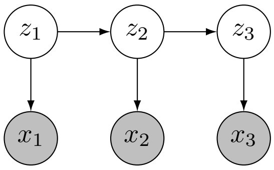
Answers to PMR tutorial sheet 7

November 3, 2011

1(i).



$$P(\mathbf{X}) = \sum_{\mathbf{z}_n \in \{1,2,3\}} \alpha(\mathbf{z}_n)\beta(\mathbf{z}_n) = \sum_{\mathbf{z}_3 \in \{1,2,3\}} \alpha(\mathbf{z}_3)\beta(\mathbf{z}_3) \qquad \text{(choose } n\text{=3)}$$

We have that $\beta(\mathbf{z}_3) = 1$ for all values of $\mathbf{z}_3$; we just need to calculate $\alpha(\mathbf{z}_3)$ using the recursion formula.

$$\begin{aligned}
\alpha(\mathbf{z}_1) &= p(\mathbf{x}_1, \mathbf{z}_1) = p(\mathbf{x}_1|\mathbf{z}_1)p(\mathbf{z}_1) \\
&= \begin{pmatrix} 0.7 \times 0.9 \\ 0.4 \times 0.1 \\ 0.8 \times 0.0 \end{pmatrix} = \begin{pmatrix} 0.63 \\ 0.04 \\ 0.0 \end{pmatrix} \\
\alpha(\mathbf{z}_2) &= \sum_{\mathbf{z}_1 \in \{1,2,3\}} \alpha(\mathbf{z}_1) a_{\mathbf{z}_1\mathbf{z}_2} p(\mathbf{x}_2|\mathbf{z}_2) \\
&= \begin{pmatrix} (0.63 \times 0.5 + 0.04 \times 0.0 + 0.0 \times 0.0) \times 0.3 \\ (0.63 \times 0.3 + 0.04 \times 0.6 + 0.0 \times 0.0) \times 0.6 \\ (0.63 \times 0.2 + 0.04 \times 0.4 + 0.0 \times 1.0) \times 0.2 \end{pmatrix} = \begin{pmatrix} 0.0945 \\ 0.1278 \\ 0.0284 \end{pmatrix} \\
\alpha(\mathbf{z}_3) &= \sum_{\mathbf{z}_2 \in \{1,2,3\}} \alpha(\mathbf{z}_2) a_{\mathbf{z}_2\mathbf{z}_3} p(\mathbf{x}_3|\mathbf{z}_3) \\
&= \begin{pmatrix} (0.0945 \times 0.5 + 0.1278 \times 0.0 + 0.0284 \times 0.0) \times 0.3 \\ (0.0945 \times 0.3 + 0.1278 \times 0.6 + 0.0284 \times 0.0) \times 0.6 \\ (0.0945 \times 0.2 + 0.1278 \times 0.4 + 0.0284 \times 1.0) \times 0.2 \end{pmatrix} = \begin{pmatrix} 0.014175 \\ 0.063018 \\ 0.019684 \end{pmatrix}
\end{aligned}$$

$$\therefore P(\mathbf{X}) = 0.014175 + 0.063018 + 0.019684 = 0.096877.$$

1(ii). $P(\mathbf{z}_2|\mathbf{X}) = \frac{\alpha(\mathbf{z}_2)\beta(\mathbf{z}_2)}{P(\mathbf{X})}$. Hence, we require $\beta(\mathbf{z}_2)$. Using the recursion formula, we have

$$\begin{aligned}
\beta(\mathbf{z}_3) &= 1 \qquad \forall \mathbf{z}_3 && \text{(by definition)} \\
\beta(\mathbf{z}_2) &= \sum_{\mathbf{z_3} \in \{1,2,3\}} \beta(\mathbf{z}_3) a_{\mathbf{z_2}\mathbf{z_3}} p(\mathbf{x}_3|\mathbf{z}_3) = \sum_{\mathbf{z_3} \in \{1,2,3\}} a_{\mathbf{z_2}\mathbf{z_3}} p(\mathbf{x}_3|\mathbf{z}_3) && (\beta(\mathbf{z}_3) = 1 \text{ constant}) \\
&= \begin{pmatrix} 0.5 \times 0.3 + 0.3 \times 0.6 + 0.2 \times 0.2 \\ 0.0 \times 0.3 + 0.6 \times 0.6 + 0.4 \times 0.2 \\ 0.0 \times 0.3 + 0.0 \times 0.6 + 1.0 \times 0.2 \end{pmatrix} = \begin{pmatrix} 0.37 \\ 0.44 \\ 0.20 \end{pmatrix} \\
\therefore \gamma(\mathbf{z_2}) &= P(\mathbf{z}_2|\mathbf{X}) \\
&= \frac{1}{0.096877} \times \begin{pmatrix} 0.0945 \times 0.37 \\ 0.1278 \times 0.44 \\ 0.0284 \times 0.20 \end{pmatrix} = \begin{pmatrix} 0.361 \\ 0.580 \\ 0.059 \end{pmatrix}
\end{aligned}$$

Note: $\beta(\mathbf{z}_2) = P(\mathbf{x}_3|\mathbf{z}_2)$, and hence does not sum to 1.

Note: In rectrospect, one can compute $P(\mathbf{X})$ with $\sum_{\mathbf{z}_2 \in \{1,2,3\}} \alpha(\mathbf{z}_2)\beta(\mathbf{z}_2)$, i.e. chosing n=2, so that there is actually no need to compute $\alpha(\mathbf{z_3})$ in (i).

Note: Getting the correct numbers depends on interpreting A correctly. One way is to remember that $a_{\mathbf{z}_n \mathbf{z}_{n+1}} = P(\mathbf{z}_{n+1}|\mathbf{z}_n)$, i.e. transition probability, so that we need $\sum_{\mathbf{z}_{n+1}} a_{\mathbf{z}_n \mathbf{z}_{n+1}} = 1$.

2. We have $\beta_0 = 1$, $\beta_1 = 0.75$, $\beta_2 = -0.2$. Hence

$$\begin{aligned}
\gamma_0 &= (1 + \beta_1^2 + \beta_2^2) = 1 + (0.75)^2 + (-0.2)^2 = 1.6025, \\
\gamma_1 &= \beta_1\beta_0 + \beta_2\beta_1 = 0.75 + 0.75 \times (-0.2) = 0.6, \\
\gamma_2 &= \beta_2\beta_0 = -0.2
\end{aligned}$$

and $\gamma_k = 0$ for $k \geq 3$. This is the "cut off" in the autocorrelation function for a MA process. Note also that $\gamma_{-k} = \gamma_k$.

3. The model is defined by the following distributions

$$\begin{aligned}
\mathbf{z}_1 &\sim N(\boldsymbol{\mu}_0, P_0) && \text{(Initialisation; Not } V_0 \text{ as in Bishop 13.77)} \\
\mathbf{z}_{n+1}|\mathbf{z}_n &\sim N(A\mathbf{z}_n, \Gamma) && \text{(Dynamic model)} \\
\mathbf{x}_n|\mathbf{z}_n &\sim N(C\mathbf{z}_n, \Sigma) && \text{(Observation model)}
\end{aligned}$$

giving the following update equations

$$\begin{aligned}
P_{n-1} &= AV_{n-1}A^{\mathrm{T}} + \Gamma && \text{(time update of state covariance)} \\
\boldsymbol{\mu}_n &= A\boldsymbol{\mu}_{n-1} + K_n(\mathbf{x}_n - CA\boldsymbol{\mu}_{n-1}) && \text{(measurement update of state mean)} \\
V_n &= (I - K_nC)P_{n-1} && \text{(measurement update of state covariance)}
\end{aligned}$$

where $K_n = P_{n-1}C^{\mathrm{T}}(CP_{n-1}C^{\mathrm{T}} + \Sigma)^{-1}$.

Our system has

$$\boldsymbol{\mu}_0 = 0 \qquad P_0 = \sigma_0^2 \qquad A = 1 \qquad \Gamma = 1 \qquad C = 1 \qquad \Sigma = 1$$

for $n = 1$,

$$\begin{aligned}
K_1 &= \frac{\sigma_0^2}{\sigma_0^2 + 1} \\
\mu_1 &= \frac{\sigma_0^2}{\sigma_0^2 + 1} x_1 && \lim_{\sigma_0^2 \to \infty} \mu_1 = x_1 \\
V_1 &= \left(1 - \frac{\sigma_0^2}{\sigma_0^2 + 1}\right)\sigma_0^2 = \frac{\sigma_0^2}{\sigma_0^2 + 1} && \lim_{\sigma_0^2 \to \infty} V_1 = 1
\end{aligned}$$

for $n = 2$,

$$\begin{aligned}
P_1 &= \frac{\sigma_0^2}{\sigma_0^2 + 1} + 1 = \frac{2\sigma_0^2 + 1}{\sigma_0^2 + 1} && \lim_{\sigma_0^2 \to \infty} P_1 = 2 \\
K_2 &= \frac{2\sigma_0^2 + 1}{\sigma_0^2 + 1}\left(\frac{2\sigma_0^2 + 1}{\sigma_0^2 + 1} + 1\right)^{-1} = \frac{2\sigma_0^2 + 1}{3\sigma_0^2 + 2} && \lim_{\sigma_0^2 \to \infty} K_2 = \frac{2}{3} \\
\mu_2 &= \mu_1 + K_2(x_2 - \mu_1) = (1 - K_2)\mu_1 + K_2x_2 \\
&= \frac{\sigma_0^2}{3\sigma_0^2 + 2}x_1 + \frac{2\sigma_0^2 + 1}{3\sigma_0^2 + 2}x_2 && \lim_{\sigma_0^2 \to \infty} \mu_2 = \frac{1}{3}x_1 + \frac{2}{3}x_2 \\
V_2 &= \left(1 - \frac{2\sigma_0^2 + 1}{3\sigma_0^2 + 2}\right)\frac{2\sigma_0^2 + 1}{\sigma_0^2 + 1} = \frac{2\sigma_0^2 + 1}{3\sigma_0^2 + 2} && \lim_{\sigma_0^2 \to \infty} V_2 = \frac{2}{3}
\end{aligned}$$

for $n = 3$,

$$P_2 = \frac{2\sigma_0^2 + 1}{3\sigma_0^2 + 2} + 1 = \frac{5\sigma_0^2 + 3}{3\sigma_0^2 + 2} \qquad \lim_{\sigma_0^2 \to \infty} P_2 = \frac{5}{3}$$

$$K_3 = \frac{5\sigma_0^2 + 3}{3\sigma_0^2 + 2} \left(\frac{5\sigma_0^2 + 3}{3\sigma_0^2 + 2} + 1 \right)^{-1} = \frac{5\sigma_0^2 + 3}{8\sigma_0^2 + 5} \qquad \lim_{\sigma_0^2 \to \infty} K_3 = \frac{5}{8}$$

$$\mu_3 = \mu_2 + K_3(x_3 - \mu_2) = (1 - K_3)\mu_2 + K_3 x_3$$

$$= \frac{(\sigma_0^2 x_1 + (2\sigma_0^2 + 1)x_2 + (5\sigma_0^2 + 3)x_3}{8\sigma_0^2 + 5} \qquad \lim_{\sigma_0^2 \to \infty} \mu_3 = \frac{x_1 + 2x_2 + 5x_3}{8}$$

$$V_3 = \left(1 - \frac{5\sigma_0^2 + 3}{8\sigma_0^2 + 5} \right) \frac{5\sigma_0^2 + 3}{3\sigma_0^2 + 2} = \frac{5\sigma_0^2 + 3}{8\sigma_0^2 + 5} \qquad \lim_{\sigma_0^2 \to \infty} V_3 = \frac{5}{8}$$

Note: Present is weighted more than the past
Note: Can work in the limits directly instead of using σ_0^2 all the way.