Exercise 1 (Monte Carlo for Gaussians)

1. Let's prove that $E[\phi(X)] = E[\phi(X+\theta)exp(\frac{-1}{2}\theta^T\theta - \theta^T X)]$.

$$E[\phi(X+\theta)exp(\frac{-1}{2}\theta^T\theta - \theta^T X)] = \int_{\mathbb{R}^d}\phi(x+\theta)exp(\frac{-1}{2}\theta^T\theta - \theta^T X)\pi(x)dx_1...dx_d =$$

$$= \int_{\mathbb{R}^d}\phi(x+\theta)exp\left(\frac{-1}{2}\theta^T\theta - \theta^T X\right)exp(-x^T x/2)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d =$$

$$\int_{\mathbb{R}^d}\phi(x+\theta)exp\left(\frac{-1}{2}(x-\theta)^T(x-\theta)\right)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d$$

Finally, making $x-\theta = y$,

$$\int_{\mathbb{R}^d}\phi(y)exp\left(\frac{-1}{2}(y)^T(y)\right)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d = E[\phi(Y)]$$

$\square$

2. Let's show that

$$\sigma^2(\theta) = E\left[\phi^2(X)exp\left(\frac{-1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta)\right)\right] - (E[\phi(X)]^2$$

Note that, using the result in the previous item we have:

$$\sigma^2(\theta) = V\left[\phi(X+\theta)exp\left(\frac{-1}{2}\theta^T\theta - \theta^T X\right)\right] =$$

$$= E\left[\left(\phi(X+\theta)exp\left(\frac{-1}{2}\theta^T\theta - \theta^T X\right)\right)^2\right] - E\left[\phi(X+\theta)exp\left(\frac{-1}{2}\theta^T\theta - \theta^T X\right)\right]^2 =$$

$$= E\left[\left(\phi(X+\theta)exp\left(\frac{-1}{2}\theta^T\theta - \theta^T X\right)\right)^2\right] - E\left[\phi(X)\right]^2$$

Now, let's rearrange the first term in the variance.

$$\sigma^2(\theta) = \int_{\mathbb{R}^d}\phi(x+\theta)^2 exp\left(-\theta^T\theta - 2\theta^T X\right)exp(-x^Tx/2)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d =$$

Make $X+\theta = Y$, then:

$$\int_{\mathbb{R}^d} \phi(y)^2 exp\left(-\theta^T\theta - 2\theta^T(y-\theta)\right) exp(-(y-\theta)^T(y-\theta)/2)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d =$$

$$= \int_{\mathbb{R}^d} \phi(y)^2 exp\left(\frac{1}{2}(y-\theta)^T(y-\theta) - \frac{y^Ty}{2}\right) exp\left(\frac{-y^Ty}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx_1...dx_d =$$

$$= E\left[\phi^2(X) exp\left(\frac{-1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta)\right)\right]$$

Therefore,

$$\sigma^2(\theta) = E\left[\phi^2(X) exp\left(\frac{-1}{2}X^TX + \frac{1}{2}(X-\theta)^T(X-\theta)\right)\right] - (E[\phi(X)]^2$$

$\square$

3. Let's calculate $\nabla^2\sigma^2(\theta) = H(\theta)$.

$$\frac{\partial\sigma^2(\theta)}{\partial\theta_i} = \frac{E[\phi(X)^2 exp(\frac{-X^TX+(X-\theta)^T(X-\theta)}{2})]}{\partial\theta_i} =$$

$$= \int_\chi \phi(x)^2 exp(-x^Tx)\frac{\partial}{\partial\theta_i} exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= \int_\chi \phi(x)^2 exp(-x^Tx)(\theta_i - x_i) exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx$$

We calculated the gradient, let's now calculate the second derivative. First the diagonal.

$$\frac{\partial}{\partial\theta_i}\int_\chi \phi(x)^2 exp(-x^Tx)(\theta_i - x_i) exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= E[\phi(X)^2] + \int_\chi \phi(x)^2 exp(-x^Tx) exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)(x_i-\theta_i)(x_i-\theta_i)\frac{1}{(\sqrt{2\pi})^d}dx$$

Now the rest:

$$\frac{\partial}{\partial\theta_j}\int_\chi \phi(x)^2 exp(-x^Tx)(\theta_i - x_i) exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= \int_\chi \phi(x)^2 exp(-x^Tx) exp\left(\frac{(x-\theta)^T(x-\theta)}{2}\right)(x_i-\theta_i)(x_j-\theta_j)\frac{1}{(\sqrt{2\pi})^d}dx$$

$\square$

4. We already know that the Hessian is positive definite. Hence, we only need to show that the derivative is equal to zero at θ^*.

$$\nabla\sigma^2(\theta) = \int_\chi \phi(x)^2 exp(-x^T x)(\theta - x)exp\left(\frac{-(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= \int_\chi \phi(x)^2 exp(-x^T x)(\theta - x)exp\left(\frac{-(x-\theta)^T(x-\theta)}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= \int_\chi \phi(x)^2(\theta - x)exp\left(\frac{-x^T x}{2} - \theta^T x + \frac{-\theta^T\theta}{2}\right)\frac{1}{(\sqrt{2\pi})^d}dx =$$

$$= \int_\chi \phi(x)^2(\theta - x)exp(-\theta^T x)exp(-x^T x/2)\frac{exp(\theta^T\theta/2)}{(\sqrt{2\pi})^d}dx = 0$$

Finally, since the last term doens't depend on X, we can eliminate it, obtaining:

$$E[\phi(X)^2(\theta - X)exp(-\theta^T X)] = 0$$

$\square$

5. Couldn't solve.

Exercise 2 (Metropolis-Hastings)

1. First, generate $Y \sim q(X_t, \cdot)$. With probability equal to $\alpha(x, y)$, make $X_{t+1} = Y$, and with probability $1 - \sum_{z\in\mathbb{Z}} \alpha(x, z)q(x, z)$ make $X_{t+1} = X_t$. Then, repeat the process.

2. We want to show that $\pi(x)T(x, y) = \pi(y)T(y, x)$. For $y = x$, this is trivial. Now, assume $y \neq x$. Therefore:

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

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

$\square$

3. Using Metropolis-Hastings, one has:

$$\alpha = min\left\{1, \frac{\pi(x^*)q(x_{t-1}) \mid x^*}{\pi(x_{t-1})q(x^* \mid x_{t-1})}\right\}$$

So, $\alpha(x, y) = \frac{\gamma(x,y)}{\pi(x)q(x,y)}$, for $x = x_{t-1}, x^* = y, q(x, y) = q(x^* \mid x_{t-1})$.
Make:

$$\gamma(x,y) = max\left\{\pi(y)q(y,x), \pi(x)q(x,y)\right\}$$

$$\therefore$$

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

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

$\square$

4. First, let's rewrite the estimate as a function of Y. Note that $Y^{(k)} = X^{(\tau_k)}$, hence:

$$\frac{1}{\tau_k - 1}\sum_{t=1}^{\tau_k - 1}\phi(X^{(t)}) = \frac{1}{k-1}\sum_{t=1}^{k-1}\phi(Y^{(t)})$$

Now, let's prove it has the transition desired transition kernel.

$$K(x,y) = P(Y^{(k)} = y \mid Y^{(k-1)} = x) = P(X^{(\tau_k)} = y \mid X^{(\tau_k - 1)} = x)$$

The event $X^{(\tau_k)} = y \mid X^{(\tau_k-1)} = x$ is equivalent to selecting y starting from x and then accepting y, hence $q(x, y) \cdot \alpha(x, y) = P(X^{(\tau_k)} = y \mid X^{(\tau_k-1)} = x)$.

Therefore, $P(Y^{(k)} = y \mid Y^{(k-1)} = x) = \frac{q(x,y)\alpha(x,y)}{\sum_{z\in\mathbb{Z}} \alpha(x,z)q(x,y)}$. $\square$

5. Let's show that $\tilde{\pi}(x)K(x,y) = \tilde{\pi}(y)K(y,x)$.

$$\frac{\pi(x)m(x)}{\sum_{z\in\mathbb{Z}}\pi(z)m(z)} \cdot \frac{q(x,y)\alpha(x,y)}{\sum_{z\in\mathbb{Z}}\alpha(x,z)q(x,z)} = \frac{\pi(x)q(x,y)\alpha(x,y)}{\sum_{z\in\mathbb{Z}}\pi(z)m(z)}$$

Note that $\alpha(x,y) = \frac{\gamma(x,y)}{\pi(x)q(x,y)} = \frac{\gamma(y,x)}{\pi(x)q(x,y)}$. Therefore:

$$\frac{\pi(x)q(x,y)\gamma(x,y)}{\alpha(x,z)q(x,z)\sum_{z\in\mathbb{Z}}\pi(z)m(z)} = \frac{\gamma(y,x)}{\sum_{z\in\mathbb{Z}}\pi(z)m(z)} = \frac{\alpha(y,x)\pi(y)q(y,x)}{\sum_{z\in\mathbb{Z}}\pi(z)m(z)} = \tilde{\pi}(y)K(y,x)$$

$\square$

6. Couldn't solve.

Exercise 3 (Metropolis-Hastings)

1. Let's show that $\int_\chi \pi(x)(\alpha(x)q(y)(1-\alpha(x)))\delta_x(y)dx = \pi(y)$. Note that, $\pi(x) = \frac{q(x)}{\alpha(x)Z_\pi}$, hence:

$$\int_\chi \pi(x)(\alpha(x)q(y)(1-\alpha(x)))\delta_x(y)dx = \int_\chi \frac{q(x)}{\alpha(x)Z_\pi}(\alpha(x)q(y)(1-\alpha(x)))\delta_x(y)dx =$$

$$= \int_\chi \frac{q(x)}{\alpha(x)Z_\pi}\alpha(x)q(y)dx + \pi(y)(1-\alpha(y)) = \frac{q(y)}{Z_\pi}\int_\chi q(x)dx + \pi(y)(1-\alpha(y)) =$$

$$= \pi(y)\alpha(y) + \pi(y)(1-\alpha(y)) = \pi(y)$$

□

2. Couldn't solve.

Exercise 4 (Gibbs Sampler)

1. $$\pi(x \mid y) = \frac{\pi(x,y)}{\pi(y)} \propto \frac{exp((x-1)^2(y-2)^2)/2}{(y-2)^{-1}} \sim N(1,(y-2)^{-2})$$

   Now, follow the same procedure to for $\pi(y \mid x)$, hence:

   $$\pi(y \mid x) \sim N(2,(x-1)^{-2})$$

   □

2. The sampler doesn't make sense, because:

   $$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \pi(x,y)dxdy \propto \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} exp((x-1)^2(y-2)^2)/2dxdy =$$

   $$= \int_{-\infty}^{\infty} \sqrt{2\pi}((y-2)^{-2})^{\frac{1}{2}}dy = \infty$$

   □

Exercise 5 (Gibbs Sampler)

1.

$$p(Z_i = z \mid \theta_1, \theta_2) = p(X_i+Y_i = z \mid \theta_1, \theta_2) = \sum_{y=0}^{z} p(X_i+Y_i = z \mid Y_i = y, \theta_1, \theta_2)p(Y_i = y \mid \theta_2) =$$

$$= \sum_{y=0}^{z} p(X_i = z-y \mid \theta_1)p(Y_i = y \mid \theta_2) = \sum_{y=0}^{z} \binom{m_i}{z-y}\theta_1^{z-y}(1-\theta_1)^{m_i-(z-y)}\binom{n_i}{y}\theta_2^{y}(1-\theta_2)^{n_i-y} \therefore$$

$$p(Z_1, ..., Z_T \mid \theta_1, \theta_2) = \prod_{i=1}^{T}\left(\sum_{y_i=0}^{z_i}\binom{m_i}{z_i-y_i}\theta_1^{z_i-y_i}(1-\theta_1)^{m_i-(z_i-y_i)}\binom{n_i}{y_i}\theta_2^{y_i}(1-\theta_2)^{n_i-y_i}\right)$$

$\square$

2. To sample from $p(\theta_1, \theta_2 \mid Z_1, ..., Z_T)$ we will use auxiliary varibles $X_{1:T}, Y_{1:T}$.

$$p(\theta_1 \mid \theta_2, X_{1:T}, Y_{1:T}, Z_{1:T}) \propto \pi(\theta_1 \mid \theta_2, Y_{1:T})p(X_{1:T} \mid \theta_1, \theta_2, Y_{1:T}) \propto$$

$$\propto \pi(\theta_1)p(X_{1:T} \mid \theta_1) \propto \theta_1^{\sum x_i}(1-\theta_1)^{\sum m_i - x_i} \therefore$$

$$\theta_1 \mid X_{1:T} \sim Beta\left(1 + \sum_{i=1}^{T} x_i, 1 + \sum_{i=1}^{T} m_i - x_i\right)$$

Similarly to θ_2:

$$\theta_2 \mid Y_{1:T} \sim Beta\left(1 + \sum_{i=1}^{T} y_i, 1 + \sum_{i=1}^{T} n_i - y_i\right)$$

Finally,

$$p(X_{1:T}, Y_{1:T} \mid \theta_1, \theta_2, Z_{1:T}) \propto \prod_{i=1}^{T}\binom{m_i}{x_i}\theta_1^{x_i}(1-\theta_1)^{m_i-x_i}\binom{n_i}{y_i}\theta_2^{y_i}(1-\theta_2)^{n_i-y_i}\mathbb{I}_{\{x_i+y_i=z_i\}}$$

Note that we can now use a Gibbs sampler by first sampling $\theta_1^{(0)}$ and $\theta_2^{(0)}$ from independent $U_{[0,1]}$. Secondly, we can sample $X^{(0)}$ and $Y^{(0)}$ from $p(X_{1:T}, Y_{1:T} \mid \theta_1^{(0)}, \theta_2^{(0)}, Z_{1:T})$, since the distribution found is discrete with finite support, hence, we can calculate the probabilty for each possible value of $X^{(0)}$ and $Y^{(0)}$. With the values for the auxiliary variables, we sample $\theta_1^{(1)}$ and $\theta_2^{(1)}$ from their posteriors, which is possible since they are Beta distributions. Again we sample X, Y using the updated θ_1, θ_2 and repeat the process.