EM Algorithm Tutorial

李修成

0.1 公式

$$l(\theta) = \log p(x;\theta) = \log \Sigma_z p(x,z;\theta) \tag{1}$$

$$\begin{aligned} l(\theta) = \log p(x;\theta) &= \log \Sigma_z p(x,z;\theta) \\ &= \log \Sigma_z Q(z) \frac{p(x,z;\theta)}{Q(z)} \\ &\geq \Sigma_z \log \frac{p(x,z;\theta)}{Q(z)} \end{aligned} \tag{2}$$

$$\begin{aligned} l(\theta) - L(x,z;\theta) &= \log p(x;\theta) - \Sigma_z Q(z) \log \frac{p(x,z;\theta}{Q(z)} \\ &= \Sigma_z Q(z) \log p(x;\theta) - \Sigma_z Q(z) \log \frac{p(x,z;\theta)}{Q(z)} \\ &= -\Sigma_z Q(z) \log \frac{p(z|x)}{Q(z)} = D(Q(z) \parallel p(z|x;\theta)) \end{aligned} \tag{3}$$

0.2 算法

Repeat until convergence {
    E-step:
    For each i, set
        $Q_i(z^{(i)}) := p(z^{(i)}|x^{(i)};\theta)$
    M-step:
    Set
        $\theta := arg\ max_\theta \Sigma_i \Sigma_z Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)};\theta)}{Q_i(z^{(i)};\theta)}$
}

0.3 收敛性的证明

$$l(\theta) \geq \Sigma_i \Sigma_z Q_i(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)};\theta)}{Q_i(z^{(i)})} \tag{4}$$

$$\begin{aligned} l(\theta^{(t+1)}) &\geq \Sigma_i \Sigma_z Q_i^{(t)}(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)};\theta^{(t+1)})}{Q_i^{(t)}(z^{(i)})} \\ &\geq \Sigma_i \Sigma_z Q_i^{(t)}(z^{(i)}) \log \frac{p(x^{(i)}, z^{(i)};\theta^{(t)})}{Q_i^{(t)}(z^{(i)})} \\ &= l(\theta^{(t)}) \end{aligned} \tag{5}$$

0.4 EM 算法求解 GMM（高斯混合模型）

E-step: 令

$$
\begin{aligned}
w_j^{(i)} &= Q_i(z^{(i)} = j) \\
&= p(z^{(i)} = j | x^{(i)}; \phi, \mu, \Sigma) \\
&= \frac{p(x^{(i)} | z^{(i)} = j; \mu, \Sigma) p(z^{(i)} = j; \phi)}{\Sigma_{l=1}^{k} p(x^{(i)} | z^{(i)} = l; \mu, \Sigma) p(z^{(i)} = l; \phi)}
\end{aligned} \tag{6}
$$

M-step:

$$
\begin{aligned}
l(\theta) &= \Sigma_{i=1}^{m} \Sigma_z Q_i(z^{(i)}) \log p(x^{(i)}, z^{(i)}; \mu, \Sigma) \\
&= \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} Q_i(z^{(i)} = j) \log p(x^{(i)} | z^{(i)} = j; \mu, \Sigma) p(z^{(i)} = j; \phi) \\
&= \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} w_j^{(i)} \log \frac{1}{(2\pi)^{n/2} |\Sigma_j|^{1/2}} exp(-\frac{1}{2}(x^{(i)} - \mu_j)^T \Sigma_j^{-1} (x^{(i)} - \mu_j)) \phi_j
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
\nabla_{\mu_j} l(\theta) &= \nabla_{\mu_j} \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} w_j^{(i)} \frac{1}{2} (x^{(i)} - \mu_j)^T \Sigma_j^{-1} (x^{(i)} - \mu_j) \\
&= \Sigma_{i=1}^{m} w_j^{(i)} \Sigma_j^{-1} (x^{(i)} - \mu_j) = 0 \\
\mu_j &= \frac{\Sigma_{i=1}^{m} w_j^{(i)} x^{(i)}}{\Sigma_{i=1}^{m} w_j^{(i)}}
\end{aligned} \tag{8}
$$

$$
\begin{aligned}
\nabla_{\phi_j} l(\theta) &= \nabla_{\phi_j} \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} w_j^{(i)} \log \phi_j \\
&= \Sigma_{i=1}^{m} \frac{w_j^{(i)}}{\phi_j}
\end{aligned} \tag{9}
$$

$$
\begin{aligned}
L(\phi) &= \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} w_j^{(i)} \log \phi_j + \beta(\Sigma_{j=1}^{k} \phi_j - 1) \\
\frac{\partial L}{\partial \phi_j} &= \Sigma_{i=1}^{m} \frac{w_j^{(i)}}{\phi_j} + \beta \Rightarrow -\beta \phi_j = \Sigma_{i=1}^{m} w_j^{(i)} \\
-\beta \Sigma_{j=1}^{k} \phi_j &= \Sigma_{i=1}^{m} \Sigma_{j=1}^{k} w_j^{(i)} = m \Rightarrow -\beta = m \\
\phi_j &= \Sigma_{i=1}^{m} \frac{w_j^{(i)}}{m}
\end{aligned} \tag{10}
$$

令 $S_j = \Sigma_j^{-1}, b = (x^{(i)} - \mu_j)$

$$
\begin{aligned}
\nabla_{s_j} l(\theta) &= \nabla_{s_j} \Sigma_{i=1}^{m} w_j^{(i)} (\frac{1}{2} \log |S_j| - \frac{1}{2} b^T S_j b) \\
&= \frac{1}{2} \Sigma_j^{(i)} (S_j^{-1} - bb^T) = 0 \\
S_j^{-1} &= \frac{\Sigma_{j=1}^{m} bb^T}{\Sigma_{i=1}^{m} w_j^{(i)}} \\
\Sigma_j &= \frac{\Sigma_{j=1}^{m} (x^{(i)} - \mu_j)(x^{(i)} - \mu_j)^T}{\Sigma_{i=1}^{m} w_j^{(i)}}
\end{aligned} \tag{11}
$$

(12)