## 2 Muestreo de Gibbs

Al igual que el algoritmo de Metropolis, el algoritmo de Gibbs permite simular una cadena de Markov $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots$ con distribución de equilibrio $p(\boldsymbol{\theta}|\boldsymbol{x})$. En este caso, sin embargo, cada nuevo valor de la cadena se obtiene a través de un proceso iterativo que sólo requiere generar muestras de distribuciones cuya dimensión es menor que $d$ y que en la mayoría de los casos tienen una forma más sencilla que la de $p(\boldsymbol{\theta}|\boldsymbol{x})$.

Sea $\boldsymbol{\theta} = (\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_k)$ una partición del vector $\boldsymbol{\theta}$, donde $\boldsymbol{\theta}_i \in \mathbb{R}^{d_i}$ y $\sum_{i=1}^k d_i = d$. Las densidades

$$\begin{array}{c} p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_k, \boldsymbol{x}) \\ \vdots \\ p(\boldsymbol{\theta}_i|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_{i+1}, \ldots, \boldsymbol{\theta}_k, \boldsymbol{x}) \qquad (i = 2, \ldots, k-1) \\ \vdots \\ p(\boldsymbol{\theta}_k|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{k-1}, \boldsymbol{x}) \end{array}$$

se conocen como *densidades condicionales completas* y en general pueden identificarse fácilmente al inspeccionar la forma de la distribución final $p(\boldsymbol{\theta}|\boldsymbol{x})$. De hecho, para cada $i = 1, \ldots, k$,

$$p(\boldsymbol{\theta}_i|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{i-1}, \boldsymbol{\theta}_{i+1}, \ldots, \boldsymbol{\theta}_k, \boldsymbol{x}) \propto p(\boldsymbol{\theta}|\boldsymbol{x}),$$

donde $p(\boldsymbol{\theta}|\boldsymbol{x}) = p(\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_k|\boldsymbol{x})$ es vista sólo como función de $\boldsymbol{\theta}_i$.

Dado un valor inicial $\boldsymbol{\theta}^{(0)} = (\boldsymbol{\theta}_1^{(0)}, \ldots, \boldsymbol{\theta}_k^{(0)})$, el algoritmo de Gibbs simula una cadena de Markov en la que $\boldsymbol{\theta}^{(t+1)}$ se obtiene a partir de $\boldsymbol{\theta}^{(t)}$ de la siguiente manera:

generar una observación $\boldsymbol{\theta}_1^{(t+1)}$ de $p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2^{(t)}, \boldsymbol{\theta}_3^{(t)}, \ldots, \boldsymbol{\theta}_k^{(t)}, \boldsymbol{x})$;

generar una observación $\boldsymbol{\theta}_2^{(t+1)}$ de $p(\boldsymbol{\theta}_2|\boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_3^{(t)}, \ldots, \boldsymbol{\theta}_k^{(t)}, \boldsymbol{x})$;

$\vdots$

generar una observación $\boldsymbol{\theta}_k^{(t+1)}$ de $p(\boldsymbol{\theta}_k|\boldsymbol{\theta}_1^{(t+1)}, \boldsymbol{\theta}_2^{(t+1)}, \ldots, \boldsymbol{\theta}_{k-1}^{(t+1)}, \boldsymbol{x})$.

La sucesión $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots$ así obtenida es entonces una realización de una cadena de Markov cuya distribución de transición está dada por

$$P(\boldsymbol{\theta}^{(t+1)}|\boldsymbol{\theta}^{(t)}) = \prod_{i=1}^{k} p(\boldsymbol{\theta}_i^{(t+1)}|\boldsymbol{\theta}_1^{(t+1)}, \ldots, \boldsymbol{\theta}_{i-1}^{(t+1)}, \boldsymbol{\theta}_{i+1}^{(t)}, \ldots, \boldsymbol{\theta}_k^{(t)}, \boldsymbol{x}).$$

*Comentario.* En ocasiones la distribución final implica cierta estructura de independencia condicional entre algunos de los elementos del vector $\boldsymbol{\theta}$. Es estos casos es común que muchas de las densidades condicionales completas se simplifiquen.

**Ejemplo 5.1** Consideremos el modelo jerárquico definido por

I. $p(\boldsymbol{x}|\boldsymbol{\omega}) = \prod_{i=1}^{m} p(x_i|\omega_i)$;

II. $p(\boldsymbol{\omega}|\phi) = \prod_{i=1}^{m} p(\omega_i|\phi)$;

III. $p_0(\phi)$.

Esta estructura define un modelo para $\boldsymbol{x}$ parametrizado por $\boldsymbol{\theta} = (\boldsymbol{\omega}, \phi) = (\omega_1, \ldots, \omega_m, \phi)$y con distribución inicial $p(\boldsymbol{\theta}) = p_0(\phi)\, p(\boldsymbol{\omega}|\phi)$, de manera que la distribución final está dada por

$$p(\boldsymbol{\theta}|\boldsymbol{x}) \propto p_0(\phi) \prod_{i=1}^{m} \{p(x_i|\omega_i)\, p(\omega_i|\phi)\}.$$

Entonces $k = m + 1$ y las densidades condicionales completas toman la forma

$$\begin{aligned}
p(\boldsymbol{\theta}_1|\boldsymbol{\theta}_2, \ldots, \boldsymbol{\theta}_{k-1}, \boldsymbol{\theta}_k, \boldsymbol{x}) &= p(\omega_1|\phi, x_1) \\
&\vdots \\
p(\boldsymbol{\theta}_{k-1}|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{k-2}, \boldsymbol{\theta}_k, \boldsymbol{x}) &= p(\omega_m|\phi, x_m) \\
p(\boldsymbol{\theta}_k|\boldsymbol{\theta}_1, \ldots, \boldsymbol{\theta}_{k-2}, \boldsymbol{\theta}_{k-1}, \boldsymbol{x}) &= p_0(\phi|\boldsymbol{\omega}),
\end{aligned}$$

donde $p_0(\phi|\boldsymbol{\omega}) \propto p_0(\phi)\, p(\boldsymbol{\omega}|\phi)$. ◁