

Analisi Fattoriale

Analisi Esplorativa

Aldo Solari



Introduzione

- Nelle scienze sociali, in particolare in psicologia, spesso è problematico misurare le variabili di interesse direttamente. Ad esempio:
 - Intelligenza
 - Classe sociale
- Queste variabili non osservabili (variabili latenti) sono chiamate *fattori comuni*
- E' possibile esaminare queste variabili indirettamente, misurando variabili osservabili che sono ad esse collegate. Ad esempio
 - Punteggio in varie prove di intelligenza, etc.
 - Occupazione, Tasso di istruzione, Casa di proprietà, etc.
- L'obiettivo dell'*analisi fattoriale* è studiare le relazioni tra variabili osservabili e fattori comuni



Outline

① Il modello fattoriale

② Metodi di stima



Il modello fattoriale

$$\begin{aligned}x_1 &= \lambda_{11}f_1 + \dots + \lambda_{1k}f_k + u_1 \\x_2 &= \lambda_{21}f_1 + \dots + \lambda_{2k}f_k + u_2 \\&\vdots = \vdots \\x_p &= \lambda_{p1}f_1 + \dots + \lambda_{pk}f_k + u_p\end{aligned}$$

dove

- $x_{p \times 1} = (x_1, \dots, x_p)'$ sono le *variabili osservate* (variabili casuali)
- $f_{k \times 1} = (f_1, \dots, f_k)'$ sono i *fattori comuni* (var. casuali non oss.)
- $u_{p \times 1} = (u_1, \dots, u_p)'$ sono i *fattori specifici* (var. casuali non oss.)
- λ_{ij} sono i *pesi fattoriali* (costanti incognite)



Il modello fattoriale (in forma matriciale)

$$\boxed{\begin{matrix} x \\ p \times 1 \end{matrix} = \begin{matrix} \Lambda \\ p \times k \end{matrix} \begin{matrix} f \\ k \times 1 \end{matrix} + \begin{matrix} u \\ p \times 1 \end{matrix}}$$

Assunzioni

- Variabili osservate: $\mathbb{E}\left(\begin{matrix} x \\ p \times 1 \end{matrix}\right) = \begin{matrix} 0 \\ p \times 1 \end{matrix}$ (altrimenti centrare sullo 0)
- Fattori comuni: $\mathbb{E}\left(\begin{matrix} f \\ k \times 1 \end{matrix}\right) = \begin{matrix} 0 \\ k \times 1 \end{matrix}$, $\mathbb{Cov}\left(\begin{matrix} f \\ k \times 1 \end{matrix}\right) = \mathbb{E}\left(\begin{matrix} f & f' \\ k \times 1 & 1 \times k \end{matrix}\right) = \begin{matrix} I \\ k \times k \end{matrix}$
- Fattori specifici: $\mathbb{E}\left(\begin{matrix} u \\ p \times 1 \end{matrix}\right) = \begin{matrix} 0 \\ p \times 1 \end{matrix}$,
 $\mathbb{Cov}\left(\begin{matrix} u \\ p \times 1 \end{matrix}\right) = \mathbb{E}\left(\begin{matrix} u & u' \\ p \times 1 & 1 \times p \end{matrix}\right) = \begin{matrix} \Psi \\ p \times p \end{matrix} = \text{diag}(\psi_1, \dots, \psi_p)$
- Incorrelazione tra f e u : $\mathbb{Cov}\left(\begin{matrix} u \\ p \times 1 \end{matrix}, \begin{matrix} f \\ k \times 1 \end{matrix}\right) = \mathbb{E}\left(\begin{matrix} u & f' \\ p \times 1 & 1 \times k \end{matrix}\right) = \begin{matrix} 0 \\ p \times k \end{matrix}$



Matrice di varianza/covarianza Σ di x

$$\Sigma_{p \times p} = \Lambda_{p \times k} \Lambda'_{k \times p} + \Psi_{p \times p}$$

Dimostrazione:

$$\begin{aligned}\Sigma_{p \times p} &= \text{Cov}\left(\begin{matrix} x \\ p \times 1 \end{matrix}\right) = \mathbb{E}\left(\begin{matrix} x & x' \\ p \times 1 & 1 \times p \end{matrix}\right) \\ &= \mathbb{E}[(\Lambda f + u)(\Lambda f + u)'] \\ &= \mathbb{E}[\Lambda f(\Lambda f)' + u(\Lambda f)' + (\Lambda f)u' + uu'] \\ &= \Lambda \mathbb{E}(ff')\Lambda' + \mathbb{E}(uf')\Lambda' + \Lambda \mathbb{E}(fu') + \mathbb{E}(uu') \\ &= \Lambda \text{Cov}(f)\Lambda' + \text{Cov}(u, f)\Lambda' + \Lambda \text{Cov}(f, u) + \text{Cov}(u) \\ &= \Lambda\Lambda' + \Psi\end{aligned}$$



Numero di parametri

- Il modello fattoriale ipotizza che

$$p(p+1)/2$$

parametri corrispondenti alle p varianze e alle $p(p-1)/2$ covarianze di $\Sigma_{p \times p}$ possano essere espressi con

$$(p+1)k$$

parametri corrispondenti ai pk pesi fattoriali di $\Lambda_{p \times k}$ e le p varianze specifiche di $\Psi_{p \times p}$

- Per esempio, se abbiamo $p = 12$ variabili osservabili $x_{p \times 1}$ e un modello fattoriale con $k = 2$ fattori, allora i $p(p-1)/2 = 78$ parametri di $\Sigma_{p \times p}$ devono essere ridotti ai $(p+1)k = 36$ parametri di

$$\Lambda_{p \times k} \text{ e } \Psi_{p \times k}$$



Scomposizione della varianza di x_i

$$\begin{aligned}\sigma_{ii} &= \text{Var}(x_i) = \{\Sigma_{p \times p}\}_{ii} = \{\Lambda\Lambda'\}_{ii} + \{\Psi\}_{ii} \\ &= \sum_{j=1}^k \lambda_{ij}^2 + \psi_i \\ &= \underbrace{h_i^2}_{\text{comunalità}} + \underbrace{\psi_i}_{\text{var. specifica}}\end{aligned}$$

- $h_i^2 = \lambda_{i1}^2 + \dots + \lambda_{ik}^2$ è la comunalità, ovvero la varianza dovuta ai k fattori comuni
- ψ_i è la varianza specifica di x_i non attribuibile ai fattori comuni



Covarianza tra x_i e x_j

$$\begin{aligned}\sigma_{ij} &= \text{Cov}(x_i, x_j) = \{\Sigma_{p \times p}\}_{ij} = \{\Lambda\Lambda'\}_{ij} + \{\Psi\}_{ij} \\ &= \sum_{l=1}^k \lambda_{il}\lambda_{jl} \\ &= \lambda_{i1}\lambda_{j1} + \dots + \lambda_{ik}\lambda_{jk}\end{aligned}$$



Covarianza tra x e f

$$\begin{aligned}\mathbb{Cov}_{\substack{p \times 1 \\ k \times 1}}(x, f) &= \mathbb{E}_{\substack{p \times 1 \\ 1 \times k}}(x f') \\ &= \mathbb{E}[(\Lambda f + u) f'] \\ &= \Lambda \mathbb{E}(f f') + \mathbb{E}(u f') \\ &= \Lambda_{\substack{p \times k}}\end{aligned}$$



quindi il peso fattoriale λ_{ij} rappresenta la covarianza tra x_i e f_j :

$$\mathbb{Cov}(x_i, f_j) = \{\Lambda_{\substack{p \times k}}\}_{ij} = \lambda_{ij}$$



Trasformazioni di scala

- Assumiamo il modello fattoriale per x :

$$\underset{p \times 1}{x} = \underset{p \times k}{\Lambda} \underset{k \times 1}{f} + \underset{p \times 1}{u}$$

- Consideriamo una trasformazione di scala per x :

$$\underset{p \times 1}{y} = \underset{p \times p}{A} \underset{p \times 1}{x}$$

dove $\underset{p \times p}{A} = \text{diag}(a_1, \dots, a_p)$ è una trasformazione di scala

- Il modello fattoriale è ancora valido per y ?



Invarianza rispetto a trasformazioni di scala

- Abbiamo

$$\begin{aligned} \underset{p \times 1}{y} &= \underset{p \times p}{A} \underset{p \times 1}{x} \\ &= \underset{p \times p}{A} \left(\underset{p \times k}{\Lambda} \underset{k \times 1}{f} + \underset{p \times 1}{u} \right) \\ &= \underset{p \times p}{A} \underset{p \times k}{\Lambda} \underset{k \times 1}{f} + \underset{p \times p}{A} \underset{p \times 1}{u} \\ &= \underset{p \times k}{\Lambda_y} \underset{k \times 1}{f} + \underset{p \times 1}{u_y} \end{aligned}$$

e

$$\begin{aligned} \text{Cov}(y) &= \text{Cov}(Ax) = A \text{Cov}(x) A' = A \Sigma A' \\ &= A \Lambda \Lambda' A' + A \Psi A' \\ &= \Lambda_y \Lambda_y' + \Psi_y \end{aligned}$$

quindi il modello fattoriale è ancora valido per y con pesi fattoriali $\Lambda_y = A\Lambda$ e varianze specifiche $\Psi_y = A\Psi A'$



Modello fattoriale per la matrice di correlazione

- Il risultato precedente mostra che il modello fattoriale rimane essenzialmente inalterato se effettuiamo una trasformazione di scala
- La standardizzazione

$$\underset{p \times 1}{z} = \underset{p \times p}{D}^{-1/2} \underset{p \times 1}{x}$$

è un caso particolare di trasformazione di scala dove

$$D^{-1/2} = \text{diag}(1/\sqrt{\sigma_{11}}, \dots, 1/\sqrt{\sigma_{pp}})$$

- Questo significa che, invece di considerare la decomposizione della matrice di varianze/covarianze di x , $\mathbb{Cov}(x)$, possiamo considerare la decomposizione della matrice di correlazione di x , $\mathbb{Corr}(x)$, o equivalentemente, la decomposizione della matrice di varianze/covarianze di z , $\mathbb{Cov}(z) = D^{-1/2} \Sigma D^{-1/2} = \mathbb{Corr}(x)$
- Si noti che sebbene il modello fattoriale è invariante rispetto a trasformazioni di scala, la *stima* dei parametri potrebbe essere influenzata dalle trasformazioni di scala



Non-unicità dei pesi fattoriali

Sia $A_{k \times k}$ una matrice ortogonale: $AA' = A'A = I$

$$\begin{aligned} x_{p \times 1} &= \Lambda_{p \times k} f_{k \times 1} + u_{p \times 1} \\ &= \Lambda_{p \times k} A_{k \times k} A'_{k \times k} f_{k \times 1} + u_{p \times 1} \\ &= \Lambda^*_{p \times k} f^*_{k \times 1} + u_{p \times 1} \end{aligned}$$

- $\Lambda^*_{p \times k} = \Lambda_{p \times k} A_{k \times k}$
- $f^*_{k \times 1} = A'_{k \times k} f_{k \times 1}$
- $\mathbb{E}(f^*)_{k \times 1} = A'_{k \times k} \mathbb{E}(f)_{k \times 1} = 0_{k \times 1}$
- $\mathbb{Cov}(f^*)_{p \times p} = A'_{p \times k} \mathbb{Cov}(f)_{k \times k} A_{k \times p} = I_{p \times p}$
- $\mathbb{Cov}(x) = \Sigma = \Lambda \Lambda' + \Psi = \Lambda A A' \Lambda' + \Psi = \Lambda^* \Lambda^{*'} + \Psi$



Non-unicità dei pesi fattoriali

- Il risultato precedente mostra che il modello fattoriale con fattori comuni $f_{k \times 1}$ e pesi fattoriali $\Lambda_{p \times k}$, e il modello fattoriale con fattori comuni $f^*_{k \times 1}$ e pesi fattoriali $\Lambda^*_{p \times k}$ sono equivalenti per spiegare la matrice di varianza/covarianza Σ di $x_{p \times 1}$



Outline

① Il modello fattoriale

② Metodi di stima



Stima del modello fattoriale

Obiettivo

Determinare due matrici $\hat{\Lambda}$ e $\hat{\Psi}$ tali che $\widehat{\mathbb{Cov}(x)} = \hat{\Sigma} = S = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$
oppure

Determinare due matrici $\hat{\Lambda}$ e $\hat{\Psi}$ tali che $\widehat{\mathbb{Corr}(x)} = R = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$



Stima naïve

- Si consideri il seguente esempio: sulla base di un campione di voti di studenti su tre materie, x_1 (*Classics*), x_2 (*French*) e x_3 (*English*) si è ottenuta la seguente matrice di correlazione R

$$\mathbf{R} = \begin{matrix} & \begin{matrix} \text{Classics} \\ \text{French} \\ \text{English} \end{matrix} \end{matrix} \begin{pmatrix} 1.00 & & \\ 0.83 & 1.00 & \\ 0.78 & 0.67 & 1.00 \end{pmatrix}.$$

- Si consideri il modello fattoriale ad 1 fattore

$$x_1 = \lambda_1 f + u_1$$

$$x_2 = \lambda_2 f + u_2$$

$$x_3 = \lambda_3 f + u_3$$



Stima naïve

- Le sei equazioni derivanti dall'uguaglianza $R = \Lambda\Lambda' + \Psi$ sono

$$\mathbf{R} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} (\lambda_1 \ \lambda_2 \ \lambda_3) + \begin{pmatrix} \psi_1 & 0 & 0 \\ 0 & \psi_2 & 0 \\ 0 & 0 & \psi_3 \end{pmatrix},$$

are

$$\hat{\lambda}_1 \lambda_2 = 0.83,$$

$$\hat{\lambda}_1 \lambda_3 = 0.78,$$

$$\hat{\lambda}_1 \lambda_4 = 0.67,$$

$$\psi_1 = 1.0 - \hat{\lambda}_1^2,$$

$$\psi_2 = 1.0 - \hat{\lambda}_2^2,$$

$$\psi_3 = 1.0 - \hat{\lambda}_3^2.$$

The solutions of these equations are

$$\hat{\lambda}_1 = 0.99, \quad \hat{\lambda}_2 = 0.84, \quad \hat{\lambda}_3 = 0.79,$$

$$\hat{\psi}_1 = 0.02, \quad \hat{\psi}_2 = 0.30, \quad \hat{\psi}_3 = 0.38.$$



Casi di Heywood

Esempio 4 (tratto da Everitt e Dunn, 2001):

Si stimino i parametri del modello fattoriale ad un fattore per i punteggi ottenuti nelle tre materie:

x_1 : Classics, x_2 : French e x_3 : English.

$$x_1 = \lambda_1 f + u_1,$$

$$x_2 = \lambda_2 f + u_2,$$

$$x_3 = \lambda_3 f + u_3.$$

data la matrice di correlazione campionaria $R = \begin{matrix} & \begin{matrix} \text{Classics} & \text{French} & \text{English} \end{matrix} \\ \begin{matrix} \text{Classics} \\ \text{French} \\ \text{English} \end{matrix} & \begin{pmatrix} 1.00 & & \\ 0.84 & 1.00 & \\ 0.60 & 0.35 & 1.00 \end{pmatrix} \end{matrix}$

Si ottiene il sistema
$$\begin{aligned} \hat{\psi}_1 &= 1.0 - \hat{\lambda}_1^2, & \lambda_1 \lambda_2 &= 0.84 \\ \hat{\psi}_2 &= 1.0 - \hat{\lambda}_2^2, & \text{e } \lambda_3 \lambda_2 &= 0.35 \\ \hat{\psi}_3 &= 1.0 - \hat{\lambda}_3^2, & \lambda_1 \lambda_3 &= 0.6 \end{aligned}$$

da cui la soluzione
$$\begin{aligned} \hat{\lambda}_1 &= 1.2, & \hat{\lambda}_2 &= 0.7, & \hat{\lambda}_3 &= 0.5, \\ \hat{\psi}_1 &= -0.44, & \hat{\psi}_2 &= 0.51, & \hat{\psi}_3 &= 0.75 \end{aligned}$$

Stime di tipo intuitivo possono condurre a soluzioni non accettabili anche se il modello è identificato esattamente!



Modello ad un fattore: $\mathbb{C}orr(x)$

Esempio 3 (tratto da Hardle e Simar 2003)

Si stimi il modello fattoriale ad un fattore per la matrice di correlazione relativa a valutazioni fornite da 40 clienti su 25 tipologie di auto per le variabili X_1 : economicità, X_2 : accessori, X_3 : deprezzamento

$$\mathcal{R} = \begin{pmatrix} 1 & 0.975 & 0.613 \\ 0.975 & 1 & 0.620 \\ 0.613 & 0.620 & 1 \end{pmatrix}$$

Soluzione

$$\begin{pmatrix} 1 & r_{X_1X_2} & r_{X_1X_3} \\ r_{X_1X_2} & 1 & r_{X_2X_3} \\ r_{X_1X_3} & r_{X_2X_3} & 1 \end{pmatrix} = \mathcal{R} = \begin{pmatrix} \hat{\lambda}_1^2 & \hat{\lambda}_1\hat{\lambda}_2 & \hat{\lambda}_1\hat{\lambda}_3 \\ & \hat{\lambda}_2^2 & \hat{\lambda}_2\hat{\lambda}_3 \\ & & \hat{\lambda}_3^2 \end{pmatrix} + \begin{pmatrix} \hat{\psi}_1 & 0 & 0 \\ & \hat{\psi}_2 & 0 \\ & & \hat{\psi}_3 \end{pmatrix}$$

Da cui

$$\frac{r_{X_1X_2}r_{X_1X_3}}{r_{X_2X_3}} = (0.982)^2 = \hat{\lambda}_1^2, \quad \frac{r_{X_1X_2}r_{X_2X_3}}{r_{X_1X_3}} = (0.993)^2 = \hat{\lambda}_2^2, \quad \frac{r_{X_1X_3}r_{X_2X_3}}{r_{X_1X_2}} = (0.624)^2 = \hat{\lambda}_3^2$$

e

$$\psi_1 = 1 - \hat{\lambda}_1^2 = 0.035$$

$$\psi_2 = 0.014$$

$$\psi_3 = 0.610$$

Le prime due comunaltà sono ≈ 1 . X_1 e X_2 sono ben spiegate dal 1° fattore (economicità + accessori)



Modello ad un fattore: $\mathbb{C}ov(x)$

Example 11.1 Let $p = 3$ and $k = 1$, then $d = 0$ and

$$\Sigma = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{pmatrix} = \begin{pmatrix} q_1^2 + \psi_{11} & q_1 q_2 & q_1 q_3 \\ q_1 q_2 & q_2^2 + \psi_{22} & q_2 q_3 \\ q_1 q_3 & q_2 q_3 & q_3^2 + \psi_{33} \end{pmatrix}$$

with $\mathcal{Q} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix}$ and $\Psi = \begin{pmatrix} \psi_{11} & 0 & 0 \\ 0 & \psi_{22} & 0 \\ 0 & 0 & \psi_{33} \end{pmatrix}$. Note that here the constraint (11.8) is automatically verified since $k = 1$. We have

$$q_1^2 = \frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}; \quad q_2^2 = \frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}; \quad q_3^2 = \frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}$$

and

$$\psi_{11} = \sigma_{11} - q_1^2; \quad \psi_{22} = \sigma_{22} - q_2^2; \quad \psi_{33} = \sigma_{33} - q_3^2.$$

In this particular case ($k = 1$), the only rotation is defined by $\mathcal{G} = -1$, so the other solution for the loadings is provided by $-\mathcal{Q}$.



Vincoli

- Numero di parametri del modello fattoriale $\Lambda\Lambda' + \Psi : pk + p$
- Vincolo 1: $\Lambda'\Psi^{-1}\Lambda = \text{diag}(b_1, \dots, b_k)$ con $b_1 \geq \dots \geq b_k$
- Il Vincolo 1 impone $k(k-1)/2$ vincoli
- Numero di parametri del modello fattoriale $\Lambda\Lambda' + \Psi$ dato il Vincolo 1: $pk + p - k(k-1)/2$
- Come alternativa al Vincolo 1 si può considerare
Vincolo 2: $\Lambda'D^{-1}\Lambda = \text{diag}(c_1, \dots, c_k)$ con $c_1 \geq \dots \geq c_k$ e
 $D = \text{diag}(\sigma_{11}, \dots, \sigma_{pp})$



Gradi di libertà

- I *gradi di libertà* (= numero dei parametri “liberi”) sono dati dalla differenza tra i $p(p+1)/2$ parametri di $\sum_{p \times p}$ e il numero di parametri del modello fattoriale dato il Vincolo 1:

$$d = p(p+1)/2 - (pk + p - k(k-1)/2) = (p-k)^2/2 - (p+k)/2$$

- Se $d < 0$, allora il modello è indeterminato (ci sono infinite soluzioni)
- Se $d = 0$, allora la soluzione è unica (ma non necessariamente propria)
- $d > 0$, allora ci sono più equazioni che parametri: non c'è una soluzione esatta (ci si accontenta di una approssimazione)



Modello indeterminato

Example 11.2 Suppose now $p = 2$ and $k = 1$, then $d < 0$ and

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} = \begin{pmatrix} q_1^2 + \psi_{11} & q_1 q_2 \\ q_1 q_2 & q_2^2 + \psi_{22} \end{pmatrix}.$$

We have infinitely many solutions: for any α ($\rho < \alpha < 1$), a solution is provided by

$$q_1 = \alpha; \quad q_2 = \rho/\alpha; \quad \psi_{11} = 1 - \alpha^2; \quad \psi_{22} = 1 - (\rho/\alpha)^2.$$



Metodi di stima

Data $\hat{\Sigma}_{p \times p} = S_{p \times p}$ (oppure $= R_{p \times p}$), vogliamo stimare $\hat{\Psi}_{p \times p}$ e $\hat{\Lambda}_{p \times k}$ in modo tale che $\hat{\Sigma} \approx \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$ e sia rispettato il Vincolo 1 o 2

- Naïve (senza vincolo)
- *Componenti principali*
- *Fattori principali*
- *Massima Verosimiglianza* (richiede assunzione di Normalità per $x_{p \times 1}$)

Rotazione dei fattori

Dopo aver stimato il modello fattoriale, può essere utile ruotare i pesi fattoriali $\hat{\Lambda}$ per ottenere $\hat{\Lambda}^* = \hat{\Lambda}A$ (con A matrice ortogonale), al fine di trovare configurazioni più facilmente interpretabili

Numero di fattori

In pratica, dobbiamo anche determinare il valore di k



Metodo dei fattori principali

- Si parte da $\hat{\Sigma}_{p \times p} = \widehat{\text{Corr}(x)} = R_{p \times p}$ per trovare $\hat{\Psi}_{p \times p}$ e $\hat{\Lambda}_{p \times k}$ in modo tale che $R - \hat{\Psi} \approx \hat{\Lambda}\hat{\Lambda}'$ e sia rispettato il Vincolo 2
- $R^* = R - \hat{\Psi}$ è detta *matrice di correlazione ridotta*
- $\{\text{Corr}(x)\}_{ii} = 1 = h_i^2 + \psi_i$, quindi se abbiamo a disposizione una stima iniziale \hat{h}_i^2 , allora $\{R^*\}_{ii} = 1 - \hat{\psi}_i = \hat{h}_i^2$
- $R^* = R - \hat{\Psi}$ è una matrice simmetrica, quindi la sua decomposizione spettrale è $R^* = VLV'$ con $L = \text{diag}(l_1, \dots, l_p)$ e $V = [v_1, \dots, v_p]$. Se i primi k autovalori l_1, \dots, l_k sono positivi e i rimanenti $p - k$ autovalori l_{k+1}, \dots, l_p prossimi a 0, allora

$$R^* \approx V_k L_k V_k'$$

dove $V_k_{p \times k}$ contiene le prime k colonne di V e $L_k_{k \times k} = \text{diag}(l_1, \dots, l_k)$

- Segue $R^* = R - \hat{\Psi} \approx (V_k L_k^{1/2})(V_k L_k^{1/2})' \approx \hat{\Lambda}\hat{\Lambda}'$, quindi

$$\hat{\Lambda} \approx V_k L_k^{1/2}$$



Metodo dei fattori principali - inizializzazione

- Partire dalla stima R della matrice di correlazione $\text{Corr}(x)$
- Calcolare la stima iniziale \hat{h}_i^2 della comunaltà h_i^2 come
 - $\hat{h}_i^2 = \max_{j \neq i} |\widehat{\text{Corr}}(x_i, x_j)|$
 - $\hat{h}_i^2 = 1 - \frac{1}{r^{ii}}$ dove $r^{ii} = \{R^{-1}\}_{ii}$, che equivale il coefficiente di determinazione lineare multiplo tra x_i e x_{-i}
 $(p-1) \times 1$
- Ottenere la matrice di correlazione ridotta R^* da R ma sostituendo i valori 1 sulla diagonale con $\hat{h}_1^2, \dots, \hat{h}_p^2$



Metodo dei fattori principali - algoritmo iterativo

- ➊ $R^* \leftarrow R$ e poi $\{R^*\}_{ii} \leftarrow \hat{h}_i^2, i = 1, \dots, p$
- ➋ Ottenere la decomposizione spettrale $R^* = VLV'$
- ➌ Fissare k e determinare V_k e L_k
- ➍ Stimare Λ con $\hat{\Lambda} \approx V_k L_k^{1/2}$
- ➎ Aggiornare $\hat{h}_i^2 \leftarrow \sum_{j=1}^k \hat{\lambda}_{ij}^2$ e $\{R^*\}_{ii} \leftarrow \hat{h}_i^2$
- ➏ Ripetere i passi 2-5 fino a raggiungere convergenza

Output

$\hat{\Lambda}, \hat{h}_i^2$ e $\hat{\psi}_i = 1 - \hat{h}_i^2, i = 1, \dots, p$



Metodo dei fattori principali - vincolo 2

- $D = \text{diag}(\sigma_{11}, \dots, \sigma_{pp}) = I$ perchè consideriamo la matrice di correlazione
- Vincolo 2: $\hat{\Lambda}' \hat{D}^{-1} \hat{\Lambda} = \hat{\Lambda}' \hat{\Lambda} = \text{diag}(c_1, \dots, c_k)$ con $c_1 \geq \dots \geq c_k$
- Quindi $\hat{\Lambda}$ soddisfa il Vincolo 2 perchè

$$\hat{\Lambda}' \hat{\Lambda} = (V_k L_k^{1/2})' (V_k L_k^{1/2}) = L_k = \text{diag}(l_1, \dots, l_k)$$



Casi di Heywood

- Nella procedura di stima iterativa possono succedere casi di Heywood, ovvero $\hat{\psi}_i < 0$ oppure $\hat{\psi}_i > 1$
- $\hat{\psi}_i < 0$ non ha senso perchè ψ_i è una varianza, e quindi >0
- $\hat{\psi}_i > 1$ non ha senso perchè $\text{Var}(x_i) = 1$ è quindi $\psi_i \leq 1$

