State Estimation Robotics Chapter 2 June 9 2018

2.5.1

$u^T v = tr(vu^T)$

$u^T v = \sum_{i=1}^{n} u_i v_i$ where $u = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$ and $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$

$$vu^T = \begin{bmatrix} v_1 u_1 & \cdots & v_1 u_n \\ \vdots & & \\ v_n u_1 & \cdots & v_n u_n \end{bmatrix}$$

$tr(vu^T) = \sum_{i=1}^{n} u_i v_i$ Hence $u^T v = tr(vu^T)$ ∎

2.5.2

Eq 2.29 $I(x,y) = H(x) + H(y) - H(x,y)$

If x and y are independent, $I(x,y) = 0$

$0 = H(x) + H(y) - H(x,y)$

$H(x,y) = H(x) + H(y)$ ∎

2.5.3

$x \sim N(u, \Sigma)$

Show $E[xx^T] = \Sigma + \mu\mu^T$

Note that $E[x] = \mu$ and $E[(x-\mu)(x-\mu)^T] = \Sigma$

$$\begin{aligned} E[(x-\mu)(x-\mu)^T] &= E[xx^T - x\mu^T - \mu x^T + \mu\mu^T] \\ &= E[xx^T] - E[x\mu^T] - E[\mu x^T] + \mu\mu^T \\ &= E[xx^T] - \mu\mu^T - \cancel{\mu\mu^T + \mu\mu^T}^{\,0} = \Sigma \end{aligned}$$

$\Rightarrow E[xx^T] = \Sigma + \mu\mu^T$ ∎

2.5.4 $x \sim N(\mu, \Sigma)$

Show $\mu = E[x] = \int_{-\infty}^{\infty} x\, p(x)\, dx$

$$p(x) = \frac{1}{\sqrt{(2\pi)^N \det \Sigma}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

$$= N(u_1, \Sigma_{11})\, N(u_2, \Sigma_{22}) \dots N(u_n, \Sigma_{nn})$$ from 2.53 and 2.54 combined

$$x_i = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} x_i N(u_1, \Sigma_{11}) \dots N(u_n, \Sigma_{nn})\, dx_1 \dots dx_n$$

$$= \int_{-\infty}^{\infty} x_i N(u_i, \Sigma_{ii}) * \underbrace{\int_{-\infty}^{\infty} N(u_j, \Sigma_{jj}) * \dots}_{\text{for } j \neq i \in 1 \dots n}$$

$$= \mu_i * 1 * 1 \dots = \mu_i$$

$\therefore \underline{E[x] = \mu}$ //

in case you also want to solve $\int_{-\infty}^{\infty} x_i N(u_i, \Sigma_{ii})$ [scalar case]

$$\int_{-\infty}^{\infty} \frac{x_i}{\sqrt{2\pi \Sigma_{ii}}} e^{-\frac{1}{2}\frac{(x_i - u_i)^2}{\Sigma_{ii}}}$$ let $y = x_i - u_i$

$$\int_{-\infty}^{\infty} \frac{y + u_i}{\sqrt{2\pi\Sigma_{ii}}} e^{-\frac{1}{2} y^2/\Sigma_{ii}}$$ ~~let z =~~, dz

$$\underbrace{\int_{-\infty}^{\infty} \frac{y}{\sqrt{2\pi\Sigma_{ii}}} e^{-\frac{1}{2}y^2/\Sigma_{ii}} dy}_{A} + \underbrace{\int_{-\infty}^{\infty} u_i \frac{1}{\sqrt{2\pi\Sigma_{ii}}} e^{-\frac{1}{2}y^2/\Sigma_{ii}} dy}_{B}$$

Notice that for A, let $z_1 = \frac{y}{\sqrt{2\Sigma_{ii}}}$ $dz_1 = \frac{1}{\sqrt{2\Sigma_{ii}}} dy$

$$\int_{-\infty}^{\infty} \sqrt{\frac{2\Sigma_{ii}}{\pi}}\, z_1 e^{-z_1^2} dz_1 = 0$$ since $z_1 e^{-z_1^2}$ is an odd function

$$= \int_{-\infty}^{0} z_1 e^{-z_1^2} dz_1 + \int_0^{\infty} z_1 e^{-z_1^2} dz_1$$

$$= \int_0^{\infty} -z_1 e^{-z_1^2} dz_1 + \int_0^{\infty} z_1 e^{-z_1^2} dz_1 = 0$$

Notice for B, let $z_2 = \frac{y}{\sqrt{2\Sigma_{ii}}}$ $dz_2 = \frac{1}{\sqrt{2\Sigma_{ii}}} dy$

$$\int_{-\infty}^{\infty} \frac{u_i}{\sqrt{\pi}} e^{-z_2^2} dz_2 = \frac{u_i}{\sqrt{\pi}} \underbrace{\int_{-\infty}^{\infty} e^{-z_2^2} dz_2}_{\sqrt{\pi}} = \frac{u_i}{\sqrt{\pi}} \sqrt{\pi} = \underline{u_i}$$ //

$$= A + B = 0 + u_i = u_i$$

2.5.5 Show

$$\Sigma = \int_{-\infty}^{\infty} (x-\mu)(x-\mu)^T p(x)\,dx \quad ①$$

$$= \int_{-\infty}^{\infty} \begin{bmatrix} (x_1-\mu_1)(x_1-\mu_1) & \cdots & (x_1-\mu_1)(x_n-\mu_n) \\ \vdots & & \vdots \\ (x_n-\mu_n)(x_1-\mu_1) & \cdots & (x_n-\mu_n)(x_n-\mu_n) \end{bmatrix} p(x)\,dx$$

$$\Sigma_{ij} = \int_{-\infty}^{\infty} (x_i-\mu_i)(x_j-\mu_j)\, p(x)\,dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_i-\mu_i)(x_j-\mu_j)\, N(\mu_i, \Sigma_{ii})\, N(\mu_j, \Sigma_{jj})\, dx_i dx_j$$

Note that $\int_{-\infty}^{\infty} N(\mu_k, \Sigma_{kk})\, dk$ for $k \neq i,j$ $= 1$ (total prob)

$$\Sigma_{ij} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_i-\mu_i)(x_j-\mu_j)\, N(\mu_i, \Sigma_{ii})\, N(\mu_j, \Sigma_{jj})\, dx_i\, dx_j$$

is by definitions the correlation between x_i and x_j

Hence eq ① $= \Sigma$ ▨

2.5.6

① $$e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)} = e^{-\frac{1}{2}\left(x^T\Sigma^{-1}x - 2x^T\Sigma^{-1}\mu + \mu^T\mu\right)}$$

② $$n\prod_{k=1}^{\bar{K}} e^{-\frac{1}{2}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)} = n\, e^{-\frac{1}{2}\left[\sum_{k=1}^{\bar{K}}(x-\mu_k)^T\Sigma_k^{-1}(x-\mu_k)\right]}$$

$\bar{K}$ = big K

$$= n\, e^{-\frac{1}{2}\sum_{k=1}^{\bar{K}}\left(x^T\Sigma_k^{-1}x - 2x^T\Sigma_k^{-1}\mu_k + \mu_k^T\mu_k\right)} = n\, e^{-\frac{1}{2}\left(x^T\sum_{k=1}^{\bar{K}}\Sigma_k^{-1}x - 2x^T\sum_{k=1}^{\bar{K}}\Sigma_k^{-1}\mu_k + \sum_{k=1}^{\bar{K}}\mu_k^T\mu_k\right)}$$

Let $$\Sigma^{-1} = \sum_{k=1}^{\bar{K}}\Sigma_k^{-1} \quad ③$$

$$\Sigma^{-1}\mu = \sum_{k=1}^{\bar{K}}\Sigma_k^{-1}\mu_k \quad ④$$

$$\mu^T\mu = \sum_{k=1}^{\bar{K}}\mu_k^T\mu_k + \underbrace{C}_{\text{some constant dependent on ③ and ④}}$$

②: $$n\, e^{-\frac{1}{2}\left(x^T\Sigma^{-1}x - 2x^T\Sigma^{-1}\mu + \mu^T\mu - c\right)} = n e^{\frac{c}{2}}\, e^{-\frac{1}{2}\left(x^T\Sigma^{-1}x - 2x^T\Sigma^{-1}\mu + \mu^T\mu\right)}$$

$$= n e^{\frac{c}{2}}\, e^{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)}$$

$\therefore$ for $n = e^{-\frac{c}{2}}$, eq ② $=$ ① and the direct prod of K Gaussian PDFs is also a Gaussian PDF.

2.5.7 a) Since X_K for $K \in 1 \dots N$ are statistically independent

$$p(X) = p(X_1) * \dots * p(X_K) \quad \text{or} \quad \prod_{K=1}^{N} p(X_K)$$

Note that $\int_{-\infty}^{\infty} p(X_K)\, dX_K = 1$ for any $K \in 1 \dots N$

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(X)\, dX = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \prod_{K=1}^{N} p(X_K)\, dX_1 \dots dX_K = \int_{-\infty}^{\infty} p(X_1) dX_1 * \dots * \int_{-\infty}^{\infty} p(X_N) dX_N$$

$$= \underline{1} \; //$$

Hence $p(X)$ satisfies the axiom of total probability

b) Let $X = \sum_{K=1}^{N} w_K X_K$ and $E[X_K] = \mu_K$

$$E[X] = E\left[\sum w_K X_K\right] = \sum_{K=1}^{N} w_K E[X_K] = \underline{\sum_{K=1}^{N} w_K \mu_K = \mu} \; //$$

c) let $E\left[(X_K - \mu_K)(X_K - \mu_K)^T\right] = \Sigma_K$

Note that since X_i and X_j are ~~stat.~~ ind., $E\left[(X_i - \mu_i)(X_j - \mu_j)^T\right] = 0$

$$E\left[(X-\mu)(X-\mu)^T\right] = E\left[\left(\sum w_K (X_K - \mu_K)\right)\left(\sum w_K (X_K - \mu_K)\right)^T\right]$$

$$= E\left[\sum_{i=1}^{N} \sum_{j=1}^{N} w_i w_j (X_i - \mu_i)(X_j - \mu_j)^T\right]$$

$$= \sum_{i=1}^{N} \sum_{j=1}^{N} \underbrace{w_i w_j\, E\left[(X_i - \mu_i)(X_j - \mu_j)^T\right]}_{w_i w_j \Sigma_i \text{ if } j==i,\ 0 \text{ otherwise}}$$

$$\boxed{E\left[(X-\mu)(X-\mu)^T\right] = \sum_{i=1}^{N} w_i^2 \Sigma_i}$$

2.5.8 a)

$$E[y] = E[X^T X] = E\left[\sum_{i=1}^{K} x_i x_i\right] = \sum_{i=1}^{K} \underbrace{E[x_i x_i]}_{E[x_i x_i] = 1 \text{ since } X = N(0,1)} = \sum_{i=1}^{K} 1 = \underline{K} \; //$$

$= \Sigma_{ii}$

b)

$$\Sigma = E\left[(y-K)(y-K)^T\right] = E[y^2] - 2K\,\cancel{E[y]}^{K} + K^2 = E[y^2] - K^2$$

Note: $E[y^2] = E\left[\sum_{i=1}^{K} x_i x_i \sum_{j=1}^{K} x_j x_j\right] = \sum_{i=1}^{K} \sum_{j=1}^{K} E[x_i x_i x_j x_j]$

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} E[x_i x_i] E[x_j x_j] + 2 E[x_i x_j]^2$$ according to Isserlis' theorem

substitute $E[y^2]$

$$\Sigma = \sum_{i=1}^{K} \sum_{j=1}^{K} \underbrace{E[x_i x_i]}_{1} \underbrace{E[x_j x_j]}_{1} + \sum_{i=1}^{K} \sum_{j=1}^{K} 2 \underbrace{E[x_i x_j]^2}_{1 \text{ if } i==j,\ 0 \text{ otherwise}} - K^2$$

$$= \sum_{i=1}^{K} \sum_{j=1}^{K} 1 + \sum_{i=1}^{K} 2 * 1 - K^2$$

$$= K^2 + 2K - K^2 = \underline{2K} \; //$$