統計的機械学習
第十回　レポート

37-196360　森田涼介

2019 年 6 月 24 日

宿題 1

$q(\boldsymbol{u}_j)$ について次の比例式が成立する。

$$q(\boldsymbol{u}_j) \propto p(\boldsymbol{u}_j|\boldsymbol{\Lambda}_u)\exp\left\{\sum_{j,\ i\in\mathscr{O}}\int q(\boldsymbol{v}_i)\log p(r_{j,\ i}|\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i,\ \sigma^2)\,\mathrm{d}\boldsymbol{v}_i\right\} \tag{1}$$

ここで，$p(\boldsymbol{u}_j|\boldsymbol{\Lambda}_u)$ について，

$$p(\boldsymbol{u}_j|\boldsymbol{\Lambda}_u) \propto \exp\left(-\frac{1}{2}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{\Lambda}_u\boldsymbol{u}_j\right) \tag{2}$$

が成立する。また，

$$\begin{aligned}
l_{u,\ i,\ j} &\equiv \int q(\boldsymbol{v}_i)\log p(r_{j,\ i}|\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i,\ \sigma^2)\,\mathrm{d}\boldsymbol{v}_i && (3)\\
&= -\frac{1}{2\sigma^2}\int q(\boldsymbol{v}_i)(r_{j,\ i}-\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i)^2\,\mathrm{d}\boldsymbol{v}_i\\
&= -\frac{1}{2\sigma^2}\int q(\boldsymbol{v}_i)\left(r_{j,\ i}^2-2r_{j,\ i}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i+\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\right)\mathrm{d}\boldsymbol{v}_i\\
&= -\frac{1}{2\sigma^2}\left(r_{j,\ i}^2-2r_{j,\ i}\boldsymbol{u}_j^{\mathrm{T}}\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i]+\boldsymbol{u}_j^{\mathrm{T}}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]\boldsymbol{u}_j\right) && (4)
\end{aligned}$$

となる。式 (1)，(2)，(4) から，

$$\begin{aligned}
q(\boldsymbol{u}_j) &\propto \exp\left(-\frac{1}{2}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{\Lambda}_u\boldsymbol{u}_j\right)\exp\left\{\sum_{j,\ i\in\mathscr{O}} l_{u,\ i,\ j}\right\}\\
&= \exp\left(-\frac{1}{2}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{\Lambda}_u\boldsymbol{u}_j\right)\exp\left\{\sum_{j,\ i\in\mathscr{O}} -\frac{1}{2\sigma^2}\left(r_{j,\ i}^2-2r_{j,\ i}\boldsymbol{u}_j^{\mathrm{T}}\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i]+\boldsymbol{u}_j^{\mathrm{T}}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]\boldsymbol{u}_j\right)\right\}\\
&= \exp\left\{-\frac{1}{2\sigma^2}\left(-2\boldsymbol{u}_j\left(\sum_{j,\ i\in\mathscr{O}} r_{j,\ i}\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i]\right)+\boldsymbol{u}_j^{\mathrm{T}}\left(\sum_{j,\ i\in\mathscr{O}}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]+\sigma^2\boldsymbol{\Lambda}_u\right)\boldsymbol{u}_j\right)\right\} && (5)
\end{aligned}$$

いま，多次元正規分布に従う確率変数 $\boldsymbol{x}$ について次が成立する。

$$q(\boldsymbol{x}) = \mathscr{N}(\boldsymbol{x}|\boldsymbol{\mu},\ \boldsymbol{\Sigma}) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\boldsymbol{x}-2\boldsymbol{x}\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}\right)\right\} \tag{6}$$

これと式 (5) を比べると，

$$q(\boldsymbol{u}_j) = \mathscr{N}(\boldsymbol{u}_j|\boldsymbol{\mu}_{u,\ j},\ \boldsymbol{V}_{u,\ j}) \tag{7}$$

$$\boldsymbol{V}_{u,\ j}^{-1}\boldsymbol{\mu}_{u,\ j} = \frac{1}{\sigma^2}\sum_{j,\ i\in\mathscr{O}} r_{j,\ i}\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i] \tag{8}$$

$$\boldsymbol{V}_{u,\ j}^{-1} = \frac{1}{\sigma^2}\left(\sum_{j,\ i\in\mathscr{O}}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]+\sigma^2\boldsymbol{\Lambda}_u\right) \tag{9}$$

となることがわかる。これを整理すると，

$$q(\boldsymbol{u}_j) = \mathcal{N}(\boldsymbol{u}_j | \boldsymbol{\mu}_{u,\ j},\ \boldsymbol{V}_{u,\ j}) \tag{10}$$

$$\boldsymbol{\mu}_{u,\ j} = \frac{1}{\sigma^2} \boldsymbol{V}_{u,\ j} \sum_{j,\ i \in \mathcal{O}} r_{j,\ i} \mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i] \tag{11}$$

$$\boldsymbol{V}_{u,\ j} = \sigma^2 \left(\sum_{j,\ i \in \mathcal{O}} \mathbb{E}_{q(\boldsymbol{v}_i)} \left[\boldsymbol{v}_i \boldsymbol{v}_i^{\mathrm{T}} \right] + \sigma^2 \boldsymbol{\Lambda}_u \right)^{-1} \tag{12}$$

となる。

同様に，$\boldsymbol{v}_i$ について，

$$\begin{aligned} q(\boldsymbol{v}_i) &\propto p(\boldsymbol{v}_i | \boldsymbol{\Lambda}_v) \exp\left\{ \sum_{j,\ i \in \mathcal{O}} \int q(\boldsymbol{u}_j) \log p(r_{j,\ i} | \boldsymbol{u}_j^{\mathrm{T}} \boldsymbol{v}_i,\ \sigma^2)\, \mathrm{d}\boldsymbol{u}_j \right\} \qquad (13) \\ &= \exp\left(-\frac{1}{2} \boldsymbol{v}_i^{\mathrm{T}} \boldsymbol{\Lambda}_v \boldsymbol{v}_i \right) \exp\left\{ \sum_{j,\ i \in \mathcal{O}} -\frac{1}{2\sigma^2} \left(r_{j,\ i}^2 - 2 r_{j,\ i} \boldsymbol{v}_i^{\mathrm{T}} \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] + \boldsymbol{v}_i^{\mathrm{T}} \mathbb{E}_{q(\boldsymbol{u}_j)} \left[\boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}} \right] \boldsymbol{v}_i \right) \right\} \\ &= \exp\left\{ -\frac{1}{2\sigma^2} \left(-2 \boldsymbol{v}_i \left(\sum_{j,\ i \in \mathcal{O}} r_{j,\ i} \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] \right) + \boldsymbol{v}_i^{\mathrm{T}} \left(\sum_{j,\ i \in \mathcal{O}} \mathbb{E}_{q(\boldsymbol{u}_j)} \left[\boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}} \right] + \sigma^2 \boldsymbol{\Lambda}_v \right) \boldsymbol{v}_i \right) \right\} \qquad (14) \end{aligned}$$

となることから，

$$q(\boldsymbol{v}_i) = \mathcal{N}(\boldsymbol{v}_i | \boldsymbol{\mu}_{v,\ i},\ \boldsymbol{V}_{v,\ i}) \tag{15}$$

$$\boldsymbol{V}_{v,\ i}^{-1} \boldsymbol{\mu}_{v,\ i} = \frac{1}{\sigma^2} \sum_{j,\ i \in \mathcal{O}} r_{j,\ i} \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] \tag{16}$$

$$\boldsymbol{V}_{v,\ i}^{-1} = \frac{1}{\sigma^2} \left(\sum_{j,\ i \in \mathcal{O}} \mathbb{E}_{q(\boldsymbol{u}_j)} \left[\boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}} \right] + \sigma^2 \boldsymbol{\Lambda}_v \right) \tag{17}$$

を得て，結局，

$$q(\boldsymbol{v}_i) = \mathcal{N}(\boldsymbol{v}_i | \boldsymbol{\mu}_{v,\ i},\ \boldsymbol{V}_{v,\ i}) \tag{18}$$

$$\boldsymbol{\mu}_{v,\ i} = \frac{1}{\sigma^2} \boldsymbol{V}_{v,\ i} \sum_{j,\ i \in \mathcal{O}} r_{j,\ i} \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] \tag{19}$$

$$\boldsymbol{V}_{v,\ i} = \sigma^2 \left(\sum_{j,\ i \in \mathcal{O}} \mathbb{E}_{q(\boldsymbol{u}_j)} \left[\boldsymbol{u}_j \boldsymbol{u}_j^{\mathrm{T}} \right] + \sigma^2 \boldsymbol{\Lambda}_v \right)^{-1} \tag{20}$$

を得る。

宿題 2

いま，$\boldsymbol{\mu}_{u,\ j}$ の期待値と分散について，

$$\boldsymbol{\mu}_{u,\ j} = \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] \tag{21}$$

$$\boldsymbol{V}_{u,\ j} = \mathbb{E}_{q(\boldsymbol{u}_j)}\left[(\boldsymbol{u}_j - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j])(\boldsymbol{u}_j - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j])^{\mathrm{T}}\right] \tag{22}$$

$$= \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j]\mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j]^{\mathrm{T}} \tag{23}$$

$$= \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] - \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}} \tag{24}$$

が成立する。これより，

$$\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] = \boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}} \tag{25}$$

となる。ここで，

$$\boldsymbol{u} \sim \mathcal{N}(\mathbf{0},\ \boldsymbol{\Lambda}_u) \tag{26}$$

の分散を標本 $\left\{\boldsymbol{u}_j\right\}_{j=1}^{J}$ を用いて表すことを考えると，平均 $\mathbf{0}$ から，

$$\boldsymbol{\Lambda}_u = \mathbb{E}_{q(\boldsymbol{u})}\left[\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\right] \tag{27}$$

$$= \frac{1}{J}\sum_{j=1}^{J}\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] \tag{28}$$

$$= \frac{1}{J}\sum_{j=1}^{J}\left(\boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}}\right) \tag{29}$$

$\boldsymbol{\Lambda}_u$ は対角成分しか持たないため，結局，

$$\rho_{u,\ k}^2 = \frac{1}{J}\sum_{j=1}^{J}\left([\boldsymbol{V}_{u,\ j}]_{k,\ k} + \mu_{u,\ j,\ k}^2\right) \tag{30}$$

同様に，$\boldsymbol{v}_i$ の期待値と分散について次式が成立し，

$$\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right] = \boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}} \tag{31}$$

また，

$$\boldsymbol{v} \sim \mathcal{N}(\mathbf{0},\ \boldsymbol{\Lambda}_v) \tag{32}$$

から，

$$\boldsymbol{\Lambda}_v = \mathbb{E}_{q(\boldsymbol{v})}\left[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\right] \tag{33}$$

$$= \frac{1}{I}\sum_{i=1}^{I}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right] \tag{34}$$

$$= \frac{1}{I}\sum_{i=1}^{I}\left(\boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}}\right) \tag{35}$$

$\mathbf{\Lambda}_v$ は対角成分しか持たないため，結局，

$$\rho_{v,\,k}^2 = \frac{1}{I}\sum_{i=1}^{I}\left([\boldsymbol{V}_{v,\,i}]_{k,\,k} + \mu_{v,\,i,\,k}^2\right) \tag{36}$$

最後に，r について考える。

$$r \sim \mathcal{N}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v},\, \sigma^2) \tag{37}$$

から，

$$\begin{aligned}
\sigma^2 &= \mathbb{E}_{q(\boldsymbol{u})q(\boldsymbol{v})}\left[(r - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v})^2\right] && (38)\\
&= \frac{1}{|\mathscr{O}|}\sum_{(j,\,i)\in\mathscr{O}} \mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\left(r_{j,\,i} - \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\right)^2\right] && (39)\\
&= \frac{1}{|\mathscr{O}|}\sum_{(j,\,i)\in\mathscr{O}} \mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[r_{j,\,i}^2 - 2r_{j,\,i}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i + \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\right] && (40)\\
& && (41)
\end{aligned}$$

いま，

$$\begin{aligned}
&\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[r_{j,\,i}^2\right] = r_{j,\,i}^2 && (42)\\
&\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\right] = \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j^{\mathrm{T}}\right]\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i] = \boldsymbol{\mu}_{u,\,i}^{\mathrm{T}}\boldsymbol{v}_{v,\,i} && (43)
\end{aligned}$$

であり，また，

$$\begin{aligned}
\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\right] &= \mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\mathrm{Tr}\left\{\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right\}\right] && (44)\\
&= \mathrm{Tr}\left\{\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right]\right\} && (45)\\
&= \mathrm{Tr}\left\{\left(\boldsymbol{V}_{u,\,j} + \boldsymbol{\mu}_{u,\,j}\boldsymbol{\mu}_{u,\,j}^{\mathrm{T}}\right)\left(\boldsymbol{V}_{v,\,i} + \boldsymbol{\mu}_{v,\,i}\boldsymbol{\mu}_{v,\,i}^{\mathrm{T}}\right)\right\} && (46)
\end{aligned}$$

となることから，結局，

$$\sigma^2 = \frac{1}{|\mathscr{O}|}\sum_{(j,\,i)\in\mathscr{O}}\left\{r_{j,\,i}^2 - 2r_{j,\,i}\boldsymbol{\mu}_{u,\,j}^{\mathrm{T}}\boldsymbol{\mu}_{v,\,i} + \mathrm{Tr}\left\{\left(\boldsymbol{V}_{u,\,j} + \boldsymbol{\mu}_{u,\,j}\boldsymbol{\mu}_{u,\,j}^{\mathrm{T}}\right)\left(\boldsymbol{V}_{v,\,i} + \boldsymbol{\mu}_{v,\,i}\boldsymbol{\mu}_{v,\,i}^{\mathrm{T}}\right)\right\}\right\} \tag{47}$$

となる。