状态估计 hw1

Sunday, April 26, 2020 16:31

1. prove $1=\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)dx$

$$\because \iint_{-\infty}^{\infty} e^{-(x^2+y^2)} = \int_0^{2\pi}\int_0^{\infty} e^{-r^2} r\,dr\,d\theta$$

$$= \int_0^{2\pi} -\frac{1}{2}e^{-r^2}\Big|_0^{\infty} d\theta = \pi$$

and $\iint_{-\infty}^{\infty} e^{-(x^2+y^2)}dx\,dy = \left(\int_{-\infty}^{\infty} e^{-x^2}dx\right)^2$

$$\therefore \int_{-\infty}^{\infty} e^{-x^2}dx = \sqrt{\pi}$$

$\therefore$ let $z=\frac{x-\mu}{\sqrt{2}\sigma}$ $dz = dx\frac{1}{\sqrt{2}\sigma}$

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

$$= \int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}\sigma}\exp(-z^2)\,dz\cdot(\sqrt{2}\sigma)$$

$$= \frac{1}{\sqrt{\pi}}\int_{-\infty}^{\infty}\exp(-z^2)dz$$

$$= \frac{\sqrt{\pi}}{\sqrt{\pi}} = 1 = LHS$$

$\therefore$ proved

2.(1) u, v 相同维度列向量，证明 $\vec{u}^T\vec{v} = tr(\vec{v}\vec{u}^T)$

let $\vec{u} = [u_1, \dots, u_n]^T$

$\vec{v} = [v_1 \dots v_n]^T$

$$RHS = \vec{u}^T\vec{v} = u_1v_1 + u_2v_2 + \dots + u_nv_n$$

$$= \sum_{i=1}^{n} u_iv_i$$

$$\vec{v}\vec{u}^T = \begin{bmatrix} v_1u_1 & v_1u_2 & \cdots & v_1u_n \\ v_2u_1 & v_2u_2 & \cdots & v_2u_n \\ \vdots & \vdots & \ddots & \vdots \\ v_nu_1 & v_nu_2 & \cdots & v_nu_n \end{bmatrix}$$

$$= \begin{bmatrix} & \vdots & \\ \cdots & u_iv_j & \cdots \\ & \vdots & \end{bmatrix}_{i,j}$$

$\therefore$ LHS $= tr(\vec{v}\vec{u}^T) = \sum_{i=1}^{n} u_iv_i = RHS$

$\therefore$ LHS $= tr(\vec{v}\vec{u}^T) = \sum_{i=1}^{n} u_i v_i$ = RHS

$\therefore$ proved.

3.(4) $\vec{x} \sim N(\vec{\mu}, \Sigma)$, prove $\vec{\mu} = E[\vec{x}] = \int_{-\infty}^{\infty} \vec{x}\, p(\vec{x})\, d\vec{x}$

$$RHS = \int_{-\infty}^{\infty} \vec{x}\, p(\vec{x})\, d\vec{x}$$

$$= \int_{-\infty}^{\infty} \vec{x} \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})\right) d\vec{x}$$

let $\vec{y} = \vec{x} - \vec{\mu}$, $\vec{x} = \vec{y} + \vec{\mu}$, $d\vec{x} = d\vec{y}$

$$\therefore RHS = \underbrace{\frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} \vec{y} \exp\left(-\frac{1}{2}\vec{y}^T \Sigma^{-1} \vec{y}\right) d\vec{y}}_{I_1}$$

$$+ \underbrace{\frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} \vec{\mu} \exp\left(-\frac{1}{2}\vec{y}^T \Sigma^{-1} \vec{y}\right) dy}_{I_2} = I_1 + I_2$$

① let $f(\vec{y}) = \vec{y} \exp\left(-\frac{1}{2}\vec{y}^T \Sigma^{-1} \vec{y}\right)$

so $I_1 = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}} \int_{-\infty}^{\infty} f(\vec{y})\, d\vec{y}$

and $f(-\vec{y}) = -\vec{y} \exp\left(-\frac{1}{2}(-\vec{y})^T \Sigma^{-1} (-\vec{y})\right)$

$= -\vec{y} \exp\left(-\frac{1}{2}\vec{y}^T \Sigma^{-1} \vec{y}\right)$

$= -f(\vec{y})$

$\therefore$ f is an odd function

$\therefore \int_{-\infty}^{\infty} f\, d\vec{y} = 0$

$\Rightarrow I_1 = 0$

② $\because \Sigma$ is positive definite and symmetric

$\therefore$ we can write $\Sigma = U \Lambda U^T$

where U is a full rank orthogonal matrix

containing eigenvectors of Σ as its columns,

Λ is a diagonal matrix containing Σ's

eigen values.

and we define $\Lambda^{\frac{1}{2}} \in R^{n \times n}$ to be a diagonal matrix

... square roots of corresponding

and we define $\Lambda^{\frac{1}{2}} \in K$ to be a diagonal matrix whose entries are square roots of corresponding entries of Λ, so $\Lambda = \Lambda^{\frac{1}{2}}(\Lambda^{\frac{1}{2}})^T$

$$\therefore \Sigma = U\Lambda U^T = U\Lambda^{\frac{1}{2}}(\Lambda^{\frac{1}{2}})^T U^T = BB^T$$

where $B = U\Lambda^{\frac{1}{2}}$, $\Sigma^{-1} = B^{-T}B^{-1}$, $|B| = |U\Lambda^{\frac{1}{2}}| = |\Sigma|^{\frac{1}{2}}$

$$\therefore I_2 = \frac{1}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}\int_{-\infty}^{\infty} \vec{\mu} \exp\left(-\frac{1}{2}\vec{y}^T B^{-T} B^{-1}\vec{y}\right) d\vec{y}$$

$$= \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\vec{y}^T B^{-T} B^{-1}\vec{y}\right) d\vec{y}$$

let $\vec{z} = B^{-1}\vec{y}$ $\quad \vec{y} = B\vec{z}$ $\quad d\vec{y} = |B|\, d\vec{z}$ and $|B| = |\Sigma|^{\frac{1}{2}}$

(if $\vec{y} \in R^n$ is a random variable vector with joint density function $f_{\vec{y}}: R^n \to R$, $\vec{z} = H(\vec{y})$ where H is a bijective differentiable function, then $\vec{z}$ has joint density $f_{\vec{z}}$ where $f_{\vec{z}} = f_{\vec{y}}(\vec{y}) \left|\det\begin{pmatrix} \frac{\partial y_1}{\partial z_1} & \cdots & \frac{\partial y_1}{\partial z_n} \\ \vdots & & \\ \frac{\partial y_n}{\partial z_1} & \cdots & \frac{\partial y_n}{\partial z_n}\end{pmatrix}\right|$)

$$I_2 = \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\vec{z}^T\vec{z}\right)|B|\, d\vec{z}$$

$$= \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}|\Sigma|^{\frac{1}{2}}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\vec{z}^T\vec{z}\right)|\Sigma^{\frac{1}{2}}|\, d\vec{z}$$

$$= \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}}\int_{-\infty}^{\infty} \exp\left(-\frac{1}{2}\sum_{i=1}^{N} z_i^2\right) dz_i$$

let $t_i = z_i \frac{1}{\sqrt{2}}$ $\quad dt_i = dz_i \frac{1}{\sqrt{2}}$

$$I_2 = \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}}\prod_{i=1}^{N}\int_{-\infty}^{\infty} \exp\left(-t_i^2\right)\sqrt{2}\, dt_i$$

$$= \frac{\vec{\mu}}{(2\pi)^{\frac{N}{2}}}\prod_{i=1}^{N}\sqrt{2\pi}$$

$$= \vec{\mu}$$

$\therefore$ RHS $= I_1 + I_2 = \vec{\mu} =$ LHS

$\therefore$ proved!

$\therefore$ RHS = E[(x - μ ... (top line cut off)

$\therefore$ proved!

4.(5) $\vec{x} \sim N(\vec{\mu}, \Sigma)$, prove $\Sigma = E[(\vec{x}-\vec{\mu})(\vec{x}-\vec{\mu})^T] = \int_{-\infty}^{\infty} (\vec{x}-\vec{\mu})(\vec{x}-\vec{\mu})^T p(\vec{x}) d\vec{x}$

$\vec{x}\vec{x}^T - \vec{x}\vec{\mu}^T - \vec{\mu}\vec{x}^T + \vec{\mu}\vec{\mu}^T$

$x_1\mu_1 \quad x_1\mu_n$

$\vdots$

$x_n\mu_1 \quad x_n\mu_n$

$\vec{\mu}\vec{x}$

$\mu_1 x_1 \quad \mu_1 x_n$

$$\text{LHS} = \Sigma = \begin{bmatrix} \mathrm{cov}(x_1, x_1), & \mathrm{cov}(x_1, x_2) & \cdots & \mathrm{cov}(x_1, x_n) \\ \mathrm{cov}(x_2, x_1), & \mathrm{cov}(x_2, x_2) & \cdots & \mathrm{cov}(x_2, x_n) \\ \vdots \\ \mathrm{cov}(x_n, x_1) & \mathrm{cov}(x_n, x_2) & \cdots & \mathrm{cov}(x_n, x_n) \end{bmatrix}$$

$$= \begin{bmatrix} E((x_1-\mu_1)^2), & E((x_2-\mu_2)(x_1-\mu_1)), & \dots, & E((x_n-\mu_n)(x_1-\mu_1)) \\ E((x_1-\mu_1)(x_2-\mu_2)), & E((x_2-\mu_2)^2), & \dots, & E((x_n-\mu_n)(x_2-\mu_2)) \\ \vdots \\ E((x_1-\mu_1)(x_n-\mu_n)), & E((x_2-\mu_2)(x_n-\mu_n)) & \dots, & E((x_n-\mu_n)^2) \end{bmatrix}$$

$$= E\left(\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \\ \vdots \\ x_n-\mu_n \end{bmatrix} \cdot [x_1-\mu_1 \ \ x_2-\mu_2, \dots, x_n-\mu_n] \right)$$

$$= E\left((\vec{x}-\vec{\mu})(\vec{x}-\vec{\mu})^T \right) = \text{RHS}$$

5.(6) k个独立高斯变量 $\vec{x}_k \sim N(\vec{\mu}_k, \Sigma_k)$

prove $\exp(-\frac{1}{2}(\vec{x}-\vec{\mu})^T \Sigma^{-1} (\vec{x}-\vec{\mu})) \equiv \eta \prod_{k=1}^{k} \exp(-\frac{1}{2}(\vec{x}_k-\vec{\mu}_k)^T \Sigma_k^{-1} (\vec{x}_k - \vec{\mu}_k))$

where $\Sigma^{-1} = \sum_{k=1}^{k} \Sigma_k^{-1}$ $\quad \Sigma^{-1}\vec{\mu} = \sum_{k=1}^{k} \Sigma_k^{-1} \vec{\mu}_k$

$\because$ gaussian pdf can be written as

$$p(\vec{x}) = \exp\left[\xi + \eta^T \vec{x} - \frac{1}{2} \vec{x}^T \Lambda \vec{x} \right] \text{①}$$

where $\Lambda = \Sigma^{-1}$, $\vec{\eta} = \Sigma^{-1}\vec{\mu}$, $\xi = -\frac{1}{2}(N \ln(2\pi) - \ln|\Lambda| + \vec{\eta}\Lambda^{-1}\vec{\eta})$

$$\therefore \prod_{k=1}^{K} p_k(\vec{x}_k) = \exp\left[\sum_{k=1}^{K} \xi_k + \left(\sum_{k=1}^{K} \vec{\eta}_k\right)^T \vec{x}_k - \frac{1}{2} \vec{x}_k^T \left(\sum_{k=1}^{K} \Lambda_k\right) \vec{x}_k \right]$$

where $\sum_{k=1}^{K} \xi_k = -\frac{1}{2}\left(kN \ln 2\pi - \sum_{k=1}^{K} \ln|\Lambda_k| + \sum_{k=1}^{K} \vec{\eta}_k^T \Lambda_k \vec{\eta}_k \right)$

$$\therefore \prod_{k=1}^{K} p_k(\vec{x}_k) = \exp\left[\sum_{k=1}^{K} \xi_k + \xi_K - \xi_K + \left(\sum_{k=1}^{K} \vec{\eta}_k\right)^T \vec{x}_k - \frac{1}{2} \vec{x}_k^T \left(\sum_{k=1}^{K} \Lambda_k\right) \vec{x}_k \right]$$

$$= \exp\left(\sum_{k=1}^{K} \xi_k - \xi_K\right) \exp\left(\xi_K + \vec{\eta}_K^T \vec{x} - \frac{1}{2} \vec{x}^T \Lambda_K \vec{x}\right) \text{②}$$

where $\vec{\eta}_K = \sum_{k=1}^{K} \eta_k$, $\Lambda_K = \sum_{k=1}^{K} \Lambda_k$

and $\xi_k = -\frac{1}{2}\left(N\ln 2\pi - \ln|\Lambda_k| + \vec{\eta}_k^{T}\Lambda_k^{-1}\vec{\eta}_k\right)$ ③

∴ comparing ①, ② with ③, we have

$$\Lambda = \Lambda_k \qquad \vec{\eta} = \vec{\eta}_k$$

$$\Rightarrow \Sigma^{-1} = \sum_{k=1}^{K}\Sigma_k^{-1}, \qquad \Rightarrow \Sigma^{-1}\vec{\mu} = \sum_{k=1}^{K}\Sigma_k^{-1}\vec{\mu}_k$$