7.19

ノートのタイトル 2016/03/05

（w^*のα_iへの依存性について）

(7.112)を見ると w^*はα_iに依存しているように見える。

しかしw^*は元々$p(w|t,\alpha)$のモードだからαとはあまり関係ないように思われる。

(7.112)については、w^*についての陰関数表記であり、反復して得たw^*の値はα_iに（ほとんど）依らない可能性がある。

(7.112の右辺のyがw^*に依るので)

事後分布のモードは、事前分布の分散にあまり依存しないように思われるので

一応是である理由はこれ

とりあえず、w^*とα_iは独立で

$$\frac{\partial w^*}{\partial \alpha_i} = 0$$

としないと(7.115)が導けない。（若しくは$\frac{\partial w^*}{\partial \alpha_i} = 0$と近似すると言ってしまっても良い。近似の是非については別にギロンすれば良い）

(7.114)より

$$\ln p(t|\alpha) \simeq \ln p(t|w^*,\alpha) + \ln p(w^*|\alpha) + \frac{M}{2}\ln(2\pi) + \frac{1}{2}\ln|\Sigma| \quad \cdots ①$$

ここで(4.89)より

$$\ln p(t|w^*,\alpha) = \sum_{n=1}^{N}\{t_n \ln y_n^* + (1-t_n)\ln(1-y_n^*)\}, \quad y_n^* = \sigma((w^*)^T\phi(x_n))$$

よって

$$\frac{\partial}{\partial \alpha_i}\ln p(t|w^*,\alpha) = 0 \longleftarrow \frac{\partial w^*}{\partial \alpha_i} = 0 \text{ とするので、} \frac{\partial y_n^*}{\partial \alpha_i} = 0 \text{ より}$$

また、(7.80)より

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

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

$$= \sum_{i=1}^{M}\left\{\frac{M}{2}\ln(2\pi) + \frac{1}{2}\ln\alpha_i - \frac{1}{2}\alpha_i(w_i^*)^2\right\}$$

これより

$$\frac{\partial}{\partial \alpha_i}p(w^*|\alpha) = \frac{1}{2\alpha_i} - \frac{1}{2}(w_i^*)^2$$

さらに

$$\frac{\partial}{\partial \alpha_i}\ln|\Sigma| = \frac{\partial}{\partial \alpha_i}\ln\overset{(C.13)}{\frac{1}{|\Sigma^{-1}|}} = -\frac{\partial}{\partial \alpha_i}\ln|\Sigma^{-1}| = \overset{(C.22)}{-\mathrm{Tr}\left(\Sigma\frac{\partial \Sigma^{-1}}{\partial \alpha_i}\right)}$$

$$= -\mathrm{Tr}\left\{\Sigma\frac{\partial}{\partial \alpha_i}(\Phi^T B\Phi + A)\right\} \longleftarrow (7.113)$$

$$= -\mathrm{Tr}\left(\Sigma\frac{\partial A}{\partial \alpha_i}\right) \longleftarrow \frac{\partial y_n^*}{\partial \alpha_i} = 0 \text{ なので } \frac{\partial B}{\partial \alpha_i} = 0$$

$$= -\mathrm{Tr}(\Sigma J_{ii}) \longleftarrow \frac{\partial A}{\partial \alpha_1} = \begin{pmatrix}\alpha_1 & 0\\ 0 & \alpha_2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} = \overset{\text{single entry matrix}}{J_{11}}$$

$$= -\Sigma_{ii}$$

これらを使って①を微分し、0とおくと

$$\frac{\partial}{\partial \alpha_i}\ln p(t|\alpha) = \frac{1}{2\alpha_i} - \frac{1}{2}(w_i^*)^2 - \frac{1}{2}\Sigma_{ii} = 0 \quad \cdots (7.115)$$

(7.115)が得られる。

(7.115)を変形すると

$$\frac{1}{\alpha_i} - (w_i^*)^2 - \Sigma_{ii} = 0$$

$$\therefore \frac{1}{\alpha_i} - \Sigma_{ii} = (w_i^*)^2$$

$$\therefore \frac{1-\alpha_i\Sigma_{ii}}{\alpha_i} = (w_i^*)^2$$

$$\therefore \alpha_i = \frac{1-\alpha_i\Sigma_{ii}}{(w_i^*)^2} = \frac{\gamma_i}{(w_i^*)^2}, \quad \gamma_i = 1-\alpha_i\Sigma_{ii} \quad \cdots (7.116)$$ を得る。