「ベイズ統計の理論と方法」

2017 年 6 月 20 日

1. はじめに

章末問題 1

次の等式が成り立つことを示せ。

$$\inf_{\beta} F_n(\beta) = \inf_{w \in W} \left\{ -\sum_{i=1}^{n} \log p(X_i|w) \right\}.$$

$$\begin{aligned} F_n(\beta) &= -\frac{1}{\beta} \log Z_n(\beta) \quad \Leftarrow (1.13) \\ &= -\frac{1}{\beta} \log \int_W \varphi(w) \prod_{i=1}^{n} p(X_i|w)^{\beta} dw. \quad \Leftarrow (1.6) \end{aligned}$$

ここで天下り的に $L_n(w)$ を定義する。(対数尤度の n 分の一の符号を変えたもの)

$$L_n(w) = -\frac{1}{n} \sum_{i=1}^{n} \log p(X_i|w) \quad \text{すると}$$

$$\exp(-n\beta L_n(w)) = \prod_{i=1}^{n} p(X_i|w)^{\beta}$$

$L_n(w)$ を最小にするパラメータを $\hat{w}$ とすると、$nL_n(\hat{w}) \leq nL_n(w)$ ゆえ、

$$\begin{aligned} F_n(\beta) &= -\frac{1}{\beta} \log \int_W \exp(-n\beta L_n(w)) \varphi(w) dw \\ &\geq -\frac{1}{\beta} \log \int_W \exp(-n\beta L_n(\hat{w})) \varphi(w) dw \\ &= -\frac{1}{\beta} \log \left\{ \exp(-n\beta L_n(\hat{w})) \underbrace{\int_W \varphi(w) dw}_{=1} \right\} \\ &= nL_n(\hat{w}) \end{aligned}$$

一方、$W(\epsilon) = \{w \in W \mid nL_n(w) < nL_n(\hat{w}) + \epsilon\}$ とすると、被積分関数が非負で $W(\epsilon) \subset W$ だから、

$$\int_W (\dots) dw \geq \int_{W(\epsilon)} (\dots) dw, \quad \text{よって} \quad -\frac{1}{\beta}\int_W (\dots) dw \leq -\frac{1}{\beta}\int_{W(\epsilon)} (\dots) dw$$

よって

$$\begin{aligned} F_n(\beta) &= -\frac{1}{\beta}\log\int_W \exp\left(-n\beta L_n(w)\right)\varphi(w)dw \\ &\leq -\frac{1}{\beta}\log\int_{W(\epsilon)} \exp\left(-n\beta L_n(w)\right)\varphi(w)dw \end{aligned}$$

さらに $w \in W(\epsilon)$ で、

$$\begin{gathered} nL_n(w) < nL_n(\hat{w}) + \epsilon \\ -n\beta L_n(w) > -n\beta L_n(\hat{w}) - \beta\epsilon \\ \exp(-n\beta L_n(w)) > \exp(-n\beta L_n(\hat{w}))\exp(-\beta\epsilon) \\ -\frac{1}{\beta}\log\int_{W(\epsilon)} \exp(-n\beta L_n(w))\varphi(w)dw < -\frac{1}{\beta}\log\int_{W(\epsilon)} \exp(-n\beta L_n(\hat{w}))\exp(-\beta\epsilon)\varphi(w)dw \end{gathered}$$

よって

$$\begin{aligned} F_n(\beta) &\leq -\frac{1}{\beta}\log\int_{W(\epsilon)} \exp(-n\beta L_n(\hat{w}))\exp(-\beta\epsilon)\varphi(w)dw \\ &= nL_n(\hat{w}) + \epsilon - \frac{1}{\beta}\log\int_{W(\epsilon)} \varphi(w)dw \end{aligned}$$

よって、

$$nL_n(\hat{w}) \leq F_n(\beta) \leq nL_n(\hat{w}) + \underbrace{\epsilon - \frac{1}{\beta}\log\int_{W(\epsilon)} \varphi(w)dw}_{\to 0 \text{ as } \epsilon\to 0 \text{ and } \beta\to\infty}$$

よって $F_n(\beta) \to nL_n(\hat{w})$ で $\sup_\beta F_n(\beta) = \lim_{\beta\to\infty} F_n(\beta)$ であるが、

$$\inf_{w\in W}\left\{-\sum_{i=1}^n \log p(X_i|w)\right\} = nL_n(\hat{w}).$$

□

章末問題 2

$L(w) = -\mathbb{E}_X[\log p(X|w)]$ とするとき、次の不等式が成り立つことを示せ。

$$\mathbb{E}[F_n(\beta)] \leq \frac{1}{\beta}\log\int \exp(-\beta L(w))\varphi(w)dw$$

$$
\begin{aligned}
L_n(w) &= -\frac{1}{n}\sum_{i=1}^{n}\log p(X_i|w) \\
\hat{L}(w) &= L_n(w) - L(w) \quad \Rightarrow \ \mathbb{E}[\hat{L}(w)] = 0 \\
F_n(\beta) &= -\frac{1}{\beta}\log\int \exp\left(-n\beta L_n(w)\right)\varphi(w)dw \\
&= -\frac{1}{\beta}\log\underbrace{\frac{\int \exp\left(-n\beta\hat{L}(w)\right)\exp\left(-n\beta L(w)\right)\varphi(w)dw}{\int \exp\left(-n\beta L(w)\right)\varphi(w)dw}}_{\text{重み}\exp(-n\beta L(w))\varphi(w)\ \text{での}\ \mathbb{E}[\exp\left(-n\beta\hat{L}\right)]} - \frac{1}{\beta}\log\int \exp\left(-n\beta L(w)\right)\varphi(w)dw
\end{aligned}
$$

最後の変形は、$L_n = \hat{L} + L$ 及び $\frac{1}{\beta}\log\int\exp\left(-n\beta L(w)\right)\varphi(w)dw$ を足して引いてしたもの。

$f(x) = \exp(-n\beta x)$ とおくと f は凸関数。すると Jensen の不等式 $\mathbb{E}[f(x)] \geq f(\mathbb{E}[x])$ を用いて、

$$
\mathbb{E}[\exp(-n\beta\hat{L})] \geq \exp(-n\beta\underbrace{\mathbb{E}[\hat{L}]}_{=0}) = 1
$$

よって $F_n(\beta)$ の式の第 1 項は消えて、

$$
F_n(\beta) \leq -\frac{1}{\beta}\log\int\exp\left(-n\beta L(w)\right)\varphi(w)dw \quad \square
$$

章末問題 3

例 1 では分散 $\sigma^2 > 0$ を定数としている。分散もパラメータとして推測する場合を考える。$s = 1/\sigma^2$ とおくと正規分布の式 (1.25) は

$$
p(x|m,s) = \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{s}{2}x^2 + msx - \left(\frac{m^2 s}{s} - \frac{1}{2}\log s\right)\right\}
$$

となる。この確率モデルの共役事前分布を作れ。

$$
\begin{aligned}
p(x|m,s) &= \frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{s}{2}x^2 + msx - \left(\frac{m^2 s}{2} - \frac{1}{2}\log s\right)\right\} \\
&= v(x)\exp\left(f(w)\cdot g(x)\right) \\
v(x) &= \frac{1}{\sqrt{2\pi}} \\
f(w) &= \left(-\frac{s}{2},\ ms,\ -\frac{m^2 s}{2} + \frac{1}{2}\log s\right) \\
g(x) &= \left(x^2,\ x,\ 1\right) \quad \Rightarrow \phi = (\phi_1, \phi_2, \phi_3)\ \text{で置き換える}
\end{aligned}
$$

$g(x)$ を ϕ で置き換えて得られる共役事前分布は、

$$\begin{aligned}\varphi(m,s|\phi) &= \frac{1}{z(\phi)}\exp\left\{-\frac{s}{2}\phi_1 + ms\phi_2 - \left(\frac{m^2 s}{2} - \frac{1}{2}\log s\right)\phi_3\right\} \\ &= \frac{1}{z(\phi)} s^{\frac{\phi_3}{2}} \exp\left\{-\frac{s}{2}\phi_1 + ms\phi_2 - \frac{m^2 s}{2}\phi_3\right\}\end{aligned}$$

これは

- m に関しては正規分布
- s に関してはガンマ分布 $(\propto s^{\phi_3/2}e^{-s/\alpha} = \mathrm{Gamma}(\frac{\phi_3}{2},\alpha))$

の形をしている。

$$\begin{aligned}z(\phi) &= \int s^{\frac{\phi_3}{2}} ds \int \exp\left\{-\frac{s}{2}\phi_1 + ms\phi_2 - \frac{m^2 s}{2}\phi_3\right\} dm \quad \Rightarrow \textbf{先に } m \textbf{ に関するガウス積分を実行} \\ &= \int s^{\frac{\phi_3}{2}} ds \int \exp\left\{-\frac{s\phi_3}{2}\left(m^2 - 2\frac{\phi_2}{\phi_3}m\right) - \frac{s\phi_1}{2}\right\} dm \\ &= \int s^{\frac{\phi_3}{2}} ds \int \exp\left\{-\frac{s\phi_3}{2}\left(m - \frac{\phi_2}{\phi_3}m\right)^2 + \frac{s\phi_3}{2}\frac{\phi_2^2}{\phi_3^2} - \frac{s\phi_1}{2}\right\} dm \\ &= \int s^{\frac{\phi_3}{2}} \exp\left\{\frac{s\phi_3}{2}\frac{\phi_2^2}{\phi_3^2} - \frac{s\phi_1}{2}\right\} ds \underbrace{\int \exp\left\{-\frac{s\phi_3}{2}\left(m - \frac{\phi_2}{\phi_3}m\right)^2\right\} dm}_{\sqrt{2\pi/(s\phi_3)}} \\ &= \sqrt{\frac{2\pi}{\phi_3}} \int s^{\frac{\phi_3 - 1}{2}} \exp\left\{-\frac{\phi_1\phi_3 - \phi_2^2}{2\phi_3}s\right\} ds\end{aligned}$$

ここでガンマ分布が

$$\mathrm{Gamma}(x|k,\theta) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\Gamma(k)\theta^k}$$

を用いると $\int s^{k-1}\exp\{-\frac{s}{\theta}\}ds = \Gamma(k)\theta^k$ ゆえ

$$z(\phi) = \sqrt{\frac{2\pi}{\phi_3}}\left(\frac{2\phi_3}{\phi_1\phi_3 - \phi_2^2}\right)^{\frac{\phi_3+1}{2}} \Gamma(\frac{\phi_3+1}{2})$$

□