# 9.4

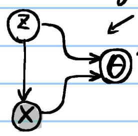
ノートのタイトル 2016/04/16

MAP推定に9.3章と同じ議論をあてはめる
最初の目的は、$\theta$の事後分布

$$\ln p(\theta|X)$$

の$\theta$についての最大化である。
ここで潜在変数の存在を仮定したモデルを考えると

$\theta$の事後分布を考える際のモデルなので矢印が$\theta$に向かっている。（図8.37）




MAP推定では$\theta$を確率変数と考える

$$\ln p(\theta|X) = \ln\left\{\sum_{Z} p(\theta, Z|X)\right\}$$

となる。右辺で$\ln$が直接$p$に作用していないため最大化する問題は解きにくいと予想される
なので 9.3章と同様に

$$\ln p(\theta, Z|X)$$

を$\theta$について最大化させることにし、その解$\theta$が$\ln p(\theta|X)$も最大化すると考える。(←これの証明なし)
しかし、$Z$は潜在変数なので、$\ln p(\theta, Z|X)$を最大化するためのデータが与えられない
なので、9.3章と同様に

$$E[\ln p(\theta, Z|X)]$$ (ただし期待値の計算に使う分布は$Z$の事後分布$p(Z|X,\theta)$とする)

この期待値を$\theta$について最大化させ、その解$\theta$が$\ln p(\theta, Z|X)$も最大化すると考える。(←これの証明なし)
ここで

$$
\begin{aligned}
E[\ln p(\theta, Z|X)] &= \sum_{Z} p(Z|X,\theta) \ln p(\theta, Z|X) \\
&= \sum_{Z} p(Z|X,\theta) \ln \frac{p(Z,X|\theta)p(\theta)}{p(X)} \quad \longleftarrow \ p(\theta, Z|X)p(X) = p(Z,X|\theta)p(\theta) \\
&= \sum_{Z} p(Z|X,\theta) \ln p(Z,X|\theta) + \underbrace{\sum_{Z} p(Z|X,\theta)}_{1} \ln p(\theta) - \underbrace{\sum_{Z} p(Z|X,\theta)}_{1} \ln p(X) \\
&= \sum_{Z} p(Z|X,\theta) \ln p(Z,X|\theta) + \ln p(\theta) - \ln p(X)
\end{aligned}
$$

Zに依らないので和の外に出してよい

であり、また$p(X)$は$\theta$に依らないので$E[\ln p(\theta, Z|X)]$の$\theta$についての最大化は

$$\sum_{Z} p(Z|X,\theta) \ln p(Z,X|\theta) + \ln p(\theta)$$

の$\theta$について最大化と同値である。
またMステップにおいては、$Z$の事後分布を$\theta^{old}$を使って計算するので、
結局Mステップにおける最大化すべき量は

(9.30)

$$\sum_{Z} p(Z|X,\theta^{old}) \ln p(Z,X|\theta) + \ln p(\theta) = Q(\theta, \theta^{old}) + \ln p(\theta)$$

となる。