Advanced probabilistic methods - Sketch

Christian Segercrantz 481056

March 17, 2022

$$\begin{aligned}
\mathcal{L}(q) =& \mathbb{E}_q[\log p(\mathbf{X},\mathbf{Z})] - \mathbb{E}_q[\log q(\mathbf{Z})] && (1)\\
\mathbb{E}_q \log p(\mathbf{X},\mathbf{Z}) =& \mathbb{E}_q\left[\log p(\tau) + \log p\theta + \log p(\mathbf{z} \mid \tau) + \log p(\mathbf{x} \mid \mathbf{z}\theta)\right] && (2)\\
=& \mathbb{E}_{q(\tau)}\left[\log p(\tau)\right] + \mathbb{E}_{q(\theta)}\left[\log p\theta\right] + \mathbb{E}_{q(\mathbf{z})q(\tau)}\left[\log p(\mathbf{z} \mid \tau)\right] + \mathbb{E}_{q(\mathbf{z})q(\theta)}\left[\log p(\mathbf{x} \mid \mathbf{z},\theta)\right] && (3)\\
\mathbb{E}_q[\log q(\mathbf{Z})] =& \mathbb{E}_q\left[-\log q(\mathbf{z}) - \log q(\tau) - \log q(\theta)\right] && (4)\\
=& -\mathbb{E}_{q(\mathbf{z})}\left[\log q(\mathbf{z})\right] - \mathbb{E}_{q(\tau)}\left[\log q(\tau)\right] - \mathbb{E}_{q(\theta)}\left[\log q(\theta)\right] && (5)\\
\mathcal{L}(q) =& \mathbb{E}_{q(\tau)}\left[\log p(\tau)\right] + \mathbb{E}_{q(\theta)}\left[\log p(\theta)\right] + \mathbb{E}_{q(\mathbf{z})q(\tau)}\left[\log p(\mathbf{z} \mid \tau)\right] + \mathbb{E}_{q(\mathbf{z})q(\theta)}\left[\log p(\mathbf{x} \mid \mathbf{z},\theta)\right] && (6)\\
& -\mathbb{E}_{q(\mathbf{z})}\left[\log q(\mathbf{z})\right] - \mathbb{E}_{q(\tau)}\left[\log q(\tau)\right] - \mathbb{E}_{q(\theta)}\left[\log q(\theta)\right] && (7)
\end{aligned}$$

$$\begin{aligned}
p(\theta) =& \mathcal{N}(\theta \mid 0, \beta_0^{-1}) && (8)\\
\mathbb{E}_{q(\theta)}\left[\log p(\theta)\right] =& \mathbb{E}_{q(\theta)}\left[\log \mathcal{N}(\theta \mid 0, \beta_0^{-1})\right] && (9)\\
\propto& \mathbb{E}_{q(\theta)}\left[-\frac{1}{2}\theta^2\beta_0\right] && (10)\\
=& -\frac{1}{2}\mathbb{E}_{q(\theta)}\left[\theta^2\right]\beta_0 && (11)\\
\mathbb{E}_{q(\theta)}\left[\theta^2\right] =& \mathrm{var}(\theta) + \mathbb{E}_{q(\theta)}\left[\theta\right]^2 && (12)\\
=& \beta_2^{-1} + m_2^2 && (13)\\
\mathbb{E}_{q(\theta)}\left[\log p(\theta)\right] \propto& -\frac{\beta_0}{2}(\beta_2^{-1} + m_2^2) && (14)
\end{aligned}$$

$$\begin{aligned}
q(\theta) =& \mathcal{N}(\theta \mid m_2, \beta_2^{-1}) && (15)\\
\mathbb{E}_{q(\theta)}\left[\log q(\theta)\right] =& \mathbb{E}_{q(\theta)}\left[\log \mathcal{N}(\theta \mid m_2, \beta_2^{-1})\right] && (16)\\
\propto& \mathbb{E}_{q(\theta)}\left[-\frac{1}{2}(\theta - m_2)^2\beta_2\right] && (17)\\
=& -\frac{1}{2}\mathbb{E}_{q(\theta)}\left[(\theta - m_2)^2\right]\beta_2 && (18)\\
=& -\frac{1}{2}\mathbb{E}_{q(\theta)}\left[(\theta^2 - 2\theta m_2 + m_2^2\right]\beta_2 && (19)\\
\propto& -\frac{1}{2}\mathbb{E}_{q(\theta)}\left[\theta^2 - 2\theta m_2\right]\beta_2 && (20)\\
=& -\frac{\beta_2}{2}(\mathbb{E}_{q(\theta)}\left[\theta^2\right] - \mathbb{E}_{q(\theta)}\left[2\theta m_2\right]) && (21)\\
=& -\frac{\beta_2}{2}(\beta_2^{-1} + m_2^2 - 2m_2^2) && (22)\\
=& -\frac{\beta_2}{2}(\beta_2^{-1} - m_2^2) && (23)\\
=& \frac{m_2^2\beta_2}{2} - \frac{1}{2} && (24)\\
& && (25)
\end{aligned}$$

$$
\begin{aligned}
p(x \mid z, \theta) &= \prod_{n=1}^{N} \mathcal{N}(x_n \mid 0, 1)^{z_{n1}} \mathcal{N}(x_n \mid \theta, 1)^{z_{n2}} && (26) \\
E_{q(z)q(\theta)} \left[\log p(x \mid z, \theta)\right] &= E_{q(z)q(\theta)} \left[\sum_{n=1}^{N} z_{n1} \log \mathcal{N}(x_n \mid 0, 1) + \sum_{n=1}^{N} z_{n2} \log \mathcal{N}(x_n \mid \theta, 1) \right] && (27) \\
&\propto E_{q(z)q(\theta)} \left[\sum_{n=1}^{N} z_{n1} \left(-\frac{1}{2} x_n^2\right) + \sum_{n=1}^{N} z_{n2} \left(-\frac{1}{2}\right) (x_n - \theta)^2 \right] && (28) \\
&= E_{q(z)q(\theta)} \left[\sum_{n=1}^{N} z_{n1} \left(-\frac{1}{2} x_n^2\right) + \sum_{n=1}^{N} z_{n2} \left(-\frac{1}{2}\right) (x_n^2 - 2x_n\theta + \theta^2) \right] && (29) \\
&= \sum_{n=1}^{N} E_{q(z)} \left[z_{n1}\right] \left(-\frac{1}{2} x_n^2\right) + \sum_{n=1}^{N} E_{q(z)} \left[z_{n2}\right] \left(-\frac{1}{2}\right) \left(x_n^2 - 2x_n E_{q(\theta)} \left[\theta\right] + E_{q(\theta)} \left[\theta^2\right]\right) && (30) \\
E_{q(z)} \left[z_{nk}\right] &= r_{nk} && (31) \\
E_{q(\theta)} \left[\theta\right] &= m_2 && (32) \\
E_{q(\theta)} \left[\theta^2\right] &= \beta_2^{-1} + m_2^2 && (33) \\
E_{q(z)q(\theta)} \left[\log p(x \mid z, \theta)\right] &\propto \sum_{n=1}^{N} r_{n1} \left(-\frac{1}{2} x_n^2\right) + \sum_{n=1}^{N} r_{n2} \left(-\frac{1}{2}\right) (x_n^2 - 2x_n m_2 + \beta_2^{-1} + m_2^2) && (34) \\
&= \sum_{n=1}^{N} r_{n1} \left(-\frac{1}{2} x_n^2\right) + \sum_{n=1}^{N} r_{n2} \left(-\frac{1}{2}\right) ((x_n - m_2)^2 + \beta_2^{-1}) && (35)
\end{aligned}
$$

We denote x_b as the number of black pebbles drawn and x_w as the number of white ones. For the first model we thus get

$$
\begin{aligned}
x_i &\sim \text{Bin}(x_i \mid n, a) = \binom{n}{x_i} a^{x_i} (1-a)^{x_i} && (36) \\
a &\sim \text{Beta}(a \mid 1, 1) = 1 && (37) \\
p(x \mid M_1) &= \int_0^1 \text{Bin}(x_b \mid n, a) \text{Bin}(x_w \mid n, a) \text{Beta}(a \mid 1, 1) && (38) \\
&= \int_0^1 \binom{n}{x_b} a^{x_b} (1-a)^{n-x_b} \binom{n}{x_w} a^{x_w} (1-a)^{n-x_w} da && (39) \\
&= \binom{n}{x_b} \binom{n}{x_w} \int_0^1 a^{x_b+x_w} (1-a)^{2n-x_b-x_w} da && (40) \\
&= \binom{n}{x_b} \binom{n}{x_w} \int_0^1 a^{x_b+x_w-1+1} (1-a)^{x_b+x_w-1+1} da && (41) \\
&= \binom{n}{x_b} \binom{n}{x_w} \mathrm{B}(x_b + x_w + 1, 2n - w_b - x_w + 1) && (42)
\end{aligned}
$$

For the second model we have

$$
\begin{aligned}
x_i &\sim \text{Bin}(n, a_i) = \binom{n}{x_i} a_i^{x_i} (1-a_i)^{n-x_i} && (43) \\
b &\sim \text{Beta}(1, 1) && (44) \\
& && (45)
\end{aligned}
$$

We again note $i = b$ for black marbles and $i = w$ for the white ones.

$$
\begin{aligned}
p(x \mid M_2) &= \int_0^1 \int_0^1 \mathrm{Bin}(x_b \mid n, a_b)\mathrm{Bin}(x_w \mid n, a_w)\mathrm{Beta}(a_b \mid 1,1)\mathrm{Beta}(a_w \mid 1,1) da_b da_w && (46) \\
&= \int_0^1 \int_0^1 \binom{n}{x_b} a_b^{x_b}(1-a_b)^{n-x_b} \binom{n}{x_w} a_w^{x_w}(1-a_w)^{n-x_w} da_b da_w && (47) \\
&= \binom{n}{x_b}\binom{n}{x_w} \int_0^1 a_b^{x_b}(1-a_b)^{n-x_b} da_b \int_0^1 a_w^{x_w}(1-a_w)^{n-x_w} da_w && (48) \\
&= \binom{n}{x_b}\binom{n}{x_w} \int_0^1 a_b^{x_b-1+1}(1-a_b)^{n-x_b-1+1} da_b \int_0^1 a_w^{x_w-1+1}(1-a_w)^{n-x_w-1+1} da_w && (49) \\
&= \binom{n}{x_b}\binom{n}{x_w} \mathrm{B}(x_b+1, n-x_b+1)\mathrm{B}(x_w+1, n-x_w+1) && (50) \\
& && (51)
\end{aligned}
$$

We can now formulate the posterior odds as

$$
\begin{aligned}
\frac{p(x \mid M_1)}{p(x \mid M_2)} &= \frac{\binom{n}{x_b}\binom{n}{x_w}\mathrm{B}(x_b + x_w + 1, 2n - w_b - x_w + 1)}{\binom{n}{x_b}\binom{n}{x_w}\mathrm{B}(x_b+1, n-x_b+1)\mathrm{B}(x_w+1, n-x_w+1)} && (52) \\
&= \frac{\mathrm{B}(x_b + x_w + 1, 2n - w_b - x_w + 1)}{\mathrm{B}(x_b+1, n-x_b+1)\mathrm{B}(x_w+1, n-x_w+1)} && (53)
\end{aligned}
$$