El modelo factorial

El modelo factorial

- Supondremos que observamos un vector de variables $\mathbf{x}$, de dimensiones $p \times 1$, en n elementos de una población.

- El modelo de análisis factorial establece que este vector de datos observados se genera mediante la relación:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u}$$

- $\mathbf{f}$: es un vector $m \times 1$ de variables latentes o factores no observados. Supondremos que siguen una distribución $N_m(\mathbf{0}, I)$, es decir, los factores son variables de media cero e independientes entre sí y con distribución Normal.

- Λ: es una matriz de tamaño $p \times m$ de constantes desconocidas $(m < p)$. Contiene los coeficientes que describen cómo los factores $\mathbf{f}$ afectan a las variables observadas $\mathbf{x}$ y se denomina matriz de carga (loading matrix).

El modelo factorial

- Supondremos que observamos un vector de variables $\mathbf{x}$, de dimensiones $p \times 1$, en n elementos de una población.

- El modelo de análisis factorial establece que este vector de datos observados se genera mediante la relación:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u}$$

- $\mathbf{u}$: es un vector de tamaño $p \times 1$ de perturbaciones no observadas. Recoge el efecto de todas las variables distintas de los factores que influyen sobre $\mathbf{x}$. Supondremos que $\mathbf{u} \sim N_p(\mathbf{0}, \psi)$ donde ψ es una matriz diagonal, y también que las perturbaciones están incorreladas con los factores $\mathbf{f}$, es decir, $cov[\mathbf{f}, \mathbf{u}] = \mathbf{0}$.

El modelo factorial

Con estas tres hipótesis deducimos que:

- $\boldsymbol{\mu}$ es la media de las variables de las variables $\mathbf{x}$, ya que tanto los factores como las perturbaciones tienen media cero.
- $\mathbf{x}$ tiene distribución Normal, al ser suma de variables Normales, y llamando S a su matriz de covarianzas

$$\mathbf{x} \sim N_p(\boldsymbol{\mu}, S)$$

- La ecuación del modelo implica que dada una muestra aleatoria simple de n elementos generada por el modelo factorial, cada dato x_{ij} puede escribirse como:

$$x_{ij} = \mu_j + \lambda_{j1} f_{1i} + \cdots + \lambda_{jm} f_{mi} + u_{ij}$$

que descompone x_{ij}, el valor observado en el individuo i de la variable j, como suma de $m + 2$ términos.

El modelo factorial

- El primer término es la media de la variable j, μ_j
- Del segundo término al término $m+1$ recogen el efecto de los m factores.
- Y el último término, $m+2$, es una perturbación específica de cada observación, u_{ij}.
- Los efectos de los factores sobre x_{ij} son el producto de los coeficientes $\lambda_{j1}, \ldots, \lambda_{jm}$ que dependen de la relación entre cada factor y la variable j (y que son los mismos para todos los elementos de la muestra), por los valores de los m factores en el elemento muestral i, $f_{1i}, \ldots, f_{mi}$.
- Poniendo juntas las ecuaciones para todas las n observaciones, la matriz de datos $\mathbf{x}$ de tamaño $n \times p$ puede escribirse como:

$$\mathbf{x} = \mathbf{1}\boldsymbol{\mu}^t + \mathbf{F}\Lambda^t + \mathbf{U}$$

donde $\mathbf{1}$ es un vector $n \times 1$ de unos, $\mathbf{F}$ es una matriz $n \times m$ que contiene los m factores para los n elementos, Λ^t es la transpuesta de la matriz de carga que resulta ser $m \times p$ cuyos coeficientes constantes relacionan las variables y los factores, y $\mathbf{U}$ es una matriz $n \times p$ de perturbaciones.

Propiedades

- La matriz de carga Λ contiene las covarianzas entre los factores y las variables observadas.

- En efecto, la matriz de covarianzas Λ entre las variables y los factores se obtiene multiplicando por $\mathbf{f}^t$ por la derecha y tomando esperanzas en el modelo factorial:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u}$$

$$\mathbf{x} - \boldsymbol{\mu} = \Lambda\mathbf{f} + \mathbf{u}$$

$$(\mathbf{x} - \boldsymbol{\mu})\mathbf{f}^t = \Lambda\mathbf{f}\mathbf{f}^t + \mathbf{u}\mathbf{f}^t$$

$$\boldsymbol{E}[(\mathbf{x} - \boldsymbol{\mu})\mathbf{f}^t] = \Lambda\mathrm{E}[\mathbf{f}\mathbf{f}^t] + \boldsymbol{E}[\mathbf{u}\mathbf{f}^t] = \Lambda$$

- Ya que por hipótesis, los factores están incorrelados, es decir, $\mathrm{E}[\mathbf{f}\mathbf{f}^t] = I$, y tienen media cero y están incorrelados con las perturbaciones, es decir, $cov[\mathbf{u},\mathbf{f}] = \boldsymbol{E}[\mathbf{u}\mathbf{f}^t] = 0$.

Propiedades

- Entonces Λ se define como:

$$\Lambda = \boldsymbol{E}[(\mathbf{x} - \boldsymbol{\mu})\mathbf{f}^t]$$

- Esta ecuación indica que los términos λ_{ij} de la matriz de carga Λ, representan la covarianza entre la variable $\mathbf{x}_i$ y el factor $\mathbf{f}_j$ y al tener los factores varianza unidad, son los coeficientes de regresión cuando explicamos las variables observadas por los factores.

- En el caso particular en que las variables $\mathbf{x}$ estén estandarizadas, los términos λ_{ij} son también las correlaciones entre las variables y los factores.

Propiedades

- La matriz de covarianzas entre las observaciones verifica, según

$$S = E[(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^{\boldsymbol{t}}] = \Lambda E[\mathbf{f}\mathbf{f}^{\mathrm{t}}]\Lambda^{\mathrm{t}} + E[\mathbf{u}\mathbf{u}^{t}]$$

- Ya que $E[\mathbf{f}\mathbf{u}^{t}] = 0$ al estar incorrelados los factores y el ruido.
- Entonces, se obtiene la propiedad fundamental:

$$S = \Lambda\Lambda^{t} + \psi$$

Que establece que la matriz de covarianzas de los datos observados admite una descomposición como suma de dos matrices:

1. La primera $\Lambda\Lambda^{t}$ es una matriz simétrica de rango $m < p$. Esta matriz contiene la parte común al conjunto de las variables y depende de las covarianzas entre las variables y los factores.
2. La segunda ψ, es una matriz diagonal que contiene la parte específica de cada variable, que es independiente del resto.

Propiedades

- Esta descomposición $S = \Lambda\Lambda^t + \psi$ implica que las varianzas de las variables observadas pueden descomponerse como:

$$\sigma_i^2 = \sum_{j=1}^{m} \lambda_{ij}^2 + \psi_i^2$$

con $i = 1, \dots, p$, donde el primer término es la suma de los efectos de los factores y el segundo término es el efecto de la perturbación.

- Llamando:

$$h_i^2 = \sum_{j=1}^{m} \lambda_{ij}^2$$

a la suma de los efectos de los factores que llamaremos "comunalidad", tenemos que:

$$\sigma_i^2 = h_i^2 + \psi_i^2$$

para $i = 1, \dots, p$

Propiedades

- En esta descomposición:

$$\sigma_i^2 = h_i^2 + \psi_i^2$$

- h_i^2 se llama la i −ésima *"comunalidad".*

- ψ_i^2 es la varianza del i −ésimo elemento de la perturbación $\mathbf{u}$ y se llaman *"unicidades". O especificidades*

Propiedades

- Esta igualdad puede interpretarse como una descomposición de la varianza en:

$$\sigma_i^2 = h_i^2 + \psi_i^2$$

$$Varianza\ observada = Variabilidad\ común\ (Comunalidad) + Variabilidad\ específica\ (Unicidad)$$

- Esto es análoga a la descomposición clásica de la variabilidad de los datos en una parte explicada y otra no explicada que se realiza en el análisis de la varianza.

- En el modelo factorial la parte explicada es debida a los factores y la no explicada es debido al ruido o componente aleatorio.

Unicidad del modelo

- En el modelo factorial, ni la matriz de carga, Λ, ni los factores, $\mathbf{f}$, son observables.

- Esto plantea un problema de indeterminación: dos representaciones $(\Lambda, \mathbf{f})$ y $(\Lambda^*, \mathbf{f}^*)$ serán equivalentes si

$$\Lambda\mathbf{f} = \Lambda^*\mathbf{f}^*$$

- Esta situación conduce a dos tipos de indeterminación.
  1. Un conjunto de datos puede explicarse con la misma precisión con factores incorrelados o correlados.
  2. Los factores no quedan determinados de manera única.

Unicidad del modelo

- Vamos a analizar estas dos indeterminaciones. Para mostrar la primera, si H es cualquier matriz no singular, la representación usual puede también escribirse como

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda H H^{-1}\mathbf{f} + \mathbf{u}$$

- Y llamando $\Lambda^{*} = \Lambda H$ a la nueva matriz de carga y $\mathbf{f}^{*} = H^{-1}\mathbf{f}$ a los nuevos factores:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda^{*}\mathbf{f}^{*} + \mathbf{u}$$

- Donde los nuevos factores tienen ahora una distribución:

$$\mathbf{f}^{*} \sim N(\mathbf{0}, H^{-1}(H^{-1})^{t})$$

- Y por lo tanto están correlados.

Unicidad del modelo

- Análogamente, partiendo de factores correlados $\mathbf{f} \sim N(\mathbf{0}, S_f)$ siempre podemos encontrar una expresión equivalente de las variables mediante un modelo con factores incorrelados.

- En efecto, sea A una matriz tal que $S_f = AA^t$, que siempre existe si S_f es definida positiva, entonces $\mathrm{A}^{-1} S_f \left(\boldsymbol{A}^{-\mathbf{1}}\right)^{\boldsymbol{t}} = \boldsymbol{I}$.
- Entonces:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda A A^{-1} \mathbf{f} + \mathbf{u}$$

- Y tomando $\Lambda^* = \Lambda A$ como la nueva matriz de coeficientes de los factores y $\mathbf{f}^* = A^{-1}\mathbf{f}$ como los nuevos factores, el modelo es equivalente a otro con factores incorrelados.

- Esta indeterminación se ha resuelto en las hipótesis del modelo tomando siempre los factores como incorrelados.

Unicidad del modelo

- En segundo lugar, si H es ortogonal, los modelos:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u}$$

$$\text{y}$$

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda H H^{\mathrm{t}}\mathbf{f} + \mathbf{u}$$

son indistinguibles porque $H = H^t = H^{-1}$ y $HH^{\mathrm{t}} = I$

- Ambos contienen factores incorrelados, con matriz de covarianzas la identidad.

- En este sentido, decimos que el modelo factorial está indeterminado ante rotaciones.

- Esta indeterminación se puede resolver imponiendo restricciones sobre los componentes de la matriz de carga.

Ejemplo

- Supongamos que $\mathbf{x} = [\mathbf{x_1}, \mathbf{x_2}, \mathbf{x}_3]$ y el modelo factorial M_1 siguiente:

$$\mathbf{x} = \boldsymbol{\mu} + \Lambda\mathbf{f} + \mathbf{u} = \begin{pmatrix} \mathbf{0} \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ 0 & \mathbf{1} \\ 1 & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{f_1} \\ \mathbf{f_2} \end{pmatrix} + \begin{pmatrix} \mathbf{u_1} \\ \mathbf{u}_2 \\ \mathbf{u_3} \end{pmatrix}$$

- Supongamos que los factores están incorrelados.

- Vamos a escribirlo como otro modelo equivalente de factores también incorrelados.

- Tomando $H = 1/\sqrt{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, esta matriz es ortogonal porque $H = H^t = H^{-1}$.

Ejemplo

- Entonces:

$$\mathbf{x} = \begin{pmatrix} \mathbf{1} & \mathbf{1} \\ 0 & \mathbf{1} \\ 1 & \mathbf{0} \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{f_1} \\ \mathbf{f_2} \end{pmatrix} + \begin{pmatrix} \mathbf{u_1} \\ \mathbf{u}_2 \\ \mathbf{u_3} \end{pmatrix}$$

- Llamando a este modelo M_2, puede escribirse como:

$$\mathbf{x} = \begin{pmatrix} \frac{\mathbf{2}}{\sqrt{2}} & \mathbf{0} \\ \frac{\mathbf{1}}{\sqrt{2}} & -\frac{\mathbf{1}}{\sqrt{2}} \\ \frac{\mathbf{1}}{\sqrt{2}} & \frac{\mathbf{1}}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} \mathbf{g_1} \\ \mathbf{g_2} \end{pmatrix} + \begin{pmatrix} \mathbf{u_1} \\ \mathbf{u}_2 \\ \mathbf{u_3} \end{pmatrix}$$

Ejemplo

- Los nuevos factores $\mathbf{g}$ están relacionados con los anteriores por:

$$\begin{pmatrix} \mathbf{g_1} \\ \mathbf{g_2} \end{pmatrix} = \frac{\mathbf{1}}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} \mathbf{f_1} \\ \mathbf{f_2} \end{pmatrix}$$

- Y son por lo tanto una rotación de los factores iniciales.
- Comprobemos que estos nuevos factores están también incorrelados. Su matriz de varianzas es:

$$S_g = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} S_f \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

- Y si $S_f = I$, entonces $S_g = I$ porque $HH^{-1} = I$.
- Entonces, los modelos M_1 y M_2 son indistinguibles.