9.22

(7.81) $p(w|t,X,\alpha,\beta)=N(w|m,\Sigma)$

(7.82) $m=\beta\Sigma\Phi^Tt$

(7.83) $\Sigma=(A+\beta\Phi^T\Phi)^{-1}$

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

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

(7.76) $p(t|x,w,\beta)=N(t|y(x),\beta^{-1})$

(7.77) $y(x)=\sum_{i=1}^{M}w_i\phi(x)_i$

なので

$$
\begin{aligned}
E_w[\ln p(t|X,w,\beta)p(w|\alpha)] &= \int p(w|t)\ln p(t|X,w,\beta)p(w|\alpha)\,dw \\
&= \int p(w|t)\ln \prod_{n=1}^{N}\overset{(7.79),(7.76)}{N(t_n|y_n,\beta^{-1})}\prod_{i=1}^{M}\overset{(7.80)}{N(w_i|0,\alpha_i^{-1})}\,dw \\
&= \int p(w|t)\left\{\sum_n \ln N(t_n|y_n,\beta^{-1})+\sum_i \ln N(w_i|0,\alpha_i^{-1})\right\}dw \\
&= \int p(w|t)\left[\sum_n\left\{\frac{1}{2}\ln\frac{\beta}{2\pi}-\frac{1}{2}\beta(t_n-y_n)^2\right\}+\sum_i\left(\frac{1}{2}\ln\frac{\alpha_i}{2\pi}-\frac{1}{2}\alpha_i w_i^2\right)\right]dw \\
&= \frac{N}{2}\ln\frac{\beta}{2\pi}-\frac{\beta}{2}\sum_n E_w[(t_n-y_n)^2]+\sum_i\frac{1}{2}\ln\frac{\alpha_i}{2\pi}-\frac{1}{2}\sum_i\alpha_i E_w[w_i^2] \quad\cdots ①
\end{aligned}
$$

ここで

$$E_w[w_i]=(E_w[w])_i=\left(\int p(w|t)w\,dw\right)_i\overset{(7.81)}{=}\left(\int N(w|m,\Sigma)w\,dw\right)_i=m_i$$

$$E_w[w_iw_j]=(E_w[ww^T])_{ij}=\left(\int p(w|t)ww^T dw\right)_{ij}\overset{(7.81)}{=}\left(\int N(w|m,\Sigma)ww^T dw\right)_{ij}\overset{(2.62)}{=}(mm^T+\Sigma)_{ij}$$

$$E_w[y_n]=E_w\left[\sum_i w_i\phi_{ni}\right]=\sum_i E_w[w_i]\phi_{ni}=\sum_i m_i\phi_{ni}=m^T\phi_n$$

$$
\begin{aligned}
E_w[y_n^2] &= E_w\left[\left(\sum_i w_i\phi_{ni}\right)^2\right]=E_w\left[\sum_i\sum_j w_i\phi_{ni}w_j\phi_{nj}\right]=\sum_i\sum_j E_w[w_iw_j]\phi_{ni}\phi_{nj}=\sum_i\sum_j(mm^T+\Sigma)_{ij}\phi_{ni}\phi_{nj} \\
&= \sum_i\sum_j(mm^T)_{ij}\phi_{ni}\phi_{nj}+\sum_i\sum_j\Sigma_{ij}\phi_{ni}\phi_{nj} \\
&= \phi_n^Tmm^T\phi_n+\phi_n^T\Sigma\phi_n
\end{aligned}
$$

$$
\left\{
\begin{aligned}
\phi_n^Tmm^T\phi_n &= (\phi_{n1}\ \phi_{n2})\begin{pmatrix}m_1\\m_2\end{pmatrix}(m_1\ m_2)\begin{pmatrix}\phi_{n1}\\\phi_{n2}\end{pmatrix} \\
&= (\phi_{n1}m_1+\phi_{n2}m_2)(m_1\phi_{n1}+m_2\phi_{n2}) \\
&= m_1m_1\phi_{n1}\phi_{n1}+m_1m_2\phi_{n1}\phi_{n2}+m_2m_1\phi_{n2}\phi_{n1}+m_2m_2\phi_{n2}\phi_{n2} \\
\phi_n^T\Sigma\phi_n &= (\phi_{n1}\ \phi_{n2})\begin{pmatrix}\Sigma_{11}&\Sigma_{12}\\\Sigma_{21}&\Sigma_{22}\end{pmatrix}\begin{pmatrix}\phi_{n1}\\\phi_{n2}\end{pmatrix}=(\phi_{n1}\ \phi_{n2})\begin{pmatrix}\Sigma_{11}\phi_{n1}+\Sigma_{12}\phi_{n2}\\\Sigma_{21}\phi_{n1}+\Sigma_{22}\phi_{n2}\end{pmatrix} \\
&= \Sigma_{11}\phi_{n1}\phi_{n1}+\Sigma_{12}\phi_{n1}\phi_{n2}+\Sigma_{21}\phi_{n2}\phi_{n1}+\Sigma_{22}\phi_{n2}\phi_{n2}
\end{aligned}
\right.
$$

$$
\begin{aligned}
E_w[(t_n-y_n)^2] &= E_w[t_n^2-2t_ny_n+y_n^2]=t_n^2-2t_nm^T\phi_n+\phi_n^Tmm^T\phi_n+\phi_n^T\Sigma\phi_n \\
&= (t_n-m^T\phi_n)^2+\phi_n^T\Sigma\phi_n
\end{aligned}
$$

である。

① を α で微分して 0 とおくと

$$\frac{\partial E}{\partial \alpha_i} = \frac{1}{2}\frac{1}{\alpha_i} - \frac{1}{2}E_w[w_i^2] = 0$$

$$\therefore \frac{1}{\alpha_i} - (mm^T + \Sigma)_{ii} = 0$$

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

となり、(9.67) を得る。

① を β で微分して 0 とおくと

$$\frac{\partial E}{\partial \beta} = \frac{N}{2}\frac{1}{\beta} - \frac{1}{2}\sum_n E_w\left[(t_n - y_n)^2\right] = 0$$

スカラーなので $m^T\phi_n = \phi_n^T m$

$$\therefore \frac{1}{\beta} = \frac{1}{N}\sum_n\left\{(t_n - \phi_n^T m)^2 + \phi_n^T \Sigma \phi_n\right\}$$

$$= \frac{1}{N}\left\{\|t - \Phi m\|^2 + \mathrm{Tr}(\Phi\Sigma\Phi^T)\right\}$$

$$= \frac{1}{N}\left\{\|t - \Phi m\|^2 + \frac{1}{\beta}\sum_i(1 - \alpha_i\Sigma_{ii})\right\}$$

となり (9.68) を得る。

$$\Phi = \begin{pmatrix}\phi_1^T \\ \phi_2^T \\ \vdots\end{pmatrix} \cdots (3.16) \text{ より } \Phi m = \begin{pmatrix}\phi_1^T m \\ \phi_2^T m \\ \vdots\end{pmatrix}$$

$$\therefore (t - \Phi m)^T(t - \Phi m) \overset{\text{ベクトルの内積}}{=} \sum_n (t_n - \phi_n^T m)^2$$

$$\Phi\Sigma\Phi^T = \begin{pmatrix}\phi_1^T \\ \phi_2^T\end{pmatrix}\Sigma(\phi_1\ \phi_2) \overset{\text{ブロック行列の積}}{=} \begin{pmatrix}\phi_1^T\Sigma \\ \phi_2^T\Sigma\end{pmatrix}(\phi_1\ \phi_2) = \begin{pmatrix}\phi_1^T\Sigma\phi_1 & \phi_1^T\Sigma\phi_2 \\ \phi_2^T\Sigma\phi_1 & \phi_2^T\Sigma\phi_2\end{pmatrix}$$

$$\therefore \mathrm{Tr}(\Phi\Sigma\Phi^T) = \phi_1^T\Sigma\phi_1 + \phi_2^T\Sigma\phi_2 + \cdots = \sum_n \phi_n^T\Sigma\phi_n$$

$$(7.83)\quad \Sigma = (A + \beta\Phi^T\Phi)^{-1}$$

より

$$(A + \beta\Phi^T\Phi)\Sigma = I$$

$$\therefore \Phi^T\Phi\Sigma = \frac{1}{\beta}(I - A\Sigma)$$

$$\therefore \mathrm{Tr}(\Phi^T\Phi\Sigma) = \frac{1}{\beta}\mathrm{Tr}(I - A\Sigma) = \frac{1}{\beta}\sum_i(1 - \alpha_i\Sigma_{ii})$$

よって

$$\mathrm{Tr}(\Phi\Sigma\Phi^T) \overset{(C.9)}{=} \mathrm{Tr}(\Phi^T\Phi\Sigma) = \frac{1}{\beta}\sum_i(1 - \alpha_i\Sigma_{ii})$$