene i complementi a 1 della correlazione parzializzata fra due variabili rispetto a tutte e altre.
- Valori alti indicano correlazioni parziali basse e viceversa

Comunalità

Comunalità

	Iniziale	Estrazione
I1	0,776	**0,005**
I2	0,941	0,828
I3	0,963	0,770
I4	0,943	0,859
I5	0,953	0,878
I6	0,886	0,722
I7	0,934	0,831
I8	0,962	0,826
I9	0,961	0,928
I10	0,948	0,800
I11	0,919	0,674
I12	0,938	0,949

- La comunalità è la % di varianza della variabile spiegata dai fattori comuni
- Se i fattori comuni non spiegano nulla, meglio eliminare

Quanti fattori estrarre

- Teoria (analisi della letteratura)
- Rango della matrice (solo teorico)
- Criterio di Kaiser (1959) o di Guttman (1954) ovvero autovalori maggiori di 1 (sovrastima)
- Almeno il 60-75% di varianza spiegata
- Scree-test di Cattell (1966)
- Test statistici di alcuni metodi di estrazione
- Il buon senso
- Analisi parallela

Gorsuch (1983) ritiene necessario effettuare più analisi e tenere quei fattori che si mantengono nelle varie soluzioni

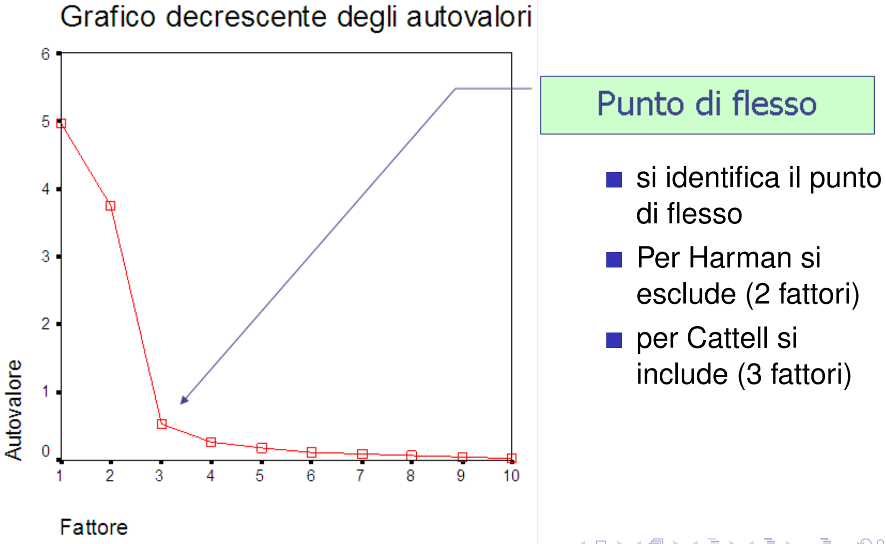
% di varianza spiegata

Varianza totale spiegata

Fattore	Autovalori iniziali		
	Totale	% di varianza	% cumulata
1	4,961	49,611	49,611
2	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

Metodo di estrazione: Fattorizzazione dell'asse principale.

Scree-test



- si identifica il punto di flesso
- Per Harman si esclude (2 fattori)
- per Cattell si include (3 fattori)

Analisi parallela

- Horn (1965) ha proposto una strategia per decidere il numero dei fattori chiamata analisi parallela
- L'idea iniziale è stata apprezzata e alcuni autori la giudicano "la miglio trategia possibile"

	Attuali	Casuali	
1	2,776	1,756	si tiene
2	0,828	1,082	si scarta
3	0,276	0,797	
4	0,119	0,364	
	4,00	4,00	

- Ci sono diverse soluzioni possibili
- Partendo da una matrice (stesse dimensioni) di dati casuali si calcolano gli autovalori e si tengono gli autovalori *attuali* se maggiori di quelli *casuali*

Come estrarre i fattori

Ci sono diversi metodi per estrarre i fattori

- Basati sulle componenti principali
  - Componenti principali (1 sulla diagonale, ma non è un'AFE)
  - Fattori principali o assi principali (stima iniziale della comunalità sulla diagonale)
- Massima verosimiglianza (permette un test sui fattori)
- Minimi quadrati (permette un test sui fattori)
- Altri metodi poco usati (alfa factoring, image factoring...)

Con un numero di casi sufficientemente grande, si equivalgono tutti, in quanto portano alla stessa soluzione

Metodi di estrazione

- Componenti principali: assume che i punteggi delle variabili siano misurati senza errore ovvero siano affidabili ($\mathbf{R} = \mathbf{A}\mathbf{A}'$), quindi stima direttamente il campione e non la popolazione. Se il campione è ampio e rappresentativo della popolazione da cui è stato estratto, tende a dare gli stessi risultati degli assi principali. L'estrazione avviene una componente alla volta, cercando di estrarre quella che spiega la maggior varianza possibile. Dopo l'estrazione di una componente, viene calcolata la matrice dei residui e su questa viene ripetuta l'estrazione.
  Quindi, questo metodo, tende a privilegiare il primo fattore (come se fosse un fattore generale) che quindi include più item degli altri e tende a spiegare tutta la varianza.

Metodi di estrazione

- Assi principali o fattori principali: matematicamente è la stessa procedura delle componenti principali con la differenza che viene usata la comunalità ($\mathbf{R} = \mathbf{A}\mathbf{A}' + \mathbf{U}^{\mathbf{2}}$) come stima della correlazione di una variabile con se stessa (nella popolazione). La comunalità è anche la stima più bassa dell'affidabilità di una variabile.
  Anche questo metodo tende a privilegiare il primo fattore.
  È possibile usare altri valori al posto della comunalità (ad. es. valori di affidabilità pubblicati se si usa un test standardizzato; la correlazione più alta di quella variabile...).
  In ogni caso, alla fine del processo, il software calcolerà la comunalità dovuta alla soluzione trovata.

Metodi di estrazione

- Massima verosimiglianza e minimi quadrati: cercano di trovare dei fattori che riproducano al meglio la matrice di correlazione di partenza. Solo alla fine calcolano la comunalità.
- Alpha factor: cerca di massimizzare l’affidabilità
- Image factor: cerca di creare fattori che minimizzino l’unicità (o fattore specifico)

Campione vs. popolazione

- a volte, l’analisi fattoriale serve per “capire” le idee o il comportamento di un campione e non ci serve di generalizzare i risultati
- altre volte, vogliamo ottenere attori che siano utilizzabili con altri campioni nell’ipotesi che valgano per tutta la popolazione
- componenti principali e assi principali assumono che il campione utilizzato sia una popolazione. I risultati fanno quindi riferimento al campione e non dovrebbero essere estesi alla popolazione se non dopo aver verificato che i fattori ottenuti restano stabili nel tempo e con l’uso di altri campioni
- massima verosimiglianza e image factoring assumono che le variabili misurate siano le miglior rappresentanti delle variabili misurabili sulla popolazione. I risultati si possono quindi estendere alla popolazione.

Test sui fattori

- Massima verosimiglianza e minimi quadrati permettono di calcolare una statistica di significatività (un chi-quadro) sull'adattamento del modello fattoriale in base al numero dei fattori
- Se il chi-quadro è **non significativo**, possiamo dire che la soluzione con q fattori è accettabile
- Siccome è in un χ^2 tende ad essere significativo con molte variabili e tanti soggetti. Si può quindi usare per calcolare un RMSEA.

Campione vs. popolazione

- a volte, l'analisi fattoriale serve per "capire" le idee o il comportamento di un campione e non ci serve di generalizzare i risultati
- altre volte, vogliamo ottenere attori che siano utilizzabili con altri campioni nell'ipotesi che valgano per tutta la popolazione
- componenti principali e assi principali assumono che il campione utilizzato sia una popolazione. I risultati fanno quindi riferimento al campione e non dovrebbero essere estesi alla popolazione se non dopo aver verificato che i fattori ottenuti restano stabili nel tempo e con l'uso di altri campioni
- massima verosimiglianza e image factoring assumono che le variabili misurate siano le miglior rappresentanti delle variabili misurabili sulla popolazione. I risultati si possono quindi estendere alla popolazione.

Come ruotare

- Metodi ortogonali (i fattori non correlano)
  - Varimax: lavora cercando di massimizzare le saturazioni alte, inimizzando quelle basse, all'interno dei singoli fattori. E' consigliato se si vuole ottenere una netta separazione fra i fattori o se non si hanno criteri precisi da seguire
  - Quartimax: lavora cercando di sparpagliare la varianza entro le singole variabili. Questo dovrebbe facilitare la lettura delle variabili e non dovrebbe produrre fattori generale. Spesso però non ci riesce.
  - Equamax: lavora equamente sulle variabili e sui fattori mantenendo costante la varianza spiegata dall'intera soluzione. Spesso non riesce ad ottenere la soluzione semplice.

  Le saturazioni sono correlazioni fra le variabili e i fattori.

La rotazione si può fare solo con due o più fattori. Anche se i fattori non correlano fra loro, i punteggi fattoriali possono correlare.

Come ruotare

- Metodi obliqui (i fattori correlano fra loro)
  - Oblimin: il grado di associazione fra i fattori è determinato da un parametro δ generalmente posto a 0; valori positivi (fino a .8) aumentano la correlazione, negativi (fino a -.8) diminuiscono la correlazione. Tutti consigliano di non toccarlo.
  - Promax: è sostanzialmente uguale a oblimin, ma pensato per essere più veloce e lavorare con grandi quantità di dati.

Entrambi i metodi si basano una prima rotazione Varimax, su cui poi si innesta un procedimento interattivo che cerca di aumentare le saturazioni alte e ridurre quelle basse avvicinando gli assi.
Le saturazioni finali non sono correlazioni ma parametri di regressione

Significato della rotazione

Non ruotata

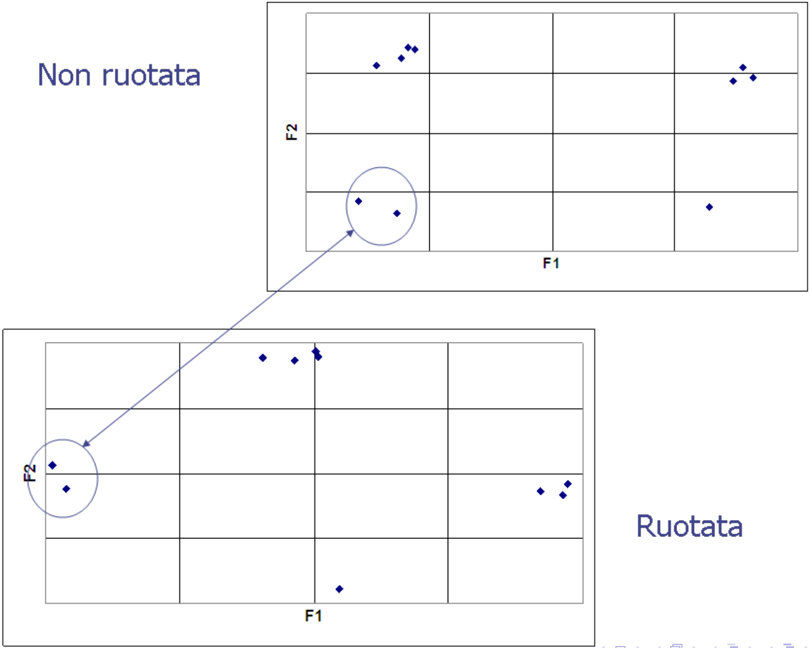
	1	2
x1	,716	,563
x2	,612	,625
x3	-,819	,464
x4	,633	-,681
x5	-,739	,431
x6	,588	,719
x7	,561	,698
x8	-,780	,541
x9	-,642	-,624
x10	,788	-,571

Ruotata

	1	2
x1	-,191	,890
x2	-,071	,872
x3	,926	-,166
x4	-,922	-,119
x5	,843	-,141
x6	,007	,929
x7	,014	,895
x8	,946	-,083
x9	,096	-,891
x10	-,970	,064

Tentativo di trovare una “struttura semplice” (Thurstone, 1947): saturazioni elevate in un fattore e il più basse possibili negli altri (focalizziamoci su X4 e X10)

Non ruotata



Ruotata

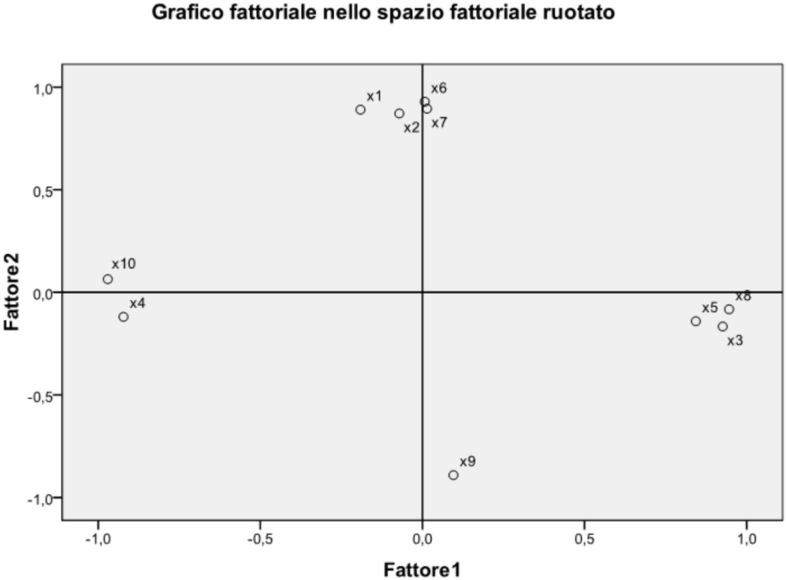


Grafico fattoriale nello spazio fattoriale ruotato
Fattore2
Fattore1
1,0
0,5
0,0
-0,5
-1,0
-1,0
-0,5
0,0
0,5
1,0
x1
x2
x6
x7
x10
x4
x5
x8
x3
x9

Ortogonale o obliquo?

- La scelta principale è se si desiderano fattori correlati oppure no
- La soluzione ortogonale è più facile da interpretare perché produce una sola matrice di saturazioni e perché le saturazioni sono anche correlazioni
- La soluzione obliqua produce due matrici:
  - la matrice dei modelli (pattern matrix): parametri di regressione usati per calcolare i punteggi fattoriali (la matrice che va interpretata)
  - la matrice di struttura (structure matrix): correlazioni (parziali) fra fattori e variabili
- Normalmente si cercano fattori il più possibili “netti” (cioè che misurano qualcosa di diverso dagli altri), quindi ortogonali
- D’altra parte, è possibile che i costrutti psicologici non siano correlati?
- Qualcuno suggerisce: usate obliqua e se i fattori correlano poco, usate ortogonale

Saturazioni e correlazioni

- Nella soluzione non ruotata e in quelle ortogonali, le saturazioni sono la correlazione fra le variabili e il fattore
- In tal caso il loro quadrato corrisponde alla proporzione di varianza spiegata dal fattore per quella variabile
- La somma dei quadrati corrisponde alla comunalità ovvero alla percentuale di varianza spiegata dai fattori comuni per quella variabile
- Nella soluzione ruotata, le comunalità si calcolano come somma dei prodotti di ogni valore della pattern matrix per ogni valore corrispondente della structure matrix

Varianza, comunalità, unicità

- L'osservata X ha varianza 1 (è standardizzata!)
- Può essere suddivisa in una parte dovuta ai fattori unici (U) e una parte dovuta ai fattori comuni (F):

$$var(X) = var(F) + var(U)$$

- Il rapporto fra $var(F)$ e $var(X)$ si chiama ”comunalità” (h^2), mentre $var(U)$ si chiama “unicità” (u^2)
- Essendo la $var(x) = 1 = h^2 + u^2$
- L'unicità può essere ulteriormente suddivisa in varianza specifica dell'item e varianza d'errore, ma l'AF non riesce a distinguere fra le due
- Può capitare che $h^2 + u^2 > 1$ (Heywood case); indica un problema nella matrice iniziale

Soluzione non ruotata: Lisrel

	Factor 1	Factor 2	Unique Var
	--------	--------	----------
VAR1	0.333	**0.843**	0.178
VAR2	0.208	**0.846**	0.240
VAR3	**-0.916**	-0.036	0.160
VAR4	**0.919**	-0.256	0.089
VAR5	**-0.842**	-0.022	0.290
VAR6	0.141	**0.918**	0.138
VAR7	0.102	**0.901**	0.178
VAR8	**-0.941**	0.054	0.112
VAR9	-0.197	**-0.872**	0.200
VAR10	**0.976**	-0.076	0.041

$h^2=$
.333^2
+
.843^2)

$u^2=.178$

1=
$h^2+u^2=$
.822+178

Soluzione ruotata Varimax: Lisrel

	Factor 1	Factor 2	Unique Var
	--------	--------	----------
VAR1	0.209	**0.882**	0.178
VAR2	0.085	**0.867**	0.240
VAR3	**-0.901**	-0.166	0.160
VAR4	**0.946**	-0.122	0.089
VAR5	**-0.831**	-0.142	0.290
VAR6	0.009	**0.929**	0.138
VAR7	-0.027	**0.906**	0.178
VAR8	**-0.939**	-0.081	0.112
VAR9	-0.071	**-0.892**	0.200
VAR10	**0.977**	0.064	0.041

1-
(.209^2
+.882^2)
=.178

Soluzione ruotata Promax: Lisrel

	Factor 1	Factor 2	Unique Var
	--------	--------	----------
VAR1	**0.875**	0.141	0.178
VAR2	**0.869**	0.017	0.240
VAR3	-0.106	**-0.895**	0.160
VAR4	-0.188	**0.963**	0.089
VAR5	-0.086	**-0.826**	0.290
VAR6	**0.936**	-0.065	0.138
VAR7	**0.916**	-0.099	0.178
VAR8	-0.017	**-0.940**	0.112
VAR9	**-0.894**	-0.001	0.200
VAR10	-0.003	**0.980**	0.041

1-
(.875^2
+.141^2)
diverso da
.178

Factor Correlations

	Factor 1	Factor 2
	--------	--------
Factor 1	1.000	
Factor 2	0.070	1.000

Soluzione non ruotata: Spss

	Fattore			
	1	2	Comunalità	Unicità
x1	,716	,563	0,829	0,171
x2	,612	,625	0,765	0,235
x3	-,819	,464	0,885	0,115
x4	,633	-,681	0,865	0,135
x5	-,739	,431	0,731	0,269
x6	,588	,719	0,863	0,137
x7	,561	,698	0,801	0,199
x8	-,780	,541	0,902	0,098
x9	-,642	-,624	0,802	0,198
x10	,788	-,571	0,946	0,054

$$0,716^2 + ,563^2 = 0,829 \qquad 1 - 0,829 = 0,171$$

Soluzione ruotata Varimax: Spss

Matrice fattoriale ruotata[a]

	Fattore	
	1	2
X10	-,970	
X8	,946	
X3	,926	
X4	-,922	
X5	,843	
X6		,929
X7		,895
X9		-,891
X1		,890
X2		,872

Metodo estrazione: fattorizzazione dell'asse principale.
Metodo rotazione: Varimax con normalizzazione di Kaiser.

a. La rotazione ha raggiunto i criteri di convergenza in 3 iterazioni.

Spss permette di non visualizzare le saturazioni ritenute non significative e di riordinare i valori in modo decrescente per fattore

In questo caso:

- **Fattore 1**: item x10, x8, x3, x4, x5
- **Fattore 2**: item x6, x7, x9, x1, x2

Soluzione ruotata Promax: spss

	Mat. modelli		**Mat. struttura**		
	Fattore		Fattore		
	1	2	1	2	h^2
I2	-,128	,883	-,254	,902	0,829
I3	-,009	,873	-,133	,875	0,765
I4	,921	-,101	,936	-,233	0,885
I5	-,938	-,186	-,911	-,053	0,865
I6	,840	-,081	,851	-,201	0,731
I7	,075	,937	-,059	,926	0,863
I8	,079	,903	-,049	,892	0,801
I9	,947	-,015	,949	-,150	0,902
I10	,032	-,891	,159	-,895	0,802
I12	-,973	-,005	-,973	,134	0,946

$$-,128 \times -,254+$$

$$+,883 \times ,902 =,829$$

Punteggi fattoriali

- Punteggio che ogni caso statistico (soggetto) assume in un certo fattore
- Tutti i programmi possono calcolare i punteggi fattoriali, usando varie forme di regressione multipla (difficili da interpretare)
  - Regressione
  - Bartlett
  - Anderson-Rubin
- Metodi manuali
  - Metodo delle saturazioni
  - Metodo congenerico (punteggi fattoriali compositi)
  - Metodo dei punteggi fattoriali normalizzati non centrati

Punteggi fattoriali automatici

- **Regressione**: la matrice delle saturazioni (la matrice di configurazione nelle oblique) viene moltiplicata per l'inversa delle correlazioni ($\mathbf{F} = \mathbf{A}\mathbf{R}^{-1}$). I punteggi hanno media 0 e varianza pari al quadrato della correlazione multipla fra i punteggi grezzi e quelli stimati (per cui non è facile conoscere a priore i valori minimo e massimo). I diversi punteggi fattoriali tendono a correlare fra loro.
- **Bartlett** (1937): variazione della tecnica precedente che cerca di minimizzare l'importanza dei fattori unici (ovvero le variabili che non rientrano nel fattore). I punteggi fattoriali possono correlare fra loro.
- **Anderson-Rubin** (1956): variazione del metodo di Bartlett con standardizzazione (media 0, varianza 1) e ortogonalità (i punteggi fattoriali dei diversi fattori non correlano fra loro).

Punteggi fattoriali manuali

- **Saturazioni**: le saturazioni delle variabili di un fattore vengono moltiplicate per il punteggio grezzo che il soggetto ha dato in quelle variabili. Non sono facili da interpretare perché non si conosce a priori la metrica.
- **Compositi**: si sommano (o si fa la media) i punteggi grezzi delle sole osservate che fanno parte del fattore, usando le variabili monofattoriali più sature (se la saturazione è negativa, si inverte l'item, ad es. con `COMPUTE x=(max+1)-x.`). Se si usa la media e se le variabili hanno la stessa metrica, è semplice interpretare i punteggi ottenuti. Usando la somma vi è maggior variabilità, ma è si conosce comunque la gamma di oscillazione teorica.

Punteggi fattoriali manuali

- **standardizzati non centrati** (standardized, non centered factor scores; Thompson, 1993):
  1. Trasformare in punti z le variabili (standardizzati)
  2. Sommare ai punti z di ogni variabile, la media (rispetto ai casi) dei punti grezzi delle variabili (non centrati)
  3. Costruire i punteggi fattoriali ponderando i punti z non centrati con i valori della “Matrice dei coefficienti di punteggio fattoriale” ($\mathbf{W}$)

$$PF_1 = W_{11}x_1 + W_{21}x_2$$

Punteggi fattoriali manuali

Matrice dei coefficienti di punteggio fattoriale

	Fattore	
	1	2
I2	,001	,250
I3	-,065	,101
I4	,254	-,045
I5	,084	-,016

```
des I2 to I10 I12 / save.
* uso le medie di I2, I3 e così via.
comp cI2=zI2 + 5.30.
...
comp fcF1=(.001*cI2) + (-.065*cI3) + ...
comp fcF2=(.250*cI2) + (.101*cI3) + ...
```

Uso dei punteggi fattoriali

- I punteggi fattoriali si usano come qualsiasi altra variabile continua
- In particolare si usano proprio al posto delle singole variabile che li hanno prodotti per studiare
  - studiare l'affidabilità della scala (alfa di Cronbach)
  - differenze di medie (t-test, anove)
  - classificare i soggetti (cluster)
  - per spiegare altre variabili (regressione)

Interpretazione

- Per l'interpretazione dei fattori si usano le saturazioni (o pesi fattoriali)
- Tanto più una variabile ha una saturazione "alta" tanto più è influenzata dal fattore, tanto più il suo significato si avvicina al significato del fattore
- Non esistono regole "statistiche" per interpretare le saturazioni
- In una soluzione ortogonale, le saturazioni sono "correlazioni" e quindi si possono usare gli stessi criteri
- Overall e Klett (1972) suggeriscono un valore minimo di .35
- Tutte le variabili che saturano sul fattore (oltre un certo livello) si utilizzano
- variabili che non saturano nessun fattore possono essere ignorate
- anche variabili che saturano su più fattori possono essere ignorate se si vuole ottenere latenti "pulite"

Interpretazione

- L'interpretazione è facilitata che ci sono saturazioni negative perché la dimensione rappresentata dal fattore diventa bipolare
- Le saturazioni negative indicano solo che la variabile partecipa al significato del fattore in senso opposto
- "In diversi casi, l'unione delle variabili osservate non mette l'analista in grado di denominare il fattore, per cui si pone l'esigenza di astrarsi, cercando con la mente il 'fattore latente' sotto la combinazione delle variabili" (Fabbris, 1997, p. 196)