Bayesian Factor Model

J. C. Martínez-Ovando

1. Estructura y especificación

Suponemos que los datos $\{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_T\}$ toman valores en $\mathbb{R}^p$ (con $p < \infty$).

Para cualquier numero $k \leq p$ se tiene que el modelo de factores queda especificado como

$$\begin{aligned}
\boldsymbol{y}_t | \boldsymbol{f}_t &\sim \mathrm{N}\left(\boldsymbol{y}_t | \boldsymbol{B}\boldsymbol{f}_t, \boldsymbol{\Sigma}\right) \\
\boldsymbol{f}_t &\sim \mathrm{N}\left(\boldsymbol{f}_t | \boldsymbol{0}, \boldsymbol{I}\right),
\end{aligned}$$

para $t = 1, \ldots, T$, donde

- $\boldsymbol{f}_t$ es el vector de factores k-dimensional asociado con $\boldsymbol{y}_t$
- $\boldsymbol{B}$ es una matriz de dimensión $(p \times k)$ con los vectores de cargas asociados con los factores $\boldsymbol{f}_t$ para $\boldsymbol{y}_t$
- $\boldsymbol{\Sigma}$ es una matriz $\mathrm{diag}\{\sigma_1^2, \ldots, \sigma_p^2\}$ positivo definida de dimensión $(p \times p)$
- $\boldsymbol{I}$ es la matriz identidad de dimensión $(k \times k)$.

De la especificación anterior se sigue que la varianza no condicional de las observaciones es,

$$var(\boldsymbol{y}_t | \boldsymbol{B}, \boldsymbol{\Sigma}) = \boldsymbol{B}\boldsymbol{B}' + \boldsymbol{\Sigma}.$$

El modelo implica que, condicional a los factores comunes, las variables observables no estn correlacionadas: por lo tanto, los factores comunes explican toda la estructura de dependencia entre las p variables. lo anterior motiva para considerar:

- la sucesión de factores $\{\boldsymbol{f}_t\}_{t=1}^T$ como un conjunto de variables latentes, y
- al conjunto $\{\boldsymbol{B}, \boldsymbol{\Sigma}\}$ como los parámetros del modelo.

La verosimilitud extendida del modelo (i.e. verosimilitud para parámetros y variables latentes), se define como

$$\begin{aligned}
lik\left(\boldsymbol{B}, \boldsymbol{\Sigma}, \{\boldsymbol{f}_t\}_{t=1}^T | \{\boldsymbol{y}_t\}_{t=1}^T\right) &= p\left(\{\boldsymbol{y}_t\}_{t=1}^T | \boldsymbol{B}, \boldsymbol{\Sigma}, \{\boldsymbol{f}_t\}_{t=1}^T\right) \\
&\propto \det(\boldsymbol{\Sigma})^{-T/2} \mathrm{etr}\left\{-(1/2)\boldsymbol{\Sigma}^{-1}\boldsymbol{e}\boldsymbol{e}'\right\},
\end{aligned}$$

donde $\mathrm{etr}\{\boldsymbol{A}\} = \exp\{\mathrm{tr}(A)\}$ para toda matriz $\boldsymbol{A}$.

1.1. Distribución inicial

1.2. Distribución final

Algorithm 1: Gibbs sampler para el modelo de factores

Inicialización: $\boldsymbol{B}^{(0)}$, $\boldsymbol{\Sigma}^{(0)}$, Y $\{\boldsymbol{f}_t^{(0)}\}_{t=1}^{T}$

for k *en* $1 : M$ **do**

  simular $\boldsymbol{B}^{(k)}$ de

$$p\left(\boldsymbol{B}|\boldsymbol{\Sigma}^{(k-1)}, \{\boldsymbol{f}_t^{(k-1)}\}_{t=1}^{T}, \{\boldsymbol{y}_t\}_{t=1}^{T}\right)$$

  simular $\boldsymbol{\Sigma}^{(k)}$ de

$$p\left(\boldsymbol{\Sigma}|\boldsymbol{B}^{(k)}, \{\boldsymbol{f}_t^{(k-1)}\}_{t=1}^{T}, \{\boldsymbol{y}_t\}_{t=1}^{T}\right)$$

  simular $\{\boldsymbol{f}_t^{(k)}\}_{t=1}^{T}$ de

$$p\left(\{\boldsymbol{f}_t\}_{t=1}^{T}|\boldsymbol{B}^{(k)}, \boldsymbol{\Sigma}^{(k)}, \{\boldsymbol{y}_t\}_{t=1}^{T}\right)$$

end

return $\left\{\boldsymbol{B}^{(k)}, \boldsymbol{\Sigma}^{(k)}, \{\boldsymbol{f}_t^{(k)}\}_{t=1}^{T}\right\}_{k=1}^{M}$