Cómputo científico para probabilidad y estadística. Tarea 8. MCMC: MH con Kerneles Híbridos y Gibbs Sampler

Juan Esaul González Rangel

Noviembre 2023

1. Aplique el algoritmo de Metropolis-Hastings considerando como función objetivo la distribución normal bivariada

$$f_{X_1,X_2}(\bar{x}) = \frac{1}{2\pi}|\Sigma|^{-1/2}\exp\left\{-\frac{1}{2}(\bar{x}-\mu)'\Sigma^{-1}(\bar{x}-\mu)\right\}$$

donde,

$$\mu = \begin{pmatrix}\mu_1\\ \mu_2\end{pmatrix} \quad \Sigma = \begin{pmatrix}\sigma_1^2 & \rho\sigma_1\sigma_2\\ \rho\sigma_1\sigma_2 & \sigma_2^2\end{pmatrix}$$

Así, se tienen las siguientes distribuciones condicionales:

$$X_1|X_2 = x_2 \sim N\left(\mu_1 + \rho\frac{\sigma_1}{\sigma_2}(x_2-\mu_2), \sigma_1^2(1-\rho^2)\right)$$

$$X_2|X_1 = x_1 \sim N\left(\mu_2 + \rho\frac{\sigma_2}{\sigma_1}(x_1-\mu_1), \sigma_2^2(1-\rho^2)\right)$$

Considere las siguientes propuestas:

$$q_1((x_1', x_2')|(x_1, x_2)) = f_{X_1|X_2}(x_1'|x_2)\mathbb{1}_{(x_2'=x_2)}$$

$$q_2((x_1', x_2')|(x_1, x_2)) = f_{X_2|X_1}(x_2'|x_1)\mathbb{1}_{(x_1'=x_1)}$$

A partir del algoritmo MH usando Kerneles híbridos simule valores de la distribución normal bivariada, fijando $\sigma_1 = \sigma_2 = 1$, considere los casos $\rho = 0.8$ y $\rho = 0.95$[1].

2. Considere los tiempos de falla $t1, \dots, t_n$ con distribución *Weibull*(α, λ):

$$f(t_i|\alpha,\lambda) = \alpha\lambda t_i^{\alpha-1}e^{-t_i^{\alpha}\lambda}$$

Se asumen como a priori $\alpha \sim \exp(c)$ y $\lambda|\alpha \sim$Gama(α, b), por lo tanto, $f(\alpha,\lambda) = f(\lambda|\alpha)f(\alpha)$[2]. Así, para la disitribución posterior se tiene:

$$f(\alpha,\lambda|\bar{t}) \propto f(\bar{t}|\alpha,\lambda)f(\alpha,\lambda)$$

[1] Ver la tesis de Cricelio Montesinos para una explicación más extensa del Gibbs, Montesinos, C (2016) "Distribución de Direcciones en el Gibbs Sampler Generalizad", MSc Dissertation, CIMAT. `https://www.cimat.mx/es/Tesis_digitales/`. También vean la Enciclopedia de Estadística de Wiley, la entrada de *Gibbs Sampler*: `https://www.cimat.mx/~jac/2016WileytStatsRef_GibbsSampling.pdf.`

[2] Este ejemplo aparece en Kundu, D. (2008), "Bayesian Inference and Life Testing Plan for the Weibull Distribution in Presence of Progressive Censoring", Technometrics, 50(2), 144–154.

A partir del algoritmo MH usando Kerneles híbridos simule valores de la distribución posterior $f(\alpha, \lambda|\bar{t})$, considerando las siguientes propuestas:

Propuesta 1:

$$\lambda_p|\alpha, \bar{t} \sim Gama\left(\alpha + n, b + \sum_{i=1}^{n} t_i^{\alpha}\right) \quad \text{y dejando } \alpha \text{ fijo.}$$

Propuesta 2:

$$\alpha_p|\lambda, \bar{t} \sim Gama(n+1, -\log(b) - \log(r_1) + c), \text{ con } r_1 = \prod_{i=1}^{n} t_i \text{ y dejando } \lambda \text{ fijo.}$$

Propuesta 3:

$\alpha_p \sim \exp(c)$ y $\lambda_p|\alpha_p \sim Gama(\alpha_p, b)$.

Propuesta 4 (RWMH):

$\alpha_p = \alpha + \varepsilon$, con $\varepsilon \sim N(0, \sigma)$ y dejando λ fijo. Simular datos usando $\alpha = 1$ y $\lambda = 1$ con $n = 20$. Para la a priori usar $c = 1$ y $b = 1$.

3. Considere el ejemplo referente al número de fallas de bombas de agua en una central nuclear[3], donde p_i representa el número de fallas en el tiempo de operación t_i, con $i = 1, \ldots, n$.

   Se considera el modelo $p_i \sim$Poisson$(\lambda_i t_i)$, (las λ_i son independientes entre si), con distribuciones a priori $\lambda_i|\beta \sim Gama(\alpha, \beta)$ y $\beta \sim Gama(\gamma, \delta)$, por lo tanto:

$$f(\lambda_1, dots, \lambda_n, \beta) = f(\lambda_1|\beta) f(\lambda_2|\beta) \ldots f(\lambda_n|\beta) f(\beta)$$

   Para la distribución posterior se tiene:

$$f(\lambda_1, \ldots, \lambda_n, \beta|\bar{p}) \propto L(\bar{p}, \bar{\lambda}, \beta) f(\lambda_1, \ldots, \lambda_n, \beta)$$

   Simule valores de la distribución posterior $f(\lambda_1, \ldots, \lambda_n, \beta|\bar{p})$, usando un kernel híbrido, considerando las propuestas:

$$\lambda_i|\bar{\lambda}_{-i}, \beta, \bar{t} \sim Gama(p_i + \alpha, \beta + t_i)$$

Bomba (i)	1	2	3	4	5	6	7	8	9	10
T. de uso (t_i)	94.32	15.72	62.88	125.76	5.24	31.44	1.05	1.05	2.1	10.48
# de fallas (p_i)	5	1	5	14	3	17	1	1	4	22

Tabla 1: Datos de bombas de agua en centrales nucleares (Robert y Casella, p. 385) para el ejemplo 8.3.

$$\beta|\bar{\lambda}, \bar{t} \sim Gama\left(n\alpha + \gamma, \delta + \sum_{i=1}^{n} \lambda_i\right).$$

   Verifique que estas son propuestas Gibbs.

   Use los datos del Cuadro 1 con los parámetros a priori $\alpha = 1.8, \gamma = 0.01$ y $\delta = 1$.

[3]Este ejemplo fue usado en el artículo original del Gibbs sampler del Gelfand y Smith (1990). Vea también Norton, R.A., Christen, J.A. y Fox, C. (2017), "Sampling hyperparameters in hierarchical models: improving on Gibbs for high-dimensional latent fields and large data set" Communications in Statistics - Simulation and Computation, `http://dx.doi.org/10.1080/03610918.2017.1353618`