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 $x - \theta = y$,

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

□

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.

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

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,

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

□

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:

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

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

□

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:

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

□

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:

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

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

□

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:

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

$$\therefore$$

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

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

□

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