13.24

$\mathbf{z}_n' = \begin{pmatrix} \mathbf{z}_n \\ 1 \end{pmatrix}$, $\mathbf{A}' = \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{0} & 1 \end{pmatrix}$, $(\mathbf{\Gamma}^{-1})' = \begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix}$, $\mathbf{C}' = \begin{pmatrix} \mathbf{C} & \mathbf{c} \end{pmatrix}$ とすると

$$(\mathbf{z}_n' - \mathbf{A}'\mathbf{z}_{n-1}')^T (\mathbf{\Gamma}^{-1})' (\mathbf{z}_n' - \mathbf{A}'\mathbf{z}_{n-1}') = \left(\begin{pmatrix} \mathbf{z}_n \\ 1 \end{pmatrix} - \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}_{n-1} \\ 1 \end{pmatrix} \right)^T \begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \left(\begin{pmatrix} \mathbf{z}_n \\ 1 \end{pmatrix} - \begin{pmatrix} \mathbf{A} & \mathbf{a} \\ \mathbf{0} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{z}_{n-1} \\ 1 \end{pmatrix} \right)$$

$$= \begin{pmatrix} \mathbf{z}_n - (\mathbf{A}\mathbf{z}_{n-1} + \mathbf{a}) \\ 0 \end{pmatrix}^T \begin{pmatrix} \mathbf{\Gamma}^{-1} & \mathbf{0} \\ \mathbf{0} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{z}_n - (\mathbf{A}\mathbf{z}_{n-1} + \mathbf{a}) \\ 0 \end{pmatrix} = \{\mathbf{z}_n - (\mathbf{A}\mathbf{z}_n + \mathbf{a})\}^T \mathbf{\Gamma}^{-1} \{\mathbf{z}_n - (\mathbf{A}\mathbf{z}_n + \mathbf{a})\}$$

なので、(13.127)は

$$p(\mathbf{z}_n | \mathbf{z}_{n-1}) = \mathcal{N}(\mathbf{z}_n | \mathbf{A}\mathbf{z}_{n-1} + \mathbf{a}, \mathbf{\Gamma})$$

$$= \frac{1}{(2\pi)^{\frac{D}{2}}} \frac{1}{|\mathbf{\Gamma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\mathbf{z}_n' - \mathbf{A}'\mathbf{z}_{n-1}')^T (\mathbf{\Gamma}^{-1})' (\mathbf{z}_n' - \mathbf{A}'\mathbf{z}_{n-1}') \right\} \cdots (13.127)'$$

↖ 正定値でない共分散行列を持つガウス分布の次元を上げたガウス分布になっている。
こういう分布に対して(2.115)などが成立するかどうか良く分かりません。

となる。また(13.128)は

$$p(\mathbf{x}_n | \mathbf{z}_n) = \mathcal{N}(\mathbf{x}_n | \mathbf{C}\mathbf{z}_n + \mathbf{c}, \mathbf{\Sigma})$$

$$= \frac{1}{(2\pi)^{\frac{M}{2}}} \frac{1}{|\mathbf{\Sigma}|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2} (\mathbf{x}_n - \mathbf{C}'\mathbf{z}_n')^T \mathbf{\Sigma}^{-1} (\mathbf{x}_n - \mathbf{C}'\mathbf{z}_n') \right\} \cdots (13.128)'$$

となる。

(13.127)′, (13.128)′ の平均の部分が (13.75)、(13.76) に似ているので 13.3.1節の (13.89), (13.90), (13.91) がそのまま使える。

らしいが、(13.127)′、(13.128)′ と (13.75)、(13.76) は似て非なるものなので、13.3.1節の式がそのまま成り立つかどうか良く分からない。

なので、(13.127)、(13.128) に対する $\boldsymbol{\mu}_n$, $\mathbf{V}_n$, c_n の更新式を求めて、これが $\mathbf{z}_n'$, $\mathbf{A}'$, $(\mathbf{\Gamma}^{-1})'$, $\mathbf{C}'$ を用いてどう表現されるか確認することにする。

$p(z_n|z_{n-1}) = N(z_n|Az_{n-1}+a, \Gamma)$ ‥ (13.127)

$p(x_n|z_n) = N(x_n|Cz_n+c, \Sigma)$ ‥ (13.128)

(13.85)に代入する

$$c_n N(z_n|\mu_n, V_n) = N(x_n|Cz_n+c, \Sigma)\times\int N(z_n|Az_{n-1}+a, \Gamma)N(z_{n-1}|\mu_{n-1}, V_{n-1})\,dz_{n-1}$$

(2.115)より

$$\int N(z_n|Az_{n-1}+a, \Gamma)N(z_{n-1}|\mu_{n-1}, V_{n-1})\,dz_{n-1} = N(z_n|A\mu_{n-1}+a, \Gamma+AV_{n-1}A^T)$$

よって

$$c_n N(z_n|\mu_n, V_n) = N(x_n|Cz_n+c, \Sigma)N(z_n|A\mu_{n-1}+a, \overbrace{\Gamma+AV_{n-1}A^T}^{P_{n-1}\ (13.88)})$$

$$= N(x_n|Cz_n+c, \Sigma)N(z_n|A\mu_{n-1}+a, P_{n-1})$$

(2.116)より

$$p(z_n|x_n) = N(z_n|Y\{C^T\Sigma^{-1}(x_n-c)+P_{n-1}^{-1}(A\mu_{n-1}+a)\}, Y)$$

$$Y = (P_{n-1}^{-1}+C^T\Sigma^{-1}C)^{-1}$$

(2.115)より

$$p(x_n) = N(x_n|C(A\mu_{n-1}+a)+c, \Sigma+CP_{n-1}C^T)$$

となる。

$c_n N(z_n|\mu_n, V_n) = p(x_n|z_n)p(z_n) = p(z_n|x_n)p(x_n)$ で、

x_nは観測値に固定されているので

$$c_n = N(x_n|C(A\mu_{n-1}+a)+c, \Sigma+CP_{n-1}C^T) \quad \cdots (13.91)'$$

$$V_n = Y = (P_{n-1}^{-1}+C^T\Sigma^{-1}C)^{-1} \quad \leftarrow (C.7)$$

$$= P_{n-1}-P_{n-1}C^T(\Sigma+CP_{n-1}C^T)^{-1}CP_{n-1} = (I-\overset{(13.92)}{K_n}C)P_{n-1} \quad \cdots (13.90)'$$

$$\mu_n = Y\{C^T\Sigma^{-1}(x_n-c)+P_{n-1}^{-1}(A\mu_{n-1}+a)\} = YC^T\Sigma^{-1}(x_n-c)+YP_{n-1}^{-1}(A\mu_{n-1}+a)$$

$$= (P_{n-1}^{-1}+C^T\Sigma^{-1}C)^{-1}C^T\Sigma^{-1}(x_n-c)+(P_{n-1}^{-1}+C^T\Sigma^{-1}C)^{-1}P_{n-1}^{-1}(A\mu_{n-1}+a)$$

$$= \underbrace{P_{n-1}C^T(CP_{n-1}C^T+\Sigma)^{-1}}_{\text{(C.5)より}}(x_n-c)+\underbrace{\{P_{n-1}-P_{n-1}C^T(\Sigma+CP_{n-1}C^T)^{-1}CP_{n-1}\}}_{\text{(C.7)より}}P_{n-1}^{-1}(A\mu_{n-1}+a)$$

$$= \overset{(13.92)}{K_n}(x_n-c)+(I-\overset{(13.92)}{K_n}C)(A\mu_{n-1}+a)$$

$$= (A\mu_{n-1}+a)+K_n\{(x_n-c)-C(A\mu_{n-1}+a)\} \quad \cdots (13.89)'$$

となる。

以前と同様 $x_n'=\begin{pmatrix} x_n\\ 1\end{pmatrix}$, $A'=\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}$, $\Gamma'=\begin{pmatrix}\Gamma & 0\\ 0 & 0\end{pmatrix}$, $C'=(C\ \ c)$ とし

また $\mu_n'=\begin{pmatrix}\mu_n\\ 1\end{pmatrix}$, $V_n'=\begin{pmatrix}V_n & 0\\ 0 & 0\end{pmatrix}$, $P_{n-1}'=\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}$, $K_n'=\begin{pmatrix}K_n\\ 0\end{pmatrix}$ とすると

$$A'V_{n-1}'A'^T+\Gamma'=\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}\begin{pmatrix}V & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}^T+\begin{pmatrix}\Gamma & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}AV_{n-1}A^T+\Gamma & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}$$ なので

$$P_{n-1}'=A'V_{n-1}'A'^T+\Gamma' \quad \cdots (13.88)''$$

である。また

$$P_{n-1}'C'^T(C'P_{n-1}'C'^T+\Sigma)^{-1}=\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}C^T\\ c^T\end{pmatrix}\left((C\ \ c)\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}C^T\\ c^T\end{pmatrix}+\Sigma\right)^{-1}$$

$$=\begin{pmatrix}P_{n-1}C^T\\ 0\end{pmatrix}(CP_{n-1}C^T+\Sigma)^{-1}=\begin{pmatrix}P_{n-1}C^T(CP_{n-1}C^T+\Sigma)^{-1}\\ 0\end{pmatrix}=\begin{pmatrix}K_n\\ 0\end{pmatrix}$$ なので

$$K_n'=P_{n-1}'C'^T(C'P_{n-1}'C'^T+\Sigma)^{-1} \quad \cdots (13.92)''$$

である。また $(13.89)'$を用いて

$$A'\mu_{n-1}'+K_n'(x_n-C'A'\mu_{n-1}')=\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}\begin{pmatrix}\mu_{n-1}\\ 1\end{pmatrix}+\begin{pmatrix}K_n\\ 0\end{pmatrix}\left(x_n-(C\ \ c)\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}\begin{pmatrix}\mu_{n-1}\\ 1\end{pmatrix}\right)$$

$$=\begin{pmatrix}A\mu_{n-1}+a\\ 1\end{pmatrix}+\begin{pmatrix}K_n\\ 0\end{pmatrix}(x_n-C(A\mu_{n-1}+a)-c)$$

$$=\begin{pmatrix}A\mu_{n-1}+a+K_n\{x_n-C(A\mu_{n-1}+a-c)\}\\ 1\end{pmatrix}\overset{(13.89)'}{=}\begin{pmatrix}\mu_n\\ 1\end{pmatrix}$$

なので

$$\mu_n'=A'\mu_{n-1}'+K_n'(x_n-C'A'\mu_{n-1}') \quad \cdots (13.89)''$$

となる。また $(13.90)'$を用いて

$$(I-K_n'C')P_{n-1}'=\left(\begin{pmatrix}I & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}K_n\\ 0\end{pmatrix}(C\ \ c)\right)\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}$$

$$=\left(\begin{pmatrix}I & 0\\ 0 & 1\end{pmatrix}-\begin{pmatrix}K_nC & K_nc\\ 0 & 0\end{pmatrix}\right)\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}=\begin{pmatrix}I-K_nC & -K_nc\\ 0 & 1\end{pmatrix}\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}$$

$$=\begin{pmatrix}(I-K_nC)P_{n-1} & 0\\ 0 & 0\end{pmatrix}\overset{(13.90)'}{=}\begin{pmatrix}V_n & 0\\ 0 & 0\end{pmatrix}$$

なので

$$V_n'=(I-K_n'C')P_{n-1}' \quad \cdots (13.90)''$$

となる。また

$$C'A'\mu_{n-1}'=(C\ \ c)\begin{pmatrix}A & a\\ 0 & 1\end{pmatrix}\begin{pmatrix}\mu_{n-1}\\ 1\end{pmatrix}=(C\ \ c)\begin{pmatrix}A\mu_{n-1}+a\\ 1\end{pmatrix}=C(A\mu_{n-1}+a)+c$$

であり、また

$$\Sigma+C'P_{n-1}'C'^T=\Sigma+(C\ \ c)\begin{pmatrix}P_{n-1} & 0\\ 0 & 0\end{pmatrix}\begin{pmatrix}C^T\\ c^T\end{pmatrix}=\Sigma+(C\ \ c)\begin{pmatrix}P_{n-1}C^T\\ 0\end{pmatrix}=\Sigma+CP_{n-1}C^T$$

である。これより $(13.91)'$は

$$c_n=N(x_n\mid C(A\mu_{n-1}+a)+c,\ \Sigma+CP_{n-1}C^T)$$

$$=N(x_n\mid C'A'\mu_{n-1}',\ \Sigma+C'P_{n-1}'C'^T) \quad \cdots (13.91)''$$

となる。

以上より $(13.89)''$, $(13.90)''$, $(13.91)''$ は、(13.89), (13.90), (13.91) を $\mathbb{x}'_n$, $\mathbb{A}'$, $\mathbb{P}'$, $\mathbb{C}'$, $\mathbb{M}'_n$, $\mathbb{V}'_n$ で置き換えたものになっていることが確認できた。

(13.100), (13.101), (13.103), (13.104), (13.113), (13.114), (13.115), (13.116) の確認も必要 ⇒ やってない

たぶん正定値でないガウス分布の次数を下げたガウス分布について
(2.115), (2.116) の成立だけが怪しい。
他の形式的式変形は成立しそうな気がするので
(2.115), (2.116) だけ確認すればいいのかもしれない ⇒ やってない。