Chapter 23. Monte Carlo inference

23.1.
The cdf of a standard Cauchy is given by:

$$F(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}.$$

Therefore, we have:

$$F^{-1}(u) = \tan(\pi(u - \frac{1}{2})).$$

Now we sample $u \sim U(0,1)$ and get $x = F^{-1}(u)$.

23.2.
We have a pdf of a Gamma distribution as:

$$p(x) = Ga(x; a, b) = \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx}.$$

Now we can derive an unnormalized Gamma distribution by setting $b = 1$, to get:

$$\tilde{p}(x) = \frac{1}{\Gamma(a)} x^{a-1} e^{-x}.$$

The Cauchy distribution has density:

$$q(x) = \mathcal{T}(x|\mu, \sigma, 1) = \frac{1}{\pi\sigma(1 + (\frac{x-\mu}{\sigma})^2)}.$$

To get the minimum rejection rate, we have to make two modes the same. We need $a > 1$ to the mode of $\tilde{p}(x)$ exist. Now we have:

$$\mu = a - 1$$

Moreover, we should have $Mq(a-1) = \tilde{p}(a-1)$ to make the bound tightest, to get:

$$M = \frac{1}{\Gamma(a)} \pi\sigma (a-1)^{a-1} e^{-(a-1)}.$$

We should ensure

$$\frac{\tilde{p}(x)}{Mq(x)} = x^{a-1} e^{-x} (1 + (\frac{x-(a-1)}{\sigma})^2)(a-1)^{1-a} e^{a-1} \leq 1.$$

Let

$$r(x) = x^{a-1} e^{-x} (1 + (\frac{x-(a-1)}{\sigma})^2) \propto \frac{\tilde{p}(x)}{Mq(x)}.$$

Then

$$\frac{\partial r(x)}{\partial x} = -\frac{1}{\sigma^2} e^{-x} x^{a-2} (x-(a-1))((x-a)^2 + \sigma^2 - (2a-1)).$$

This first derivative must have only one point, $a-1$, where it changes its sign, because otherwise it would change its sign at three points, ruining the fact that 1 is the unique global maximum of $\frac{\tilde{p}(x)}{Mq(x)}$. Hence we have $\sigma^2 \geq 2a-1$. To maximize $\frac{\tilde{p}(x)}{Mq(x)}$ with respect to σ, we must minimize σ. Combining these, we can conclude that the optimal value is $\sigma = \sqrt{2a-1}$.

23.3.
By equation 4.125 and 4.126,

$$\begin{aligned}
p(\mathbf{y}_t|\mathbf{z}_{t-1}) = \int p(\mathbf{y}_t|\mathbf{z}_t)p(\mathbf{z}_t|\mathbf{z}_{t-1})d\mathbf{z}_t &= \int \mathcal{N}(\mathbf{y}_t|\mathbf{H}_t\mathbf{z}_t, \mathbf{R}_t)\mathcal{N}(\mathbf{z}_t|f_t(\mathbf{z}_{t-1}), \mathbf{Q}_{t-1})d\mathbf{z}_t \\
&= \mathcal{N}(\mathbf{y}_t|\boldsymbol{\mu}_{\mathbf{y},t}, \boldsymbol{\Sigma}_{\mathbf{y},t}) \\
\boldsymbol{\mu}_{\mathbf{y},t} &= \mathbf{H}_t f_t(\mathbf{z}_{t-1}) \\
\boldsymbol{\Sigma}_{\mathbf{y},t} &= \mathbf{R}_t + \mathbf{H}_t\mathbf{Q}_{t-1}\mathbf{H}_t^T \\
p(\mathbf{z}_t|\mathbf{z}_{t-1}, \mathbf{y}_t) = \frac{p(\mathbf{y}_t|\mathbf{z}_t)p(\mathbf{z}_t|\mathbf{z}_{t-1})}{p(\mathbf{y}_t|\mathbf{z}_{t-1})} &\propto \mathcal{N}(\mathbf{y}_t|\mathbf{H}_t\mathbf{z}_t, \mathbf{R}_t)\mathcal{N}(\mathbf{z}_t|f_t(\mathbf{z}_{t-1}), \mathbf{Q}_{t-1}) \\
&= \mathcal{N}(\mathbf{z}_t|\boldsymbol{\mu}_{\mathbf{z},t}, \boldsymbol{\Sigma}_{\mathbf{z},t}) \\
\boldsymbol{\mu}_{\mathbf{z},t} &= (\mathbf{Q}_t^{-1} + \mathbf{H}_t^T\mathbf{R}_t^{-1}\mathbf{H}_t)^{-1}(\mathbf{H}_t^T\mathbf{R}_t^{-1}\mathbf{y}_t + \mathbf{Q}_t^{-1}f_t(\mathbf{z}_{t-1})) \\
\boldsymbol{\Sigma}_{\mathbf{z},t} &= (\mathbf{Q}_t^{-1} + \mathbf{H}_t^T\mathbf{R}_t^{-1}\mathbf{H}_t)^{-1}.
\end{aligned}$$