# Exercise 1 (Gibbs Sampler)

1. First, let $X' = X^{(t)}$, $X = X^{(t-1)}$, $Y' = Y^{(t)}$ and $Y = Y^{(t-1)}$ . Then:
$$K^S((x,y),(x',y')) = \pi_{Y|X}(y' \mid x)\pi_{X|Y}(x' \mid y')$$
Then, to show that it is not reversible:
$$\pi(x,y)K((x,y),(x',y')) = \pi(x,y)\pi(y' \mid x)\pi(x' \mid y')$$
$$\pi(x',y')K((x',y'),(x,y)) = \pi(x',y')\pi(y \mid x')\pi(x \mid y)$$
$$\therefore$$
$$\frac{\pi(x,y)K((x,y),(x',y'))}{\pi(x',y')K((x',y'),(x,y))} = \frac{\pi(x,y)\pi(y' \mid x)\pi(x' \mid y')}{\pi(x',y')\pi(y \mid x')\pi(x \mid y)} = \frac{\pi(y)\pi(y' \mid x)}{\pi(y')\pi(y \mid x')} \neq 1$$
Therefore, it is not reversible. $\square$

2. First, the kernel expression:
$$K(x,x') = \int \pi(y' \mid x)\pi(x' \mid y')dy'$$
Now, let's show that it is $\pi_X$-reversible.
$$\pi(x')K(x',x) = \pi(x')\int \pi(y \mid x')\pi(x \mid y)dy = \int \pi(x')\frac{\pi(y,x')}{\pi(x')}\pi(x \mid y)dy =$$
$$= \int \pi(y,x')\pi(x \mid y)dy = \int \pi(y,x')\frac{\pi(x,y)}{\pi(y)}dy = \int \pi(x,y)\frac{\pi(x',y)}{\pi(y)}dy =$$
$$= \pi(x)\int \pi(y \mid x)\pi(x' \mid y)dy = \pi(x)K(x,x')$$
$\square$

3. First, the kernel expression is:
$$K^R((x,y),(x',y')) = \pi(y' \mid x)\pi(x' \mid y')0.5 + \pi(x' \mid y)\pi(y' \mid x')0.5$$
Note that it is half the density of sampling first from $y$ plus half the density of sampling first from $x$.

Now, let's show that it is reversible:
$$\frac{\pi(x,y)[\pi(y' \mid x)\pi(x' \mid y')0.5 + \pi(x' \mid y)\pi(y' \mid x')0.5]}{\pi(x',y')[\pi(y \mid x')\pi(x \mid y')0.5 + \pi(x \mid y')\pi(y \mid x)0.5]} =$$
$$= \frac{\frac{\pi(y'|x)}{\pi(y')} + \frac{\pi(x'|y)}{\pi(x')}}{\frac{\pi(y|x')}{\pi(y)} + \frac{\pi(x|y')}{\pi(x)}} = 1$$
$\square$

# Exercise 2 (Metropolis-within-Gibbs)

1. Note that:

$$\alpha(X_1 \mid X_1^{(t-1)}, X_2^{(t-2)}) = min\left\{1, \frac{\pi(X_1', X_2^{(t-1)})\pi(X_1^{(t-1)} \mid X_2^{(t-1)})}{\pi(X_1^{(t-1)}, X_2^{(t-1)})\pi(X_1' \mid X_2^{(t-1)})}\right\} = min\{1, 1\}$$

Therefore, we get a systematic scan Gibbs sampler, where one samples $X_1^t \sim \pi(\cdot \mid X_2^{(t-1)})$, then we accept, since $\alpha = 1$, and finally sample $X_2^t \sim \pi(\cdot \mid X_1^{(t)})$. □

2. First, let's write the kernel. Since we only accept or reject the variabel $X_1$, the kernel is the M-H kernel multiplied by the probability density function of $\pi_{X_2|X_1}(X_2 \mid X_1)$. Let $X_1^t, X_2^t = Y_1, Y_2$:

$$K((x_1, x_2), (y_1, y_2)) = (q(y_1 \mid x_1, x_2)\alpha(y_1 \mid x_1, x_2)) + (1 - a(y_1 \mid x_1, x_2))\delta_{y_1}(x_1)\pi_{(Y_2|Y_1}(y_2 \mid y_1)$$

Note that $\alpha = 1$. With that, we show that the kernel is invarant:

$$\int\int K((x_1, x_2), (y_1, y_2))\pi(x_1, x_2)dx_1dx_2 = \int\int \pi(y_1 \mid x_2)\pi(y_2 \mid y_1)\pi(x_1, x_2)dx_1dx_2 =$$

$$= \int \pi(y_1 \mid x_2)\pi(y_2 \mid y_1)\pi(x_2)dx_2 = \int \pi(y_1, x_2)\pi(y_2 \mid y_1)dx_2 = \pi(y_1, y_2)$$

□

# Exercise 3 (Metropolis-Hastings and Gibbs Sampler)

1. Let's show that the chain is reversible. If $x = y$, it is trivially reversible. If $x \neq y$, then:

$$T(x,y)\pi(x) = \alpha(x,y)q(x,y)\pi(x) = \frac{\gamma(x,y)}{\pi(x)q(x,y)}q(x,y)\pi(x) = \gamma(y,x) =$$

$$= \alpha(y,x)q(y,x)\pi(y) = T(y,x)\pi(y)$$

$\square$

2. First, let's verify that it is the M-H algorithm:

$$\alpha = \frac{\gamma(x,y)}{\pi(x)q(x,y)} = \frac{min\{\pi(x)q(x,y), \pi(y)q(y,x)\}}{\pi(x)q(x,y)} = min\left\{1, \frac{\pi(y)q(y,x)}{\pi(x)q(x,y)}\right\}$$

Now, let's give the Barker acceptance ratio:

$$\alpha(x,y) = \frac{\pi(x)q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)} = \frac{\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

$\square$

3. Let's consider $x \neq y$. Note that:

$$\frac{1}{\pi(x)q(x,y)} \geq \frac{1}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

$$\therefore$$

$$\frac{\pi(y)q(y,x)q(x,y)}{\pi(x)q(x,y)} \geq \frac{q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

Finally, if $\frac{\pi(y)q(y,x)q(x,y)}{\pi(x)q(x,y)} \leq 1$, then:

$$min\left\{1, \frac{\pi(y)q(y,x)q(x,y)}{\pi(x)q(x,y)}\right\} \geq \frac{q(x,y)\pi(y)q(y,x)}{\pi(x)q(x,y) + \pi(y)q(y,x)}$$

We conclude that the M-H algorithm provides estimators with lower asymptotic variance than Barker's algorithm.

4. First, since the probability of chosing and index is the same for both algorithms, it is not important when showing which kernel is bigger of equal than the other. Now, consider $x \neq y$.

Modified: $T(x_k^{(t-1)}, x_k^*) = min\left\{1, \frac{1-\pi(x_k^{(t-1)}|x_{-k})}{1-\pi(x_k^*|x_{-k})}\right\} \frac{\pi(x_k^*|x_{-k})}{1-\pi(x_k^{(t-1)}|x_{-k})}$

Standard: $T(x_k^{(t-1)}, x_k^*) = \pi(x_k^* \mid x_{-k})$

Let's consider both cases for the Modified kernel, hence, when $\alpha = 1$ and $\alpha = \frac{1-\pi(x_k^{(t-1)}|x_{-k})}{1-\pi(x_k^*|x_{-k})}$ First, since $1 - \pi(x_k^{(t-1)} \mid x_{-k}) \leq 1$, we have:

$$1 \cdot \frac{\pi(x_k^* \mid x_{-k})}{1 - \pi(x_k^{(t-1)} \mid x_{-k})} \geq \pi(x_k^* \mid x_{-k})$$

Also:

$$\frac{1 - \pi(x_k^{(t-1)} \mid x_{-k})}{1 - \pi(x_k^* \mid x_{-k})} \cdot \frac{\pi(x_k^* \mid x_{-k})}{1 - \pi(x_k^{(t-1)} \mid x_{-k})} \geq \pi(x_k^* \mid x_{-k})$$

Therefore, the modified kernel provides estimators with lower assymptotic variance. $\square$

# Exercise 4 (Metropolis-Hastings)

1. Let $Y = X'$, $\theta = v'$:

$$q(y \mid x) = \int f(v \mid x) g(y \mid v) dv$$

$$\alpha_{MH} = min\left\{1, \frac{\pi(y) \int f(v \mid y) g(x \mid v) dv}{\pi(x) \int f(v \mid x) g(y \mid v) dv}\right\}$$

Since we the density function $f(v \mid x)$ is unknown, then we cannot evaluate this $\alpha_{MH}$.

2. Let's show that the chain is invariant:

$$\int \int \pi(x, v) g(y \mid v) f(\theta \mid y) \cdot min\left\{1, \frac{\pi(y) g(x \mid \theta)}{\pi(x) g(y \mid v)}\right\} dvdx =$$

If $\alpha = 1$, then

$$\int \pi(v) g(y \mid v) f(\theta \mid y) \cdot 1dv = f(\theta \mid y)\pi(y) = \bar{\pi}(y, \theta)$$

Else:

$$\int \int \pi(x, v) g(y \mid v) f(\theta \mid y) \cdot \frac{\pi(y) g(x \mid \theta)}{\pi(x) g(y \mid v)} dvdx = \int \pi(y) f(\theta \mid y) g(x \mid \theta) dx =$$

$$= \pi(y) f(\theta \mid y) = \bar{\pi}(y, \theta)$$

Finally, $\bar{\pi}(y) = \int \bar{\pi}(y, \theta) d\theta = \int \pi(y) f(\theta \mid y) d\theta = \pi(y)$.

$\square$

3. First, note that $min(U, V) = \frac{U+V-|U-V|}{2}$, hence:

$$E\left[\frac{U + V - \mid U - V \mid}{2}\right] = \frac{E[U] + E[V] - E[\mid U - V \mid]}{2}$$

Without loss of generality, assume $E[U] \geq E[V]$. Then:

$$E[\mid U - V \mid] \geq E[U] - E[V]$$

$$\therefore$$

$$\frac{E[U] + E[V] - E[\mid U - V \mid]}{2} \leq \frac{E[U] + E[V] - E[U] - E[V]}{2} = E[V]$$

$\square$

4. Couldn't solve.

# Exercise 4 (Metropolis-Hastings)

1. $$E[y^2] + \alpha^2 E[Z^2] + 2\alpha|E[YZ]| \geq 0$$

   Therefore, we have a second degree polynomial in $\alpha$ that is greater or equal than 0. Hence, we know that $b^2 - 4ac \leq 0$, which means:

   $$(2E[YZ])^2 - 4E[Z^2]E[Y^2] \leq 0 \therefore E[Z^2]E[Y^2] \geq |E[YZ]|$$

   $\square$

2. Using the inequality from the previouse question, we have:

   $$Cov(Y,Z) = E[(Y - E[Y])(Z - E[Z])] \therefore Cov(Y,Z)^2 \leq E[(Y - E[Y])^2]E[(Z - E[Z])^2]$$

   $$\therefore$$

   $$\sqrt{Var(Y)Var(Z)} \geq Cov(Y,Z)$$

   Since $Var(Y) = Var(Z)$ by hypothesis, then:

   $$Var(Y) \geq Cov(Z,Y)$$

   $\square$

3. Couldn't prove the inequality. But, given that it is true, it implies that thinning doesn't improve the results, in other words, it doesn't lower the variance.