(13.89)-(13.91) について、意外とやぐらしかったので、まとめておく。

やり方は異なるが、詳解　確率ロボティクス (上田隆一 著)*[1]の 2 章や付録 B がよくまとめられている。

基本的な方針としては本に記載されているように、(2.115),(2.116),(C.5),(C.7) を使う。求める順番は (13.91),(13.90),(13.89) としたほうが良いと感じる。

(13.85) を改めて見てみると、(13.55) を考慮し、マルコフ過程であることに注意すると、

$$\begin{aligned} c_n p(\mathbf{z}_n|\mathbf{x}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1}) = c_n\hat{\alpha}(\mathbf{z}_n) = p(\mathbf{x}_n|\mathbf{z}_n)\int\hat{\alpha}(\mathbf{z}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1})d\mathbf{z}_{n-1} \\ = p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1})\int p(\mathbf{z}_{n-1}|\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_1,...,\mathbf{x}_{n-1})d\mathbf{z}_{n-1} \\ = p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1})\int p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1})d\mathbf{z}_{n-1} \\ = p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) = p(\mathbf{x}_n,\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) \end{aligned} \tag{1}$$

これを踏まえると、

$$c_n = \frac{p(\mathbf{x}_n,\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1})}{p(\mathbf{z}_n|\mathbf{x}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1})} = p(\mathbf{x}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) \tag{2}$$

になる。また、積分部分は、(1) を参考に (13.86) を考えると、

$$\begin{aligned} p(\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) = \int p(\mathbf{z}_{n-1}|\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_1,...,\mathbf{x}_{n-1})d\mathbf{z}_{n-1} \\ = \int p(\mathbf{z}_{n-1}|\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_{n-1})d\mathbf{z}_{n-1} = \int\hat{\alpha}(\mathbf{z}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1})d\mathbf{z}_{n-1} \end{aligned} \tag{3}$$

まず、(13.87) を考える。(3) にて、すべて、$\mathbf{x}_1,...,\mathbf{x}_{n-1}$ が条件づけされているので、**ひとまず**、ないものとして考える。すると、(13.84) を考えて、$\hat{\alpha}(\mathbf{z}_{n-1}) = p(\mathbf{z}_{n-1}|\mathbf{x}_{n-1},\mathbf{x}_{n-1}) = \mathcal{N}(\mathbf{z}_{n-1}|\mu_{n-1},\mathbf{V}_{n-1})$ が (2.113) 相当になり、(13.75) の $p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_{n-1}) = \mathcal{N}(\mathbf{z}_n|\mathbf{A}\mathbf{z}_{n-1},\mathbf{\Gamma})$ が (2.114) 相当、(2.115) の左辺が $p(\mathbf{z}_n|\mathbf{x}_{n-1})$ 相当と考えられる。(上田先生の本では (2.115) に相当する式が (2.113),(2.114) の式の周辺化として、書かれているため、このあたりのアナロジーが成り立つことがわかりやすい。)

その結果、記号を比較して、積分の部分は

$$\begin{aligned} p(\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) = \int p(\mathbf{z}_{n-1}|\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{x}_1,...,\mathbf{x}_{n-1})d\mathbf{z}_{n-1} \\ = \mathcal{N}(\mathbf{z}_n|\mathbf{A}\mu_{n-1},\mathbf{A}\mathbf{V}_{n-1}\mathbf{A}^T+\mathbf{\Gamma}) = \mathcal{N}(\mathbf{z}_n|\mathbf{A}\mu_{n-1},{\mathbf{P}_{n-1}}^{-1}) \end{aligned} \tag{4}$$

となり、(13.87) が求まる。($\mathbf{P}_{n-1}$ は (13.88) による。)

さて、(13.91) は

$$c_n = p(\mathbf{x}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1}) = \int p(\mathbf{z}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1})p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1})d\mathbf{z}_n \tag{5}$$

そのため、今度は (4) が (2.113) 相当になり、(13.76) より $p(\mathbf{x}_n|\mathbf{z}_n,\mathbf{x}_1,...,\mathbf{x}_{n-1}) = p(\mathbf{x}_n|\mathbf{z}_n) = \mathcal{N}(\mathbf{x}_n|\mathbf{C}\mathbf{z}_n,\mathbf{\Sigma})$ が、(2.114) 相当になり、$c_n = p(\mathbf{x}_n|\mathbf{x}_1,...,\mathbf{x}_{n-1})$ が (2.115) の左辺となる。記号を比較して、

$$c_n = \mathcal{N}(\mathbf{x}_n|\mathbf{C}\mathbf{A}\mu_{n-1},\mathbf{C}\mathbf{P}_{n-1}\mathbf{C}^T+\mathbf{\Sigma}) \tag{6}$$

となり、(13.91) が成立している。

*[1] この本などによって、2020 年度日本機械学会教育賞を受賞されている。

(1) より、$\hat{\alpha}(\mathbf{z}_n) = p(\mathbf{z}_n|\mathbf{x}_n, \mathbf{x}_1, ..., \mathbf{x}_{n-1}) = \mathcal{N}(\mathbf{z}_n|\mu_n, \mathbf{V}_n)$ は、$c_n = p(\mathbf{x}_n|\mathbf{x}_1, ..., \mathbf{x}_{n-1})$ と同じ条件で、(2.116) の記号と比較すると、、

$$\hat{\alpha}(\mathbf{z}_n) = \mathcal{N}(\mathbf{z}_n|\mu_n, \mathbf{V}_n) = \mathcal{N}(\mathbf{z}_n|(\mathbf{P}_{n-1}+\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}(\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{x}_n+\mathbf{P}_{n-1}{}^{-1}\mathbf{A}\mu_{n-1}), (\mathbf{P}_{n-1}{}^{-1}+\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}) \tag{7}$$

係数を比較することで、

$$\mathbf{V}_n = (\mathbf{P}_{n-1}{}^{-1} + \mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1} \tag{8}$$

右辺に (C.7) の左辺を適用し、(13.92) を考えると、

$$\begin{aligned}\mathbf{V}_n = \mathbf{P}_{n-1} - \mathbf{P}_{n-1}\mathbf{C}^T(\mathbf{\Sigma} + \mathbf{C}\mathbf{P}_{n-1}\mathbf{C}^T)^{-1}\mathbf{C}\mathbf{P}_{n-1} &= \mathbf{I}\mathbf{P}_{n-1} - (\mathbf{P}_{n-1}\mathbf{C}^T(\mathbf{\Sigma} + \mathbf{C}\mathbf{P}_{n-1}\mathbf{C}^T)^{-1})\mathbf{C}\mathbf{P}_{n-1} \\ &= \mathbf{I}\mathbf{P}_{n-1} - \mathbf{K}_n\mathbf{C}\mathbf{P}_{n-1} = (\mathbf{I} - \mathbf{K}_n\mathbf{C})\mathbf{P}_{n-1}\end{aligned} \tag{9}$$

となり、(13.90) が求まる。
同様に

$$\begin{aligned}\mu_n &= (\mathbf{P}_{n-1}{}^{-1} + \mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}(\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{x}_n + \mathbf{P}_{n-1}{}^{-1}\mathbf{A}\mu_{n-1}) \\ &= (\mathbf{P}_{n-1}{}^{-1} + \mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{x}_n + \mathbf{V}_n\mathbf{P}_{n-1}{}^{-1}\mathbf{A}\mu_{n-1} \\ &= (\mathbf{P}_{n-1}{}^{-1} + \mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{x}_n + (\mathbf{I} - \mathbf{K}_n\mathbf{C})\mathbf{P}_{n-1}\mathbf{P}_{n-1}{}^{-1}\mathbf{A}\mu_{n-1} \\ &= (\mathbf{P}_{n-1}{}^{-1} + \mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{C})^{-1}\mathbf{C}^T\mathbf{\Sigma}^{-1}\mathbf{x}_n + (\mathbf{I} - \mathbf{K}_n\mathbf{C})\mathbf{A}\mu_{n-1}\end{aligned} \tag{10}$$

これに (C.5) の左辺を適用して、

$$\begin{aligned}\mu_n &= \mathbf{P}_{n-1}\mathbf{C}^T(\mathbf{C}\mathbf{P}_{n-1}\mathbf{C}^T + \mathbf{\Sigma})^{-1}\mathbf{x}_n + (\mathbf{I} - \mathbf{K}_n\mathbf{C})\mathbf{A}\mu_{n-1} \\ &= \mathbf{K}_n\mathbf{x}_n + \mathbf{A}\mu_{n-1} - \mathbf{K}_n\mathbf{C}\mathbf{A}\mu_{n-1} = \mathbf{A}\mu_{n-1} + \mathbf{K}_n(\mathbf{x}_n - \mathbf{C}\mathbf{A}\mu_{n-1})\end{aligned} \tag{11}$$

(13.89) が成立することが確認できる。