演習 13.31 について、早めにベイズの定理の適用を適用することで簡単にしてみる。

(13.12) も利用し、(13.14) より、

$$\begin{aligned}\xi(\mathbf{z}_{n-1},\mathbf{z}_n) = p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{X}) = p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)\\ = p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)p(\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)\\ = p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)\gamma(\mathbf{z}_n)\end{aligned} \tag{1}$$

$p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)$ は演習 13.29 の (12) にて求めている。ただし、インデックスが少し異なることに注意。

$$\begin{aligned}p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_N) = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{L}^{-1}(\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{z}_n + {\mathbf{V}_{n-1}}^{-1}\mu_{n-1}),\mathbf{L}^{-1})\\ = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{z}_n + \mathbf{L}^{-1}{\mathbf{V}_{n-1}}^{-1}\mu_{n-1},\mathbf{L}^{-1})\end{aligned} \tag{2}$$

この式に演習 13.29 の (16),(13) を適用すると、

$$\begin{aligned}p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_N) = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{L}^{-1}\mathbf{A}^T\mathbf{\Gamma}^{-1}\mathbf{z}_n + \mathbf{L}^{-1}{\mathbf{V}_{n-1}}^{-1}\mu_{n-1},\mathbf{L}^{-1})\\ = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{J}_{-1}n\mathbf{z}_n + (\mathbf{V}_{n-1} - \mathbf{J}_{n-1}\mathbf{A}\mathbf{V}_{n-1}){\mathbf{V}_{n-1}}^{-1}\mu_{n-1},\mathbf{V}_{n-1} - \mathbf{J}_{n-1}\mathbf{A}\mathbf{V}_{n-1})\\ = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{J}_{n-1}\mathbf{z}_n + (I - \mathbf{J}_{n-1}\mathbf{A})\mathbf{V}_{n-1}{\mathbf{V}_{n-1}}^{-1}\mu_{n-1},(I - \mathbf{J}_{n-1}\mathbf{A})\mathbf{V}_{n-1})\\ = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{J}_{n-1}\mathbf{z}_n + (I - \mathbf{J}_{n-1}\mathbf{A})\mu_{n-1},(I - \mathbf{J}_{n-1}\mathbf{A})\mathbf{V}_{n-1})\end{aligned} \tag{3}$$

(2) に (3),(13.98) を代入し、

$$\begin{aligned}\xi(\mathbf{z}_{n-1},\mathbf{z}_n) = p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_n,\cdots,\mathbf{x}_N)\gamma(\mathbf{z}_n)\\ = \mathcal{N}(\mathbf{z}_{n-1}|\mathbf{J}_{n-1}\mathbf{z}_n + (I - \mathbf{J}_{n-1}\mathbf{A})\mu_{n-1},(I - \mathbf{J}_{n-1}\mathbf{A})\mathbf{V}_{n-1})\mathcal{N}(\mathbf{z}_n|\hat{\mu}_n,\hat{\mathbf{V}}_n)\end{aligned} \tag{4}$$

参考までに (13.103)((13.65)) が (2) と等価であることを確認する。(13.56),(13.55) を踏まえ、マルコフ性があることに注意すると、(この操作は演習 13.29 の資料でも行っている。)

$$\begin{aligned}\xi(\mathbf{z}_{n-1},\mathbf{z}_n) = (c_n)^{-1}\hat{\alpha}(\mathbf{z}_{n-1})p(\mathbf{x}_n|\mathbf{z}_n)p(\mathbf{z}_n|\mathbf{z}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = (p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1}))^{-1}p(\mathbf{z}_{n-1}|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = (p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1}))^{-1}p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = (p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1}))^{-1}p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = (p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1}))^{-1}p(\mathbf{x}_n,\mathbf{z}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = (p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1}))^{-1}p(\mathbf{z}_n|\mathbf{x}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{x}_n|\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = p(\mathbf{z}_n|\mathbf{x}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n)\\ = \hat{\alpha}(\mathbf{z}_n)p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_{n-1})\hat{\beta}(\mathbf{z}_n) = p(\mathbf{z}_{n-1}|\mathbf{z}_n,\mathbf{x}_1,\cdots,\mathbf{x}_N)\gamma(\mathbf{z}_n)\end{aligned} \tag{5}$$

よって、等価になる。

(以下、解答と同様)

$\xi(\mathbf{z}_{n-1},\mathbf{z}_n)$ は上巻 P88 にあるように、ガウス分布である、周辺分布と条件付き分布の積になっている。$\mathbf{z} = [\mathbf{z}_n,\mathbf{z}_{n-1}]^T$ を (2.101) に相当する式として、(2.99),(2.100),(2.108),(13.100) を考慮すると、平均は

$$\begin{aligned}\mathbf{E}([\mathbf{z}_n,\mathbf{z}_{n-1}]^T) = [\hat{\mu}_n,\mathbf{J}_{n-1}\hat{\mu}_{n-1} + (I - \mathbf{J}_{n-1}\mathbf{A})\mu_{n-1}]^T\\ = [\hat{\mu}_n,\mu_{n-1} + \mathbf{J}_{n-1}(\hat{\mu}_{n-1} - \mathbf{A}\mu_{n-1}]^T = [\hat{\mu}_n,\hat{\mu}_{n-1}]^T\end{aligned} \tag{6}$$

また、共分散 $cov[\mathbf{z}_{n-1}, \mathbf{z}_n]$ は (1.42) を考えると

$$cov[\mathbf{z}_{n-1}, \mathbf{z}_n] = \mathbb{E}_{\mathbf{z}_{n-1}, \mathbf{z}_n}[(\mathbf{z}_{n-1} - \mathbb{E}[\mathbf{z}_{n-1}])({\mathbf{z}_n}^T - \mathbb{E}[{\mathbf{z}_n}^T])] \tag{7}$$

同様に

$$cov[\mathbf{z}] = \mathbb{E}_{\mathbf{z}_{n-1}, \mathbf{z}_n}[\begin{pmatrix} \mathbf{z}_n - \mathbb{E}[\mathbf{z}_n] \\ \mathbf{z}_{n-1} - \mathbb{E}[\mathbf{z}_{n-1}] \end{pmatrix} \begin{pmatrix} {\mathbf{z}_n}^T - \mathbb{E}[{\mathbf{z}_n}^T] & {\mathbf{z}_{n-1}}^T - \mathbb{E}[{\mathbf{z}_{n-1}}^T] \end{pmatrix} = \begin{pmatrix} cov[\mathbf{z}_n] & cov[\mathbf{z}_n, \mathbf{z}_{n-1}] \\ cov[\mathbf{z}_{n-1}, \mathbf{z}_n] & cov[\mathbf{z}_{n-1}] \end{pmatrix} \tag{8}$$

(2.105) を考慮すると

$$cov[\mathbf{z}_{n-1}, \mathbf{z}_n] = \mathbf{J}_{n-1}\hat{\mathbf{V}}_n \tag{9}$$

になる