Homework 2: Gaussian Distribution

Trần Minh Khoa

September 21, 2021

1. Multivariate Gaussian Distribution

$$p(x|\mu,\sigma^2) = \frac{1}{(2\pi)^{\frac{D}{2}}|\Sigma|^{\frac{1}{2}}}\ e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

$$\triangle^2 = (x-\mu)^T\Sigma^{-1}(x-\mu)$$

prove SVD
Λ is a diagonal matrix and U is an orthogonal matrix
ΣU = UΛ. Since U is orthogonal, we have $U^{-1} = U^T$,
so $\Sigma = U\Lambda U^T$
Then we have:
$\Sigma x = U\Lambda U^T x = U\Lambda \begin{bmatrix} u_1^T x \\ . \\ . \\ u_n^T x \end{bmatrix} = U \begin{bmatrix} \lambda_1 u_1^T x \\ . \\ . \\ \lambda_n u_n^T x \end{bmatrix} = \sum_{i=1}^{D}(\lambda_i u_i^T x)u_i =$
$\sum_{i=1}^{D} \lambda_i u_i u_i^T x = (\sum_{i=1}^{D} \lambda_i u_i u_i^T)x$

$$\Rightarrow \Sigma = \sum_{i=1}^{D} \lambda_i u_i u_i^T$$

Since $\Sigma U = U\Lambda$, inverting both sides gives $U^T\Sigma^{-1} = \Lambda^{-1}U^T$, and hence $\Sigma^{-1} = U\Lambda^{-1}U^T$. Applying the above result to Σ^{-1}, noting that Λ^{-1} is just the diagonal matrix of the inverses of the diagonal elements of Λ, we have

$$\Sigma^{-1} = \sum_{i=1}^{D} \frac{1}{\lambda_i} u_i u_i^T$$

$$\Rightarrow \triangle^2 = \sum_{i=1}^{D} \frac{1}{\lambda_i}(x-\mu)^T u_i u_i^T (x-\mu)$$

$$= \sum_{i=1}^{D} \frac{y_i^2}{\lambda_i} \qquad , with\ y_i = u_i^T(x-\mu)$$

$$|\Sigma|^{1/2} = \prod_{j=1}^{D} \lambda_j^{1/2}$$

$$\Rightarrow p(y) = \prod_{j=1}^{D} {\frac{1}{2\pi\lambda_j}}^{1/2} e^{-\frac{y_j^2}{2\lambda_j}}$$

$$\Rightarrow \int_{-\infty}^{\infty} p(y)dy = \prod_{j=1}^{D} \int_{-\infty}^{\infty} {\frac{1}{2\pi\lambda_j}}^{1/2} e^{-\frac{y_j^2}{2\lambda_j}} dy_j \ = 1$$

2. Conditional Gaussian Distribution

x = $\begin{pmatrix} x_a \\ x_b \end{pmatrix}$; μ= $\begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$

Σ= $\begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix} \Rightarrow A = \Sigma^{-1} = \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix}$

$$-\tfrac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu) = (x-\mu)^T A(x-\mu)$$

$$= -\tfrac{1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \tfrac{1}{2}(x_a-\mu_a)^T A_{ab}(x_b-\mu_b)$$

$$-\tfrac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a) - \tfrac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b)$$

$$= -\tfrac{1}{2}x_a^T A_{aa} x_a + x_a^T(A_{aa}\mu_a - A_{ab}(x_b-\mu_b)) + const$$

$\triangle^2 \ = -\tfrac{1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu + const$

$\Rightarrow \Sigma_{a|b} = A_{aa}^{-1}$

$\Sigma_{a|b}^{-1}\ \mu_{a|b} = A_{aa}\mu_a - A_{ab}(x_b-\mu_b)$

$\Rightarrow \mu_{a|b} = \Sigma_{a|b}(A_{aa}\mu_a - A_{ab}(x_b - \mu_b))$

$= A_{aa}^{-1}(A_{aa}\mu_a - A_{ab}(x_b - \mu_b)) = \mu_a - A_{aa}^{-1}A_{ab}(x_b - \mu_b)$

By using Schur complement, we have:

$A_{aa} = (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}$

$A_{ab} = -(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1}$

$\Rightarrow \Sigma_{a|b} = \Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba}$

$\mu_{a|b} = \mu_a - (\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})(-(\Sigma_{aa} - \Sigma_{ab}\Sigma_{bb}^{-1}\Sigma_{ba})^{-1}\Sigma_{ab}\Sigma_{bb}^{-1})(x_b - \mu_b) = \mu_a + \Sigma_{ab}\Sigma_{bb}^{-1}(x_b - \mu_b)$