1、

假設 target $t_1, t_2, \dots, t_N$ 與輸入 x 是 Independent

則可表達成：

$$p(t_{N+1} \mid T) = N(t_{N+1} \mid m(x_{N+1}), \sigma(x_{N+1}) I)$$

而 T 為一 $N \times D$ 的矩陣 ⇒ $t_1^T, t_2^T, \dots, t_N^T$

$$m(x_{N+1})^T = k^T C_N T$$

$$\sigma^2(x_{N+1}) = c - k^T C_N^{-1} k \quad \Rightarrow \text{from 課本 } 6.67$$

C_N 僅與輸入有關

#

2.

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

$$= \prod_{n=1}^{N} \sqrt{\frac{\beta}{2\pi}}\, e^{-(t_n - w^T\phi)\frac{\beta}{2}(t_n - \phi w)}$$

↓ 根據 7.77 ，得知 mean 爲 $w^T\phi$

根據 7.80

$$p(w|\alpha) = \left(\frac{\alpha_1}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{\alpha_1 w_1^2}{2}} \times \left(\frac{\alpha_2}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{\alpha_2 w_2^2}{2}} \times \cdots \times \left(\frac{\alpha_M}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{\alpha_M w_M^2}{2}}$$

$$= \prod_{i=1}^{M} \left(\frac{\alpha_i}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{\alpha_i w_i^2}{2}}$$

$$= \frac{\sqrt{\alpha_i}}{(2\pi)^{\frac{M}{2}}} e^{-\frac{1}{2} w^T A w}$$

$$w^T = (w_1, w_2, w_3) \quad , \quad A = \begin{bmatrix} \alpha_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \alpha_M \end{bmatrix}$$

$p(t|X,w,\beta)\, p(w|\alpha) \Rightarrow$ 指數項相加：

$$T_1 = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} (\alpha_j \delta_{ij} + \beta \phi_i^n \phi_j^n) w_i w_j + \sum_{i=1}^{M} \beta w_i \phi_i^n t_n$$

↓ t_n^2 先不管，後面再加回來

$$= -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \sigma_{ij}^{-1} w_i w_j + \sum_{i=1}^{M} \beta w_i \phi_i^n t_n$$

where $\sigma_{ij}^{-1} = \alpha_j \delta_{ij} + \beta \phi_i^n \phi_j^n$

$$T_1 = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \sigma_{ij}^{-1} (w_i - \sigma_{ij}\beta\phi_i^n t_n)(w_j - \sigma_{ij}\beta\phi_j^n t_n) +$$

$$\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \beta^2 \sigma_{ij}\phi_i^n \phi_j^n t_n^2$$

$$= T_{11} + T_{12}$$

$$T_{11} = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \sigma_{ij}^{-1} (w_i - \sigma_{ij}\beta\phi_i^n t_n)(w_j - \sigma_{ij}\beta\phi_j^n t_n)$$

$$T_{12} = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \beta^2 \sigma_{ij}\phi_i^n \phi_j^n t_n^2$$

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

$$= \int \prod_{n=1}^{N} \sqrt{\frac{\beta}{2\pi}} \frac{\sqrt{|A|}}{(2\pi)^{M/2}} e^{(T_{11} + T_{12} - \frac{\beta}{2} t_n^2)}$$

$$= \sqrt{|A||\Sigma|} \prod_{n=1}^{N} \sqrt{\frac{\beta}{2\pi}} e^{(\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M} \beta^2 \sigma_{ij}\phi_i^n\phi_j^n t_n^2 - \frac{\beta}{2}t_n^2)}$$

$$\ln(p(t|X,\alpha,\beta)) = \frac{1}{2}\ln|A||\Sigma| + \frac{1}{2}N\ln\left(\frac{\beta}{2\pi}\right) - \sum_{n=1}^{N}\frac{1}{2}(\beta t_n^2 -$$

$$\sum_{i=1}^{M}\sum_{j=1}^{M}\beta^2\sigma_{ij}\phi_i^n\phi_j^n t_n^2)$$

$$= -\frac{1}{2}N\ln(2\pi) - \frac{1}{2}\ln\left(\frac{\beta^{-1}}{|A\Sigma|}\right) - \sum_{n=1}^{N}\frac{1}{2}(\beta t_n^2 -$$

$$\sum_{i=1}^{M}\sum_{j=1}^{M}\beta^2\sigma_{ij}\phi_i^n\phi_j^n t_n^2)$$

from 7.83

$$= -\frac{1}{2}\Big\{N\ln(2\pi) + \ln(\beta^{-1}|(A+\beta\phi^T\phi)A^{-1}|)$$

$$+ \sum_{n=1}^{N}\frac{1}{2}(\beta t_n^2 - \beta^2\phi^n\Sigma(\phi^T)^n t_n^2)\Big\}$$

$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(\beta^{-1}|I + \beta\phi^T\phi A^{-1}|) +$$
$$t^T(\beta I - \beta^2\phi\Sigma(\phi^T))t\}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T(\beta I - \beta^2\phi(A + \beta\phi^T\phi)^{-1}\phi^T)t\}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T(\beta I - \beta^2((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1}t\}$$
$$\because \beta I - \beta^2((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1}$$
$$= \beta((\phi^T)^{-1}A\phi^{-1} + \beta I)((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1} - \beta^2((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T(\beta(\phi^T)^{-1}A\phi^{-1}((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1}t\}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T\beta(\phi A^{-1}\phi^T)^{-1}((\phi^T)^{-1}A\phi^{-1} + \beta I)^{-1}t\}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T\beta((\phi^T)^{-1}A\phi^{-1}(\phi A^{-1}\phi^T)^{-1} + \beta I(\phi A^{-1}\phi^T))^{-1}t\}$$
$$= -\frac{1}{2}\{N\ln(2\pi) + \ln(|\beta^{-1}I + \phi A^{-1}\phi^T|) +$$
$$t^T(\beta^{-1}I + (\phi A^{-1}\phi^T))^{-1}t$$

根據 7.86，$C = \beta^{-1}I + \phi A^{-1}\phi^T$

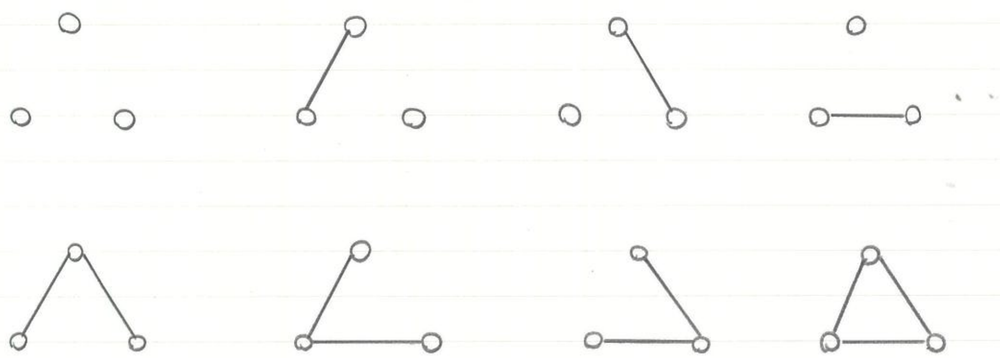
$$\ln p(t \mid X, \alpha, \beta) = -\frac{1}{2}\left\{N \ln(2\pi) + \ln(|C|) + t^T C^{-1} t\right\}$$

得證 #

3.

若空間中有 M 個點，那一共有 $C^M_2 = \frac{M(M-1)}{2}$ 種組合能組合成線（因為兩點才可以連成一條線）；而 Graph 的定義成要取空間中哪些線，故每條線都是選或不選的可能性，所以就會有 $2^{\frac{M(M-1)}{2}}$ 個 Graphs.

For $M = 3$, $2^{\frac{3(3-1)}{2}} = 8$



#

4.

$$(\beta_{k+1})^{-1} = \frac{\|t - \phi m_N\|^2 + (\beta_k)^{-1} \Sigma_i \gamma_i}{N}$$

若要使其收斂 (convergence) :

$$\beta_{k+1} = \beta_k = \beta$$

故

$$N\beta^{-1} = \|t - \phi m_N\|^2 + \beta^{-1} \Sigma_i \gamma_i$$

$$\Rightarrow \beta^{-1}(N - \Sigma_i \gamma_i) = \|t - \phi m_N\|^2$$

$$\Rightarrow \beta^{-1} = \frac{\|t - \phi m_N\|^2}{N - \Sigma_i \gamma_i} \quad \#$$

$$\alpha_i^{k+1} = \frac{1}{m_i^2 + \Sigma_{ii}}$$

$$\Rightarrow \alpha_i^{k+1}(m_i^2 + \Sigma_{ii}) = 1$$

$$\Rightarrow \alpha_i^{k+1} m_i^2 = 1 - \alpha_i^{k+1} \Sigma_{ii}$$

$$\Rightarrow \because \gamma_i = 1 - \alpha_i \Sigma_{ii} \quad \text{from } 7.89$$

$$\therefore \alpha_i^{k+1} m_i^2 = \gamma_i$$

$$\Rightarrow \alpha_i = \frac{\gamma_i}{m_i^2} \quad \#$$