Machine Learning Worksheet 4

Tomas Ladek, Michael Kratzer
3602673, 3612903
tom.ladek@tum.de, mkratzer@mytum.de

Problem 1

The Inverse of a diagonal atrix is just a diagonal matrix with each of the elements inverted.

$$\Lambda^{-1} = \begin{pmatrix} \frac{1}{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_d} \end{pmatrix}, U = \begin{pmatrix} u_{1,1} & \dots & u_{1,d} \\ \vdots & \ddots & \vdots \\ u_{d,1} & \dots & u_{d,d} \end{pmatrix} \tag{1}$$

Just multiplying $U\Lambda U^T$ results in what should be proven.

$$U\Lambda U^T = \begin{pmatrix} u_{1,1} & \dots & u_{1,d} \\ \vdots & \ddots & \vdots \\ u_{d,1} & \dots & u_{d,d} \end{pmatrix} \begin{pmatrix} \frac{1}{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \frac{1}{\lambda_d} \end{pmatrix} \begin{pmatrix} u_{1,1} & \dots & u_{d,1} \\ \vdots & \ddots & \vdots \\ u_{1,d} & \dots & u_{d,d} \end{pmatrix} \tag{2}$$

$$= \begin{pmatrix} \frac{u_{1,1}}{\lambda_1} & \dots & \frac{u_{1,d}}{\lambda_d} \\ \vdots & \ddots & \vdots \\ \frac{u_{d,1}}{\lambda_1} & \dots & \frac{u_{d,d}}{\lambda_d} \end{pmatrix} \begin{pmatrix} u_{1,1} & \dots & u_{d,1} \\ \vdots & \ddots & \vdots \\ u_{1,d} & \dots & u_{d,d} \end{pmatrix} \tag{3}$$

$$= \begin{pmatrix} \frac{u_{1,1}u_{1,1}}{\lambda_1} + \dots + \frac{u_{1,d}u_{1,d}}{\lambda_d} & \dots & \frac{u_{1,1}u_{d,1}}{\lambda_1} + \dots + \frac{u_{1,d}u_{d,d}}{\lambda_d} \\ \vdots & \ddots & \vdots \\ \frac{u_{d,1}u_{1,1}}{\lambda_1} + \dots + \frac{u_{1,d}u_{d,d}}{\lambda_d} & \dots & \frac{u_{d,1}u_{d,1}}{\lambda_1} + \dots + \frac{u_{d,d}u_{d,d}}{\lambda_d} \end{pmatrix} \tag{4}$$

$$= \frac{1}{\lambda_1} \begin{pmatrix} u_{1,1}u_{1,1} & \dots & u_{1,1}u_{d,1} \\ \vdots & \ddots & \vdots \\ u_{d,1}u_{1,1} & \dots & u_{d,1}u_{d,1} \end{pmatrix} + \dots + \frac{1}{\lambda_d} \begin{pmatrix} u_{1,d}u_{1,d} & \dots & u_{1,d}u_{d,d} \\ \vdots & \ddots & \vdots \\ u_{1,d}u_{d,d} & \dots & u_{d,d}u_{d,d} \end{pmatrix} \tag{5}$$

$$= \frac{1}{\lambda_1} u_1 u_1^T + \dots + \frac{1}{\lambda_d} u_d u_d^T \tag{6}$$

$$= \sum_{i=1}^{d} \frac{1}{\lambda_i} u_i u_i^T \tag{7}$$

Problem 2

$L = \mathbb{R}^{nxn}$ and is invertible and $Y = LX$ with $X \sim \mathcal{N}(\mu_X, \Sigma_X)$ then $X = L^{-1}Y$. So

$$p_Y(Y) = p_X(X)|det(\frac{\partial x}{\partial y})| = p_X(X)|det(J_{Y\to X})| = p_X(L^{-1}Y)|det(J_{Y\to X})| \tag{8}$$

with

$$J_{Y \to X} = \frac{\partial x}{\partial y} = \frac{\partial L^{-1}Y}{\partial Y} = L^{-1} \tag{9}$$

Then

$$\begin{aligned}
p_y(Y) &= p_x(L^{-1}Y)|det(L^{-1}) && (10)\\
&= p_x(L^{-1}Y)\frac{1}{|det(L)|} && (11)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(\Sigma_X|)}}exp(-\frac{1}{2}(L^{-1}Y-\mu_X)^T\Sigma_X^{-1}(L^{-1}Y-\mu_X))\frac{1}{|det(L)|} && (12)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(\Sigma_X|)}|det(L)|}exp(-\frac{1}{2}(L^{-1}Y-\mu_X)^T\Sigma_X^{-1}(L^{-1}Y-\mu_X)) && (13)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(\Sigma_X)||det(L)|^2}}exp(-\frac{1}{2}(L^{-1}Y-\mu_X)^T\Sigma_X^{-1}(L^{-1}Y-\mu_X)) && (14)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(\Sigma_X)||det(LL^T)|}}exp(-\frac{1}{2}(L^{-1}Y-\mu_X)^T\Sigma_X^{-1}(L^{-1}Y-\mu_X)) && (15)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(L\Sigma_X L^T)|}}exp(-\frac{1}{2}(L^{-1}Y-\mu_X)^T\Sigma_X^{-1}(L^{-1}Y-\mu_X)) && (16)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(L\Sigma_X L^T)|}}exp(-\frac{1}{2}(Y-L\mu_X)^T\Sigma_X^{-1}(Y-L\mu_X) && (17)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(L\Sigma_X L^T)|}}exp(-\frac{1}{2}L^{-1}(Y-L\mu_X)^T\Sigma_X^{-1}L^{-1}(Y-L\mu_X) && (18)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(L\Sigma_X L^T)|}}exp(-\frac{1}{2}(Y-L\mu_X)^T L^{-T}\Sigma_X^{-1}L^{-1}(Y-L\mu_X) && (19)\\
&= \frac{1}{\sqrt{(2\pi)^d|det(L\Sigma_X L^T)|}}exp(-\frac{1}{2}(Y-L\mu_X)^T(L\Sigma_X L^T)^{-1}(Y-L\mu_X) && (20)\\
Y &\sim \mathcal{N}(L\mu_X, L\Sigma_X L^T) && (21)\\
Y &\sim \mathcal{N}(\mu_Y, \Sigma_Y) && (22)
\end{aligned}$$

Problem 3

With $p(x) = \mathcal{N}(0, \Sigma_X)$ and $Y = TX + Z$ with $Z \sim \mathcal{N}(0, \Sigma_{Y|X})$ then

$$\begin{aligned}
\mu_Y &= E[TX+Z] = TE[X] + E[Z] = TE[X] && (23)\\
\Sigma_Y &= T\Sigma_X T^T + \Sigma_Z = T\Sigma_X T^T + \Sigma_{Y|X} && (24)\\
&= T\Sigma_X T^T + \Sigma_Y - \Sigma_{Y,X}\Sigma_X^{-1}\Sigma_{X,Y} && (25)
\end{aligned}$$

This means for T it must hold that

$$T\Sigma_X T^T = \Sigma_{Y,X}\Sigma_X^{-1}\Sigma_{X,Y} \tag{26}$$

Problem 4

HW4