Department of Statistics and Data Science at SUSTech

MAT7035: Computational Statistics

Tutorial 9: MCMC Methods (I): IBF and Its Generalizations

A. Inverse Bayes Formulae (IBF)

A.1 Three forms of IBF

Given two conditional densities $f_{(X|Y)}(x|y)$ and $f_{(Y|X)}(y|x)$, find the marginal density $f_X(x)$ under the assumption of $\mathcal{S}_{(X,Y)} = \mathcal{S}_X \times \mathcal{S}_Y$, where $\mathcal{S}_X$, $\mathcal{S}_Y$ and $\mathcal{S}_{(X,Y)}$ denote the marginal supports of X, Y and the joint support of (X, Y).

(a) Point-wise IBF

$$f_X(x) = \left\{ \int_{\mathcal{S}_Y} \frac{f_{(Y|X)}(y|x)}{f_{(X|Y)}(x|y)} \mathrm{d}y \right\}^{-1}, \tag{9.1}$$

for any $x \in \mathcal{S}_X$.

(b) Function-wise IBF

$$f_X(x) = \left\{ \int_{\mathcal{S}_X} \frac{f_{(X|Y)}(x|y_0)}{f_{(Y|X)}(y_0|x)} \mathrm{d}x \right\}^{-1} \frac{f_{(X|Y)}(x|y_0)}{f_{(Y|X)}(y_0|x)}, \tag{9.2}$$

for any $x \in \mathcal{S}_X$ and an arbitrarily fixed $y_0 \in \mathcal{S}_Y$.

(c) Sampling IBF

$$f_X(x) \propto \frac{f_{(X|Y)}(x|y_0)}{f_{(Y|X)}(y_0|x)}, \tag{9.3}$$

for any $x \in \mathcal{S}_X$ and an arbitrarily fixed $y_0 \in \mathcal{S}_Y$.

A.2 Discrete versions of (9.1) & (9.3)

$$\Pr(X = x) = \left\{ \sum_{y \in \mathcal{S}_Y} \frac{\Pr(Y = y|X = x)}{\Pr(X = x|Y = y)} \right\}^{-1}, \quad \text{for any } x \in \mathcal{S}_X. \tag{9.4}$$

$$\Pr(X = x) \propto \frac{\Pr(X = x|Y = y_0)}{\Pr(Y = y_0|X = x)}, \tag{9.5}$$

for any $x \in \mathcal{S}_X$ and an arbitrarily fixed $y_0 \in \mathcal{S}_Y$.

A.3 Why do they have the name of IBF?

- In Bayesian statistics, $\pi(\theta)$ is the prior density θ, $p(\theta|x)$ is the posterior density of θ after that the x was observed, and the **Bayes formula** is as follow:

$$p(\theta|x) = \frac{f(x|\theta)\pi(\theta)}{f(x)} = \frac{f(x|\theta)\pi(\theta)}{\int f(x|\theta)\pi(\theta)\mathrm{d}\theta}, \tag{9.6}$$

- (9.6) states that given $f(x|\theta)$ and $\pi(\theta)$, the posterior density $p(\theta|x)$ can be determined uniquely;

- while (9.1) – (9.5) state that given both conditional densities, the marginal density can also be determined uniquely. This is why the name of IBF is taken.

Example T9.1 (Bivariate discrete distribution). Let X be a discrete random variable with probability mass function (pmf) $p_i = \Pr(X = x_i)$ for $i = 1, 2, 3$ and Y be a discrete random variable with pmf $q_j = \Pr(Y = y_j)$ for $j = 1, 2, 3, 4$. Given two conditional distribution matrices

$$\boldsymbol{A} = \begin{pmatrix} 1/7 & 1/4 & 3/7 & 1/7 \\ 2/7 & 1/2 & 1/7 & 2/7 \\ 4/7 & 1/4 & 3/7 & 4/7 \end{pmatrix} \quad \text{and} \quad \boldsymbol{B} = \begin{pmatrix} 1/6 & 1/6 & 1/2 & 1/6 \\ 2/7 & 2/7 & 1/7 & 2/7 \\ 1/3 & 1/12 & 1/4 & 1/3 \end{pmatrix},$$

where the (i, j) element of $\boldsymbol{A}$ is $a_{ij} = \Pr(X = x_i|Y = y_j)$ and the (i, j) element of $\boldsymbol{B}$ is $b_{ij} = \Pr(Y = y_j|X = x_i)$.

(a) Find the marginal distribution of X.

(b) Find the marginal distribution of Y.

(c) Find the joint distribution of (X, Y).

Solution: (a) The support of X and Y are $\mathcal{S}_X = \{x_1, x_2, x_3\}$ and $\mathcal{S}_Y = \{y_1, y_2, y_3, y_4\}$. By using

(9.5) with $y_0 = y_3$, the X-marginal is given by

$$
\begin{aligned}
p_1 &\hat{=} \Pr(X = x_1) = f_X(x_1) \\
&\propto \frac{f_{(X|Y)}(x_1|y_0)}{f_{(Y|X)}(y_0|x_1)} = \frac{\Pr(X = x_1|Y = y_3)}{Pr(Y = y_3|X = x_1)} \\
&= \frac{a_{13}}{b_{13}} = \frac{3/7}{1/2} = \frac{6}{7}, \\
p_2 &\hat{=} \Pr(X = x_2) = f_X(x_2) \\
&\propto \frac{f_{(X|Y)}(x_2|y_0)}{f_{(Y|X)}(y_0|x_2)} = \frac{\Pr(X = x_2|Y = y_3)}{Pr(Y = y_3|X = x_2)} \\
&= \frac{a_{23}}{b_{23}} = \frac{1/7}{1/7} = \frac{7}{7}, \\
p_3 &\hat{=} \Pr(X = x_3) = f_X(x_3) \\
&\propto \frac{f_{(X|Y)}(x_3|y_0)}{f_{(Y|X)}(y_0|x_3)} = \frac{\Pr(X = x_3|Y = y_3)}{Pr(Y = y_3|X = x_3)} \\
&= \frac{a_{33}}{b_{33}} = \frac{3/7}{1/4} = \frac{12}{7}.
\end{aligned}
$$

Note that $p_1 + p_2 + p_3 = 1$, we obtain

$$
\begin{aligned}
p_1 &= \frac{6/7}{6/7 + 7/7 + 12/7} = \frac{6}{6 + 7 + 12} = \frac{6}{25} = 0.24, \\
p_2 &= \frac{7/7}{6/7 + 7/7 + 12/7} = \frac{7}{6 + 7 + 12} = \frac{7}{25} = 0.28, \\
p_3 &= \frac{12/7}{6/7 + 7/7 + 12/7} = \frac{12}{6 + 7 + 12} = \frac{12}{25} = 0.48,
\end{aligned}
$$

which are summarized into

X	x_1	x_2	x_3
$p_i = \Pr(X = x_i)$	0.24	0.28	0.48

(b) Similarly, letting $x_0 = x_3$ in (9.5) yields the following Y-marginal

Y	y_1	y_2	y_3	y_4
$q_j = \Pr(Y = y_j)$	0.28	0.16	0.28	0.28

(c) The joint distribution of (X, Y) is given by

$$
\boldsymbol{P} = \begin{pmatrix} 0.04 & 0.04 & 0.12 & 0.04 \\ 0.08 & 0.08 & 0.04 & 0.08 \\ 0.16 & 0.04 & 0.12 & 0.16 \end{pmatrix}. \qquad \|
$$

B. Generalizations of IBF

B.1 Monte Carlo versions of IBF

(a) Harmonic mean formula

For any given $x \in \mathcal{S}_X$, we have

$$f_X(x) \doteq \hat{f}_{1,X}(x) = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{f_{(X|Y)}(x|y^{(i)})} \right\}^{-1},$$

where $\{y^{(i)}\}_{i=1}^{n} \overset{\text{iid}}{\sim} f_{(Y|X)}(y|x)$.

Remark: The harmonic mean HM of the positive real numbers $x_1, x_2, \ldots, x_n$ is defined to be

$$\text{HM} = \frac{n}{\frac{1}{x_1} + \cdots + \frac{1}{x_n}} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{1}{x_i} \right\}^{-1}.$$

(b) Weighted point-wise IBF

- Let the weight $w(y)$ be a density function with the same support as $f_Y(y)$. We have

$$w(y) = \frac{f_{(Y|X)}(y|x)w(y)}{f_{(X|Y)}(x|y)f_Y(y)} \cdot f_X(x), \tag{9.7}$$

for all $x \in \mathcal{S}_X$ and $y \in \mathcal{S}_Y$. Integrating (9.7) wth respect to y on $\mathcal{S}_Y$ gives

$$f_X(x) = \left\{ \int_{\mathcal{S}_Y} \frac{f_{(Y|X)}w(y)}{f_{(X|Y)}(x|y)f_Y(y)\mathrm{d}y} \right\}^{-1}$$

for any given $x \in \mathcal{S}_X$, which is called weighted point-wise IBF.

- First Monte Carlo version:

$$f_X(x) \doteq \hat{f}_{2,X}(x) = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{w(y^{(i)})}{f_{(X|Y)}(x|y^{(i)})f_Y(y^{(i)})} \right\}^{-1}$$

where x is a given point in $\mathcal{S}_X$, $\{y^{(i)}\}_1^n \overset{\text{iid}}{\sim} f_{(Y|X)}(y|x)$, and $f_Y(\cdot)$ is calculated by the point-wise IBF (9.1) via interchanging x and y.

- Second Monte Carlo version:

$$f_X(x) \doteq \hat{f}_{3,X}(x) = \left\{ \frac{1}{n} \sum_{i=1}^{n} \frac{f_{(Y|X)}(y^{(j)}|x)}{f_{(X|Y)}(x|y^{(i)})f_Y(y^{(i)})} \right\}^{-1}$$

where x is a given point in $\mathcal{S}_X$, $\{y^{(i)}\}_1^n \overset{\text{iid}}{\sim} w(y)$, and $f_Y(\cdot)$ is calculated by the point-wise IBF (9.1) via interchanging x and y.

B.2 IBF for three vectors

To extend the IBF from two vectors to three vectors, consider three random vectors X_1, X_2 and X_3. Let $S_{(X_1,X_2,X_3)} = S_{X_1} \times S_{X_2} \times S_{X_3}$. Assume that three conditional densities

$$f_1(x_1|x_2,x_3), \quad f_2(x_2|x_1,x_3) \quad \text{and} \quad f_3(x_3|x_1,x_2)$$

are given and positive. We want to find the joint density as

$$f(x_1,x_2,x_3) = f_{X_1}(x_1) f_{(X_2|X_1)}(x_2|x_1) f_3(x_3|x_1,x_2).$$

Thus we only need to derive $f_{X_1}(x_1)$ and $f_{(X_2|X_1)}(x_2|x_1)$.

(a) By the point-wise IBF (9.1),

$$\begin{aligned}
f_{(X_2|X_1)}(x_2|x_1) &= \left\{ \int_{\mathcal{S}_{X_3}} \frac{f_3(x_3|x_1,x_2)}{f_2(x_2|x_1,x_3)} \mathrm{d}x_3 \right\}^{-1}, \\
f_{(X_1|X_2)}(x_1|x_2) &= \left\{ \int_{\mathcal{S}_{X_3}} \frac{f_3(x_3|x_1,x_2)}{f_1(x_1|x_2,x_3)} \mathrm{d}x_3 \right\}^{-1}, \quad \text{and} \\
f_{X_1}(x_1) &= \left\{ \int_{\mathcal{S}_{X_2}} \frac{f_{(X_2|X_1)}(x_2|x_1)}{f_{(X_1|X_2)}(x_1|x_2)} \mathrm{d}x_2 \right\}^{-1},
\end{aligned}$$

for any $x_1 \in \mathcal{S}_{X_1}$ and any $x_2 \in \mathcal{S}_{X_2}$

(b) By the function-wise IBF (9.2),

$$\begin{aligned}
f_{(X_2|X_1)}(x_2|x_1) &= \left\{ \int_{\mathcal{S}_{X_2}} \frac{f_2(x_2|x_1,x_{3,0})}{f_3(x_{3,0}|x_1,x_2)} \mathrm{d}x_2 \right\}^{-1} \frac{f_2(x_2|x_1,x_{3,0})}{f_3(x_{3,0}|x_1,x_2)}, \quad \text{and} \\
f_{(X_1|X_2)}(x_1|x_2) &= \left\{ \int_{\mathcal{S}_{X_1}} \frac{f_1(x_1|x_2,x_{3,0})}{f_3(x_{3,0}|x_1,x_2)} \mathrm{d}x_1 \right\}^{-1} \frac{f_1(x_1|x_2,x_{3,0})}{f_3(x_{3,0}|x_1,x_2)},
\end{aligned}$$

for any $x_1 \in \mathcal{S}_{X_1}$, any $x_2 \in \mathcal{S}_{X_2}$ and $x_{3,0}$ is arbitrarily fixed in $\mathcal{S}_{X_3}$. Then

$$f_{X_1}(x_1) = \left\{ \int_{\mathcal{S}_{X_1}} \frac{f_{(X_1|X_2)}(x_1|x_{2,0})}{f_{(X_2|X_1)}(x_{2,0}|x_1)} \mathrm{d}x_1 \right\}^{-1} \frac{f_{(X_1|X_2)}(x_1|x_{2,0})}{f_{(X_2|X_1)}(x_{2,0}|x_1)},$$

for any $x_1 \in \mathcal{S}_{X_1}$ and $x_{2,0}$ is arbitrarily fixed in $\mathcal{S}_{X_2}$.

(c) By the sampling IBF (9.3),

$$
\begin{aligned}
f_{(X_2|X_1)}(x_2|x_1) &\propto \frac{f_2(x_2|x_1, x_{3,0})}{f_3(x_{3,0}|x_1, x_2)}, \quad \text{and} \\
f_{(X_1|X_2)}(x_1|x_2) &\propto \frac{f_1(x_1|x_2, x_{3,0})}{f_3(x_{3,0}|x_1, x_2)},
\end{aligned}
$$

for any $x_1 \in \mathcal{S}_{X_1}$, any $x_2 \in \mathcal{S}_{X_2}$ and $x_{3,0}$ is arbitrarily fixed in $\mathcal{S}_{X_3}$. Then

$$
f_{X_1}(x_1) \propto \frac{f_{(X_1|X_2)}(x_1|x_{2,0})}{f_{(X_2|X_1)}(x_{2,0}|x_1)},
$$

for any $x_1 \in \mathcal{S}_{X_1}$ and $x_{2,0}$ is arbitrarily fixed in $\mathcal{S}_{X_2}$.