(13.17) に関して、本のとおりに考える。

求めたい式は (13.12) の

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

本にあるように (13.10) と (13.7),(13.8) より、

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

なお、ここで、(13.8) と同様に、

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

であることに注意する。

さて、$p(\mathbf{Z}|\mathbf{X},\theta^{old})$ を考える。マルコフ性から、

$$p(\mathbf{Z}|\mathbf{X},\theta^{old})=p(\mathbf{z}_1|\mathbf{X},\theta^{old})\prod_{n=2}^{N}p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{X},\theta^{old})=p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})\prod_{n=m+1}^{N}p(\mathbf{z}_n|\mathbf{z}_{n-1},\mathbf{X},\theta^{old}) \tag{4}$$

また、これは、

$$p(\mathbf{Z}|\mathbf{X},\theta^{old})=p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1},...,\mathbf{z}_N|\mathbf{z}_m,\mathbf{X},\theta^{old}) \tag{5}$$

$$p(\mathbf{Z}|\mathbf{X},\theta^{old})=p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1}|\mathbf{z}_m,\mathbf{X},\theta^{old})p(\mathbf{z}_{m+2},...,\mathbf{z}_N|\mathbf{z}_{m+1},\mathbf{X},\theta^{old}) \tag{6}$$

ここで、

$$p(\mathbf{Z}|\mathbf{X},\theta^{old})=p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1}|\mathbf{z}_m,\mathbf{X},\theta^{old})p(\mathbf{z}_{m+2},...,\mathbf{z}_N|\mathbf{z}_{m+1},\mathbf{X},\theta^{old}) \tag{7}$$

このとき、周辺化を検討すると、

$$\sum_{\mathbf{z}_1,...,\mathbf{z}_{m-1}}p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})=p(\mathbf{z}_m|\mathbf{X},\theta^{old})=\gamma(\mathbf{z}_m) \tag{8}$$

$$\sum_{\mathbf{z}_{m+1},...,\mathbf{z}_N}p(\mathbf{z}_{m+1},...,\mathbf{z}_N|\mathbf{z}_{m+1},\mathbf{X},\theta)=1 \tag{9}$$

$$\sum_{\mathbf{z}_1,...,\mathbf{z}_{m-1}}p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1}|\mathbf{z}_m,\mathbf{X},\theta^{old})=p(\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1}|\mathbf{z}_m,\mathbf{X},\theta^{old})=p(\mathbf{z}_{m+1},\mathbf{z}_m|\mathbf{X},\theta^{old}) \tag{10}$$

この結果、$\sum_{\mathbf{z}_1,...,\mathbf{z}_{m-1}}p(\mathbf{z}_1,...,\mathbf{z}_m|\mathbf{X},\theta^{old})p(\mathbf{z}_{m+1}|\mathbf{z}_m,\mathbf{X},\theta^{old})=\xi(\mathbf{z}_{m+1},\mathbf{z}_m)$ となっている。

(9) については m+2 から始めても同様の結果になることに注意する。

いよいよ、(1) を検討する。

$$Q(\theta,\theta^{old})=\sum_{\mathbf{z}_1,...,\mathbf{z}_N}p(\mathbf{Z}|\mathbf{X},\theta^{old})(\sum_{k=1}^{K}z_{1k}\,ln\,\pi_k+\sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}z_{n-1}z_n\,ln\,A_{jk}+\sum_{n=1}^{N}\sum_{k=1}^{K}z_{nk}\,ln\,p(\mathbf{x}_n|\mathbf{z}_n,\phi_k)) \tag{11}$$

カッコの 1 項目を考える。

$$\begin{aligned}\sum_{\mathbf{z}_1,...,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old})\sum_{k=1}^{K} z_{1k}\, ln\, \pi_k &= \sum_{k=1}^{K}\sum_{\mathbf{z}_1,...,\mathbf{z}_N} p(\mathbf{z}_1|\mathbf{X},\theta^{old})p(\mathbf{z}_2,...,\mathbf{z}_N|\mathbf{z}_1,\mathbf{X},\theta^{old})z_{1k}\, ln\, \pi_k \\ &= \sum_{k=1}^{K}\sum_{\mathbf{z}_1}\sum_{\mathbf{z}_2,...,\mathbf{z}_N} p(\mathbf{z}_1|\mathbf{X},\theta^{old})p(\mathbf{z}_2,...,\mathbf{z}_N|\mathbf{z}_1,\mathbf{X},\theta^{old})z_{1k}\, ln\, \pi_k \\ &= \sum_{k=1}^{K}\sum_{\mathbf{z}_1}(p(\mathbf{z}_1|\mathbf{X},\theta^{old})\sum_{\mathbf{z}_2,...,\mathbf{z}_N} p(\mathbf{z}_2,...,\mathbf{z}_N|\mathbf{z}_1,\mathbf{X},\theta^{old}))z_{1k}\, ln\, \pi_k \\ &= \sum_{k=1}^{K}\sum_{\mathbf{z}_1} p(\mathbf{z}_1|\mathbf{X},\theta^{old})z_{1k}\, ln\, \pi_k = \sum_{k=1}^{K}(\sum_{\mathbf{z}_1}\gamma(\mathbf{z}_1)z_{1k})\, ln\, \pi_k = \sum_{k=1}^{K}\gamma(z_{1k})\, ln\, \pi_k\end{aligned} \tag{12}$$

同様に 2 項目を考えると、

$$\begin{aligned}\sum_{\mathbf{z}_1,...,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old})\sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K} z_{n-1}z_n\, ln\, A_{jk} &= \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{\mathbf{z}_1,...,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old})z_{n-1}z_n\, ln\, A_{jk} \\ &= \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{\mathbf{z}_1,...,\mathbf{z}_{n-2}}\sum_{\mathbf{z}_{n-1},\mathbf{z}_n}\sum_{\mathbf{z}_{n+1},...,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old})z_{n-1}z_n\, ln\, A_{jk} \\ = \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{\mathbf{z}_{n-1},\mathbf{z}_n}\sum_{\mathbf{z}_1,...,\mathbf{z}_{n-2}}\sum_{\mathbf{z}_{n+1},...,\mathbf{z}_N} & p(\mathbf{z}_1,...,\mathbf{z}_{n-2}|\mathbf{X},\theta^{old})p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{X},\theta^{old}) \\ & p(\mathbf{z}_{n+1},...,\mathbf{z}_N|\mathbf{z}_{n+1},\mathbf{X},\theta^{old})z_{n-1}z_n\, ln\, A_{jk} \\ &= \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{\mathbf{z}_{n-1},\mathbf{z}_n}((\sum_{\mathbf{z}_1,...,\mathbf{z}_{n-2}} p(\mathbf{z}_1,...,\mathbf{z}_{n-2}|\mathbf{X},\theta^{old})p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{X},\theta^{old})) \\ & \sum_{\mathbf{z}_{n+1},...,\mathbf{z}_N} p(\mathbf{z}_{n+1},...,\mathbf{z}_N|\mathbf{z}_{n+1},\mathbf{X},\theta^{old}))z_{n-1}z_n\, ln\, A_{jk} \\ = \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\sum_{\mathbf{z}_{n-1},\mathbf{z}_n} p(\mathbf{z}_{n-1},\mathbf{z}_n|\mathbf{X},\theta^{old})z_{n-1}z_n\, ln\, A_{jk} &= \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}(\sum_{\mathbf{z}_{n-1},\mathbf{z}_n}\xi(\mathbf{z}_{n-1},\mathbf{z}_n)z_{n-1}z_n)\, ln\, A_{jk} \\ &= \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\xi(\mathbf{z}_{n-1,j},\mathbf{z}_{n,k})\, ln\, A_{jk}\end{aligned} \tag{13}$$

3 項目は

$$
\begin{aligned}
&\sum_{\mathbf{z}_1,\dots,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old}) \sum_{n=1}^{N}\sum_{k=1}^{K} z_{mk}\, ln\, p(\mathbf{x}_m|\mathbf{z}_m,\phi_k) = \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\mathbf{z}_1,\dots,\mathbf{z}_N} p(\mathbf{Z}|\mathbf{X},\theta^{old}) z_{nk}\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k) \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\mathbf{z}_1,\dots,\mathbf{z}_{n-1}}\sum_{\mathbf{z}_n}\sum_{\mathbf{z}_{n+1},\dots,\mathbf{z}_N} p(\mathbf{z}_1,...,\mathbf{z}_n|\mathbf{X},\theta^{old}) p(\mathbf{z}_{n+1},...,\mathbf{z}_N|\mathbf{z}_n,\mathbf{X},\theta^{old}) z_{nk}\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k) \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\mathbf{z}_n}((\sum_{\mathbf{z}_1,\dots,\mathbf{z}_{n-1}} p(\mathbf{z}_1,...,\mathbf{z}_n|\mathbf{X},\theta^{old}))\sum_{\mathbf{z}_{n+1},\dots,\mathbf{z}_N} p(\mathbf{z}_{n+1},...,\mathbf{z}_N|\mathbf{z}_n,\mathbf{X},\theta^{old})) z_{nk}\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k) \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K}\sum_{\mathbf{z}_n} p(\mathbf{z}_n|\mathbf{X},\theta^{old}) z_{nk}\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k) = \sum_{n=1}^{N}\sum_{k=1}^{K}(\sum_{\mathbf{z}_n}\gamma(\mathbf{z}_n) z_{nk})\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k) \\
&= \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(\mathbf{z}_{nk})\, ln\, p(\mathbf{x}_n|\mathbf{z}_n,\phi_k)
\end{aligned}
\tag{14}
$$

上記より、

$$
Q(\theta,\theta^{old}) = \sum_{k=1}^{K}\gamma(z_{1k})\, ln\, \pi_k + \sum_{n=2}^{N}\sum_{j=1}^{K}\sum_{k=1}^{K}\xi(z_{n-1,j}, z_{n,k})\, ln\, A_{jk} + \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\, ln\, p(x_n|\phi_k) \tag{15}
$$

となる。