模式识别：作业 #2

201928014629008
牛李金梁

2019 年 10 月 20 日

Question 1

(a)

$$
\begin{aligned}
\sigma_{ij} &= \mathsf{E}\left[(x_i-\mu_i)(x_j-\mu_j)\right]\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x_i,x_j)(x_i-\mu_i)(x_j-\mu_j)\,dx_i x_j\\
&= \int_{-\infty}^{\infty} p(x_i)(x_i-\mu_i)\,dx_i \int_{-\infty}^{\infty} p(x_j)(x_j-\mu_j)\,dx_j\\
&= \left[\int_{-\infty}^{\infty} x_i p(x_i)\,dx_i - \mu_i\int_{-\infty}^{\infty} p(x_i)\,dx_i\right]\left[\int_{-\infty}^{\infty} x_j p(x_j)dx_j - \mu_j\int_{-\infty}^{\infty} p(x_j)dx_j\right]\\
&= (\mu_i-\mu_i)(\mu_j-\mu_j)\\
&= 0
\end{aligned}
$$

(b)

设有一个 d 维的高斯分布 $\boldsymbol{x}\sim N(\boldsymbol{\mu},\boldsymbol{\Sigma})$

其中，$\boldsymbol{x}=\begin{bmatrix} x_1\\ \vdots \\ x_d\end{bmatrix}$，$\boldsymbol{\mu}=\begin{bmatrix} \mu_1\\ \vdots \\ \mu_d\end{bmatrix}$，$\boldsymbol{\Sigma}=\begin{bmatrix} \sigma_1^2 & \sigma_1\sigma_2 & \cdots & \sigma_1\sigma_d\\ \sigma_2\sigma_1 & \sigma_2^2 & \cdots & \sigma_2\sigma_d\\ \vdots & \vdots & \ddots & \vdots\\ \sigma_d\sigma_1 & \sigma_d\sigma_2 & \cdots & \sigma_d^2\end{bmatrix}$

因为 $\sigma_{ij}=0$，所以 $|\boldsymbol{\Sigma}|=\sigma_1^2\sigma_2^2\cdots\sigma_d^2$，$\boldsymbol{\Sigma}^{-1}=diag\left(\frac{1}{\sigma_1^2},\frac{1}{\sigma_2^2},\cdots,\frac{1}{\sigma_d^2}\right)$

$$
\begin{aligned}
p(x_1-\mu_1,x_2-\mu_2,\cdots,x_d-\mu_d)=p(\boldsymbol{x}-\boldsymbol{\mu}) &= \frac{1}{(2\pi)^{d/2}|\boldsymbol{\Sigma}|^{1/2}}\exp\left[-\frac{1}{2}\boldsymbol{x}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{x}\right]\\
&= \frac{1}{(2\pi)^{d/2}\sigma_1\sigma_2\cdots\sigma_d}\exp\left[-\frac{1}{2}\sum_{i=1}^{d}\frac{{x_i}^2}{\sigma_i^2}\right]\\
&= \prod_{i=1}^{d}\frac{1}{\sqrt{2\pi}\sigma_i}\exp\left[-\frac{1}{2}\left(\frac{x_i}{\sigma_i^2}\right)^2\right]\\
&= \prod_{i=1}^{d} p(x_i-\mu_i)
\end{aligned}
$$

根据上式容易知，对于高斯分布，如果 $\sigma_{ij}=0$，那么 $x_i-\mu_i$ 和 $x_j-\mu_j$ 是统计独立的 $(i\neq j)$。

Question 2

(a)

$$
P\left[z_{ik}=1|P(\omega_i)\right]=P(\omega_i)
$$

$$P\left[z_{ik}=0|P\left(\omega_i\right)\right]=1-P\left(\omega_i\right)$$

因此，$P\left[z_{ik}|P\left(\omega_i\right)\right]=P(\omega_i)^{z_{ik}}\left[1-P\left(\omega_i\right)\right]^{1-z_{ik}}$
对于独立随机变量，有

$$\begin{aligned}P\left[z_{i1},\cdots,z_{in}|P\left(\omega_i\right)\right]&=\prod_{k=1}^{n}P\left[z_{ik}|P\left(\omega_i\right)\right]\\&=\prod_{k=1}^{n}P(\omega_i)^{z_{ik}}\left[1-P\left(\omega_i\right)\right]^{1-z_{ik}}\end{aligned}$$

(b)

对数似然函数

$$\begin{aligned}l\left(P\left(\omega_i\right)\right)&=\ln P\left[z_{i1},\cdots,z_{in}|P\left(\omega_i\right)\right]\\&=\ln\left\{\prod_{k=1}^{n}P(\omega_i)^{z_{ik}}\left[1-P\left(\omega_i\right)\right]^{1-z_{ik}}\right\}\\&=\sum_{k=1}^{n}\left[z_{ik}\ln P\left(\omega_i\right)+\left(1-z_{ik}\right)\ln\left[1-P\left(\omega_i\right)\right]\right]\end{aligned}$$

求梯度，

$$\nabla_{P(\omega_i)}l\left(P\left(\omega_i\right)\right)=\frac{1}{P\left(\omega_i\right)}\sum_{k=1}^{n}z_{ik}-\frac{1}{1-P\left(\omega_i\right)}\sum_{k=1}^{n}\left(1-z_{ik}\right)$$

令 $\nabla_{P(\omega_i)}l\left(P\left(\omega_i\right)\right)=0$，有

$$\left(1-\hat{P}\left(\omega_i\right)\right)\sum_{k=1}^{n}z_{ik}=\hat{P}\left(\omega_i\right)\sum_{k=1}^{n}\left(1-z_{ik}\right)$$

$$\sum_{k=1}^{n}z_{ik}-\hat{P}\left(\omega_i\right)\sum_{k=1}^{n}z_{ik}=n\hat{P}\left(\omega_i\right)-\hat{P}\left(\omega_i\right)\sum_{k=1}^{n}z_{ik}$$

$$\hat{P}\left(\omega_i\right)=\frac{1}{n}\sum_{k=1}^{n}z_{ik}$$

Question 3

$p_1\left(\boldsymbol{x}\right)\sim N\left(\boldsymbol{\mu},\boldsymbol{\Sigma}\right)$，
$p_2\left(\boldsymbol{x}\right)$ 是任意分布。

$$\begin{aligned}D_{KL}\left(p_2,p_1\right)&=\int p_2\left(\boldsymbol{x}\right)\ln\frac{p_2\left(\boldsymbol{x}\right)}{p_1\left(\boldsymbol{x}\right)}d\boldsymbol{x}\\&=\int p_2\left(\boldsymbol{x}\right)\ln p_2\left(\boldsymbol{x}\right)d\boldsymbol{x}+\frac{1}{2}\int p_2\left(\boldsymbol{x}\right)\left[d\ln\left(2\pi\right)+\ln\left|\boldsymbol{\Sigma}\right|+\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^t\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\right]d\boldsymbol{x}\end{aligned}$$

分别对 $\boldsymbol{\mu}$，$\boldsymbol{\Sigma}$ 求导有，

$$\frac{\partial D_{KL}\left(p_2,p_1\right)}{\partial \boldsymbol{\mu}} = -\int p_2\left(\boldsymbol{x}\right)\boldsymbol{\Sigma}^{-1}\left(\boldsymbol{x}-\boldsymbol{\mu}\right)d\boldsymbol{x}$$

$$\begin{aligned}\frac{\partial D_{KL}\left(p_2,p_1\right)}{\partial \boldsymbol{\Sigma}} &= \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \boldsymbol{\Sigma}} \times \frac{\partial D_{KL}\left(p_2,p_1\right)}{\partial \boldsymbol{\Sigma}^{-1}} \\ &= \frac{\partial \boldsymbol{\Sigma}^{-1}}{\partial \boldsymbol{\Sigma}} \times \frac{1}{2}\int p_2\left(\boldsymbol{x}\right)\frac{\partial}{\partial \boldsymbol{\Sigma}^{-1}}\left(\ln|\boldsymbol{\Sigma}| + \left(x-\boldsymbol{\mu}\right)^t\boldsymbol{\Sigma}^{-1}\left(x-\boldsymbol{\mu}\right)\right)d\boldsymbol{x} \\ &= \frac{1}{-\boldsymbol{\Sigma}^2} \times \frac{1}{2}\int p_2\left(\boldsymbol{x}\right)\left(-\boldsymbol{\Sigma} + \left(\boldsymbol{x}-\boldsymbol{\mu}\right)\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^t\right)d\boldsymbol{x}\end{aligned}$$

令 $\frac{\partial D_{KL}(p_2,p_1)}{\partial \boldsymbol{\mu}} = 0$，得

$$\int p_2\left(\boldsymbol{x}\right)\left(\boldsymbol{x}-\boldsymbol{\mu}\right)dx = 0$$

即，

$$E_2\left[\boldsymbol{x}-\boldsymbol{\mu}\right] = 0$$

$$\boldsymbol{\mu} = \mathsf{E}_2\left(\boldsymbol{x}\right)$$

令 $\frac{\partial D_{KL}(p_2,p_1)}{\partial \boldsymbol{\Sigma}} = 0$，得

$$\int p_2\left(\boldsymbol{x}\right)\left[-\boldsymbol{\Sigma} + \left(\boldsymbol{x}-\boldsymbol{\mu}\right)\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^t\right]d\boldsymbol{x} = 0$$

即，

$$E_2\left[-\boldsymbol{\Sigma} + \left(\boldsymbol{x}-\boldsymbol{\mu}\right)\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^t\right] = 0$$

$$\boldsymbol{\Sigma} = E_2\left[\left(\boldsymbol{x}-\boldsymbol{\mu}\right)\left(\boldsymbol{x}-\boldsymbol{\mu}\right)^t\right]$$

Question 4

(a)

$$\begin{aligned}\hat{\boldsymbol{\mu}}_{n+1} &= \frac{1}{n+1}\sum_{k=1}^{n+1}\boldsymbol{x}_k \\ &= \frac{1}{n+1}\left(\sum_{k=1}^{n}\boldsymbol{x}_k + \boldsymbol{x}_{n+1}\right) \\ &= \frac{1}{n+1}\left(n\hat{\boldsymbol{\mu}}_n + \boldsymbol{x}_{n+1}\right) \\ &= \frac{n}{n+1}\hat{\boldsymbol{\mu}}_n + \frac{1}{n+1}\boldsymbol{x}_{n+1} \\ &= \hat{\boldsymbol{\mu}}_n - \frac{1}{n+1}\left(\hat{\boldsymbol{\mu}}_n - \boldsymbol{x}_{n+1}\right) \\ &= \hat{\boldsymbol{\mu}}_n + \frac{1}{n+1}\left(\boldsymbol{x}_{n+1} - \hat{\boldsymbol{\mu}}_n\right)\end{aligned}$$

$$
\begin{aligned}
\boldsymbol{C}_{n+1} &= \frac{1}{n}\sum_{k=1}^{n+1}\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)^t \\
&= \frac{1}{n}\left[\sum_{k=1}^{n}\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)^t+\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)^t\right]
\end{aligned}
$$

其中，

$$
\begin{aligned}
\sum_{k=1}^{n}\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)^t &= \sum_{k=1}^{n}\left[\left(\boldsymbol{x}_k-\frac{1}{n+1}\sum_{k=1}^{n+1}\boldsymbol{x}_k\right)\left(\boldsymbol{x}_k-\frac{1}{n+1}\sum_{k=1}^{n+1}\boldsymbol{x}_k\right)^t\right] \\
&= \sum_{k=1}^{n}\left[\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n+\frac{\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}}{n+1}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n+\frac{\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}}{n+1}\right)^t\right] \\
&= \sum_{k=1}^{n}\left[\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n\right)^t+\frac{1}{\left(n+1\right)^2}\left(\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}\right)\left(\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}\right)^t\right] \\
&= \sum_{k=1}^{n}\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_n\right)^t+\frac{n}{\left(n+1\right)^2}\left(\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}\right)\left(\hat{\boldsymbol{\mu}}_n-\boldsymbol{x}_{n+1}\right)^t \\
&= \left(n-1\right)\boldsymbol{C}_n+\frac{n}{\left(n+1\right)^2}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t
\end{aligned}
$$

$$
\begin{aligned}
\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)^t &= \left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n-\frac{1}{n+1}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n-\frac{1}{n+1}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\right)^t \\
&= \left(\frac{n}{n+1}\right)^2\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t
\end{aligned}
$$

因而，

$$
\begin{aligned}
\boldsymbol{C}_{n+1} &= \frac{1}{n}\left[\sum_{k=1}^{n}\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_k-\hat{\boldsymbol{\mu}}_{n+1}\right)^t+\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_{n+1}\right)^t\right] \\
&= \frac{n-1}{n}\boldsymbol{C}_n+\frac{1}{\left(n+1\right)^2}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t+\frac{n}{\left(n+1\right)^2}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t \\
&= \frac{n-1}{n}\boldsymbol{C}_n+\frac{1}{n+1}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t
\end{aligned}
$$

(b)

$\hat{\boldsymbol{\mu}}_n=\frac{1}{n}\sum\limits_{k=1}^{n}\boldsymbol{x}_k$ 的时间复杂度为 $O\left(dn\right)$，$\hat{\boldsymbol{\mu}}_{n+1}=\hat{\boldsymbol{\mu}}_n+\frac{1}{n+1}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)$ 所需的时间复杂度为 $O\left(d\right)$，因此计算 $\hat{\boldsymbol{\mu}}_n$ 的时间复杂度为 $O\left(dn\right)$。

在 $\boldsymbol{C}_{n+1}=\frac{n-1}{n}\boldsymbol{C}_n+\frac{1}{n+1}\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)\left(\boldsymbol{x}_{n+1}-\hat{\boldsymbol{\mu}}_n\right)^t$ 中，在已知 $\boldsymbol{C}_n,\hat{\boldsymbol{\mu}},\boldsymbol{x}_n$ 的情况下，计算该式需要时间复杂度 $O\left(d^2\right)$。且计算 $\boldsymbol{C}_n$ 需要 $O\left(d^2n\right)$，计算计 $\hat{\boldsymbol{\mu}}_n$ 需要 $O\left(dn\right)$，因此计算 $\boldsymbol{C}_n$ 的时间复杂度为 $O\left(d^2n\right)$。

Question 5

(a)

$$\boldsymbol{\theta}=\left[\begin{array}{c}\theta_1\\ \theta_2\end{array}\right],\ \boldsymbol{\theta}^0=\left[\begin{array}{c}2\\ 4\end{array}\right]$$

$$\begin{aligned}
Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^0\right)&=\varepsilon_{x_{32}}\left[\ln p\left(\boldsymbol{x}_g,\boldsymbol{x}_b;\boldsymbol{\theta}\right)|\boldsymbol{\theta}^0;D_g\right]\\
&=\ln p\left(x_1|\boldsymbol{\theta}\right)+\ln p\left(x_2|\boldsymbol{\theta}\right)+\int_{-\infty}^{\infty}\ln p\left(x_{32}|\boldsymbol{\theta}\right)p\left(x_{32}|\boldsymbol{\theta}^0,x_{31}=2\right)dx_{32}\\
&=\ln p\left(x_1|\boldsymbol{\theta}\right)+\ln p\left(x_2|\boldsymbol{\theta}\right)+\int_{-\infty}^{\infty}\ln p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}\right)\frac{p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}^0\right)}{\int_{-\infty}^{\infty}p\left(\left(\begin{array}{c}2\\ x'_{32}\end{array}\right)|\boldsymbol{\theta}^0\right)dx'_{32}}dx_{32}\\
&=\ln p\left(x_1|\boldsymbol{\theta}\right)+\ln p\left(x_2|\boldsymbol{\theta}\right)+2e^4\int_{-\infty}^{\infty}\ln p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}\right)p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}^0\right)dx_{32}
\end{aligned}$$

设

$$\begin{aligned}
K&=2e^4\int_{-\infty}^{\infty}\ln p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}\right)p\left(\left(\begin{array}{c}2\\ x_{32}\end{array}\right)|\boldsymbol{\theta}^0\right)dx_{32}\\
&=\int_{-\infty}^{\infty}\ln\left[p\left(2|\theta_1\right)p\left(x_{32}|\theta_2\right)\right]p\left(x_{32}|\theta_2=4\right)dx_{32}\\
&=\frac{1}{4}\int_0^4\ln\left[\frac{1}{\theta_1}e^{-2\theta_1}p\left(x_{32}|\theta_2\right)\right]dx_{32}
\end{aligned}$$

由于 $p\left(x_2\right)\sim U\left(0,\theta_2\right)$，且存在样本 $x_{22}=3$，因此 $\theta_2\geq 3$
当 $3\leq\theta_2\leq 4$ 时，

$$\begin{aligned}
K&=\frac{1}{4}\int_0^{\theta_2}\ln\left[\frac{1}{\theta_1}e^{-2\theta_1}\frac{1}{\theta_2}\right]dx_{32}\\
&=\frac{1}{4}\theta_2\ln\left(\frac{1}{\theta_1}e^{-2\theta_1}\frac{1}{\theta_2}\right)\\
&=-\frac{1}{2}\theta_1\theta_2-\frac{1}{4}\theta_2\ln\left(\theta_1\theta_2\right)
\end{aligned}$$

当 $\theta_2>4$ 时，

$$\begin{aligned}
K&=\frac{1}{4}\int_0^4\ln\left(\frac{1}{\theta_1}e^{-2\theta_1}\frac{1}{\theta_2}\right)dx_{32}\\
&=\ln\left(\frac{1}{\theta_1}e^{-2\theta_1}\frac{1}{\theta_2}\right)\\
&=-2\theta_1-\ln\left(\theta_1\theta_2\right)
\end{aligned}$$

$$Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^0\right)=\left\{\begin{array}{c}\ln p\left(x_1|\boldsymbol{\theta}\right)+\ln p\left(x_2|\boldsymbol{\theta}\right)+-\frac{1}{2}\theta_1\theta_2-\frac{1}{4}\theta_2\ln\left(\theta_1\theta_2\right),3\leq\theta_2\leq 4\\ \ln p\left(x_1|\boldsymbol{\theta}\right)+\ln p\left(x_2|\boldsymbol{\theta}\right)-2\theta_1-\ln\left(\theta_1\theta_2\right),\theta_2>4\end{array}\right.$$

(b)

由于 $\int_{-\infty}^{\infty} p\left(x_1\right) dx_1=\int_{-\infty}^{\infty} \frac{1}{\theta_1} e^{-\theta_1 x_1} dx_1=1$，因此 θ_1= 1。
当 $3 \leq \theta_2 \leq 4$ 时，

$$Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^0\right)=-4-2 \ln \theta_2-\frac{1}{2} \theta_2-\frac{1}{4} \theta_2 \ln \theta_2$$

由于 $\frac{\partial Q}{\partial \theta_2}=-\frac{2}{\theta_2}-\frac{1}{4} \ln \theta_2-\frac{3}{4}<0$，所以 θ_2= 3 时取最大。
当 θ_2> 4 时，

$$Q\left(\boldsymbol{\theta};\boldsymbol{\theta}^0\right)=-6-3 \ln \theta_2$$

由于 $\frac{\partial Q}{\partial \theta_2}=-\frac{3}{\theta_2}<0$，所以 θ_2= 4 时取最大。

综上，又由于两种情况在 θ_2= 4 时取值相同，因而使 Q 最大的 $\boldsymbol{\theta}=\begin{bmatrix} 1 \\ 3 \end{bmatrix}$

Question 6

$$\gamma_{ij}(t)=\frac{\alpha_i(t-1) a_{ij} b_{ij} \beta_i(t)}{P\left(\boldsymbol{V}^T|\boldsymbol{\theta}\right)}$$

$$\hat{a}_{ij}=\frac{\sum\limits_{t=1}^{T} \gamma_{ij}(t)}{\sum\limits_{t=1}^{T} \sum\limits_{k=1}^{c} \gamma_{ik}(t)}$$

$$\hat{b}_{jk}=\frac{\sum\limits_{k=1}^{c} \gamma_{jk}(t)}{\sum\limits_{t=1}^{T} \gamma_{jk}(t)}$$

要计算 $\gamma_{ij}(t)$，通过前向算法可以计算出 $P\left(\boldsymbol{V}^T|\boldsymbol{\theta}\right)$ 和 $\alpha_i(t)$，时间复杂度为 $O\left(c^2 T\right)$，后向算法可以算出 $\beta_i(t)$，时间复杂度同样为 $O\left(c^2 T\right)$。因此，计算 $\hat{a}_{ij}$ 和 $\hat{b}_{jk}$ 时，计算 $\gamma_{ij}(t)$ 的累加也需要 $O\left(c^2 T\right)$。综合起来所需的时间复杂度为 $O\left(c^2 T\right)$。