MAT 653: Statistical Simulation

Instructor: Dr. Wei Li Scribe: Jianchen Wei

Nov 30th, 2021

Gibbs Sampling

Recall: the block-wise Metropolis-Hastings. To sample from $f(X,Y)$. We used two candidates: $q_X(x|X,Y), q_Y(y|X,Y)$.

1. Update X:

$$\begin{aligned}\gamma_X(X^*|X^{(s)},Y^{(s)}) &= \frac{f(X^*,Y^{(s)})}{f(X^{(s)},Y^{(s)})}\cdot\frac{q_X(X^{(s)}|X^{(s)},Y^{(s)})}{q_X(X^*|X^{(s)},Y^{(s)})}\\ &= \frac{f_{X|Y}(X^*|Y^{(s)})}{f_{X|Y}(X^{(s)}|Y^{(s)})}\cdot\frac{q_X(X^{(s)}|X^{(s)},Y^{(s)})}{q_X(X^*|X^{(s)},Y^{(s)})}\end{aligned}$$

Suppose we take $q_X(x|X^{(s)},Y^{(s)})$ to be full condition distribution $f_{X|Y}(x|Y^{(s)})$, take $q_Y(y|X^{(s)},Y^{(s)})$ to be full condition distribution $f_{Y|X}(y|X^{(s)})$, then

$$\begin{aligned}\gamma_X(X^*|X^{(s)},Y^{(s)}) &= \frac{f(X^*,Y^{(s)})}{f(X^{(s)},Y^{(s)})}\cdot\frac{q_X(X^{(s)}|X^{(s)},Y^{(s)})}{q_X(X^*|X^{(s)},Y^{(s)})}\\ &= \frac{f_{X|Y}(X^*|Y^{(s)})}{f_{X|Y}(X^{(s)}|Y^{(s)})}\cdot\frac{f_{X|Y}(X^{(s)}|Y^{(s)})}{f_{X|Y}(X^*|Y^{(s)})}\\ &= 1\end{aligned}$$

2. Update Y. Similar argument goes through if we use the full conditional distributions as our candidate transition densities.

Definition: Gibbs sampling is a special case of (blockwise) MH that take $q_X(x|X_c,Y_c)$ to be $f_{X|Y}(x|Y_c)$, and take $q_Y(y|X_c,Y_c)$ to be $f_{Y|X}(y|X_c)$, which then yield $\gamma_X=\gamma_Y=1$. So following this sampling scheme, all proposals are automatically accepted.

Algorithm: Two stages Gibbs sampling

Target: $f(X,Y)$, possibly unnormalized

Take $x^{(0)}$. For $s=1,2,\cdots$, generate

- $Y^{(s)}\sim f_{Y|X}(\cdot|X^{(s-1)})$
- $X^{(s)}\sim f_{X|Y}(\cdot|Y^{(s)})$

Algorithm: Multi-stage Gibbs sampling

Target: $f(X_1,X_2,\cdots,X_d)$

Starting Values: $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \cdots, X_d^{(0)})$. Let $f_j(X_j|X_{-j})$ denote the conditional density of X_j given all the rest components $X_{-j} := \{X_i : i \neq j\}$.

The algorithm generates $X^{(s)}$ from $X^{(s-1)}$ as follows:

$$
\begin{aligned}
&(1) \quad X_1^{(s)} \sim f_1(\cdot|X_2^{(s-1)}, \cdots, X_d^{(s-1)}) \\
&(2) \quad X_2^{(s)} \sim f_2(\cdot|X_1^{(s)}, X_3^{(s-1)}, \cdots, X_d^{(s-1)}) \\
&\quad\ \vdots \\
&(d) \quad X_d^{(s)} \sim f_d(\cdot|X_1^{(s)}, \cdots, X_{d-1}^{(s)})
\end{aligned}
$$

Repeat $s \leftarrow s+1$.

Advantage: For high-dim problem, all the situation can be univariate and all probabilities are accepted.

Remark:

- The Gibbs sampler is a composition of MH moves with accept probability = 1. Each move is reversible, but the composition itself is not.
- Both (Blockwise) MH and Gibbs sampling have the target distribution as the invariant distribution (steady-state distribution).

Example: Two-stage Gibbs sampling

$$f(x) \propto \frac{e^{-x^2/20}}{(1+(z_1-x)^2)(1+(z_2-x)^2)}, \quad z_1 = -4.3, z_2 = 5.2$$

Note that: $\dfrac{1}{1+(z_i-x)^2} = \displaystyle\int_0^\infty e^{-w_i(1+(z_i-x)^2)}\, dw_i$, then we can write

$$f(x, w_1, w_2) \propto e^{-x^2/20} \prod_{i=1}^{2} e^{-w_i(1+(z_i-x)^2)}$$

so $f(x)$ is just the marginal pdf of $f(x, w_1, w_2)$.

Gibbs sampling: $\vec{w} = (w_1, w_2)$

- $X^{(s)} \sim f_{X|\vec{w}}(\cdot|\vec{w}^{(s-1)})$
- $\vec{w} \sim f_{\vec{w}|X}(\cdot|X^{(s)})$

Here

$$f_{X|\vec{w}}(x|w_1, w_2) \overset{x}{\propto} e^{-\left(\sum w_i\right)x^2 + 2x\sum w_i z_i} \cdot e^{-x^2/20} \sim N\left(\frac{\sum w_i z_i}{\sum w_i + 1/20}, \frac{1}{2\left(\sum w_i + 1/20\right)}\right)$$

$$f_{\vec{w}|X}(w_1, w_2|x) \overset{\vec{w}}{\propto} e^{-w_1(1+(z_1-x)^2)} \cdot e^{-w_2(1+(z_2-x)^2)}$$

Note from the factorization above, w_1, w_2 are independent given x, therefore,

$$f(w_1|X^{(s)}) \overset{w_1}{\propto} \text{exponential}(1 + (z_1 - X^{(s)})^2)$$
$$f(w_2|X^{(s)}) \overset{w_2}{\propto} \text{exponential}(1 + (z_2 - X^{(s)})^2)$$

Multivariate Normal

Bivariate normal

$$(X, Y) \sim N\left(\begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}, \begin{pmatrix}\sigma_X^2 & \sigma_{XY} \\ \sigma_{XY} & \sigma_Y^2\end{pmatrix}\right),$$

Let $\rho = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$ be the correlation. It is a fact that

$$f_{X|Y}(x|y) \sim N(\mu_X + \rho\frac{\sigma_X}{\sigma_Y}(y - \mu_Y), \sigma_X^2(1 - \rho^2))$$

$$f_{Y|X}(y|x) \sim N(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X), \sigma_Y^2(1 - \rho^2))$$