Multivariate Gaussian distribution

You

Ngày 22 tháng 9 năm 2021

1 $Homework_2$

1.1 Ex1: Proof p(x) is normal.

$$p(x|\mu,\sigma^2) = \frac{1}{(2\pi)^{\frac{D}{2}}|\sum|^{\frac{1}{2}}}.exp\left\{\frac{-1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right\}$$

1.2 Proof

Đặt

$$\Delta^2 = \frac{-1}{2}(x-\mu)^T.\Sigma^{-1}.(x-\mu)$$
$$= \frac{-1}{2}(x^T-\mu^T).\Sigma^{-1}.(x-\mu)$$
$$= \frac{-1}{2}x^T\Sigma^{-1}x + \frac{1}{2}x^T\Sigma^{-1}\mu + \frac{1}{2}\mu^T\Sigma^{-1}x - \frac{1}{2}\mu^T\Sigma^{-1}\mu$$

Trong đó: *

$$\frac{1}{2}\mu^T\Sigma^{-1}\mu = constant$$

vì (1*D).(D*D).(D*1) = Constant
*

$$\frac{1}{2}x^T\Sigma^{-1}\mu + \frac{1}{2}\mu^T\Sigma^{-1}x$$

Ta có: $(\frac{1}{2}x^T\Sigma^{-1}\mu)^T = \frac{1}{2}\mu^T(\Sigma^{-1})^Tx$ (since $(ABC)^T = C^TB^TA^T$)
mà + Σ đối xứng nên $\Sigma^{-1} = (\Sigma^{-1})^T = \Sigma$ + $\frac{1}{2}x^T\Sigma^{-1}\mu = constant$ (do (1*D).(D*D).(D*1)) Vậy $\Delta = \frac{-1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu + constant$
Ta có: Σ đối xứng nên:

$$\Sigma = \sum_{i=1}^{D}\lambda_iu_iu_i^T -> \Sigma^{-1} = \sum_{i=1}^{D}\frac{1}{\lambda_i}u_iu_i^T$$

Proof: vì u_i là các eigenvecto nên theo tính chất $u^T.u = 0$ -> $\sum_{i=1}^{D}u_i.u_i^T = I$ and: Σ đối xứng nên các eigenvalues sẽ là số thực. and: $\Sigma^T = \Sigma^{-1}$

$$\Delta^2 = \frac{-1}{2}(x-\mu)^T.\Sigma^{-1}.(x-\mu)$$
$$(1) = \frac{-1}{2}(x-\mu)^T\sum_{i=1}^{D}\frac{1}{\lambda_i}u_iu_i^T(x-\mu)$$

Set: $y_i = u_i^T(x-\mu)$ Since y_i có giá trị là một số thực (do (1*D)(D*1) nên $y_i = y_i^T = (u_i^T(x-\mu))^T = (x-\mu)^Tu_i$

$$(1) <=> \Delta^2 = \frac{-1}{2}\sum_{i=1}^{D}\frac{1}{\lambda_i}(x-\mu)^Tu_iu_i^T(x-\mu)$$

$$= \sum_{i=1}^{D} \frac{-y_i^2}{2\lambda_i}$$

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

(trong đó λ_j là cac eigenvalues va det(A) bằng tich cac eigenvalues)

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

$$= \prod_{j=1}^{D} (\frac{1}{2\pi\lambda_j})^{\frac{1}{2}} e^{\sum_{j=1}^{D} \frac{-y_j^2}{2\lambda_j}}$$

$$= \prod_{j=1}^{D} (\frac{1}{2\pi\lambda_j})^{\frac{1}{2}} . \prod_{j=1}^{D} e^{\frac{-y_j^2}{3\lambda_i}}$$

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

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

+) pick j = 1 and set $\lambda = \sigma^2$-> $\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} .e^{\frac{-y^2}{2\sigma^2}} = 1$. Thus:

$$\int_{-\infty}^{\infty} p(y)dy = 1 => Normality$$

1.3 Conditional Gaussion distribution

Set $x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$ and $\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$ and $\Sigma = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}$ Đặt: matrix $A = \Sigma^{-1} = \begin{bmatrix} ccA_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{bmatrix}$
Since Σ đối xứng nên Σ_{aa} và Σ_{bb} cũng đối xứng Tìm $p(x_a|x_b)$ Ta có:

$$\Delta^2 = \frac{-1}{2}(x-\mu)^T.\Sigma^{-1}.(x-\mu)$$

$$= \frac{-1}{2}. \begin{pmatrix} x_a - \mu_a \\ x_b - \mu_b \end{pmatrix}^T . \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix} . \begin{pmatrix} x_a - \mu_a \\ x_b - \mu_b \end{pmatrix}$$

Set: $(x_a - \mu_a)^T = C_1$ $(x_b - \mu_b)^T = C_2$ $(x_a - \mu_a) = D_1$ $(x_b - \mu_b) = D_2$ thay vào Δ^2 có:

$$= \frac{-1}{2}. \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}^T . \begin{pmatrix} A_{aa} & A_{ab} \\ A_{ba} & A_{bb} \end{pmatrix} . \begin{pmatrix} D_1 \\ D_2 \end{pmatrix}$$

$$= \frac{-1}{2}(\begin{pmatrix} C_1A_{aa} + C_2A_{ba} & C_1A_{ab} + C_2A_{bb} \end{pmatrix} . \begin{pmatrix} D_1 \\ D_2 \end{pmatrix})$$

$$= \frac{-1}{2} [(C_1A_{aa} + C_2A_{ba}).D_1 + (C_1A_{ab} + C_2A_{bb}).D_2]$$

$$= \frac{-1}{2}(C_1A_{aa}D_1 + C_2A_{ba}D_1 + C_1A_{ab}D_2 + C_2A_{bb}D_2)$$

$$\Delta^2 = \frac{-1}{2}(x_a-\mu_a)^T A_{aa}(x_a-\mu_a) - \frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a) - \frac{1}{2}.(x_a-\mu_a)^T A_{ab}(x_b-\mu_b) - \frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b)$$

trong đó; $-\frac{1}{2}(x_b-\mu_b)^T A_{bb}(x_b-\mu_b) = constant$ do (1*b).(b*b).(b*1) thuộc R $\frac{1}{2}(x_b-\mu_b)^T A_{ba}(x_a-\mu_a) = \frac{1}{2}.(x_a-\mu_a)^T A_{ab}(x_b-\mu_b)$ do đây là 2 ma trận chuyển vị. Vậy :

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

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

trong đó: $\frac{1}{2}\mu_a^T A_{aa}x_a = \frac{1}{2}x_a^T A_{aa}\mu_a$ and $\frac{1}{2}\mu_a^T A_{aa}\mu_a = constant$

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

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

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

Compare with Gaussian distribution: $\Delta = \frac{-1}{2}x^T\Sigma^{-1}x + x^T\Sigma^{-1}\mu + constant$

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

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

Nhân 2 vế với $\Sigma = A_{aa}^{-1}$ ta được:

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

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

By using Schur complement:

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

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

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

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

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

$$p(x_a|x_b) = N(x_{a|b}|\mu_{a|b}, \Sigma_{a|b})$$