Ngô Đoàn Kiên - 11201976

Bài tập VN ML Buổi 2

1. Prove That :

a) Gaussian distribution is normalize

Để chứng minh ta phải tính ra rằng:

$$\int_{-\infty}^{+\infty} \mathcal{N}(x|\mu,\sigma^2)dx = 1$$

hay $$\frac{1}{\sqrt{2\pi\sigma^2}} \underbrace{\int_{-\infty}^{\infty} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}dx}_{(1)} = 1 \quad (*)$$

Ta đi tìm (1):

giả sử: $\mu = 0$ và đặt (1) = I

Ta có: $$I = \int_{-\infty}^{\infty} \exp\left(-\frac{1}{2\sigma^2}x^2\right)dx$$

Bình phương biểu thức I, ta được:

$$I^2 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left(-\frac{x^2}{2\sigma^2} - \frac{y^2}{2\sigma^2}\right)dx\,dy$$

Đổi biến hệ tọa độ (x, y) sang hệ tọa độ (r, θ)

$\rightarrow$ Đặt $\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$ $\quad (x^2 + y^2 = r^2)$

Vì đổi biến nên ta phải tìm sự thay đổi của biểu thức trong tích phân sau khi đổi $dx\,dy \rightarrow d(r)d(\theta)$

$\rightarrow$ Dùng Jacobian

$$\frac{\partial(x,y)}{\partial(r,\theta)} = \begin{vmatrix} \frac{\partial(x)}{\partial(r)} & \frac{\partial(x)}{\partial(\theta)} \\ \frac{\partial(y)}{\partial(r)} & \frac{\partial(y)}{\partial(\theta)} \end{vmatrix} = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix}$$

Thứ
Ngày . . No.

$= r\cos^2(\theta) + r\sin^2(\theta) = r$ (do $\sin^2(\theta) + \cos^2(\theta) = 1$)

căn của I^2 được thay đổi vì góc theta chỉ đi từ $0 \to 2\pi$ trên vòng tròn lượng giác, bán kính r luôn >0

$$\Rightarrow I^2 = \int_0^{2\pi}\int_0^{\infty} \exp\left(\frac{-r^2}{2\sigma^2}\right) r\, d(r)\, d(\theta)$$ (Dùng $\sin^2 + \cos^2 = 1$ để rút gọn)

$$= 2\pi \int_0^{\infty} \exp\left(-\frac{r^2}{2\sigma^2}\right) r\, d(r)$$

Đặt $u = r^2 \Rightarrow du = 2r\, d(r) \Leftrightarrow \frac{d(u)}{2r} = d(r)$

$$\Rightarrow I^2 = 2\pi \int_0^{\infty} \exp\left(-\frac{u}{2\sigma^2}\right) \cancel{r} \cdot \frac{1}{2\cancel{r}}\, d(u)$$

$$= \pi \int_0^{\infty} \exp\left(-\frac{u}{2\sigma^2}\right) d(u)$$

$$= \pi \left[(-2\sigma^2) \cdot \exp\left(-\frac{u}{2\sigma^2}\right)\right]_0^{\infty}$$

$$= \pi \cdot \left[(-2\sigma^2) \cdot e^{-\infty} - (-2\sigma^2) \cdot e^0\right]$$

$$= \pi \cdot 2\sigma^2 = 2\pi\sigma^2$$

$$\Rightarrow I = \sqrt{2\pi\sigma^2}$$

Thay $I = \sqrt{2\pi\sigma^2}$ vào biểu thức (*) đầu bài chứng minh ta được

$$(*) \Leftrightarrow \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \sqrt{2\pi\sigma^2} = 1$$

$\Leftrightarrow 1 = 1 \Rightarrow$ proved với $\mu = 0$ (**)

Trở lại tổng quát:

$$\int_{-\infty}^{+\infty} \mathcal{N}(x \mid \mu, \sigma^2)\, dx = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot \int_{-\infty}^{+\infty} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) d(x)$$

Đặt $y = x - \mu$

$\Rightarrow$ biểu thức sẽ giống hệt (**) đã chứng minh

$\Rightarrow$ proved

b) $X \sim N(\mu, \sigma^2)$

prove : $E(X) = \mu$.

we have:

$$p(x|\mu,\sigma^2) = f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and

$$E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, d(x)$$

$$\Rightarrow E(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right) dx$$

Đặt $t = \frac{x-\mu}{\sqrt{2\sigma^2}} \Rightarrow d(t) = d(x) \cdot \frac{1}{\sqrt{2\sigma^2}}$

$\Leftrightarrow d(x) = d(t) \cdot \sqrt{2\sigma^2}$

$$\Rightarrow E(X) = \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} \underline{(\sqrt{2\sigma^2} \cdot t + \mu)} \cdot \exp(-t^2)\, dt$$

do: $x = \sqrt{2\sigma^2} \cdot t + \mu$

$$= \frac{1}{\sqrt{\pi}} \left(\sqrt{2\sigma^2} \cdot \underbrace{\int_{-\infty}^{\infty} t \cdot \exp(-t^2)\, dt}_{(1)} + \mu \underbrace{\int_{-\infty}^{\infty} \exp(-t^2)\, dt}_{2} \right)$$

$\cancel{= \frac{1}{\sqrt{\pi}} \sqrt{2\sigma^2}}$

với tích phân (1) : đặt $u = e^{-x^2} \Rightarrow du = -2x e^{-x^2} dx$

$\Rightarrow \frac{-1}{2} du = x \cdot e^{-x^2} dx$

với tích phân (2) : dùng chứng minh ở phần a đã tính có:

$\Rightarrow (2) = \sqrt{\pi}$

$$\Rightarrow E(X) = \frac{1}{\sqrt{\pi}} \left(\sqrt{2\sigma^2} \cdot \int_{-\infty}^{\infty} \frac{-1}{2} du + \mu\sqrt{\pi} \right)$$

$$= \frac{1}{\sqrt{\pi}} \left(\sqrt{2\sigma^2} \cdot \underbrace{\left[\frac{-1}{2} \exp(-t^2)\right]_{-\infty}^{\infty}}_{= 0 - 0} + \mu\sqrt{\pi} \right)$$

$$= \frac{1}{\sqrt{\pi}} \cdot \mu\sqrt{\pi} = \mu \Rightarrow \text{proved}$$

c) Variance$(X) = \sigma^2$

$$f_X(x) = \frac{1}{\sqrt{\sigma^2 2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

$$Var(X) = E(X^2) - [E(X)]^2 = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx - \mu^2$$

$$\Rightarrow Var(X) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} x^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx - \mu^2$$

Đặt $t = \frac{x-\mu}{\sqrt{2\sigma^2}} \Rightarrow d(t) = d(x) \cdot \frac{1}{\sqrt{2\sigma^2}}$

$$\Rightarrow Var(X) = \frac{\sqrt{2\sigma^2}}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} (\sqrt{2\sigma^2} \cdot t + \mu)^2 \cdot \exp(-t^2)\,dt - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} \left[(2\sigma^2 + 2\sqrt{2\sigma^2} \cdot t \cdot \mu + \mu^2) \cdot \exp(-t^2)\right] dt - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \cdot \Bigg[\underbrace{2\sigma^2 \int_{-\infty}^{\infty} t^2 \cdot \exp(-t^2)\,dt}_{(1)} + \underbrace{2\sqrt{2\sigma^2} \cdot \mu \int_{-\infty}^{\infty} t \cdot \exp(-t^2)\,dt}_{(2)} + \underbrace{\mu^2 \int_{-\infty}^{\infty} \exp(-t^2)\,dt}_{(3)}\Bigg] - \mu^2$$

Làm từng tích phân (1); (2); (3)

(1): tích phân từng phần:

$$\text{đặt } \begin{cases} u = t \\ dv = t \cdot \exp(-t^2)\,dt \end{cases} \Rightarrow \begin{cases} du = dt \\ v = \frac{-1}{2} \exp(-t^2) \end{cases}$$

$$\Rightarrow (1) \Leftrightarrow t \cdot \left(\frac{-1}{2} \exp(-t^2)\right)\Bigg|_{-\infty}^{+\infty} + \frac{1}{2} \int_{-\infty}^{+\infty} e^{-t^2}\,dt$$

$$= 0 + \frac{1}{2}\sqrt{\pi} = \frac{\sqrt{\pi}}{2}$$

Thứ

Ngày . . No.

(2) : = 0 (đã tính từ các phần trước)

(3) = $\sqrt{\pi}$ (đã tính từ các phần trước

Từ (1) ; (2) ; (3) ⇒

$$\cancel{E(}\ Var(X) = \frac{1}{\sqrt{\pi}} \cdot \left(2\sigma^2 \cdot \frac{\sqrt{\pi}}{2} + 0 + \mu^2 \cdot \sqrt{\pi}\right) - \mu^2$$

$$= \sigma^2 + \mu^2 - \mu^2 = \sigma^2 \Rightarrow \text{proved}$$

d) Multivariate gaussian dist. is normalized

D dimension vector x, The multivariate gaussian dist. takes the form:

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

The functional dependence of the gaussian on x is through the quadratic form:

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

we consider eigenvector equation for covariance matrix:

$$\Sigma u_i = \lambda_i u_i \quad (i=1,2,\dots,D)$$

Because Σ is a real, symmetric matrix $\Rightarrow$ its eigenvalues will be real, and its eigenvectors can be chosen to form an orthonormal set

$$\Rightarrow u_i^T.u_j = I_{ij} \quad (2) \quad \left(I_{ij} = \begin{cases} 1, & \text{if } i=j \\ 0, & \text{otherwise} \end{cases}\right)$$

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

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

From (1) $\Rightarrow$

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

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

($y_i = u_i^T (x-\mu)$ and because of symmetrix $\Rightarrow u_i^T (x-\mu) = (x-\mu)^T . u_i$)

~~F~~

$\Rightarrow$ we form a vector y

$$y = U.(x-\mu)$$

Thứ
Ngày . .
No.

From (2) $\Rightarrow U$ is orthogonal

* Change from x to the y coordinate system, we have a Jacobian matrix J:

$$J_{ij} = \frac{\partial x_i}{\partial y_j} = U_{ji}$$

we have: $|J|^2 = |U^T|^2 = |U^T| \cdot |U| = |U^T \cdot U| = |I| = 1$ (orthogonal property)

$$\Rightarrow |J| = 1$$

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

With y coordinate system, Gauss dist. take the form:

$$p(y) = p(x) \cdot |J| = \prod_{j=1}^{D} \frac{1}{(2\pi \lambda_j)^{1/2}} \cdot \exp\left(-\frac{y_j^2}{2\lambda_j}\right)$$

$$\Rightarrow \int_{-\infty}^{+\infty} p(y)\, dy = \prod_{j=1}^{D} \int_{-\infty}^{+\infty} \frac{1}{(2\pi \lambda_j)^{1/2}} \exp\left(-\frac{y_i^2}{2\lambda_i}\right) dy_i \quad (3)$$

have the same form as the univariate that has been proved in the 1.a problem $\Rightarrow (3) = 1 \Rightarrow$ proved

KOKUYO

2. Calculate:

a). The conditional Gauss. dist.

- suppose x is D-dimensional vector with Gauss dist $\mathcal{N}(x \mid \mu, \Sigma)$

- We partition x into two disjoint subsets x_a; x_b:

x_a comprises of M component

x_b comprises the remaining (D-M) components.

$$\Rightarrow x = \begin{pmatrix} x_a \\ x_b \end{pmatrix}$$

similarly with μ and Σ

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

Note: Σ is symmetric $\Rightarrow \Sigma = \Sigma^T$

$\Rightarrow \Sigma_{aa}$ and Σ_{bb} are symmetric, whil $\Sigma_{ba} = \Sigma_{ab}^T$

set $\Lambda = \Sigma^{-1}$ (precision matrix)

$$= \begin{pmatrix} \Lambda_{aa} & \Lambda_{ab} \\ \Lambda_{ba} & \Lambda_{bb} \end{pmatrix} \quad \left(\begin{array}{l} \text{also symmetric} \\ \text{because } \Sigma \text{ is sym.} \end{array} \right)$$

$\Rightarrow \Lambda_{aa}$ and Λ_{bb} are symmetric, while $\Lambda_{ab}^T = \Lambda_{ba}$

From the quadratic form ~~in 1.d.(1)~~ and x, μ above, we obtain:

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

~~Compare (1) to the Δ quadratic form, we have~~

from (1) $\Rightarrow$ (1) $= -\frac{1}{2}(x-\mu)^T \cdot \Lambda (x-\mu) = -\frac{1}{2} x_a^T \Lambda_{aa}^{-1} x_a + x_a^T \left(\Lambda_{aa}\mu_a - \Lambda_{ab} (x_b - \mu_b) \right) + C$

Thứ
Ngày . . No.

$$\Sigma_{a|b} \doteq \Lambda_{aa}^{-1}$$

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

using schur complement:

we have:

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{pmatrix}$$

$$\text{with } M = (A - BD^{-1}C)^{-1}$$

and with $\Sigma^{-1} = \Lambda$

$$\Rightarrow \left\{ \begin{aligned} \Lambda_{aa} &= (\Sigma_{aa} - \Sigma_{ab} \cdot \Sigma_{bb}^{-1} \Sigma_{ba})^{-1} \\ \Lambda_{ab} &= -(\Sigma_{aa} - \Sigma_{ab} \cdot \Sigma_{bb}^{-1} \Sigma_{ba})^{-1} \cdot \Sigma_{ab} \Sigma_{bb}^{-1} \end{aligned} \right.$$

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

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

Ngày . . No.

b) The marginal distribution

$$p(x_a) = \int p(x_a, x_b)\, dx_b$$

integrate out x_b:

$$\frac{-1}{2} x_b^T \Lambda_{bb} x_b + x_b^T m$$

$$= \frac{-1}{2} (x_b - \Lambda_{bb}^{-1} m)^T \cdot \Lambda_{bb} (x_b - \Lambda_{bb}^{-1} m) + \frac{1}{2} m^T \Lambda_{bb}^{-1} m \quad (1)$$

with $m = \Lambda_{bb} \cdot \mu_b - \Lambda_{ba} \cdot (x_a - \mu_a)$

Take the exponential of right-hand side of (1) we have the integration over x_b:

$$\int \exp\left\{ -\frac{1}{2} (x_b - \Lambda_{bb}^{-1} m)^T \Lambda_{bb} (x_b - \Lambda_{bb}^{-1} m) \right\} dx_b \quad (2)$$

since (2) is the integral over an unnormalized Gauss, The result will be the reciprocal of The normalization coefficient.

– Combine the left-hand-side of (1) depend on x_a with m above,

we obtain:

$$\frac{1}{2} [\Lambda_{bb} \cdot \mu_b - \Lambda_{ba} (x_a - \mu_a)]^T \cdot \Lambda_{bb}^{-1} \cdot [\Lambda_{bb} \cdot \mu_b - \Lambda_{ba} \cdot (x_a - \mu_a)]$$

$$- \frac{1}{2} x_a^T \Lambda_{aa} \cdot x_a + x_a^T (\Lambda_{aa} \cdot \mu_a + \Lambda_{ab} \mu_b) + C$$

comparison with the quadratic form from section 2.a) The covariance of marginal dist. of $p(x_a)$ is given by:

Scanned with CamScanner

Thứ
Ngày . . No.

$$\Sigma_a = (\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})^{-1} \quad (3)$$

mean is given by:

$$\Sigma_a(\Lambda_{aa} - \Lambda_{ab}\Lambda_{bb}^{-1}\Lambda_{ba})\mu_a = \mu_a$$

the covariance (3) is the same form as we did in the 2.a) conditional dist.

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

$\Rightarrow$ The marginal dist. $p(x_a)$ has mean and covar:

$$\begin{cases} E[x_a] = \mu_a \\ cov[x_a] = \Sigma_{aa} \end{cases}$$