13.8

出力分布を離散多項分布(13.22)とすると

$$p(x_n|z_n,\phi)=\prod_{i=1}^{D}\prod_{k=1}^{K}\mu_{ik}^{x_{ni}z_{nk}}=\prod_{k=1}^{K}\left(\prod_{i=1}^{D}\mu_{ik}^{x_{ni}}\right)^{z_{nk}} \longleftarrow \phi=\{\mu\}\text{である}$$

これを(13.9)

$$p(x_n|z_n,\phi)=\prod_{k=1}^{K}p(x_n|\phi_k)^{z_{nk}}$$

と見比べて、観測変数の分布

$$p(x_n|\phi_k)=\prod_{i=1}^{D}\mu_{ik}^{x_{ni}}$$

を得る。これを(13.17)に入れて

$$Q(\theta,\theta^{old})=C+\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\ln\prod_{i=1}^{D}\mu_{ik}^{x_{ni}}=C+\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\sum_{i=1}^{D}x_{ni}\ln\mu_{ik}$$

を得る。

μ_{ik}は制約条件 $\sum_{i=1}^{D}\mu_{ik}=1$ を持つのでラグランジアンは

$$L=C+\sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\sum_{i=1}^{D}x_{ni}\ln\mu_{ik}+\lambda_1\left(\sum_{i=1}^{D}\mu_{i1}-1\right)+\cdots+\lambda_K\left(\sum_{i=1}^{D}\mu_{iK}-1\right)$$

となる。これよりμ_{ik}についての停留条件は

$$0=\frac{\partial L}{\partial\mu_{ik}}=\sum_{n=1}^{N}\gamma(z_{nk})x_{ni}\frac{1}{\mu_{ik}}+\lambda_k$$

$$\therefore\ \mu_{ik}=\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_{ni}}{-\lambda_k}$$

となる。μ_{ik}の制約より

$$1=\sum_{i=1}^{D}\mu_{ik}=\sum_{i=1}^{D}\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_{ni}}{-\lambda_k}$$

$$\therefore\ -\lambda_k=\sum_{i=1}^{D}\sum_{n=1}^{N}\gamma(z_{nk})x_{ni}=\sum_{n=1}^{N}\gamma(z_{nk}) \longleftarrow x\text{は離散多項分布なので}\ \sum_{i=1}^{D}x_{ni}=1$$

よって

$$\mu_{ik}=\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_{ni}}{\sum_{n=1}^{N}\gamma(z_{nk})}$$

を得る。

観測変数がベルヌーイ分布 (B.1) とすると

$$p(x_n|\mu_k) = \mu_k^{x_n}(1-\mu_k)^{1-x_n}$$

これを (13.9) に入れて出力分布は

$$p(x_n|z_n,\phi) = \prod_{k=1}^{K} p(x_n|\phi_k)^{z_{nk}} = \prod_{k=1}^{K}\left\{\mu_k^{x_n}(1-\mu_k)^{1-x_n}\right\}^{z_{nk}}$$

となる。

また (13.17) に入れて

$$Q(\theta,\theta^{old}) = C + \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\ln\mu_k^{x_n}(1-\mu_k)^{1-x_n}$$

$$= C + \sum_{n=1}^{N}\sum_{k=1}^{K}\gamma(z_{nk})\left\{x_n\ln\mu_k + (1-x_n)\ln(1-\mu_k)\right\}$$

μ_k についての停留条件は

$$0 = \frac{\partial Q}{\partial\mu_k} = \sum_{n=1}^{N}\gamma(z_{nk})\left(\frac{x_n}{\mu_k} - \frac{1-x_n}{1-\mu_k}\right)$$

これより

$$\frac{\sum_{n=1}^{N}\gamma(z_{nk})x_n}{\mu_k} = \frac{\sum_{n=1}^{N}\gamma(z_{nk})(1-x_n)}{1-\mu_k}$$

$$\therefore \frac{1}{\mu_k} = \frac{\sum_{n=1}^{N}\gamma(z_{nk})}{\sum_{n=1}^{N}\gamma(z_{nk})x_n}$$

$$\therefore \mu_k = \frac{\sum_{n=1}^{N}\gamma(z_{nk})x_n}{\sum_{n=1}^{N}\gamma(z_{nk})}$$

を得る。