

# 1 Modelo

Para un conjunto de observaciones  $p$ -dimensionales,  $(y_t)_{t=1}^n$ , organizados en forma matricial en  $Y$  matriz de dimensión  $n \times p$ , se especifica el modelo bayesiano de factores como,

$$y_t | \beta, f_t, \Sigma \sim N_p(y | \beta' f_t, \Sigma) \quad (1)$$

donde  $\beta$  es la matriz ortogonal (columnas) de  $k$  cargas de dimensión  $p \times k$ ,  $f_t$  es un vector de factores latentes  $k$ -dimensional, para  $t = 1, \dots, n$  y  $\Sigma$  es una matriz positiva definida simétrica de dimensión  $p \times p$ . Adicionalmente, el modelo supone,

$$f_f \sim N_k(f | 0, I_k),$$

donde  $I_k$  denota la matriz diagonal de dimensión  $k \times k$ .

En este modelo, la varianza marginal de  $y_t$  se calcula como  $\Omega = \Sigma + \beta\beta'$ .

Los parámetros del modelo son  $\beta$  y  $\Sigma$ , mientras que las variables latentes con  $(f_t)_{t=1}^n$ .

## 1.1 Prior

Para hacer inferencia bayesiana, sobre la clase de parámetros, se necesita complementar el modelo con la distribución inicial. Tradicionalmente se supone que  $\beta$  y  $\Sigma$  son mutuamente independientes entre sí, por lo que la prior puede descomponerse como

$$\pi(\beta, \Sigma) = \pi(\beta)\pi(\Sigma)$$

donde

$$\pi(\beta) = \prod_{i=1}^p \prod_{j=1}^k \pi(\beta_{ij}),$$

donde  $\pi(\beta_{ij}) = N(\beta_{ij} | 0, c)$ , si  $i \neq j$ , y  $\pi(\beta_{ij}) = N(\beta_{ij} | 0, c) \mathbb{I}(\beta_{ij} > 0)$ , si  $i = j$ , con  $c > 0$  dado, y

$$\pi(\Sigma) = Wi - Inv(\Sigma | n_0, S_0),$$

con  $n_0$  y  $S_0$  dados.

## 1.2 Posterior

Para un conjunto de datos observados,  $(y_t)_{t=1}^n$ , la distribución final sobre parámetros y variables latentes tiene el siguiente kernel asociado,

$$\pi(\beta, \Sigma, (f_t)_{t=1}^n | \text{datos}) \propto \prod_{t=1}^n N_p(y_t | \beta f_t, \Sigma) N_k(f_t | 0, I_k) \times \prod_{i=1}^p \prod_{j=1}^k \pi(\beta_{ij}) Wi - Inv(\Sigma | n_0, S_0),$$

para la cual la constante de normalización no es obtenible analíticamente.

---

## 2 Algoritmo

En esta sección describimos un algoritmo de Gibbs sampler para obtener datos simulados de la distribución posterior del modelo para parámetros y variables latentes.

Los parámetros del algoritmo son los siguientes:

1.  $M$  que el número de datos por simular en el Gibbs sampler
2.  $c$ ,  $n_0$  y  $S_0$  que son los hiperparámetros de la *prior*

El input del algoritmo, son los datos  $Y$ .

El output del algoritmo es la colección de  $M$  simulaciones de la distribución posterior, i.e.

$$\left\{ \beta^{(m)}, \Sigma^{(m)}, (f_t^{(m)})_{t \geq 1} \right\}_{m=1}^M,$$

las cuales se obtienen del siguiente proceso iterativo:

1. En el paso inicial, para  $m = 0$ , fijar valores iniciales  $\beta^{(0)}$ ,  $\Sigma^{(0)}$ , y  $(f_t^{(0)})_{t \geq 1}$ .  
–En teoría, el desempeño del algoritmo debe ser irrestricto a la elección de estos valores iniciales.–
2. En las iteraciones subsecuentes del algoritmo, para  $m = 1, 2, \dots, M$ , se simulan los datos por bloques, para los parámetros,

$$\beta^{(m)} | \Sigma^{(m-1)}, (f_t^{(m-1)})_{t \geq 1} \sim \prod_{ij} \pi(\beta_{ij} | \dots) \quad (2)$$

$$\Sigma^{(m)} | \beta^{(m)}, (f_t^{(m-1)})_{t \geq 1} \sim \prod_{ij} \pi(\sigma_{ij} | \dots), \quad (3)$$

las cuales pueden factorisarse de manera simple, y para las variables latentes, incorporamos un proceso iterativo adicional sobre las  $ts$ , como

$$f_t | \dots \sim N_k(f_t | At(\beta^{(m)}) \text{solve}(\Sigma^{(m)}) y_t, A) \quad (4)$$

donde

$$A = (I_k + t(\beta^{(m)}) \text{solve}(\Sigma^{(m)}) \beta^{(m)})^{-1},$$

con  $t(\beta^{(m)})$  denotando la transpuesta de  $\beta^{(m)}$  y  $\text{solve}(\Sigma^{(m)})$  la inversa de  $\Sigma^{(m)}$ .

En cada iteración es recomendable almacenar las variables simuladas en repositorios independientes, para posteriormente hacer inferencia y/o predicción con ellas.