

(13.17) に関して、本のとおりを考える。

求めたい式は (13.12) の

$$Q(\theta, \theta^{old}) = \sum_{\mathbf{Z}} p(\mathbf{Z}|\mathbf{X}, \theta^{old}) \ln p(\mathbf{X}, \mathbf{Z}|\theta) = \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{z}_1, \dots, \mathbf{z}_N|\mathbf{X}, \theta^{old}) \ln p(\mathbf{X}, \mathbf{z}_1, \dots, \mathbf{z}_N|\theta) \quad (1)$$

本にあるように (13.10) と (13.7), (13.8) より、

$$p(\mathbf{X}, \mathbf{Z}|\theta) = p(\mathbf{z}_1|\pi) \left[\prod_{n=2}^N p(\mathbf{z}_n|\mathbf{z}_{n-1}, A) \right] \prod_{m=1}^N p(\mathbf{x}_m|\mathbf{z}_m, \phi) = \prod_{k=1}^K \pi_k^{z_{1k}} \left[\prod_{n=2}^N \prod_{j=1}^K \prod_{k=1}^K A_{jk}^{z_{n-1} z_n} \right] \prod_{m=1}^N \prod_{k=1}^K p(\mathbf{x}_m|\mathbf{z}_m, \phi_k)^{z_{mk}} \quad (2)$$

なお、ここで、(13.8) と同様に、

$$p(\mathbf{x}_m|\mathbf{z}_m, \phi) = \prod_{k=1}^K p(\mathbf{x}_m|\mathbf{z}_m, \phi_k)^{z_{mk}} \quad (3)$$

であることに注意する。

(1) を検討する。

$$Q(\theta, \theta^{old}) = \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{Z}|\mathbf{X}, \theta^{old}) \left(\sum_{k=1}^K z_{1k} \ln \pi_k + \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K z_{n-1} z_n \ln A_{jk} + \sum_{n=1}^N \sum_{k=1}^K z_{nk} \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k) \right) \quad (4)$$

カッコの 1 項目を考える。 $\mathbf{z}_2, \dots, \mathbf{z}_2$ について、周辺化する。 z_{1k} は \mathbf{z}_1 に依存するので、 \mathbf{z}_1 は周辺化できない。

$$\begin{aligned} \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{Z}|\mathbf{X}, \theta^{old}) \sum_{k=1}^K z_{1k} \ln \pi_k &= \sum_{k=1}^K \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_N|\mathbf{X}, \theta^{old}) z_{1k} \ln \pi_k \\ &= \sum_{k=1}^K \sum_{\mathbf{z}_1} p(\mathbf{z}_1|\mathbf{X}, \theta^{old}) z_{1k} \ln \pi_k = \sum_{k=1}^K \left(\sum_{\mathbf{z}_1} \gamma(\mathbf{z}_1) z_{1k} \right) \ln \pi_k = \sum_{k=1}^K \gamma(z_{1k}) \ln \pi_k \end{aligned} \quad (5)$$

ここでは、(13.13), (13.15) も利用している。

同様に 2 項目を考えると、 n に依存して、 $\mathbf{z}_{n-1}, \mathbf{z}_n$ が周辺化できない。

$$\begin{aligned} \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{Z}|\mathbf{X}, \theta^{old}) \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K z_{n-1, j} z_{n, k} \ln A_{jk} &= \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{z}_1, \dots, \mathbf{z}_N|\mathbf{X}, \theta^{old}) z_{n-1} z_n \ln A_{jk} \\ &= \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \sum_{\mathbf{z}_{n-1}, \mathbf{z}_n} p(\mathbf{z}_{n-1}, \mathbf{z}_n|\mathbf{X}, \theta^{old}) z_{n-1} z_n \ln A_{jk} = \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \left(\sum_{\mathbf{z}_{n-1}, \mathbf{z}_n} \xi(\mathbf{z}_{n-1}, \mathbf{z}_n) z_{n-1} z_n \right) \ln A_{jk} \\ &= \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \xi(\mathbf{z}_{n-1, j}, \mathbf{z}_{n, k}) \ln A_{jk} \end{aligned} \quad (6)$$

3 項目は

$$\begin{aligned}
\sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{Z}|\mathbf{X}, \theta^{old}) \sum_{n=1}^N \sum_{k=1}^K z_{nk} \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k) &= \sum_{n=1}^N \sum_{k=1}^K \sum_{\mathbf{z}_1, \dots, \mathbf{z}_N} p(\mathbf{z}_1, \dots, \mathbf{z}_N|\mathbf{X}, \theta^{old}) z_{nk} \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k) \\
&= \sum_{n=1}^N \sum_{k=1}^K \sum_{\mathbf{z}_n} p(\mathbf{z}_n|\mathbf{X}, \theta^{old}) z_{nk} \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k) = \sum_{n=1}^N \sum_{k=1}^K \left(\sum_{\mathbf{z}_n} \gamma(\mathbf{z}_n) z_{nk} \right) \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k) \\
&= \sum_{n=1}^N \sum_{k=1}^K \gamma(\mathbf{z}_{nk}) \ln p(\mathbf{x}_n|\mathbf{z}_n, \phi_k)
\end{aligned} \tag{7}$$

上記より、

$$Q(\theta, \theta^{old}) = \sum_{k=1}^K \gamma(z_{1k}) \ln \pi_k + \sum_{n=2}^N \sum_{j=1}^K \sum_{k=1}^K \xi(z_{n-1,j}, z_{n,k}) \ln A_{jk} + \sum_{n=1}^N \sum_{k=1}^K \gamma(z_{nk}) \ln p(x_n|\phi_k) \tag{8}$$

となる。