演習 13.29 について、解答があるがもう少し見通しよくできそうなので、やってみる。
まず、(13.99) を記載する。

$$c_{n+1}\hat{\beta}(\mathbf{z}_n) = \int \hat{\beta}(\mathbf{z}_{n+1}) p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}) p(\mathbf{z}_{n+1}|\mathbf{z}_n) d\mathbf{z}_{n+1} \tag{1}$$

本にあるようにこの両辺に $\hat{\alpha}(\mathbf{z}_n)$ をかける。$\hat{\alpha}(\mathbf{z}_n)$ は $\mathbf{z}_{n+1}$ の関数でないので、積分の中の任意の位置に入れても良い。

$$c_{n+1}\hat{\alpha}(\mathbf{z}_n)\hat{\beta}(\mathbf{z}_n) = \int \hat{\beta}(\mathbf{z}_{n+1}) p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}) p(\mathbf{z}_{n+1}|\mathbf{z}_n) \hat{\alpha}(\mathbf{z}_n) d\mathbf{z}_{n+1} \tag{2}$$

(13.56) より、

$$c_n = p(\mathbf{x}_n|\mathbf{x}_1, \cdots, \mathbf{x}_{n-1}) \tag{3}$$

(13.55) より、

$$\hat{\alpha}(\mathbf{z}_n) = p(\mathbf{z}_n|\mathbf{x}_1, \cdots, \mathbf{x}_n) \tag{4}$$

(13.13)(13.64) より

$$\gamma(\mathbf{z}_n) = p(\mathbf{z}_n|\mathbf{x}_1, \cdots, \mathbf{x}_N) = \hat{\alpha}(\mathbf{z}_n)\hat{\beta}(\mathbf{z}_n) \tag{5}$$

(2) の右辺の一部、$p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1})p(\mathbf{z}_{n+1}|\mathbf{z}_n)\hat{\alpha}(\mathbf{z}_n)$ について考える。ここで (3), (4)、マルコフ性も考慮し、ベイズの定理を適用する。

$$\begin{aligned} p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1})p(\mathbf{z}_{n+1}|\mathbf{z}_n)\hat{\alpha}(\mathbf{z}_n) &= p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_{n+1}|\mathbf{z}_n, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_n|\mathbf{x}_1, \cdots, \mathbf{x}_n) \\ &= p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_n, \mathbf{z}_{n+1}|\mathbf{z}_n, \mathbf{x}_1, \cdots, \mathbf{x}_n) \\ &= p(\mathbf{x}_{n+1}|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_{n+1}|\mathbf{x}_1, \cdots, \mathbf{x}_n) \\ &= p(\mathbf{x}_{n+1}, \mathbf{z}_{n+1}|\mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) \\ &= p(\mathbf{z}_{n+1}|\mathbf{x}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{x}_{n+1}|\mathbf{x}_1, \cdots, \mathbf{x}_n) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) \\ &= \hat{\alpha}(\mathbf{z}_{n+1}) c_{n+1} p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) \end{aligned} \tag{6}$$

これを、(2) に代入し、(5) を考慮すると、

$$\begin{aligned} c_{n+1}\gamma(\mathbf{z}_n) &= \int \hat{\beta}(\mathbf{z}_{n+1})\hat{\alpha}(\mathbf{z}_{n+1}) c_{n+1} p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) d\mathbf{z}_{n+1} \\ &= c_{n+1} \int \gamma(\mathbf{z}_{n+1}) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) d\mathbf{z}_{n+1} \end{aligned} \tag{7}$$

となり、

$$\gamma(\mathbf{z}_n) = \int \gamma(\mathbf{z}_{n+1}) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) d\mathbf{z}_{n+1} \tag{8}$$

となる。なお、この式は (5) を考慮すると、マルコフ性もあるので、

$$\begin{aligned} p(\mathbf{z}_n|\mathbf{x}_1, \cdots, \mathbf{x}_N) &= \int p(\mathbf{z}_{n+1}|\mathbf{x}_1, \cdots, \mathbf{x}_N) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_n) d\mathbf{z}_{n+1} \\ &= \int p(\mathbf{z}_{n+1}|\mathbf{x}_1, \cdots, \mathbf{x}_N) p(\mathbf{z}_n|\mathbf{z}_{n+1}, \mathbf{x}_1, \cdots, \mathbf{x}_N) d\mathbf{z}_{n+1} \end{aligned} \tag{9}$$

となっており、周辺化の式になっていて、(2.115) の式が使えることが示唆されている。

さて、(8) などの式に出てきている、$p(\mathbf{z}_n|\mathbf{z}_{n+1},\mathbf{x}_1,\cdots,\mathbf{x}_N)$ について、考える。
(4) と (13.84) を考えると、

$$\hat{\alpha}(\mathbf{z}_n)=p(\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_n)=\mathcal{N}(\mathbf{z}_n|\mu_n,\mathbf{V}_n) \tag{10}$$

(13.75) より、マルコフ性があるので、

$$p(\mathbf{z}_{n+1}|\mathbf{z}_n)=p(\mathbf{z}_{n+1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_n)=\mathcal{N}(\mathbf{z}_{n+1}|\mathbf{A}\mathbf{z}_n,\mathbf{\Gamma}) \tag{11}$$

(ここから先は解答と同様)
(10) が (2.113) に、(11) が (2.114) に対応するとしたら、(2.116) より、

$$\begin{aligned}p(\mathbf{z}_n|\mathbf{z}_{n+1},\mathbf{x}_1,\cdots,\mathbf{x}_N)&=\mathcal{N}(\mathbf{z}_n|({\mathbf{V}_n}^{-1}+\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{A})^{-1}(\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{z}_{n+1}+{\mathbf{V}_n}^{-1}\mu_n),({\mathbf{V}_n}^{-1}+\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{A})^{-1})\\&=\mathcal{N}(\mathbf{z}_n|\mathbf{L}^{-1}(\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{z}_{n+1}+{\mathbf{V}_n}^{-1}\mu_n),\mathbf{L}^{-1})\end{aligned} \tag{12}$$

ここで、(C.7),(13.88),(13.102) も考えると、

$$\begin{aligned}\mathbf{L}^{-1}&=({\mathbf{V}_n}^{-1}+\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{A})^{-1}=\mathbf{V}_n-\mathbf{V}_n\mathbf{A}^T(\mathbf{\Gamma}+\mathbf{A}\mathbf{V}_n\mathbf{A}^T)^{-1}\mathbf{A}\mathbf{V}_n\\&=\mathbf{V}_n-\mathbf{V}_n\mathbf{A}^T{\mathbf{P}_n}^{-1}\mathbf{A}\mathbf{V}_n=\mathbf{V}_n-\mathbf{J}_n\mathbf{A}\mathbf{V}_n=\mathbf{V}_n-\mathbf{J}_n\mathbf{P}_n{\mathbf{J}_n}^T\end{aligned} \tag{13}$$

なお、(13.102) に関して、両辺に右から $\mathbf{P}_n$ をかけて、

$$\mathbf{J}_n\mathbf{P}_n=\mathbf{V}_n\mathbf{A}^T \tag{14}$$

$\mathbf{V}_n,\mathbf{P}_n$ が正規分布の精度行列になるので、（正定値）対称行列になっているため、(14) の両辺の転置を取ると

$$\mathbf{P}_n{\mathbf{J}_n}^T={\mathbf{P}_n}^T{\mathbf{J}_n}^T=(\mathbf{J}_n\mathbf{P}_n)^T=(\mathbf{V}_n\mathbf{A}^T)^T=\mathbf{A}{\mathbf{V}_n}^T=\mathbf{A}\mathbf{V}_n \tag{15}$$

更に、(C.5),(14) を考えると、

$$\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}=({\mathbf{V}_n}^{-1}+\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{A})^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}=\mathbf{V}_n\mathbf{A}^T(\mathbf{\Gamma}+\mathbf{A}\mathbf{V}_n\mathbf{A}^T)^{-1}=\mathbf{V}_n\mathbf{A}^T{\mathbf{P}_n}^{-1}=\mathbf{J}_n\mathbf{P}_n{\mathbf{P}_n}^{-1}=\mathbf{J}_n \tag{16}$$

(9) に対して、上記の式が (2.114) 相当になり、(2.113) に対応するのは、(13.98) より、

$$\gamma(\mathbf{z}_{n+1})=\mathcal{N}(\mathbf{z}_{n+1}|\hat{\mu}_{n+1},\hat{\mathbf{V}}_{n+1}) \tag{17}$$

(9) の右辺は (2.115) を適用して、

$$\begin{aligned}&\int p(\mathbf{z}_{n+1}|\mathbf{x}_1,\cdots,\mathbf{x}_N)p(\mathbf{z}_n|\mathbf{z}_{n+1},\mathbf{x}_1,\cdots,\mathbf{x}_N)d\mathbf{z}_{n+1}\\&=\mathcal{N}(\mathbf{z}_{n+1}|\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mu}_{n+1}+\mathbf{L}^{-1}{\mathbf{V}_n}^{-1}\mu_{n+1},\mathbf{L}^{-1}+\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mathbf{V}}_{n+1}^{-1}(\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1})^T)\end{aligned} \tag{18}$$

また、左辺は (13.98) より、

$$\gamma(\mathbf{z}_n)=\mathcal{N}(\mathbf{z}_n|\hat{\mu}_n,\hat{\mathbf{V}}_n) \tag{19}$$

よって、

$$\hat{\mathbf{V}}_n=\mathbf{L}^{-1}+\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mathbf{V}}_{n+1}^{-1}(\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1})^T \tag{20}$$

$$\hat{\mu}_n=\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mu}_{n+1}+\mathbf{L}^{-1}{\mathbf{V}_n}^{-1}\mu_{n+1} \tag{21}$$

(20) を考える。

$$
\begin{aligned}
\hat{\mathbf{V}}_n &= \mathbf{L}^{-1} + \mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mathbf{V}}_{n+1}^{-1}(\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1})^T \\
&= \mathbf{V}_n - \mathbf{J}_n\mathbf{P}_n\mathbf{J}_n + \mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mathbf{V}}_{n+1}^{-1}(\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1})^T((13)\text{ より}) \\
&= \mathbf{V}_n - \mathbf{J}_n\mathbf{P}_n\mathbf{J}_n + \mathbf{J}_n\hat{\mathbf{V}}_{n+1}^{-1}{\mathbf{J}_n}^T((16)\text{ より}) \\
&= \mathbf{V}_n + \mathbf{J}_n(\hat{\mathbf{V}}_{n+1}^{-1} - \mathbf{P}_n){\mathbf{J}_n}^T
\end{aligned} \tag{22}
$$

となり、(13.101) が成り立つ。

(21) を考える。

$$
\begin{aligned}
\hat{\mu}_n &= \mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\hat{\mu}_{n+1} + \mathbf{L}^{-1}{\mathbf{V}_n}^{-1}\mu_{n+1} \\
&= \mathbf{J}_n\hat{\mu}_{n+1} + \mathbf{L}^{-1}{\mathbf{V}_n}^{-1}\mu_{n+1}((16)\text{ より}) \\
&= \mathbf{J}_n\hat{\mu}_{n+1} + (\mathbf{V}_n - \mathbf{J}_n\mathbf{A}\mathbf{V}_n){\mathbf{V}_n}^{-1}\mu_{n+1}((13)\text{ より}) \\
&= \mathbf{J}_n\hat{\mu}_{n+1} + (I - \mathbf{J}_n\mathbf{A})\mu_{n+1} = \mathbf{J}_n\hat{\mu}_{n+1} + \mu_{n+1} - \mathbf{J}_n\mathbf{A}\mu_{n+1} \\
&= \mu_{n+1} + \mathbf{J}_n(\hat{\mu}_{n+1} - \mathbf{A}\mu_{n+1})
\end{aligned} \tag{23}
$$

よって、(13.100) が成り立つ。