HW3

111061590

林學律

1)

We have the training set $\underline{X} = [\underline{x}_1, \underline{x}_2 \dots, \underline{x}_N]$ and target we observed $\underline{T} = [\underline{t}_1, \underline{t}_2 \dots, \underline{t}_N]$

We want to know t_{N+1} conditional distribution given known $\underline{T}$

$$p(t_{N+1} | \underline{T}) = N(\underline{t_{N+1}} \mid m(\underline{x_{N+1}}), \sigma^2(\underline{x_{N+1}}))$$

From the problem, we know that target variable has D dimenality, so we can know that $t_i = 1 \times D$ dimension

From the Bishop books, Ch6, (6.66) and (6.67)

$$m(\underline{x_{N+1}}) = k^T C_N^{-1} \underline{t}, \quad \sigma^2(\underline{x_{N+1}}) = c - k^T C_N^{-1} k$$

Now, we define C_N covariance

Univariate

$$C_N = k(\underline{x}) + \beta^{-1} \delta$$

Multivariate

$$C_N = k(\underline{x_n}, \underline{x_m}) + \beta^{-1} \delta_{nm}$$

2)

$$p(\underline{t}|\underline{X},\underline{\alpha},\beta) = \int p(\underline{t}|\underline{X},\underline{w},\beta)\,p(\underline{w}|\underline{\alpha})\,dw$$

From books (7.79)

$$p(\underline{t}|\underline{X},\underline{w},\beta) = \prod_{n=1}^{N} p(t_n|\underline{x_n},\underline{w},\beta^{-1})$$

(7.80)

$$p(\underline{w}|\underline{\alpha}) = \prod_{i=1}^{M} \mathcal{N}(w_i|0,\alpha_i^{-1})$$

$$E_D(w) = \frac{\beta}{2}\sum_{n=1}^{N}(t_n - w^T\phi(x_n))^2$$

We can get the information from (3.11)

$$p(t|\underline{X},\underline{W},\beta) = \mathcal{N}(t_n|w^T\phi(x_n),\beta^{-1}) = \frac{1}{(2\pi\beta^{-1})^{\frac{N}{2}}}\cdot\exp\left(-\frac{1}{2}\sum_{n}\frac{(t_n - w^T\phi(x_n))^2}{\beta^{-1}}\right)$$

$$= \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}}\int \exp(-E_D(w))\cdot dw$$

$$p(\underline{w}|\underline{\alpha}) = \frac{1}{(2\pi)^{\frac{M}{2}}}\prod_{i=1}^{M}\alpha_i^{\frac{1}{2}}\exp\left(-\frac{1}{2}\sum_{i=1}^{M}\frac{(w_i)^2}{\alpha_i^{-1}}\right)$$

$$p(\underline{t}|\underline{X},\underline{\alpha},\beta) = \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}}\left(\frac{1}{2\pi}\right)^{\frac{M}{2}}\prod_{i=1}^{M}\alpha_i^{\frac{1}{2}}\int\exp\left(-E_D(w) - \frac{1}{2}\sum_{i=1}^{M}\alpha_i w_i^2\right)dw$$

處理 $\exp\left(-E_D(w) - \frac{1}{2}\sum_{i=1}^{M}\alpha_i w_i^2\right)$

$$K(w) = \exp\left(-E_D(w) - \frac{1}{2}\sum_{i=1}^{M}\alpha_i w_i^2\right)$$

$$-\left(E_D(w) + \frac{1}{2}\sum_{i=1}^{M}\alpha_i w_i^2\right) = -\left(\frac{\beta}{2}\sum_{n=1}^{N}(t_n - w^T\phi(x_n))^2 + \frac{1}{2}\sum_{i=1}^{M}\alpha_i w_i^2\right)$$

$$\text{Now} = \frac{1}{2}\left\{ w^T(A+\beta\Phi^T\Phi)w - 2\beta t^T(\Phi w) + \beta t^T t \right\}$$

$$= \frac{1}{2}\left\{ w^T\Sigma^{-1}w - 2m^T\Sigma^{-1}w + \beta t^T t \right\}$$

$$= -\left(\frac{1}{2}(w-m)^T\Sigma^{-1}(w-m) + \frac{1}{2}\left(\beta t^T t - m^T\Sigma^{-1}m\right)\right)$$

Σ 為 covariance $\quad \Sigma = \left(\mathrm{diag}(\alpha_i) + \beta\Phi^T\Phi\right)^{-1}$, $\Phi \Rightarrow \phi_i(x_n)$

m 為 mean $\quad m = \beta\Sigma\Phi^T t$

$$\int \exp\left[-\left(\frac{1}{2}(w-m)^T\Sigma^{-1}(w-m) + \frac{1}{2}\left(\beta t^T t - m\Sigma^{-1}m\right)\right)\right] \cdot dw$$

$$= \exp\left[-\frac{1}{2}\left(\beta t^T t - m\Sigma^{-1}m\right)\right]\cdot(2\pi)^{\frac{N}{2}}\cdot|\Sigma|^{\frac{1}{2}}$$

$$= \frac{1}{2}\left(\beta t^T t - m^T\Sigma^{-1}m\right)$$, m and Σ 代入

$$= \frac{1}{2}\left(\beta t^T t - \beta t^T\Phi\Sigma\Phi^T t\beta\right)$$. Use woodbury identity.

$$= \frac{1}{2}t^T\left(\beta^{-1} + \Phi A^{-1}\Phi^T\right)^{-1}t$$ 與課本上的 $C = \beta^{-1} + \Phi A^{-1}\Phi^T$ 一樣

改寫

$$= \frac{1}{2}t^T C^{-1} t$$

$$p(\underline{t} \mid \underline{X}, \underline{\alpha}, \beta) = \left(\frac{1}{2\pi}\right)^{\frac{N}{2}} \left(\frac{1}{2\pi}\right)^{\frac{M}{2}} \prod_{i=1}^{M} \alpha_i^{\frac{1}{2}} \cdot (2\pi)^{\frac{M}{2}} \cdot |\Sigma|^{\frac{1}{2}} \exp\left[-\frac{1}{2} t^T C^{-1} t\right]$$

$$= \left(\frac{\beta}{2\pi}\right)^{\frac{N}{2}} |\Sigma|^{\frac{1}{2}} \prod_{i=1}^{M} \alpha_i^{\frac{1}{2}} \exp\left[-\frac{1}{2} t^T C^{-1} t\right]$$

If $\alpha_i = 1$ and $\beta = 1$, $i = 1, 2, \ldots, M$

Show $\ln p(\underline{t} \mid \underline{X}, \underline{\alpha}, \beta) = -\left(\ln (2\pi)^{\frac{N}{2}} + \ln |\Sigma|^{\frac{1}{2}} + \ln \exp\left(\frac{1}{2} t^T C^{-1} t\right)\right)$

$$= -\left\{\frac{N}{2} \ln(2\pi) + \frac{1}{2} \ln|\Sigma| + \frac{1}{2} t^T C^{-1} t\right\} \#$$

3)

We have to show that $2^{n(n-1)/2}$

① If there is one vertics,

only one possible

② Two vertics

C^1_0 $\qquad\qquad$ C^1_1

Because there is only one edge between two vertics

so if there are two edges, it means that $n=3$ at least.

No edge : C^3_0

one edge : C^3_1

two edges : C^3_2

three edges : C^3_3

$\Rightarrow$ We can derive that

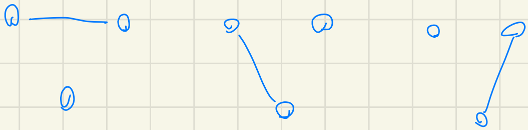
$$A = C^{n(n-1)/2}_0 + C^{n(n-1)/2}_1 + C^{n(n-1)/2}_2 + \cdots + C^{n(n-1)/n}_n$$

Now, we can show that it's equal to $2^{n(n-1)/2}$

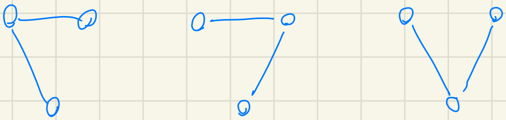
$$A = C^{n(n-1)/2}_0 \cdot 1^{n(n-1)/2} \cdot 1^0 + C^{n(n-1)/2}_1 \cdot 1^{(n(n-1)/2 - 1)} \cdot 1^1$$
$$+ \cdots + C^{n(n-1)/2}_{n(n-1)/2} \cdot 1^0 \cdot 1^{n(n-1)/2}$$
$$= (1+1)^{n(n-1)/2} = 2^{n(n-1)/2} \quad \#$$

C^3_0

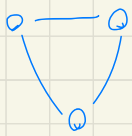
C^3_1



C^3_2



C^3_3



4)

We have to show in Sec 7 and Sec 9, α_i^{new} and $(\beta^{new})^{-1}$ is equivalent form in EM algorithms.

① $\alpha_i = \frac{\gamma_i}{m_i^2}$, $\gamma_i = 1 - \alpha_i \Sigma_{ii}$ (Sec. 7)

$\Rightarrow \alpha_i = \frac{1 - \alpha_i \Sigma_{ii}}{m_i^2} \Rightarrow \alpha_i^{new} = \frac{1 - \alpha_i^{new} \Sigma_{ii}}{m_i^2}$

$\Rightarrow \alpha_i^{new} (m_i^2 + \Sigma_{ii}) = 1 \Rightarrow \alpha_i^{new} = \frac{1}{(m_i^2 + \Sigma_{ii})}$ # 得証 (Sec. 9)

② This time, we used function in EM to get the same function in Section 7

$$(\beta^{new})^{-1} = \frac{\|t - \Phi m_N\|^2}{N} + \frac{\Sigma_i \gamma_i}{\beta^{new} N} \quad \text{(Sec. 9.)}$$

$$\Rightarrow \frac{N - \Sigma_i \gamma_i}{\beta^{new} N} = \frac{\|t - \Phi m_N\|^2}{N}$$

$$\Rightarrow \left(\frac{1}{\beta^{new}}\right) = \frac{\|t - \Phi m_N\|^2}{N} \cdot \frac{N}{N - \Sigma_i \gamma_i}$$

$$= \frac{\|t - \Phi m_N\|^2}{N - \Sigma_i \gamma_i}$$ # 得証 (Sec. 7)