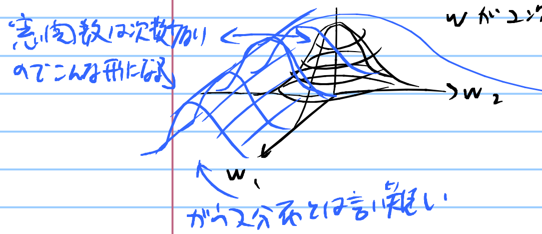
7.10

ノートのタイトル　　2016/02/22

本文で (7.84) が「2つのガウス分布の畳み込み積分」になってると言うが、2つのガウス分布の次数が異なるし、窓関数の中心の動きも $t-w$ ではなく $t-w\phi$ となっており簡単に2つのガウス分布の畳み込みとは言い難い状況である。

→ もしガウス分布の畳み込みなら結果もガウス分布で平均、分散はそれぞれの分布のものの和でいいはずだが。



w が2次のときこんな感じ

→中心の移動も $t-w_2$ でなく $t-w_2\phi(x)$ となり良く分からん動きである

なので (7.84) の積分は平方完成で解くしかない。

(7.79), (7.76), (7.77) より

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

$$=N\left(\begin{pmatrix}t_1\\ \vdots\\ t_N\end{pmatrix}\middle|\begin{pmatrix}\phi(x_1)^T w\\ \vdots\\ \phi(x_N)^T w\end{pmatrix},\begin{pmatrix}\beta^{-1}& &0\\ &\ddots& \\ 0& &\beta^{-1}\end{pmatrix}\right)=N(t|\Phi w,B^{-1}),\quad B=\beta I_N$$

$w^T\phi(x)=\phi(x)^T w$ (スカラーなので)

P.141 の Φ の定義より (注) (errata から ϕ_i とおかしいと思われる。このため訳もおかしくなっている)

$$\Phi=\begin{pmatrix}\phi(x_1)^T\\ \vdots\\ \phi(x_N)^T\end{pmatrix}$$ ← 多分この定義でいい筈

対角行列の逆行列は要素を逆数にしたもの

(7.80) より

$$p(w|\alpha)=\prod_{i=1}^{M}N(w_i|0,\alpha_i^{-1})=N\left(\begin{pmatrix}w_1\\ \vdots\\ w_M\end{pmatrix}\middle|\begin{pmatrix}0\\ \vdots\\ 0\end{pmatrix},\begin{pmatrix}\alpha_1^{-1}&&0\\&\ddots&\\0&&\alpha_M^{-1}\end{pmatrix}\right)=N(w|0,A^{-1}),\quad A=\mathrm{diag}(\alpha_i)$$

これらより (7.84) の積分は

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

$$=\int N(t|\Phi w,B^{-1})N(w|0,A^{-1})dw,\quad B=\beta I_N,\ A=\mathrm{diag}(\alpha_i)$$

$$=\int\frac{1}{(2\pi)^{\frac{N}{2}}}\frac{1}{|B^{-1}|^{\frac{1}{2}}}\exp\left\{-\frac{1}{2}(t-\Phi w)^T B(t-\Phi w)\right\}\frac{1}{(2\pi)^{\frac{M}{2}}}\frac{1}{|A^{-1}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}w^T A w\right)dw$$

$$=\frac{1}{(2\pi)^{\frac{N}{2}}}\frac{1}{|B^{-1}|^{\frac{1}{2}}}\frac{1}{(2\pi)^{\frac{M}{2}}}\frac{1}{|A^{-1}|^{\frac{1}{2}}}\int\exp\left[-\frac{1}{2}\left\{(t-\Phi w)^T B(t-\Phi w)+w^T A w\right\}\right]dw$$

となる。$\exp(\cdot)$ の中を w について平方完成する

(二次形式の平方完成の仕方)

$$x^T A x+x^T B y+y^T C x+y^T D y=\underbrace{(x+\square)^T A(x+\square)}_{\text{これが目標の形}}+\cdots$$

□を出会い系で埋めていく

$$(t-\Phi w)^T B(t-\Phi w)+w^T A w=$$

$$=t^T Bt-t^T B\Phi w-w^T\Phi^T Bt+w^T\Phi^T B\Phi w+w^T A w$$

$$=w^T(A+\Phi^T B\Phi)w-w^T\Phi^T Bt-t^T B\Phi w+t^T Bt$$

$$=\left[w-\{t^T B\Phi(A+\Phi^T B\Phi)^{-1}\}^T\right]^T(A+\Phi^T B\Phi)\left[w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\right]-t^T B\Phi(A+\Phi^T B\Phi)^{-1}(A+\Phi^T B\Phi)(A+\Phi^T B\Phi)^{-1}\Phi^T Bt+t^T Bt$$

$$=\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}^T(A+\Phi^T B\Phi)\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}+t^T\{-B\Phi(A+\Phi^T B\Phi)^{-1}\Phi^T B+B\}t$$

$$=\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}^T(A+\Phi^T B\Phi)\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}+t^T(B^{-1}+\Phi A^{-1}\Phi^T)^{-1}t$$

(C.7) より $(B^{-1}+\Phi A^{-1}\Phi^T)^{-1}=(B^{-1})^{-1}-(B^{-1})^{-1}\Phi(A+\Phi^T B\Phi)^{-1}\Phi^T B$

$=B-B\Phi(A+\Phi^T B\Phi)^{-1}\Phi^T B$

これを元の積分に入れると

$$\int \exp\left[-\frac{1}{2}\{(t-\Phi w)^T B (t-\Phi w) + w^T A w\}\right] dw$$

$$= \int \exp\left[-\frac{1}{2}\left[\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}^T (A+\Phi^T B\Phi)\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\} + t^T(B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right]\right] dw$$

$$= (2\pi)^{\frac{M}{2}} |(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}} \underbrace{\int \frac{1}{(2\pi)^{\frac{M}{2}}} \frac{1}{|(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}^T (A+\Phi^T B\Phi)\{w-(A+\Phi^T B\Phi)^{-1}\Phi^T Bt\}\right] dw}_{\text{ガウス分布の積分なので }1} \exp\left\{-\frac{1}{2} t^T (B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right\}$$

$$= (2\pi)^{\frac{M}{2}} |(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2} t^T (B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right\}$$

を得る。これより

$$p(t|X,\alpha,\beta) = \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{(2\pi)^{\frac{M}{2}}} \frac{1}{|A^{-1}|^{\frac{1}{2}}} (2\pi)^{\frac{M}{2}} |(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}} \exp\left\{-\frac{1}{2} t^T (B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right\}$$

$$= \frac{1}{(2\pi)^{\frac{N}{2}}} \underset{\uparrow\, N\times N}{\frac{1}{|B^{-1}|^{\frac{1}{2}}}} \underset{\uparrow\, M\times M}{\frac{1}{|A^{-1}|^{\frac{1}{2}}}} \underset{\uparrow\, M\times M}{|(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} t^T (B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right\}$$

ここで

$$\frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{|A^{-1}|^{\frac{1}{2}}} |(A+\Phi^T B\Phi)^{-1}|^{\frac{1}{2}} = \frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{|A^{-1}(A+\Phi^T B\Phi)|^{\frac{1}{2}}} \quad \leftarrow (C.12)\ |AB|=|A||B|,\ (C.13)\ |A^{-1}| = \frac{1}{|A|} \text{ より}$$

$$= \frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{|I_M + A^{-1}\Phi^T B\Phi|^{\frac{1}{2}}}$$

$$= \frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{|I_N + (A^{-1}\Phi^T)^T (B\Phi)^T|^{\frac{1}{2}}} \quad \leftarrow (C.14)\ |I_N + AB^T| = |I_M + A^T B|$$

$$= \frac{1}{|B^{-1}|^{\frac{1}{2}}} \frac{1}{|I_N + \Phi A^{-1}\Phi^T B|^{\frac{1}{2}}} \quad \leftarrow (A^{-1})^T = A^{-1},\ B^T = B$$

$$= \frac{1}{|I_N + \Phi A^{-1}\Phi^T B|^{\frac{1}{2}}} \frac{1}{|B^{-1}|^{\frac{1}{2}}}$$

$$= \frac{1}{|(I_N + \Phi A^{-1}\Phi^T B)B^{-1}|^{\frac{1}{2}}} \quad \leftarrow (C.12)\ |AB|=|A||B|$$

$$= \frac{1}{|B^{-1} + \Phi A^{-1}\Phi^T|^{\frac{1}{2}}}$$

なので

$$p(t|X,\alpha,\beta) = \frac{1}{(2\pi)^{\frac{N}{2}}} \frac{1}{|B^{-1}+\Phi A^{-1}\Phi^T|^{\frac{1}{2}}} \exp\left\{-\frac{1}{2} t^T (B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\right\}$$

$$= \mathcal{N}(t|0, B^{-1}+\Phi A^{-1}\Phi^T), \quad B = \beta I_N,\ A = \mathrm{diag}(\alpha_i)$$

を得る。両辺の対数を取って

$$\ln p(t|X,\alpha,\beta) = -\frac{1}{2}\{N\ln(2\pi) + \ln|B^{-1}+\Phi A^{-1}\Phi^T| + t^T(B^{-1}+\Phi A^{-1}\Phi^T)^{-1} t\} \quad \cdots (7.85)$$

を得る。