

机器人中的状态估计课后习题答案

完成人：高明

联系方式：

知乎：高明

微信：gaoming0901

2.概率论基础

2.5.1

假设 u, v 是相同维度向量, 请证明下面等式: $u^T v = \text{tr}(vu^T)$

solution:

$$u = (x_1, x_2, \dots, x_n)^T$$

$$v = (y_1, y_2, \dots, y_n)^T$$

$$u^T v = x_1 y_1 + x_2 y_2 + \dots + x_n y_n = \sum_{i=1}^n x_i y_i$$

$$uv^T = \begin{bmatrix} x_1 y_1 & \dots & \dots & \dots \\ \dots & x_2 y_2 & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \dots & \dots & \dots & x_n y_n \end{bmatrix}$$

$$\text{tr}(uv^T) = \sum_{i=1}^n x_i y_i = u^T v$$

2.5.2

如果有两个相互独立的随机变量 x, y , 它们的联合分布为 $p(x, y)$, 请证明它们概率的香浓信息等于各自独立香浓信息的和:

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

solution:

$$H(x, y)$$

$$= -E_{(x,y)}(\ln(f(x,y)))$$

$$= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) \ln(f(x,y)) dx dy$$

因为 x,y 独立

$$H(x,y)$$

$$= - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x)f(y) [\ln(f(x)) + \ln(f(y))] dx dy$$

$$= - [\int_{-\infty}^{\infty} (f(x) \ln(f(x)) dx) * \int_{-\infty}^{\infty} f(y) dy - [\int_{-\infty}^{\infty} f(y) \ln(f(y)) dy] * \int_{-\infty}^{\infty} f(x) dx$$

$$= - \int_{-\infty}^{\infty} (f(x) \ln(f(x)) dx - \int_{-\infty}^{\infty} f(y) \ln(f(y)) dy$$

$$= H(x) + H(y)$$

2.5.3

对于高斯分布的随机变量, $x \sim N(\mu, \Sigma)$, 请证明下面的等式:

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

solution:

$$\Sigma$$

$$= E[(x - \mu)(x - \mu)^T]$$

$$= E(xx^T - x\mu^T - \mu x^T + \mu\mu^T)$$

$$= E(xx^T) - E(x)\mu^T - \mu E(x^T) + \mu\mu^T$$

$$\text{因为 } E(x) = \mu$$

$$\Sigma = E(xx^T) - \mu\mu^T$$

因此

$$E(xx^T) = \Sigma + \mu\mu^T$$

2.5.4

对于高斯分布的随机变量, $x \sim N(\mu, \Sigma)$, 请证明下面的等式:

$$\mu = E(x) = \int_{-\infty}^{\infty} xp(x) dx$$

solution:

$$E(x)$$

$$= \int_{-\infty}^{\infty} \frac{x}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) dx$$

做变换:

$$y = x - \mu$$

可得:

$$x = y + \mu$$

$$E(x)$$

$$= \int_{-\infty}^{\infty} \frac{y+\mu}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}y^T \Sigma^{-1}y\right) dy$$

$$= \int_{-\infty}^{\infty} \frac{y}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}y^T \Sigma^{-1}y\right) dy + \int_{-\infty}^{\infty} \frac{\mu}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}y^T \Sigma^{-1}y\right) dy$$

上式第一项由于奇函数在关于0对称空间积分为0

上式第二项扣除 μ 满足概率归一化条件

$$E(x) = \mu$$

2.5.5

对于高斯分布的随机变量, $x \sim N(\mu, \Sigma)$, 证明下式:

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

solution:

$$E[(x - \mu)(x - \mu)^T]$$

$$= \int_{-\infty}^{\infty} \frac{(x-\mu)(x-\mu)^T}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) dx$$

做代换 $y = x - \mu$

$$E[(x - \mu)(x - \mu)^T]$$

$$= \int_{-\infty}^{\infty} \frac{yy^T}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}(y^T \Sigma^{-1}y)\right) dy \dots \dots \dots < 0 >$$

下面参考文献【1】中公式(108)如下式:

$$\frac{\partial}{\partial X} (X^T B X) = B X + B^T X$$

上式中X是矩阵, 向量算特殊矩阵, 直接带入, 向量表达式如下:

$$\frac{d}{dx}(x^T B x) = Bx + B^T x \dots\dots\dots < 1 >$$

由于协方差矩阵是对称矩阵，根据等式< 1 >:

$$\frac{d}{dx}(x^T \Sigma^{-1} x) = \Sigma^{-1} * x + \Sigma^{-T} * x = 2 * \Sigma^{-1} * x \dots\dots < 2 >$$

对于<2>式变换:

$$\begin{aligned} \frac{d}{dx}(-\frac{1}{2}x^T \Sigma^{-1} x) &= (-\frac{1}{2})(\Sigma^{-1} * x + \Sigma^{-T} * x) = \Sigma^{-1} * x \\ &= x^T \Sigma^{-1} \dots\dots\dots < 3 > \end{aligned}$$

将<3>式带入<0>式:

$$\begin{aligned} E[(x - \mu)(x - \mu)^T] \\ &= \int_{-\infty}^{\infty} \frac{-y * \Sigma}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp(-\frac{1}{2}(y^T \Sigma^{-1} y)) d(-\frac{1}{2}(y^T \Sigma^{-1} y)) \\ &= \int_{-\infty}^{\infty} \frac{-y * \Sigma}{\sqrt{(2\pi)^N \det(\Sigma)}} d(\exp(-\frac{1}{2}(y^T \Sigma^{-1} y))) \end{aligned}$$

分步积分法:

$$\begin{aligned} E[(x - \mu)(x - \mu)^T] \\ &= \frac{y * \Sigma}{\sqrt{(2\pi)^N \det(\Sigma)}} * \exp(-\frac{1}{2}(y^T \Sigma^{-1} y)) \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{\infty} \frac{\Sigma}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp(-\frac{1}{2}(y^T \Sigma^{-1} y)) dy \\ &= 0 + \Sigma \\ &= \Sigma \end{aligned}$$

2.5.6

对于K个相互独立的高斯变量, $x_k \sim N(\mu_k, \Sigma_k)$,请证明它们的归一化积仍然是高斯分布:

$$\exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)) \equiv \eta \prod_{k=1}^K \exp(-\frac{1}{2}(x_k - \mu_k)^T \Sigma_k^{-1} (x_k - \mu_k))$$

其中:

$$\Sigma^{-1} = \sum_{k=1}^K \Sigma_k^{-1}$$

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

且 η 归一化因子。

solution:

随机变量 x_k 的概率密度函数如下:

$$f_k(x) = \frac{1}{\sqrt{(2\pi)^{N_k} \det(\Sigma_k^{-1})}} \exp(-\frac{1}{2}(x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k))$$

$$f_1(x) * f_2(x) * \dots * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k} \prod_{k=1}^K \det(\Sigma_k)}} \exp(-\frac{1}{2} \sum_{k=1}^K (x - \mu_k)^T \Sigma_k^{-1}(x - \mu_k))$$

将指数部分的求和号展开:

$$f_1(x) * f_2(x) * \dots * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k} \prod_{k=1}^K \det(\Sigma_k)}} \exp(-\frac{1}{2}(x^T (\sum_{k=1}^K \Sigma_k^{-1})x - (\sum_{k=1}^K \mu_k^T \Sigma_k^{-1})x - x^T \sum_{k=1}^K \Sigma_k^{-1} \mu_k + \sum_{k=1}^K \mu_k^T \Sigma_k^{-1} \mu_k)) \dots < 0 >$$

因为协方差矩阵是对称矩阵, <0>式中

$$(\sum_{k=1}^K \mu_k^T \Sigma_k^{-1})x = x^T \sum_{k=1}^K \Sigma_k^{-1} \mu_k$$

因此:

$$f_1(x) * f_2(x) * \dots * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k} \prod_{k=1}^K \det(\Sigma_k)}} \exp(-\frac{1}{2}(x^T (\sum_{k=1}^K \Sigma_k^{-1})x - 2x^T \sum_{k=1}^K \Sigma_k^{-1} \mu_k + \sum_{k=1}^K \mu_k^T \Sigma_k^{-1} \mu_k)) \dots < 1 >$$

在式<1>中:

$x^T (\sum_{k=1}^K \Sigma_k^{-1})x$ 为二次项

$2x^T \sum_{k=1}^K \Sigma_k^{-1} \mu_k$ 为一次项

可以凑出“完全平方形式”

$$f_1(x) * f_2(x) * \dots * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k} \prod_{k=1}^K \det(\Sigma_k)}} \exp(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu) + M)$$

上式中M为一个常数;

$$f_1(x) * f_2(x) * \dots * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^K N_k} \prod_{k=1}^K \det(\Sigma_k)}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x - \mu)\right) * \exp(M) \dots \dots \dots < 2 >$$

根据上式二次项一次项对应参数，可以得到：

$$\Sigma^{-1} = \sum_{k=1}^K \Sigma_k^{-1}$$

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

为了满足归一化条件，需要将变量指数项外的其他常数项归一到 η 中，也即证明K个独立正态分布随机变量概率密度相乘归一化之后仍为正态分布

2.5.7

假设有K个互相独立的随机变量 x_k ，它们通过加权组成一个新的随机变量：

$$x = \sum_{k=1}^K \omega_k x_k$$

其中 $\sum_{k=1}^K \omega_k = 1$ 且 $\omega_k \geq 0$ ，它们的期望表示为：

$$\mu = \sum_{k=1}^K \omega_k \mu_k$$

其中 μ_k 是 $x_{\{k\}}$ 的均值，请定义出一个计算方差的表达式，注意，这些随机变量并没有假设服从高斯分布

solution:

统计学上有公式：

对于独立随机变量 X, Y

$$D(\omega_x X + \omega_y Y) = \omega_x^2 D(X) + \omega_y^2 D(Y) \dots \dots < 0 >$$

这里假设 x_k 的方差为 σ_k^2

则方差的计算公式为：

$$\sigma^2 = \sum_{i=1}^K \omega_k^2 \sigma_k^2$$

其中 $\sum_{k=1}^K \omega_k = 1$

2.5.8

当K维随机变量x服从标准正态分布，即 $x \sim N(0,1)$,则随机变量：

$$y = x^T x$$

服从自由度为K的卡方分布，请证明该随机变量的均值为K，方差为2K(题目条件暗含每一维度随机变量独立同分布假设，远书为准确提及)

solution:

$$y = x_1^2 + x_2^2 + \dots + x_K^2$$

$$E(y) = E(x_1^2 + x_2^2 + \dots + x_K^2)$$

$$= E(x_1^2) + E(x_2^2) + \dots + E(x_K^2)$$

根据统计学：

$$E(X^2) = D(X) + (E(X))^2$$

因此对于任意 $1 \leq i \leq K$:

$$E(x_k^2) = 1 + 0 = 1$$

$$E(y) = K$$

根据Isserlis定理：

$$E[x_i x_j x_k x_\ell] = E[x_i x_j] E[x_k x_\ell] + E[x_i x_k] E[x_j x_\ell] + E[x_i x_\ell] E[x_j x_k] \dots < 0 >$$

方差：

$$D(y)$$

$$= E((x_1^2 + x_2^2 + \dots x_K^2)^2)$$

$$= E(\sum_{i=1}^K x_i^4) + 2E(\sum_{i=1, j=1, i \neq j}^K E(x_i^2 x_j^2)) \dots < 1 >$$

根据<0>,其中：

$$E(\sum_{i=1}^K x_i^4)$$

$$= \sum_{i=1}^K E(x_i^4)$$

$$= 3K$$

$$E(\sum_{i=1, j=1, i \neq j}^K E(x_i^2 x_j^2))$$

$$= \frac{K(K-1)}{2}$$

将上述两式带入<0>:

$$D(y) = 2K$$

线性高斯系统估计

3.6.1

考虑时间离散系统:

$$x_k = x_{k-1} + v_k + \omega_k, \omega \text{服从} N(0, Q) \text{正态分布}$$

$$y_k = x_k + n_k, n_k \text{服从} N(0, R) \text{正态分布}$$

这可以表达一辆沿x轴前进或者后退的汽车，初始状态 \tilde{x}_0 未知，请建立批量最小二乘的状态估计方程：

$$(H^T W^{-1} H) \hat{x} = H^T W^{-1} z$$

即推导出 H, W, z 和 \hat{x} 的详细形式。令最大时间步数为 $K = 5$,并假设所有噪声互相无关，该问题存在唯一解吗？

solution:

本题思路：根据(3.40)的做法，因为没有初始状态的先验，因此将初始状态项在计算中全部略去，也就是删除矩阵中对应的行、块。

根据已知条件任意时刻 $A_{k=0,1,2,3,4} = 1, C_{k=0,1,2,3,4,5} = 1$

根据公式(3.12):

$$z = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}$$

这里删除了初始状态项，但是保留了初始时刻的观测 y_0 ,因为机器人可以在不知道自己初始位置的情况下，进行观测

同样的方法，根据式(3.13b):

$$W = \begin{bmatrix} Q_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Q_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Q_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_5 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R_0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & R_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R_6 & 0 \end{bmatrix}$$

也即:

$$W = \text{diag}(Q_1, Q_2, Q_3, Q_4, Q_5, R_0, R_1, R_2, R_3, R_4, R_5)$$

因此其逆矩阵:

$$W^{-1} = \text{diag}(Q_1^{-1}, Q_2^{-1}, Q_3^{-1}, Q_4^{-1}, Q_5^{-1}, R_0^{-1}, R_1^{-1}, R_2^{-1}, R_3^{-1}, R_4^{-1}, R_5^{-1})$$

根据(3.32),(3.33)---相对原书中公式需要删除初始状态对应那一列数据:

$$H = \begin{bmatrix} A^{-1} \\ C \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ -A_1 & 1 & 0 & 0 & 0 \\ 0 & -A_2 & 1 & 0 & 0 \\ 0 & 0 & -A_3 & 1 & 0 \\ 0 & 0 & 0 & -A_4 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ R_1 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 \\ 0 & 0 & 0 & 0 & R_5 \end{bmatrix}$$

因此:

$$H^T W^{-1} H$$

$$= \begin{bmatrix} Q_1^{-1} + Q_2^{-1} & -Q_2^{-1} & 0 & 0 & 0 \\ -Q_2^{-1} & Q_2^{-1} + Q_3^{-1} + R_2^{-1} & -Q_3^{-1} & 0 & 0 \\ 0 & -Q_3^{-1} & Q_3^{-1} + Q_4^{-1} + R_3^{-1} & -Q_4^{-1} & 0 \\ 0 & 0 & -Q_4^{-1} & Q_4^{-1} + Q_5^{-1} + R_4^{-1} & -Q_5^{-1} \\ 0 & 0 & 0 & -Q_5^{-1} & -Q_5^{-1} + R_5^{-1} \end{bmatrix}$$

根据(3.46)(3.47)判断存在唯一解问题

此问题中 $N=1$, 因此:

观测矩阵 $O = C$

$$\text{rank}(O) = \text{rank}(C) = 6$$

$$\text{rank}(O) \neq N$$

因此不存在唯一解

3.6.2

使用第一题的系统, 令 $Q = R = 1$,证明:

$$\mathbf{H}^T \mathbf{W}^{-1} \mathbf{H} = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & 0 \\ 0 & -1 & 3 & -1 & 0 \\ 0 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix}$$

此时Cholesky因子是什么, 才能满足 $LL^T = H^T W^{-1} H$?

solution:

证明只需将数据代入上一题目的公式中即可

这里假设：

$$L = \begin{bmatrix} L_1 & 0 & 0 & 0 & 0 \\ L_{21} & L_2 & 0 & 0 & 0 \\ 0 & L_{32} & L_3 & 0 & 0 \\ 0 & 0 & L_{43} & L_4 & 0 \\ 0 & 0 & 0 & L_{54} & L_5 \end{bmatrix}$$

在本问题中，L矩阵的中的每一项都是标量(0维张量)，因此：

$$L^T = \begin{bmatrix} L_1 & L_{21} & 0 & 0 & 0 \\ 0 & L_2 & L_{32} & 0 & 0 \\ 0 & 0 & L_3 & L_{43} & 0 \\ 0 & 0 & 0 & L_4 & L_{54} \\ 0 & 0 & 0 & 0 & L_5 \end{bmatrix}$$

将矩阵相乘：

$$LL^T = \begin{bmatrix} L_1^2 & L_1L_{21} & 0 & 0 & 0 \\ L_1L_{21} & L_2^2 + L_{21}^2 & L_2L_{32} & 0 & 0 \\ 0 & L_2L_{32} & L_3^2 + L_{32}^2 & L_3L_{43} & 0 \\ 0 & 0 & L_3L_{43} & L_4^2 + L_{43}^2 & L_4L_{54} \\ 0 & 0 & 0 & L_4L_{54} & L_5^2 + L_{54}^2 \end{bmatrix}$$

从矩阵最左边开始迭代就可以结算处L(注意这个矩阵元素的结算因为涉及多个开方，存在多个解，这里仅给出一个):

$$L = \begin{bmatrix} \sqrt{2} & 0 & 0 & 0 & 0 \\ -\sqrt{\frac{1}{2}} & \sqrt{\frac{5}{2}} & 0 & 0 & 0 \\ 0 & -\sqrt{\frac{2}{5}} & \sqrt{\frac{13}{5}} & 0 & 0 \\ 0 & 0 & -\sqrt{\frac{5}{13}} & \sqrt{\frac{34}{13}} & 0 \\ 0 & 0 & 0 & -\sqrt{\frac{13}{34}} & \sqrt{\frac{55}{34}} \end{bmatrix}$$

参考文献：

1.Matrix Cookbook---Kaare Brandt Petersen