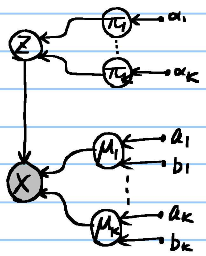
9.18

ノートのタイトル 2016/04/25

演習9.4と似ている。



目標はπ,μの事ゴ分布の対数

$\ln p(\mu,\pi|X)$, $\mu=(\mu_1,\cdots\mu_K)$, $\pi=(\pi_1,\cdots\pi_K)$, $X=\begin{pmatrix}x_1^T\\ \vdots\\ x_N^T\end{pmatrix}$

の最大化である。

ここで

$$\ln p(\mu,\pi|X)=\ln\left\{\sum_Z p(\mu,\pi,Z|X)\right\}$$

であるが、総和$\sum_Z$のため、対数が直接$p(\mu,\pi,Z|X)$に掛からず右辺の最大化は面倒と予想される。

EMアルゴリズムでは、代わりに

$$E_{Z|X}[\ln p(\mu,\pi,Z|X)]$$

を最大化させる。

ここで

$$\begin{aligned}
E_{Z|X}[\ln p(\mu,\pi,Z|X)] &= \sum_Z p(Z|X)\ln p(\mu,\pi,Z|X)\\
&= \sum_Z p(Z|X)\ln\frac{p(Z,X|\mu,\pi)\,p(\mu,\pi)}{p(X)}\\
&= \sum_Z p(Z|X)\ln p(Z,X|\mu,\pi)+\underbrace{\sum_Z p(Z|X)}_{=1}\{\ln p(\mu,\pi)-\ln p(X)\}\\
&= \sum_Z p(Z|X)\ln p(Z,X|\mu,\pi)+\ln p(\mu)+\ln p(\pi)-\ln p(X)
\end{aligned}$$

Zに依らないので$\sum_Z$の外に出す

μ,πは独立なので $p(\mu,\pi)=p(\mu)p(\pi)$

$p(X)$ は μ,π に依らないので $E_{Z|X}[\ln p(\mu,\pi,Z|X)]$ の最大化は

$$f(\mu,\pi)=\sum_Z p(Z|X)\ln p(Z,X|\mu,\pi)+\ln p(\mu)+\ln p(\pi)$$

の最大化と同値である。

ここで

$$E_{Z|X}[\ln p(Z,X|\mu,\pi)]=\sum_Z p(Z|X)\ln p(Z,X|\mu,\pi)$$

なので

$$f(\mu,\pi)=E_{Z|X}[\ln p(Z,X|\mu,\pi)]+\ln p(\mu)+\ln p(\pi)$$

となる。$E_{Z|X}[\ln p(Z,X|\mu,\pi)]$ は (9.55) で与えられる。

負担率 $\gamma(z_{nk})$ は (9.56) で与えられる。

Eステップの $\gamma(z_{nk})$ の更新式は (9.56) 式で与えられる。

(fをμ_kについて最大化)

f を μ_k で微分して 0 とおく。

演習9.15

$$\frac{\partial}{\partial \mu_k} E_{Z|X}[\ln p(Z, X|\mu, \pi)] = \sum_n \gamma(z_{nk}) \left(\frac{x_n}{\mu_k} - \frac{1-x_n}{1-\mu_k} \right)$$

また $p(\mu_k)$ としてベータ分布を選ぶと

$$p(\mu_k) = \frac{\Gamma(a_k+b_k)}{\Gamma(a_k)\Gamma(b_k)} \mu_k^{a_k-1}(1-\mu_k)^{b_k-1} \longleftarrow \text{(2.13) ベータ分布}$$

これより

$\mu_1 \cdots \mu_K$ 独立

$$\frac{\partial}{\partial \mu_k} \ln p(\mu) = \frac{\frac{\partial p(\mu)}{\partial \mu_k}}{p(\mu)} = \frac{p(\mu_1)\cdots\frac{\partial p(\mu_k)}{\partial \mu_k}\cdots p(\mu_K)}{p(\mu_1)\cdots p(\mu_k)\cdots p(\mu_K)} \longleftarrow p(\mu) = p(\mu_1, \cdots, \mu_K) = p(\mu_1)\cdots p(\mu_K)$$

$$= \frac{\frac{\partial}{\partial \mu_k} \mu_k^{a_k-1}(1-\mu_k)^{b_k-1}}{\mu_k^{a_k-1}(1-\mu_k)^{b_k-1}} = (a_k-1)\mu_k^{-1} - (b_k-1)(1-\mu_k)^{-1}$$

よって

$$\frac{\partial f}{\partial \mu_k} = \sum_n \gamma(z_{nk}) \left(\frac{x_n}{\mu_k} - \frac{1-x_n}{1-\mu_k} \right) + \frac{a_k-1}{\mu_k} - \frac{b_k-1}{1-\mu_k} = 0$$

を得る。これより

$$\frac{\sum_n \gamma(z_{nk}) x_n}{\mu_k} - \frac{\sum_n \gamma(z_{nk}) - \sum_n \gamma(z_{nk}) x_n}{1-\mu_k} + \frac{a_k-1}{\mu_k} - \frac{b_k-1}{1-\mu_k} = 0$$

$$\therefore \frac{N_k \bar{x}_k}{\mu_k} - \frac{N_k - N_k \bar{x}_k}{1-\mu_k} + \frac{a_k-1}{\mu_k} - \frac{b_k-1}{1-\mu_k} = 0 \longleftarrow \text{(9.57), (9.58)}$$

両辺に $\mu_k(1-\mu_k)$ をかけて整理すると

$$(1-\mu_k)N_k\bar{x}_k - \mu_k(N_k - N_k\bar{x}_k) + (a_k-1)(1-\mu_k) - (b_k-1)\mu_k = 0$$

$$\therefore \mu_k = \frac{N_k\bar{x}_k + a_k - 1}{N_k + a_k - 1 + b_k - 1}$$

を得る。これがMステップの μ の更新式である。

(fをπ_kについて最大化させる)

$$\sum_k \pi_k = 1$$

という条件があるので、ラグランジュ未定乗数法を使う

$$L = f + \lambda \left(\sum_k \pi_k - 1 \right)$$

をπ_kについて微分して0とおく。

(9.55)より

$$\frac{\partial}{\partial \pi_k} E_{Z|X}[\ln p(X, Z \mid M, \pi)] = \sum_{n=1}^{N} \gamma(z_{nk}) \frac{1}{\pi_k}$$

また$p(\pi)$としてディリクレ分布を選ぶと

$$p(\pi) = p(\pi_1, \cdots \pi_K) = \frac{\Gamma(\alpha_0)}{\Gamma(\alpha_1) \cdots \Gamma(\alpha_K)} \prod_{k=1}^{K} \pi_k^{\alpha_k - 1} \longleftarrow \text{(2.38) ディリクレ分布}$$

なので

$$\frac{\partial}{\partial \pi_k} \ln p(\pi) = \frac{\pi_1^{\alpha_1 - 1} \cdots \frac{\partial}{\partial \pi_k} \pi_k^{\alpha_k - 1} \cdots \pi_K^{\alpha_K - 1}}{\pi_1^{\alpha_1 - 1} \cdots \pi_k^{\alpha_k - 1} \cdots \pi_K^{\alpha_K - 1}} = (\alpha_k - 1) \pi_k^{-1}$$

よって

$$\frac{\partial L}{\partial \pi_k} = \frac{\sum_n \gamma(z_{nk})}{\pi_k} + \frac{\alpha_k - 1}{\pi_k} + \lambda = 0$$

$$\therefore \quad \frac{\overset{(9.57)}{N_k}}{\pi_k} + \frac{\alpha_k - 1}{\pi_k} + \lambda = 0$$

両辺にπ_kをかけてkについて総和すると

$$\sum_k \{ (N_k + \alpha_k - 1) + \lambda \pi_k \} = 0$$

$$\left\{ \begin{array}{l} \sum_k N_k = \sum_k \sum_n \gamma(z_{nk}) \overset{(9.56)}{=} \sum_n \sum_k \frac{\pi_k p(x_n \mid \mu_k)}{\sum_j \pi_j p(x_n \mid \mu_j)} = \sum_n \frac{\sum_k \pi_k p(x_n \mid \mu_k)}{\sum_j \pi_j p(x_n \mid \mu_j)} = N \\ \sum_{k=1}^{K} \alpha_k = \alpha_0 \quad \cdots (2.39) \\ \sum_k \pi_k = 1 \end{array} \right.$$

$$\therefore \quad \lambda = -N - \alpha_0 + K$$

λを元の式に戻してπ_kについて解くと

$$\pi_k = \frac{N_k + \alpha_k - 1}{N + \alpha_0 - K}$$

を得る。これがMステップのπの更新式となる。