宿題 2

次式を導出する。ここで，$\boldsymbol{Z} \in \mathbb{R}^d$, $\boldsymbol{\Sigma} \in \mathbb{R}^{d \times d}$, $\mu \in \mathbb{R}$ である。

$$\frac{\partial}{\partial \boldsymbol{Z}} \left\{ \boldsymbol{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \boldsymbol{Z} - \mu (\boldsymbol{Z}^{\mathrm{T}} \boldsymbol{Z} - 1) \right\} = 2\boldsymbol{\Sigma}^{-1} \boldsymbol{Z} - 2\mu \boldsymbol{Z} \tag{1}$$

$\boldsymbol{Z}$, $\boldsymbol{\Sigma}$ を次のようにおく。

$$\boldsymbol{Z} = \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} \tag{2}$$

$$\boldsymbol{\Sigma}^{-1} = \boldsymbol{\Lambda} = \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1d} \\ \vdots & \ddots & \vdots \\ \lambda_{d1} & \cdots & \lambda dd \end{bmatrix} \tag{3}$$

このとき，

$$\begin{aligned} \boldsymbol{Z}^{\mathrm{T}} \boldsymbol{Z} &= \begin{bmatrix} z_1 & \cdots & z_d \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} \\ &= {z_1}^2 + \cdots + {z_d}^2 \\ &= \sum_{i=1}^{d} {z_i}^2 \end{aligned} \tag{4}$$

$$\begin{aligned} \boldsymbol{\Sigma}^{-1} \boldsymbol{Z} &= \boldsymbol{\Lambda} \boldsymbol{Z} \\ &= \begin{bmatrix} \lambda_{11} & \cdots & \lambda_{1d} \\ \vdots & \ddots & \vdots \\ \lambda_{d1} & \cdots & \lambda_{dd} \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_d \end{bmatrix} \\ &= \begin{bmatrix} \lambda_{11} z_1 + \cdots + \lambda_{1d} z_d \\ \vdots \\ \lambda_{d1} z_1 + \cdots + \lambda_{dd} z_d \end{bmatrix} \\ &= \begin{bmatrix} \sum_{j=1}^{d} \lambda_{1j} z_j \\ \vdots \\ \sum_{j=1}^{d} \lambda_{dj} z_j \end{bmatrix} \end{aligned} \tag{5}$$

$$
\begin{aligned}
\mathbf{Z}^{\mathrm{T}}\boldsymbol{\Sigma}^{-1}\mathbf{Z} &= \mathbf{Z}^{\mathrm{T}}\boldsymbol{\Lambda}\mathbf{Z} \\
&= \begin{bmatrix} z_1 & \cdots & z_d \end{bmatrix} \begin{bmatrix} \sum_{j=1}^{d} \lambda_{1j} z_j \\ \vdots \\ \sum_{j=1}^{d} \lambda_{dj} z_j \end{bmatrix} \\
&= z_1 \sum_{j=1}^{d} \lambda_{1j} z_j + \cdots + z_d \sum_{j=1}^{d} \lambda_{dj} z_j \\
&= \sum_{i=1}^{d} z_i \sum_{j=1}^{d} \lambda_{ij} z_j \qquad (6) \\
&= \sum_{i=1}^{d} \sum_{j=1}^{d} \lambda_{ij} z_i z_j \qquad (7) \\
&= \sum_{i=1}^{d} \lambda_{ii} {z_i}^2 + \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \lambda_{ij} z_i z_j + \sum_{j=1}^{d-1} \sum_{i=j+1}^{d} \lambda_{ij} z_i z_j \qquad (8) \\
&= \sum_{i=1}^{d} \lambda_{ii} {z_i}^2 + 2 \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \lambda_{ij} z_i z_j \qquad (9)
\end{aligned}
$$

である。ここで，$\boldsymbol{\Sigma}$ は対称行列であるから，$\boldsymbol{\Lambda} = \boldsymbol{\Sigma}^{-1}$ も対称行列であり，$\lambda_{ij} = \lambda_{ji}$ となることを用いた。このとき，

$$
\frac{\partial}{\partial z_k} \left(\mathbf{Z}^{\mathrm{T}} \mathbf{Z} \right) = \frac{\partial}{\partial z_k} \left(\sum_{i=1}^{d} {z_k}^2 \right) = 2 z_k \qquad (10)
$$

$$
\begin{aligned}
\frac{\partial}{\partial z_k} \left(\mathbf{Z}^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} \mathbf{Z} \right) &= \frac{\partial}{\partial z_k} \left(\sum_{i=1}^{d} \sum_{j=1}^{d} \lambda_{ij} z_i z_j \right) \qquad (11) \\
&= \frac{\partial}{\partial z_k} \left(\sum_{i=1}^{d} \lambda_{ii} {z_i}^2 + 2 \sum_{i=1}^{d-1} \sum_{j=i+1}^{d} \lambda_{ij} z_i z_j \right) \qquad (12) \\
&= 2 \lambda_{kk} z_k + 2 \left(\sum_{j=k+1}^{d} \lambda_{kj} z_j + \sum_{i=1}^{k-1} \lambda_{ik} z_i \right) \qquad (13) \\
&= 2 \sum_{l=1}^{d} \lambda_{kl} z_l \qquad (14)
\end{aligned}
$$

であるから，

$$
\frac{\partial}{\partial \mathbf{Z}} \left(\mathbf{Z}^{\mathrm{T}} \mathbf{Z} \right) = \begin{bmatrix} \frac{\partial}{\partial z_1} \left(\mathbf{Z}^{\mathrm{T}} \mathbf{Z} \right) \\ \vdots \\ \frac{\partial}{\partial z_d} \left(\mathbf{Z}^{\mathrm{T}} \mathbf{Z} \right) \end{bmatrix} = \begin{bmatrix} 2z_1 \\ \vdots \\ 2z_d \end{bmatrix} = 2\mathbf{Z} \qquad (15)
$$

$$
\frac{\partial}{\partial \mathbf{Z}}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{Z}\right) = \begin{bmatrix} \frac{\partial}{\partial z_1}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{Z}\right) \\ \vdots \\ \frac{\partial}{\partial z_d}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{Z}\right) \end{bmatrix} \tag{16}
$$

$$
= \begin{bmatrix} 2\sum_{l=1}^{d} \lambda_{1l} z_l \\ \vdots \\ 2\sum_{l=1}^{d} \lambda_{dl} z_l \end{bmatrix} \tag{17}
$$

$$
= 2\mathbf{\Sigma}^{-1}\mathbf{Z} \tag{18}
$$

となる。また，当然 $\partial\mu \big/ \partial\mathbf{Z} = \mathbf{0}$ である。以上より，

$$
\frac{\partial}{\partial \mathbf{Z}}\left\{\mathbf{Z}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{Z} - \mu(\mathbf{Z}^{\mathrm{T}}\mathbf{Z} - 1)\right\} = \frac{\partial}{\partial \mathbf{Z}}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{Z}\right) - \mu\frac{\partial}{\partial \mathbf{Z}}\left(\mathbf{Z}^{\mathrm{T}}\mathbf{Z}\right) \tag{19}
$$

$$
= 2\mathbf{\Sigma}^{-1}\mathbf{Z} - 2\mu\mathbf{Z} \tag{20}
$$