ベイズ推論による機械学習入門
正誤表

須山敦志

2018 年 5 月 3 日 更新

第 1〜3 刷において下記の誤りがありました．お詫びして訂正いたします．

vii まえがきの記号表記

【誤】
$\mathbb{R}$ は実数の集合で，$\mathbb{R}^+$ は非負の実数の集合を表します．

【正】
$\mathbb{R}$ は実数の集合で，$\mathbb{R}^+$ は正の実数の集合を表します．

p.15 式の参照先

【誤】
また，式（1.19）の両辺を $p(y)$ で割り，条件付き分布の定義式（1.17）を用いることによって，独立性を次のように書くこともできます．

【正】
また，式（1.21）の両辺を $p(y)$ で割り，条件付き分布の定義式（1.17）を用いることによって，独立性を次のように書くこともできます．

p.69 式 (2.88)

【誤】

$$\ln \mathcal{W}(\Lambda|\nu, W) = \frac{\nu}{2}\ln\Lambda - \frac{\Lambda}{2W} + \ln C_{\mathcal{W}}(\nu, W)$$

【正】

$$\ln \mathcal{W}(\Lambda|\nu, W) = \frac{\nu-2}{2}\ln\Lambda - \frac{\Lambda}{2W} + \ln C_{\mathcal{W}}(\nu, W)$$

p.87 の本文

【誤】
ただし，パラメータ γ および a は，ここでは次のようになります．

【正】
ただし，パラメータ r および p は，ここでは次のようになります．

p.96 式 (3.86) の符号

【誤】

$$\begin{aligned}\ln p(\lambda|\mathbf{X}) =&(\frac{N}{2}+a-1)\ln\lambda \\ &+\{\frac{1}{2}(\sum_{n=1}^{N}x_n^2+\beta m^2-\hat{\beta}\hat{m}^2)+b\}\lambda+\mathrm{const.}\end{aligned}$$

【正】

$$\begin{aligned}\ln p(\lambda|\mathbf{X}) =&(\frac{N}{2}+a-1)\ln\lambda \\ &-\{\frac{1}{2}(\sum_{n=1}^{N}x_n^2+\beta m^2-\hat{\beta}\hat{m}^2)+b\}\lambda+\mathrm{const.}\end{aligned}$$

p.100 式（3.113）

【誤】

$$\ln p(\mathbf{\Lambda}|\mathbf{X}) = \sum_{n=1}^{N}\ln\mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu},\mathbf{\Lambda}^{-1})+\mathcal{W}(\mathbf{\Lambda}|\nu,\mathbf{W})+\mathrm{const.}$$

【正】

$$\ln p(\mathbf{\Lambda}|\mathbf{X}) = \sum_{n=1}^{N}\ln\mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu},\mathbf{\Lambda}^{-1})+\ln\mathcal{W}(\mathbf{\Lambda}|\nu,\mathbf{W})+\mathrm{const.}$$

p.101 式 (3.118)

【誤】

$$p(\mathbf{\Lambda}|\mathbf{x}_*) = \mathcal{W}(\mathbf{x}_*|1+\nu,\mathbf{W}(\mathbf{x}_*))$$

【正】

$$p(\mathbf{\Lambda}|\mathbf{x}_*) = \mathcal{W}(\mathbf{\Lambda}|1+\nu,\mathbf{W}(\mathbf{x}_*))$$

p.109 2 行目

【誤】
　実践

【正】
　実線

p.110 式 (3.157) の 1 行目

【誤】

$$\begin{aligned}\ln p(\mathbf{Y}|\mathbf{X}) = -\frac{1}{2}\{\lambda \sum_{n=1}^{N} y_n^2 - \ln\lambda + \ln 2\pi + \mathbf{m}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{m} - \ln|\mathbf{\Lambda}| \\ - \hat{\mathbf{m}}^{\mathrm{T}}\hat{\mathbf{\Lambda}}\hat{\mathbf{m}} + \ln|\hat{\mathbf{\Lambda}}|\}\end{aligned}$$

【正】

$$\begin{aligned}\ln p(\mathbf{Y}|\mathbf{X}) = -\frac{1}{2}\{\sum_{n=1}^{N}(\lambda y_n^2 - \ln\lambda + \ln 2\pi) + \mathbf{m}^{\mathrm{T}}\mathbf{\Lambda}\mathbf{m} - \ln|\mathbf{\Lambda}| \\ - \hat{\mathbf{m}}^{\mathrm{T}}\hat{\mathbf{\Lambda}}\hat{\mathbf{m}} + \ln|\hat{\mathbf{\Lambda}}|\}\end{aligned}$$

p.113 参考 3.3 の 8 行目

【誤】
最尤推定では事前分布 θ は導入しません.

【正】
最尤推定では事前分布 $p(\theta)$ は導入しません.

p.113 参考 3.3 の 19 行目

【誤】
一般的には次のようにパラメータの事前分布 θ を導入し事後分布を θ に関して最大化します.

【正】
一般的には次のようにパラメータの事前分布 $p(\theta)$ を導入し事後分布を θ に関して最大化します.

p.134 式 (4.48) の 1 行目のブラケットの位置

【誤】

$$\begin{aligned}\langle \ln p(x_n|\mathbf{s}_n, \boldsymbol{\lambda})\rangle_{q(\boldsymbol{\lambda})} &= \langle \sum_{k=1}^{K} s_{n,k} \ln \mathrm{Poi}(x_n|\lambda_k)\rangle_{q(\lambda_k)} \\ &= \sum_{k=1}^{K} s_{n,k}(x_n\langle\ln\lambda_k\rangle - \langle\lambda_k\rangle) + \mathrm{const.}\end{aligned}$$

【正】

$$\begin{aligned}\langle \ln p(x_n|\mathbf{s}_n, \boldsymbol{\lambda})\rangle_{q(\boldsymbol{\lambda})} &= \sum_{k=1}^{K} \langle s_{n,k} \ln \mathrm{Poi}(x_n|\lambda_k)\rangle_{q(\lambda_k)} \\ &= \sum_{k=1}^{K} s_{n,k}(x_n\langle\ln\lambda_k\rangle - \langle\lambda_k\rangle) + \mathrm{const.}\end{aligned}$$

p.134 式 (4.52) の 1 行目の期待値

【誤】

$$
\begin{aligned}
\ln q(\boldsymbol{\lambda}, \boldsymbol{\pi}) &= \langle \ln p(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda}, \boldsymbol{\pi}) \rangle_{q(\boldsymbol{\lambda}, \boldsymbol{\pi})} + \mathrm{const.} \\
&= \langle \ln p(\mathbf{X}|\mathbf{S}, \boldsymbol{\lambda}) \rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\lambda}) \\
&\quad + \langle \ln p(\mathbf{S}|\boldsymbol{\pi}) \rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\pi}) + \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
\ln q(\boldsymbol{\lambda}, \boldsymbol{\pi}) &= \langle \ln p(\mathbf{X}, \mathbf{S}, \boldsymbol{\lambda}, \boldsymbol{\pi}) \rangle_{q(\mathbf{S})} + \mathrm{const.} \\
&= \langle \ln p(\mathbf{X}|\mathbf{S}, \boldsymbol{\lambda}) \rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\lambda}) \\
&\quad + \langle \ln p(\mathbf{S}|\boldsymbol{\pi}) \rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\pi}) + \mathrm{const.}
\end{aligned}
$$

p.135 式 (4.53) の 1 行目のブラケットの位置

【誤】

$$
\begin{aligned}
\ln q(\boldsymbol{\lambda}) =& \langle \sum_{n=1}^{N} \sum_{k=1}^{K} s_{n,k} \ln \mathrm{Poi}(x_n|\lambda_k) \rangle_{q(\mathbf{s}_n)} + \sum_{k=1}^{K} \ln \mathrm{Gam}(\lambda_k|a, b) \\
&+ \mathrm{const.} \\
=& \sum_{k=1}^{K} \{ (\sum_{n=1}^{N} \langle s_{n,k} \rangle x_n + a - 1) \ln \lambda_k - (\sum_{n=1}^{N} \langle s_{n,k} \rangle + b) \lambda_k \} \\
&+ \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
\ln q(\boldsymbol{\lambda}) =& \sum_{n=1}^{N} \langle \sum_{k=1}^{K} s_{n,k} \ln \mathrm{Poi}(x_n|\lambda_k) \rangle_{q(\mathbf{s}_n)} + \sum_{k=1}^{K} \ln \mathrm{Gam}(\lambda_k|a, b) \\
&+ \mathrm{const.} \\
=& \sum_{k=1}^{K} \{ (\sum_{n=1}^{N} \langle s_{n,k} \rangle x_n + a - 1) \ln \lambda_k - (\sum_{n=1}^{N} \langle s_{n,k} \rangle + b) \lambda_k \} \\
&+ \mathrm{const.}
\end{aligned}
$$

p.150 式 (4.106) の 2 行目のブラケットの位置

【誤】

$$
\begin{aligned}
&\langle \ln p(\mathbf{x}_n|\mathbf{s}_n, \boldsymbol{\mu}, \boldsymbol{\Lambda})\rangle_{q(\boldsymbol{\mu},\boldsymbol{\Lambda})} \\
&= \langle \sum_{k=1}^{K} s_{n,k} \ln \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k)\rangle_{q(\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k)} \\
&= \sum_{k=1}^{K} s_{n,k} \{ -\frac{1}{2}\mathbf{x}_n^{\mathrm{T}}\langle\boldsymbol{\Lambda}_k\rangle\mathbf{x}_n + \mathbf{x}_n^{\mathrm{T}}\langle\boldsymbol{\Lambda}_k\boldsymbol{\mu}_k\rangle \\
&\qquad\qquad - \frac{1}{2}\langle\boldsymbol{\mu}_k^{\mathrm{T}}\boldsymbol{\Lambda}_k\boldsymbol{\mu}_k\rangle + \frac{1}{2}\langle\ln|\boldsymbol{\Lambda}_k|\rangle\} + \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
&\langle \ln p(\mathbf{x}_n|\mathbf{s}_n, \boldsymbol{\mu}, \boldsymbol{\Lambda})\rangle_{q(\boldsymbol{\mu},\boldsymbol{\Lambda})} \\
&= \sum_{k=1}^{K} \langle s_{n,k} \ln \mathcal{N}(\mathbf{x}_n|\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k)\rangle_{q(\boldsymbol{\mu}_k,\boldsymbol{\Lambda}_k)} \\
&= \sum_{k=1}^{K} s_{n,k} \{ -\frac{1}{2}\mathbf{x}_n^{\mathrm{T}}\langle\boldsymbol{\Lambda}_k\rangle\mathbf{x}_n + \mathbf{x}_n^{\mathrm{T}}\langle\boldsymbol{\Lambda}_k\boldsymbol{\mu}_k\rangle \\
&\qquad\qquad - \frac{1}{2}\langle\boldsymbol{\mu}_k^{\mathrm{T}}\boldsymbol{\Lambda}_k\boldsymbol{\mu}_k\rangle + \frac{1}{2}\langle\ln|\boldsymbol{\Lambda}_k|\rangle\} + \mathrm{const.}
\end{aligned}
$$

p.148 文中

【誤】

以上をまとめると，ポアソン混合モデルの事後分布に対するギブスサンプリングは**アルゴリズム 4.5** になります．

【正】

以上をまとめると，ガウス混合モデルの事後分布に対するギブスサンプリングは**アルゴリズム 4.5** になります．

p.150 式 (4.110) の 1 行目の期待値

【誤】

$$
\begin{aligned}
\ln q(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\pi}) &= \langle \ln p(\mathbf{X}, \mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\pi})\rangle_{q(\boldsymbol{\mu},\boldsymbol{\Lambda},\boldsymbol{\pi})} + \mathrm{const.} \\
&= \langle \ln p(\mathbf{X}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\Lambda})\rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \\
&\quad + \langle \ln p(\mathbf{S}|\boldsymbol{\pi})\rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\pi}) + \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
\ln q(\boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\pi}) &= \langle \ln p(\mathbf{X}, \mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\Lambda}, \boldsymbol{\pi})\rangle_{q(\mathbf{S})} + \mathrm{const.} \\
&= \langle \ln p(\mathbf{X}|\mathbf{S}, \boldsymbol{\mu}, \boldsymbol{\Lambda})\rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \\
&\quad + \langle \ln p(\mathbf{S}|\boldsymbol{\pi})\rangle_{q(\mathbf{S})} + \ln p(\boldsymbol{\pi}) + \mathrm{const.}
\end{aligned}
$$

p.151 式 (4.111) の 2 行目

【誤】

$$
\begin{aligned}
&\ln q(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \\
&= \langle \sum_{n=1}^{N} \sum_{k=1}^{K} s_{n,k} \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) \rangle_{q(s)} \\
&\quad + \sum_{k=1}^{K} \ln \mathrm{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \mathbf{m}, \beta, \nu, \mathbf{W}) + \mathrm{const.} \\
&= \sum_{k=1}^{K} \Big\{ \sum_{n=1}^{N} \langle s_{n,k} \rangle \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) \\
&\qquad\quad + \ln \mathrm{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \mathbf{m}, \beta, \nu, \mathbf{W}) \Big\} + \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
&\ln q(\boldsymbol{\mu}, \boldsymbol{\Lambda}) \\
&= \sum_{n=1}^{N} \langle \sum_{k=1}^{K} s_{n,k} \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) \rangle_{q(\mathbf{s}_n)} \\
&\quad + \sum_{k=1}^{K} \ln \mathrm{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \mathbf{m}, \beta, \nu, \mathbf{W}) + \mathrm{const.} \\
&= \sum_{k=1}^{K} \Big\{ \sum_{n=1}^{N} \langle s_{n,k} \rangle \ln \mathcal{N}(\mathbf{x}_n | \boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k) \\
&\qquad\quad + \ln \mathrm{NW}(\boldsymbol{\mu}_k, \boldsymbol{\Lambda}_k | \mathbf{m}, \beta, \nu, \mathbf{W}) \Big\} + \mathrm{const.}
\end{aligned}
$$

p.165 式（5.17）の中括弧

【誤】

$$
\begin{aligned}
&\langle \ln p(\mathbf{y}_n | \mathbf{x}_n, \mathbf{W}, \boldsymbol{\mu}) \rangle_{q(\mathbf{x}_n) q(\boldsymbol{\mu})} \\
&= -\frac{1}{2} \{ \sigma_y^{-2} \langle \mathbf{x}_n^{\mathrm{T}} \mathbf{W} \mathbf{W}^{\mathrm{T}} \mathbf{x}_n \rangle - 2\sigma_y^{-2} (\mathbf{y}_n - \langle \boldsymbol{\mu} \rangle)^{\mathrm{T}} \mathbf{W}^{\mathrm{T}} \langle \mathbf{x}_n \rangle + \mathrm{const.} \\
&= -\frac{1}{2} \sum_{d=1}^{D} \{ \mathbf{W}_d^{\mathrm{T}} \sigma_y^{-2} \langle \mathbf{x}_n \mathbf{x}_n^{\mathrm{T}} \rangle \mathbf{W}_d - 2\mathbf{W}_d^{\mathrm{T}} \sigma_y^{-2} (y_{n,d} - \langle \mu_d \rangle) \langle \mathbf{x}_n \rangle \\
&\quad + \mathrm{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
&\langle \ln p(\mathbf{y}_n|\mathbf{x}_n, \mathbf{W}, \boldsymbol{\mu})\rangle_{q(\mathbf{x}_n)q(\boldsymbol{\mu})} \\
&= -\frac{1}{2}\{\sigma_y^{-2}\langle \mathbf{x}_n^{\mathrm{T}}\mathbf{W}\mathbf{W}^{\mathrm{T}}\mathbf{x}_n\rangle - 2\sigma_y^{-2}(\mathbf{y}_n - \langle\boldsymbol{\mu}\rangle)^{\mathrm{T}}\mathbf{W}^{\mathrm{T}}\langle\mathbf{x}_n\rangle\} + \mathrm{const.} \\
&= -\frac{1}{2}\sum_{d=1}^{D}\{\mathbf{W}_d^{\mathrm{T}}\sigma_y^{-2}\langle \mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\rangle\mathbf{W}_d - 2\mathbf{W}_d^{\mathrm{T}}\sigma_y^{-2}(y_{n,d} - \langle\mu_d\rangle)\langle\mathbf{x}_n\rangle\} \\
&\quad + \mathrm{const.}
\end{aligned}
$$

p.171 文中

【誤】

非負値行列因子分解では，次のように各要素が非負値を持つ行列 $\mathbf{X} \in \mathbb{N}^{D\times N}$ を，同じく非負値を持つ 2 つの行列 $\mathbf{W} \in \mathbb{N}^{D\times M}$ と $\mathbf{H} \in \mathbb{N}^{M\times N}$ に分解します.

【正】

非負値行列因子分解では，次のように各要素が非負値を持つ行列 $\mathbf{X} \in \mathbb{N}^{D\times N}$ を，正の値を持つ 2 つの行列 $\mathbf{W} \in \mathbb{R^+}^{D\times M}$ と $\mathbf{H} \in \mathbb{R^+}^{M\times N}$ に近似分解します.

p.172 式（5.43）の 1 行目

【誤】

$$
\begin{aligned}
p(\mathbf{X}) &= \prod_{d=1}^{D}\prod_{n=1}^{N} p(X_{d,n}) \\
&= \prod_{d=1}^{D}\prod_{n=1}^{N} \mathrm{Del}(X_{d,n}|\sum_{m=1}^{M} S_{d,m,n})
\end{aligned}
$$

【正】

$$
\begin{aligned}
p(\mathbf{X}|\mathbf{S}) &= \prod_{d=1}^{D}\prod_{n=1}^{N} p(X_{d,n}|\mathbf{S}_{d,:,n}) \\
&= \prod_{d=1}^{D}\prod_{n=1}^{N} \mathrm{Del}(X_{d,n}|\sum_{m=1}^{M} S_{d,m,n})
\end{aligned}
$$

p.174 式 (5.51) の 3 行目

【誤】

$$
\begin{aligned}
\ln q(\mathbf{S}) &= \ln p(\mathbf{X}|\mathbf{S}) + \langle \ln p(\mathbf{S}|\mathbf{W},\mathbf{H})\rangle_{q(\mathbf{W})q(\mathbf{H})} + \text{const.} \\
&= \sum_{d=1}^{D}\sum_{n=1}^{N}\big\{\ln p(X_{d,n}|\sum_{m=1}^{M} S_{d,m,n}) \\
&\qquad\qquad + \sum_{m=1}^{M}\langle \ln p(S_{d,m,n}|W_{d,m}H_{m,n})\rangle\big\}_{q(\mathbf{W})q(\mathbf{H})} + \text{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
\ln q(\mathbf{S}) &= \ln p(\mathbf{X}|\mathbf{S}) + \langle \ln p(\mathbf{S}|\mathbf{W},\mathbf{H})\rangle_{q(\mathbf{W})q(\mathbf{H})} + \text{const.} \\
&= \sum_{d=1}^{D}\sum_{n=1}^{N}\big\{\ln p(X_{d,n}|\sum_{m=1}^{M} S_{d,m,n}) \\
&\qquad\qquad + \sum_{m=1}^{M}\langle \ln p(S_{d,m,n}|W_{d,m}H_{m,n})\rangle_{q(\mathbf{W})q(\mathbf{H})}\big\} + \text{const.}
\end{aligned}
$$

p.209 式（5.192）の 1 行目の添え字

【誤】

$$
\begin{aligned}
&\sum_{m=1}^{M}\sum_{k=1}^{K}\langle \ln \mathcal{N}(R_{n,m,k}|\sum_{d=1}^{D} U_{d,n}V_{d,m}S_{d,k}, \lambda^{-1})\rangle_{q(V_{d,m},S_{d,k},\lambda)} \\
&= -\frac{1}{2}\{\sum_{d=1}^{D}\sum_{d'=1}^{D} U_{d,n}U_{d',n}\langle\lambda\rangle \sum_{m=1}^{M}\sum_{k=1}^{K}\langle V_{d,m}V_{d',m}\rangle\langle S_{d,k}S_{d',k}\rangle \\
&\qquad - 2\sum_{d=1}^{D} U_{d,n}\langle\lambda\rangle \sum_{m=1}^{M}\sum_{k=1}^{K} R_{n,m,k}\langle V_{d,m}\rangle\langle S_{d,k}\rangle\} + \text{const.}
\end{aligned}
$$

【正】

$$
\begin{aligned}
&\sum_{m=1}^{M}\sum_{k=1}^{K}\langle \ln \mathcal{N}(R_{n,m,k}|\sum_{d=1}^{D} U_{d,n}V_{d,m}S_{d,k}, \lambda^{-1})\rangle_{q(V_{:,m},S_{:,k},\lambda)} \\
&= -\frac{1}{2}\{\sum_{d=1}^{D}\sum_{d'=1}^{D} U_{d,n}U_{d',n}\langle\lambda\rangle \sum_{m=1}^{M}\sum_{k=1}^{K}\langle V_{d,m}V_{d',m}\rangle\langle S_{d,k}S_{d',k}\rangle \\
&\qquad - 2\sum_{d=1}^{D} U_{d,n}\langle\lambda\rangle \sum_{m=1}^{M}\sum_{k=1}^{K} R_{n,m,k}\langle V_{d,m}\rangle\langle S_{d,k}\rangle\} + \text{const.}
\end{aligned}
$$