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

Il modello fattoriale

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

dove

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

Il modello fattoriale (in forma matriciale)

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

Assunzioni

- Variabili osservate: $\mathbb{E}(\underset{p\times 1}{x}) = \underset{p\times 1}{0}$ (altrimenti centrare sullo 0)
- Fattori comuni: $\mathbb{E}(\underset{k\times 1}{f}) = \underset{k\times 1}{0}$, $\mathbb{C}\text{ov}(\underset{k\times 1}{f}) = \mathbb{E}(\underset{k\times 1}{f}\underset{1\times k}{f'}) = \underset{k\times k}{I}$
- Fattori specifici: $\mathbb{E}(\underset{p\times 1}{u}) = \underset{p\times 1}{0}$,
  $\mathbb{C}\text{ov}(\underset{p\times 1}{u}) = \mathbb{E}(\underset{p\times 1}{u}\underset{1\times p}{u'}) = \underset{p\times p}{\Psi} = diag(\psi_1, \ldots, \psi_p)$
- Incorrelazione tra f e u: $\mathbb{C}\text{ov}(\underset{p\times 1}{u}, \underset{k\times 1}{f}) = \mathbb{E}(\underset{p\times 1}{u}\underset{1\times k}{f'}) = \underset{p\times k}{0}$

Matrice di varianza/covarianza Σ di x

$$\boxed{\underset{p\times p}{\Sigma} = \underset{p\times k}{\Lambda}\underset{k\times p}{\Lambda'} + \underset{p\times p}{\Psi}}$$

Dimostrazione:

$$\begin{aligned}
\underset{p\times p}{\Sigma} &= \mathbb{C}\mathrm{ov}(\underset{p\times 1}{x}) = \mathbb{E}(\underset{p\times 1}{x}\underset{1\times p}{x'}) \\
&= \mathbb{E}[(\Lambda f + u)(\Lambda f + u)'] \\
&= \mathbb{E}[\Lambda f(\Lambda f)' + u(\Lambda f)' + (\Lambda f)u' + uu'] \\
&= \Lambda\mathbb{E}(ff')\Lambda' + \mathbb{E}(uf')\Lambda' + \Lambda\mathbb{E}(fu') + \mathbb{E}(uu') \\
&= \Lambda\mathbb{C}\mathrm{ov}(f)\Lambda' + \mathbb{C}\mathrm{ov}(u,f)\Lambda' + \Lambda\mathbb{C}\mathrm{ov}(f,u) + \mathbb{C}\mathrm{ov}(u) \\
&= \Lambda\Lambda' + \Psi
\end{aligned}$$

$\square$

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 $\underset{p\times p}{\Sigma}$ possano essere espressi con

$$(p+1)k$$

parametri corrispondenti ai pk pesi fattoriali di $\underset{p\times k}{\Lambda}$ e le p varianze specifiche di $\underset{p\times p}{\Psi}$

- Per esempio, se abbiamo $p = 12$ variabili osservabili $\underset{p\times 1}{x}$ e un modello fattoriale con $k = 2$ fattori, allora i $p(p+1)/2 = 78$ parametri di $\underset{p\times p}{\Sigma}$ devono essere ridotti ai $(p+1)k = 26$ parametri di $\underset{p\times k}{\Lambda}$ e $\underset{p\times k}{\Psi}$

Scomposizione della varianza di x_i

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

- $h_i^2 = \lambda_{i1}^2 + \ldots + \lambda_{ik}^2$ 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

$$\begin{aligned}
\sigma_{ij} &= \mathbb{C}\mathrm{ov}(x_i, x_j) = \{\underset{p\times p}{\Sigma}\}_{ij} = \{\Lambda\Lambda'\}_{ij} + \{\Psi\}_{ij} \\
&= \sum_{l=1}^{k} \lambda_{il}\lambda_{jl} \\
&= \lambda_{i1}\lambda_{j1} + \ldots + \lambda_{ik}\lambda_{jk}
\end{aligned}$$

$\square$

Covarianza tra x e f

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}(\underset{p\times 1}{x}, \underset{k\times 1}{f}) &= \mathbb{E}(\underset{p\times 1}{x}\underset{1\times k}{f'}) \\
&= \mathbb{E}[(\Lambda f + u) f'] \\
&= \Lambda \mathbb{E}(f f') + \mathbb{E}(u f') \\
&= \underset{p\times k}{\Lambda}
\end{aligned}
$$

□

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

$$
\mathbb{C}\mathrm{ov}(x_i, f_j) = \{\underset{p\times k}{\Lambda}\}_{ij} = \lambda_{ij}
$$

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

- Il modello fattoriale è ancora valido per y?

Invarianza rispetto a trasformazioni di scala

- Abbiamo

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

e

$$
\begin{aligned}
\mathbb{C}\mathrm{ov}(y) &= \mathbb{C}\mathrm{ov}(Ax) = A\mathbb{C}\mathrm{ov}(x)A' = A\Sigma A' \\
&= A\Lambda\Lambda' A' + A\Psi A' \\
&= \Lambda_y \Lambda_y' + \Psi_y
\end{aligned}
$$

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

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

- Questo significa che, invece di considerare la decomposizione della matrice di varianze/covarianze di x, $\mathbb{C}\mathrm{ov}(x)$, possiamo considerare la decomposizione della matrice di correlazione di x, $\mathbb{C}\mathrm{orr}(x)$, o equivalentemente, la decomposizione della matrice di varianze/covarianze di z, $\mathbb{C}\mathrm{ov}(z) = D^{-1/2}\Sigma D^{-1/2} = \mathbb{C}\mathrm{orr}(x)$
- 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 $\underset{k\times k}{A}$ una matrice ortogonale: $AA' = A'A = I$

$$\begin{aligned}
\underset{p\times 1}{x} &= \underset{p\times k}{\Lambda}\underset{k\times 1}{f} + \underset{p\times 1}{u} \\
&= \underset{p\times k}{\Lambda}\underset{k\times k}{A}\underset{k\times k}{A'}\underset{k\times 1}{f} + \underset{p\times 1}{u} \\
&= \underset{p\times k}{\Lambda^*}\underset{k\times 1}{f^*} + \underset{p\times 1}{u}
\end{aligned}$$

- $\underset{p\times k}{\Lambda^*} = \underset{p\times k}{\Lambda}\underset{k\times k}{A}$
- $\underset{k\times 1}{f^*} = \underset{k\times k}{A'}\underset{k\times 1}{f}$
- $\mathbb{E}(f^*) = A'\mathbb{E}(f) = \underset{k\times 1}{0}$
- $\mathbb{C}\mathrm{ov}(f^*) = A'\mathbb{C}\mathrm{ov}(f)A = \underset{p\times p}{I}$
- $\mathbb{C}\mathrm{ov}(x) = \Sigma = \Lambda\Lambda' + \Psi = \Lambda A A'\Lambda' + \Psi = \Lambda^*\Lambda^{*'} + \Psi$

$\square$

Non-unicità dei pesi fattoriali

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

Outline

1. Il modello fattoriale
2. Metodi di stima

Stima del modello fattoriale

Obiettivo
Determinare due matrici $\hat{\Lambda}$ e $\hat{\Psi}$ tali che $\widehat{\mathbb{C}\mathrm{ov}(x)} = \hat{\Sigma} = S = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$
oppure
Determinare due matrici $\hat{\Lambda}$ e $\hat{\Psi}$ tali che $\widehat{\mathbb{C}\mathrm{orr}(x)} = R = \hat{\Lambda}\hat{\Lambda}' + \hat{\Psi}$

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

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

- Si consideri il modello fattoriale ad 1 fattore

$$\begin{aligned} x_1 &= \lambda_1 f + u_1 \\ x_2 &= \lambda_2 f + u_2 \\ x_3 &= \lambda_3 f + u_3 \end{aligned}$$

Stima naïve

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

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

are

$$\begin{aligned} \hat{\lambda}_1 \lambda_2 &= 0.83, \\ \hat{\lambda}_1 \lambda_3 &= 0.78, \\ \hat{\lambda}_1 \lambda_4 &= 0.67, \\ \psi_1 &= 1.0 - \hat{\lambda}_1^2, \\ \psi_2 &= 1.0 - \hat{\lambda}_2^2, \\ \psi_3 &= 1.0 - \hat{\lambda}_3^2. \end{aligned}$$

The solutions of these equations are

$$\begin{aligned} \hat{\lambda}_1 &= 0.99, \quad \hat{\lambda}_2 = 0.84, \quad \hat{\lambda}_3 = 0.79, \\ \hat{\psi}_1 &= 0.02, \quad \hat{\psi}_2 = 0.30, \quad \hat{\psi}_3 = 0.38. \end{aligned}$$

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

$$\mathbf{R} = \begin{array}{c c} & \begin{array}{ccc} \text{Classics} & \text{French} & \text{English} \end{array} \\ \begin{array}{c} \text{Classics} \\ \text{French} \\ \text{English} \end{array} & \begin{pmatrix} 1.00 & & \\ 0.84 & 1.00 & \\ 0.60 & 0.35 & 1.00 \end{pmatrix} \end{array}.$$

Si ottiene il sistema

$$\hat{\psi}_1 = 1.0 - \hat{\lambda}_1^2, \quad \hat{\psi}_2 = 1.0 - \hat{\lambda}_2^2, \quad \hat{\psi}_3 = 1.0 - \hat{\lambda}_3^2.$$

e

$$\lambda_1\lambda_2 = 0.84$$
$$\lambda_3\lambda_2 = 0.35$$
$$\lambda_1\lambda_3 = 0.6$$

da cui la soluzione

$$\hat{\lambda}_1 = 1.2, \quad \hat{\lambda}_2 = 0.7, \quad \hat{\lambda}_3 = 0.5,$$
$$\hat{\psi}_1 = -0.44, \quad \hat{\psi}_2 = 0.51, \quad \hat{\psi}_3 = 0.75$$

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

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

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

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

e

$\psi_i = 1 - \hat{\lambda}_1^2 = 0.035$ $\qquad \psi_2 = 0.014$ $\qquad \psi_3 = 0.610$

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

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

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

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

and

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

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: $\Lambda'\Psi^{-1}\Lambda = diag(b_1, \ldots, b_k)$ con $b_1 \geq \ldots \geq b_k$
- 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: $\Lambda' D^{-1}\Lambda = diag(c_1, \ldots, c_k)$ con $c_1 \geq \ldots \geq c_k$ e
  $D = diag(\sigma_{11}, \ldots, \sigma_{pp})$

Gradi di libertà

- I *gradi di libertà* (= numero dei parametri "liberi") sono dati dalla differenza tra i $p(p+1)/2$ parametri di $\underset{p \times p}{\Sigma}$ 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

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

Metodi di stima

Data $\underset{p\times p}{\hat{\Sigma}} = \underset{p\times p}{S}$ (oppure $= \underset{p\times p}{R}$), vogliamo stimare $\underset{p\times p}{\hat{\Psi}}$ e $\underset{p\times k}{\hat{\Lambda}}$ 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 $\underset{p\times 1}{x}$)

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 $\underset{p\times p}{\hat{\Sigma}} = \widehat{\mathbb{C}\mathrm{orr}(x)} = \underset{p\times p}{R}$ per trovare $\underset{p\times p}{\hat{\Psi}}$ e $\underset{p\times k}{\hat{\Lambda}}$ 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*
- $\{\mathbb{C}\mathrm{orr}(x)\}_{ii} = 1 = h_i^2 + \psi_i$, 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 $L = diag(l_1, \ldots, l_p)$ e $V = [v_1, \ldots, v_p]$. Se i primi k autovalori $l_1, \ldots, l_k$ sono positivi e i rimanenti $p - k$ autovalori $l_{k+1}, \ldots, l_p$ prossimi a 0, allora

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

  dove $\underset{p\times k}{V_k}$ contiene le prime k colonne di V e $\underset{k\times k}{L_k} = diag(l_1, \ldots, l_k)$
- 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 $\mathbb{C}\mathrm{orr}(x)$
- Calcolare la stima iniziale $\hat{h}_i^2$ della comunalità h_i^2 come
  - $\hat{h}_i^2 = \max_{j \neq i} |\widehat{\mathbb{C}\mathrm{orr}}(x_i, x_j)|$
  - $\hat{h}_i^2 = 1 - \frac{1}{r^{ii}}$ dove $r^{ii} = \{R^{-1}\}_{ii}$, che equivale il coefficiente di determinazione lineare multiplo tra x_i e $\underset{(p-1)\times 1}{x_{-i}}$
- Ottenere la matrice di correlazione ridotta R^* da R ma sostituendo i valori 1 sulla diagonale con $\hat{h}_1^2, \ldots, \hat{h}_p^2$

Metodo dei fattori principali - algoritmo iterativo

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

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

Metodo dei fattori principali - vincolo 2

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

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

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 $\mathbb{V}\mathrm{ar}(x_i) = 1$ è quindi $\psi_i \leq 1$