Homework 5

JD Gallego Posada, *University of Amsterdam* 16/05/2017

Problem 1

a) Sample $\mathbf{x}_i$ from $\tilde{q}$. Sample u_i uniformly from $[0, \tilde{q}(\mathbf{x}_i)]$. Accept $\mathbf{x}_i$ as sample if $u_i \leq p(\mathbf{x}_i)$. Else, reject and restart.

b) True. The previous samples has no influence on the current one.

c) $w_n = \frac{p(\mathbf{x}_n)}{q(\mathbf{x}_n)}$

d) $\alpha(\mathbf{x}_{t+1}, \mathbf{x}_t) = \min\left(1, \frac{\tilde{p}(\mathbf{x}_{t+1})q(\mathbf{x}_t)}{\tilde{p}(\mathbf{x}_t)q(\mathbf{x}_{t+1})}\right)$

e) Even though the proposal distribution of $\mathbf{x}_{t+1}$ is independent of $\mathbf{x}_t$, the accept probability is not.

f) $\mathbf{x}_1$, $\mathbf{x}_1$, $\mathbf{x}_3$, $\mathbf{x}_4$, $\mathbf{x}_4$

g) The case where the dimensions are independent is trivial. Now, it is harder to find a tight bounding proposal distribution in a high dimensional space. Since the discussed methods all use a proposal distribution, in general they will not work well.

Problem 2

Using Bishop's 2.140 - 2.142, $p(\mu|x,\tau) = \mathcal{N}\left(\mu|\frac{\tau^{-1}}{s_0+\tau^{-1}}\mu_0 + \frac{s_0}{s_0+\tau^{-1}}x, \left(\frac{1}{s_0}+\frac{1}{\tau^{-1}}\right)^{-1}\right)$.

Using Bishop's 2.149 - 2.151, $p(\tau|x,\mu) = \text{Gamma}\left(\tau|a+\frac{1}{2}, b+\frac{(x-\mu)^2}{2}\right)$.

Problem 4

a) $\mathbb{E}[\mathbf{x}] = [\mathbb{E}[x_1], \ldots, \mathbb{E}[x_D]]^T = \boldsymbol{\mu}$.

b) Note that if $i \neq j$, $x_i \perp\!\!\!\perp x_j$, thus $\boldsymbol{\Sigma}_{ij} = 0$. For $\boldsymbol{\Sigma}_{ii} = \mathbb{V}[x_i] = \mu_i(1-\mu_i)$.

c) $p(\mathbf{x}|\boldsymbol{\mu},\boldsymbol{\pi}) = \sum_k \pi_k p(\mathbf{x}|\boldsymbol{\mu}_k)$. By the linearity of the expectation and part a), $\mathbb{E}[\mathbf{x}] = \sum_k \pi_k \boldsymbol{\mu}_k$.

d) $\mathcal{L} = \log p(\mathbf{X}|\boldsymbol{\mu},\boldsymbol{\pi}) = \sum_n \log p(\mathbf{x}_n|\boldsymbol{\mu},\boldsymbol{\pi}) = \sum_n \log\left\{\sum_k \pi_k \prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})}\right\}$.

e) Since there is a summation inside the logarithm, there is no closed form solution.

f) $\log p(\mathbf{X},\mathbf{Z}|\boldsymbol{\mu},\boldsymbol{\pi}) = \sum_n \log p(\mathbf{x}_n,\mathbf{z}_n|\boldsymbol{\mu},\boldsymbol{\pi}) = \sum_n \sum_k z_{nk}\left\{\log \pi_k + \sum_i x_{ni}\log\mu_{ki} + (1-x_{ni})\log(1-\mu_{ki})\right\}$.

g) See below - μ and π are vectors in the PGM.

h)
$$\begin{aligned}\mathcal{L} &= \sum_n \sum_{\mathbf{z}_n} q_n(\mathbf{z}_n)\log p(\mathbf{x}_n,\mathbf{z}_n|\boldsymbol{\mu},\boldsymbol{\pi}) - \sum_n\sum_{\mathbf{z}_n} q_n(\mathbf{z}_n)\log q_n(\mathbf{z}_n)\\ &= \sum_n\sum_{\mathbf{z}_n} q_n(\mathbf{z}_n)\sum_k z_{nk}\left\{\log\pi_k + \sum_i x_{ni}\log\mu_{ki} + (1-x_{ni})\log(1-\mu_{ki})\right\} - \sum_n\sum_{\mathbf{z}_n} q_n(\mathbf{z}_n)\log q_n(\mathbf{z}_n)\end{aligned}$$

i) $\tilde{\mathcal{L}} = \mathcal{L} + \alpha\left(\sum_k \pi_k - 1\right) + \sum_n \lambda_n\left(\sum_{\mathbf{z}_n} q_n(\mathbf{z}_n) - 1\right)$

j) Let $\mathbf{e}_k$ denote the k-th column of the identity matrix $\mathbb{I}_K$.

$$\frac{\partial\tilde{\mathcal{L}}}{\partial q_n(\mathbf{e}_k)} = \log\pi_k + \log\prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})} - \log q_n(\mathbf{e}_k) - 1 + \lambda_n = 0$$

$$q_n(\mathbf{e}_k) = \exp(\lambda_n - 1)\pi_k \prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})}$$

Note that $\sum_k q_n(\mathbf{e}_k) = 1$, which implies that $\exp(1-\lambda_n) = \sum_k \pi_k \prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})}$.

$$q_n(\mathbf{e}_k) = \frac{\pi_k \prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})}}{\sum_k \pi_k \prod_i \mu_{ki}^{x_{ni}}(1-\mu_{ki})^{(1-x_{ni})}}.$$

This, of course, represents the posterior distribution $p(\mathbf{z}_n|\mathbf{x}_n, \boldsymbol{\mu}, \boldsymbol{\pi})$.

k) Recall that $\sum_k \pi_k = 1$.

$$\frac{\partial \tilde{\mathcal{L}}}{\partial \pi_k} = \frac{1}{\pi_k}\sum_n q_n(\mathbf{e}_k) + \alpha = 0 \Rightarrow -\alpha\pi_k = \sum_n q_n(\mathbf{e}_k) \Rightarrow \pi_k = \frac{1}{-\alpha}\sum_n q_n(\mathbf{e}_k)$$

However, $-\alpha = \sum_k \sum_n q_n(\mathbf{e}_k)$. This implies, $\pi_k = \dfrac{\sum_n q_n(\mathbf{e}_k)}{\sum_k \sum_n q_n(\mathbf{e}_k)}$.

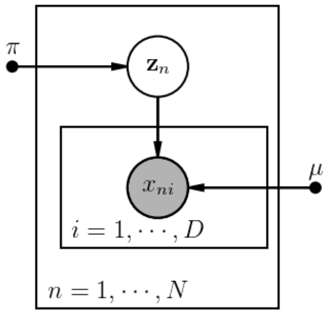

Figure 1: Exercise g)

Problem 5

Consider the discrete-time stochastic process $\{x_t\}_{t>0}$ given by $p(x_t) = 0.5\chi_0 + 0.25\chi_{-1} + 0.25\chi_1$ It is easy to see that $\mathbb{E}[x_t] = 0$ and $\mathbb{V}[x_t] = \frac{1}{2}$. Note that $x_t \perp\!\!\!\perp x_{t'}$ for $t \neq t'$. Clearly, $z^{(r)} = \sum_{t=1}^{r} x_t$.

$$\frac{r}{2} = \sum_{t=1}^{r} \mathbb{V}\left[x_t\right] = \mathbb{V}\left[\sum_{t=1}^{r} x_t\right] = \mathbb{V}[z^{(r)}] = \mathbb{E}\left[\left(z^{(r)}\right)^2\right] - \cancel{\left(\mathbb{E}[z^{(r)}]\right)^2} = \mathbb{E}\left[\left(z^{(r)}\right)^2\right]$$