Chapter 17

stevenjin8

June 12, 2021

Comments

I found this chapter to be quite confusing, mostly due to the inconsistent notation. In section 17.4.2, the author introduces $\psi_t(j) = p(\mathbf{x}_t|z_t = j)$. However, from section 17.4.2 onwards, he uses ϕ instead of ψ. I also found it confusing that in section 17.5, the author switched $\mathbf{\Psi}$ for $\mathbf{A}$.

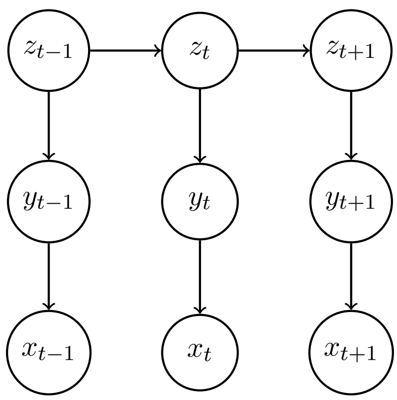
Another point of confusion was section 17.4.3.2. I spent a lot of time gazing at 17.62-17.63 and came to the conclusion that equation 17.63 is wrong (despite 17.66 being correct). I believe that the correct derivation should be

$$\begin{aligned}
\xi_{t-1,t}(i,j) &\triangleq p(z_{t-1} = i, z_t = j|\mathbf{x}_{1:T}) \\
&\propto p(\mathbf{x}_{t:T}, z_{t-1} = i, z_{t-1} = i|\mathbf{x}_{1:t-1}) \\
&= p(\mathbf{x}_t, \mathbf{x}_{t+1:T}|z_{t-1} = i, z_t = j)p(z_{t-1} = i, z_t = j|\mathbf{x}_{1:t-1}) \\
&= p(\mathbf{x}_t|z_t = j)p(\mathbf{x}_{t+1:T}|z_t = j)p(z_t = j|z_{t-1} = i)p(z_{t-1} = i|\mathbf{x}_{1:t-1}) \\
&= \phi_t(j)\beta_t(j)\psi(i,j)\alpha_{t-1}(i).
\end{aligned}$$

Exercises

Exercise 3

a.



b.

$$
\begin{aligned}
Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}}) &= \mathbb{E}\left[\log p(\mathcal{D}|\boldsymbol{\theta}) \middle| \mathcal{D}, \boldsymbol{\theta}^{\text{old}}\right] \\
&= \sum_{i=1}^{N} \sum_{t=1}^{N_i} \mathbb{E}[\log p(\mathbf{x}_{t,i}|y_i) + \log p(y_{t,i}|z_{t,i}) + \log p(z_{t,i}|z_{t,i-1})] \\
&= \sum_{i=1}^{N} \sum_{t=1}^{N_i} \mathbb{E}[\log \mathcal{N}(\mathbf{x}_{t,i}|\boldsymbol{\mu}_{y_i,z_i}, \boldsymbol{\Sigma}_{y_i,z_i}) + \log w_{y_i,z_i} + \log \Phi_{z_{t,i-1},z_{t,i}}].
\end{aligned}
$$

c.

$$
\begin{aligned}
\boldsymbol{\mu}_{j,k} &= \frac{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} \mathbf{x}_{t,i} p(y_{t,i} = j, z_{t,i} = j|\mathbf{x}_i)}{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} p(y_{t,i} = j, z_{t,i} = k|\mathbf{x}_i)} \\
\boldsymbol{\Sigma}_{j,k} &= \frac{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} \mathbf{x}_{t,i}\mathbf{x}_{t,i}^T p(y_{t,i} = j, z_{t,i} = j|\mathbf{x}_i)}{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} p(y_{t,i} = j, z_{t,i} = k|\mathbf{x}_i)} - \boldsymbol{\mu}_{j,k}\boldsymbol{\mu}_{j,k}^T \\
w_{jk} &= \frac{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} p(y_{t,i} = j, z_{t,i} = j|\mathbf{x}_i)}{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} p(z_{t,i} = k|\mathbf{x}_i)} \\
\phi_{j,k} &= \frac{\sum\limits_{i=1}^{N} \sum\limits_{t=2}^{N_i} p(z_{t-1,i} = k, z_{t,i} = j|\mathbf{x}_i)}{\sum\limits_{i=1}^{N} \sum\limits_{t=1}^{N_i} p(z_{t-1,i} = j|\mathbf{x}_i)} \\
\pi_j &\propto \sum_{i=1}^{N} p(z_{1,i} = j|\mathbf{x}_i),
\end{aligned}
$$

where the probabilities are parameterized with $\boldsymbol{\theta}^{\text{old}}$. We can find $p(z_t|\mathbf{x})$ by using results from section 17.4 and equation 17.25. Also, we can find $p(y_t|z_t, \mathbf{x})$ using Bayes' rule.

Exercise 4

The idea is the same as the previous exercise, except that

$$\boldsymbol{\mu}_m = \frac{\sum_{i=1}^{N}\sum_{t=1}^{N_i}\sum_{k=1}^{K} \mathbf{x}_{t,i} p(y_{t,i}=m, z_{t,i}=k|\mathbf{x}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{N_i}\sum_{k=1}^{K} p(y_{t,i}=m, z_{t,i}=k|\mathbf{x}_i)}$$

$$\boldsymbol{\Sigma}_m = \frac{\sum_{i=1}^{N}\sum_{t=1}^{N_i} \mathbf{x}_{t,i}\mathbf{x}_{t,i}^T p(y_{t,i}=m, z_{t,i}=j|\mathbf{x}_i)}{\sum_{i=1}^{N}\sum_{t=1}^{N_i}\sum_{m=1}^{L} p(y_{t,i}=m, z_{t,i}=k|\mathbf{x}_i)} - \boldsymbol{\mu}_m\boldsymbol{\mu}_m^T.$$