宿題 2

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

$$
\begin{aligned}
\boldsymbol{\mu}_{u,\ j} &= \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j] && (1)\\
\boldsymbol{V}_{u,\ j} &= \mathbb{E}_{q(\boldsymbol{u}_j)}\left[(\boldsymbol{u}_j - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j])(\boldsymbol{u}_j - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j])^{\mathrm{T}}\right] && (2)\\
&= \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] - \mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j]\mathbb{E}_{q(\boldsymbol{u}_j)}[\boldsymbol{u}_j]^{\mathrm{T}} && (3)\\
&= \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] - \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}} && (4)
\end{aligned}
$$

が成立する。これより，

$$
\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] = \boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}} \tag{5}
$$

となる。ここで，

$$
\boldsymbol{u} \sim \mathcal{N}(\mathbf{0},\ \boldsymbol{\Lambda}_u) \tag{6}
$$

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，

$$
\begin{aligned}
\boldsymbol{\Lambda}_u &= \mathbb{E}_{q(\boldsymbol{u})}\left[\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\right] && (7)\\
&= \frac{1}{J}\sum_{j=1}^{J}\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right] && (8)\\
&= \frac{1}{J}\sum_{j=1}^{J}\left(\boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}}\right) && (9)
\end{aligned}
$$

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

$$
\rho_{u,\ k}^2 = \frac{1}{J}\sum_{j=1}^{J}\left([\boldsymbol{V}_{u,\ j}]_{k,\ k} + \mu_{u,\ j,\ k}^2\right) \tag{10}
$$

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

$$
\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right] = \boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}} \tag{11}
$$

また，

$$
\boldsymbol{v} \sim \mathcal{N}(\mathbf{0},\ \boldsymbol{\Lambda}_v) \tag{12}
$$

から，

$$
\begin{aligned}
\boldsymbol{\Lambda}_v &= \mathbb{E}_{q(\boldsymbol{v})}\left[\boldsymbol{v}\boldsymbol{v}^{\mathrm{T}}\right] && (13)\\
&= \frac{1}{I}\sum_{i=1}^{I}\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right] && (14)\\
&= \frac{1}{I}\sum_{i=1}^{I}\left(\boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}}\right) && (15)
\end{aligned}
$$

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

$$\rho_{v,\ k}^2 = \frac{1}{I}\sum_{i=1}^{I}\left([\boldsymbol{V}_{v,\ i}]_{k,\ k} + \mu_{v,\ i,\ k}^2\right) \tag{16}$$

最後に，r について考える。

$$r \sim \mathcal{N}(\boldsymbol{u}^{\mathrm{T}}\boldsymbol{v},\ \sigma^2) \tag{17}$$

から，

$$\begin{aligned}
\sigma^2 &= \mathbb{E}_{q(\boldsymbol{u})q(\boldsymbol{v})}\left[(r - \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v})^2\right] && (18)\\
&= \frac{1}{|\mathcal{O}|}\sum_{(j,\ i)\in\mathcal{O}}\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\left(r_{j,\ i} - \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\right)^2\right] && (19)\\
&= \frac{1}{|\mathcal{O}|}\sum_{(j,\ i)\in\mathcal{O}}\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[r_{j,\ i}^2 - 2r_{j,\ i}\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i + \boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\right] && (20)\\
& && (21)
\end{aligned}$$

いま，

$$\begin{aligned}
&\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[r_{j,\ i}^2\right] = r_{j,\ i}^2 && (22)\\
&\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\right] = \mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j^{\mathrm{T}}\right]\mathbb{E}_{q(\boldsymbol{v}_i)}[\boldsymbol{v}_i] = \boldsymbol{\mu}_{u,\ i}^{\mathrm{T}}\boldsymbol{v}_{v,\ i} && (23)
\end{aligned}$$

であり，また，

$$\begin{aligned}
\mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\right] &= \mathbb{E}_{q(\boldsymbol{u}_j)q(\boldsymbol{v}_i)}\left[\mathrm{Tr}\left\{\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right\}\right] && (24)\\
&= \mathrm{Tr}\left\{\mathbb{E}_{q(\boldsymbol{v}_i)}\left[\boldsymbol{v}_i\boldsymbol{v}_i^{\mathrm{T}}\right]\mathbb{E}_{q(\boldsymbol{u}_j)}\left[\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}\right]\right\} && (25)\\
&= \mathrm{Tr}\left\{\left(\boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}}\right)\left(\boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}}\right)\right\} && (26)
\end{aligned}$$

となることから，結局，

$$\sigma^2 = \frac{1}{|\mathcal{O}|}\sum_{(j,\ i)\in\mathcal{O}}\left\{r_{j,\ i}^2 - 2r_{j,\ i}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}}\boldsymbol{\mu}_{v,\ i} + \mathrm{Tr}\left\{\left(\boldsymbol{V}_{u,\ j} + \boldsymbol{\mu}_{u,\ j}\boldsymbol{\mu}_{u,\ j}^{\mathrm{T}}\right)\left(\boldsymbol{V}_{v,\ i} + \boldsymbol{\mu}_{v,\ i}\boldsymbol{\mu}_{v,\ i}^{\mathrm{T}}\right)\right\}\right\} \tag{27}$$

となる。