

# 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:**

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

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

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

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

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

### 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 $\mu$ 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:

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

由于协方差矩阵是对称矩阵, 根据等式< 1 >:

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

对于<2>式变换:

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

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 $\eta$ 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:

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

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

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;

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

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

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 $\eta$ 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 $\mu_k$ 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 $X, Y$

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

这里假设 $x_k$ 的方差为 $\sigma_k^2$

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:**

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

$$\begin{aligned} E(y) &= E(x_1^2 + x_2^2 + \dots + x_K^2) \\ &= E(x_1^2) + E(x_2^2) + \dots + E(x_K^2) \end{aligned}$$

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：

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

方差：

$$\begin{aligned} D(y) &= E((x_1^2 + x_2^2 + \dots x_K^2)^2) \\ &= E(\sum_{i=1}^K x_i^4) + 2E(\sum_{i=1, j=1, i \neq j}^K E(x_i^2 x_j^2)) \dots \dots \dots < 1 > \end{aligned}$$

根据<0>,其中：

$$\begin{aligned} &E(\sum_{i=1}^K x_i^4) \\ &= \sum_{i=1}^K E(x_i^4) \\ &= 3K \\ &E(\sum_{i=1, j=1, i \neq j}^K E(x_i^2 x_j^2)) \\ &= \frac{K(K-1)}{2} \end{aligned}$$

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条件下，进行观测

同样的方法，根据式(3.13b):



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

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：

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

$$= \begin{bmatrix} 2Q^{-1} & -Q^{-1} & 0 & 0 & 0 \\ -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} & 0 & 0 \\ 0 & -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} & 0 \\ 0 & 0 & -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} \\ 0 & 0 & 0 & -Q^{-1} & Q^{-1} + R^{-1} \end{bmatrix}$$

$$Q > 0, R > 0$$

$W^{-1}$ 对称且正定，根据(3.37)做修改，因为删除了初始状态，只需：

$\text{rank}(H^T H) = \text{rank}(H^T) = NK = 5$ 就可以构成唯一解充分条件

显然满足条件

因此，该系统存在唯一解

## 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：

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

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):

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

### 3.6.3

使用第一题的系统，修改最小二乘解，假设噪声之间存在相关性：

$$E[y_k y_\ell] = \begin{cases} R & |k - \ell| = 0 \\ R/2 & |k - \ell| = 1 \\ R/4 & |k - \ell| = 2 \\ 0 & \text{otherwise} \end{cases}$$

此时存在唯一的最小二乘解吗？

**solution:**

$$W = \begin{bmatrix} Q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Q & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Q & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R & R/2 & R/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R/2 & R & R/2 & R/4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & R/4 & R/2 & R & R/2 & R/4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & R/4 & R/2 & R & R/2 & R/4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & R/4 & R/2 & R & R/2 & R/4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & R/4 & R/2 & R & R/2 \end{bmatrix}$$

等价于

$$W = \begin{bmatrix} Q & 0 \\ 0 & R^* \end{bmatrix}$$

其中:

$$R^* = R \begin{bmatrix} 1 & 1/2 & 1/4 & 0 & 0 & 0 \\ 1/2 & 1 & 1/2 & 1/4 & 0 & 0 \\ 1/4 & 1/2 & 1 & 1/2 & 1/4 & 0 \\ 0 & 1/4 & 1/2 & 1 & 1/2 & 1/4 \\ 0 & 0 & 1/4 & 1/2 & 1 & 1/2 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1 \end{bmatrix}$$

$R^*$ 可逆

因此 $W^{-1}$ 存在

根据(3.37):

$$\text{rank}(H^T H) = \text{rank}(H^T) = NK = 5$$

存在唯一解

### 3.6.4

使用第一题的系统, 推导卡尔曼滤波器的详细过程。本例的初始状态均值为 $\check{x}_0$ 方差为 $\check{P}_0$ , 证明: 稳态时先验和后验方差 $\check{P}$ 和 $\hat{P}$ , 当 $K \rightarrow \infty$ 为以下两个方程组的解:

$$\check{P}^2 - Q\check{P} - QR = 0$$

$$\hat{P}^2 + Q\hat{P} - QR = 0$$

此二式是离散Riccati方程的两个不同版本, 同时, 解释为什么这两个二次方程仅有一个是物理上可行的。

**solution:**

具有先验信息的条件下:

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

$$= \begin{bmatrix} \check{P}_0^{-1} + Q^{-1} + R^{-1} & -Q^{-1} & 0 & 0 & 0 & 0 \\ -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} & 0 & 0 & 0 \\ 0 & -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} & 0 & 0 \\ 0 & 0 & -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} & 0 \\ 0 & 0 & 0 & -Q^{-1} & 2Q^{-1} + R^{-1} & -Q^{-1} \\ 0 & 0 & 0 & 0 & -Q^{-1} & Q^{-1} + R^{-1} \end{bmatrix}$$

$$L = \begin{bmatrix} L_0 & 0 & 0 & 0 & 0 & 0 \\ L_{10} & L_1 & 0 & 0 & 0 & 0 \\ 0 & L_{21} & L_2 & 0 & 0 & 0 \\ 0 & 0 & L_{32} & L_3 & 0 & 0 \\ 0 & 0 & 0 & L_{43} & L_4 & 0 \\ 0 & 0 & 0 & 0 & L_{54} & L_5 \end{bmatrix}$$

$$L^T = \begin{bmatrix} L_0^T & L_{10}^T & 0 & 0 & 0 & 0 \\ 0 & L_1^T & L_{21}^T & 0 & 0 & 0 \\ 0 & 0 & L_2^T & L_{32}^T & 0 & 0 \\ 0 & 0 & 0 & L_3^T & L_{43}^T & 0 \\ 0 & 0 & 0 & 0 & L_4^T & L_{54}^T \\ 0 & 0 & 0 & 0 & 0 & L_5^T \end{bmatrix}$$

$$= \begin{bmatrix} L_0 & L_{10} & 0 & 0 & 0 & 0 \\ 0 & L_1 & L_{21} & 0 & 0 & 0 \\ 0 & 0 & L_2 & L_{32} & 0 & 0 \\ 0 & 0 & 0 & L_3 & L_{43} & 0 \\ 0 & 0 & 0 & 0 & L_4 & L_{54} \\ 0 & 0 & 0 & 0 & 0 & L_5 \end{bmatrix}$$

对于:

$$(k = 1 \dots K)$$

$$\mathbf{L}_{k-1} \mathbf{L}_{k-1}^T = \mathbf{I}_{k-1} + \mathbf{A}_{k-1}^T \mathbf{Q}_k^{-1} \mathbf{A}_{k-1}$$

$$\mathbf{L}_{k-1} \mathbf{d}_{k-1} = \mathbf{q}_{k-1} - \mathbf{A}_{k-1}^T \mathbf{Q}_k^{-1} \mathbf{v}_k$$

$$\mathbf{L}_{k,k-1} \mathbf{L}_{k-1}^T = -\mathbf{Q}_k^{-1} \mathbf{A}_{k-1}$$

$$\mathbf{I}_k = -\mathbf{L}_{k,k-1} \mathbf{L}_{k,k-1}^T + \mathbf{Q}_k^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k$$

$$\mathbf{q}_k = -\mathbf{L}_{k,k-1} \mathbf{d}_{k-1} + \mathbf{Q}_k^{-1} \mathbf{v}_k + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k$$

根据本系统修正:

$$(k = 1 \dots K)$$

$$\mathbf{L}_{k-1} \mathbf{L}_{k-1} = \mathbf{I}_{k-1} + \mathbf{Q}^{-1}$$

$$\mathbf{L}_{k-1} \mathbf{d}_{k-1} = \mathbf{q}_{k-1} - \mathbf{Q}^{-1} \mathbf{v}_k$$

$$\mathbf{L}_{k,k-1} \mathbf{L}_{k-1} = -\mathbf{Q}^{-1}$$

$$\mathbf{I}_k = -\mathbf{L}_{k,k-1} \mathbf{L}_{k,k-1} + \mathbf{Q}^{-1} + \mathbf{R}^{-1}$$

$$\mathbf{q}_k = -\mathbf{L}_{k,k-1} \mathbf{d}_{k-1} + \mathbf{Q}^{-1} \mathbf{v}_k + \mathbf{C}^T \mathbf{R}^{-1} \mathbf{y}_k$$

$$\mathbf{I}_0 = \check{\mathbf{P}}_0^{-1} + \mathbf{R}^{-1}$$

$$\mathbf{q}_0 = \check{\mathbf{P}}_0^{-1} \check{\mathbf{x}}_0 + \mathbf{R}^{-1} \mathbf{y}_0$$

$$\hat{\mathbf{x}}_K = \mathbf{L}_K^{-1} \mathbf{d}_K$$

卡尔曼滤波推导完毕

根据(3.28)

当  $K \rightarrow \infty$  处于稳态时:

$$\check{P} = \hat{P} + Q - - - - - < 0 >$$

$$K = \check{P}(\check{P} + R)^{-1} \text{-----} < 1 >$$

$$\hat{P} = (1 - K)\check{P} \text{-----} < 2 >$$

将<1><2>带入<0>:

$$\check{P} = (1 - K)\check{P} + Q = (1 - \frac{\check{P}}{\check{P}+R})\check{P} + Q$$

整理得:

$$\check{P}^2 - Q\check{P} - QR = 0$$

将<0><1>带入<2>得:

$$\hat{P} = (1 - K)(\hat{P} + Q) = (1 - \frac{\hat{P}+Q}{\hat{P}+Q+R})(\hat{P} + Q)$$

整理得:

$$\hat{P}^2 + Q\hat{P} - QR = 0$$

### 3.6.5

使用3.3.2节的MAP方法, 推导后向的卡尔曼滤波器(而非前向的)

**solution:**

反向卡尔曼滤波应该是在得到 $k$ 时刻的控制和观测之后, 对 $K - 1$ 时刻的状态进行估计

根据(3.109)式, 也即计算 $\hat{x}'_{k-1}$

原理示意图参考书中3-4图

根据(3.110):

$$(\hat{P}_{k-1}^{-1} + A_{k-1}^T Q_k^{-1} A_{k-1}) \hat{x}'_{k-1} - A_{k-1}^T Q_k^{-1} \hat{x}_k = \hat{P}_{k-1}^{-1} \hat{x}_{k-1} - A_{k-1}^T Q_k^{-1} v_k \text{-----} < 0 >$$

根据正向卡尔曼滤波器的递推公式(3.120):

带入(3.120a)(3.120b)(3.120c)(3.120e)式:

$$\begin{aligned} (\hat{P}_{k-1}^{-1} + A_{k-1}^T Q_k^{-1} A_{k-1}) \hat{x}'_{k-1} &= A_{k-1}^T Q_k^{-1} \hat{x}_k + \hat{P}_{k-1}^{-1} \hat{x}_{k-1} - A_{k-1}^T Q_k^{-1} v_k \\ &= A_{k-1}^T Q_k^{-1} A_{k-1} y_k - A_{k-1}^T Q_k^{-1} A_{k-1} C_k v_k + A_{k-1}^T Q_k^{-1} (1 - K_k C_k) A_{k-1} \hat{x}_{k-1} \end{aligned}$$

令:

$$K_{k,b} = (\hat{P}_{k-1}^{-1} + A_{k-1}^T Q_k^{-1} A_{k-1})^{-1} A_{k-1}^T Q_k^{-1}$$

$$K_{k,b,y} = K_{k,b} K_k$$

$$K_{k,b,v} = -K_{k,b} K_k C_k$$

$$K_{k,b,x} = K_{k,b} (1 - K_k C_k) A_{k-1}$$

由此反向卡尔曼滤波:

$$\hat{x}'_{k-1} = K_{k,b,y} y_k + K_{k,b,v} v_k + K_{k,b,x} \hat{x}_{k-1}$$

上式中  $y_k, v_k$  融合了  $k$  时刻的观测和控制

$\hat{x}_{k-1}$  融合了  $0 \sim k-1$  时刻的信息

## 3.5.6

证明:

$$\begin{bmatrix} 1 & & & & \\ \mathbf{A} & 1 & & & \\ \mathbf{A}^2 & \mathbf{A} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \mathbf{A}^{K-3} & \dots & 1 \\ \mathbf{A}^K & \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \dots & \mathbf{A} & 1 \end{bmatrix} = \begin{bmatrix} 1 & & & & \\ -\mathbf{A} & 1 & & & \\ & -\mathbf{A} & 1 & & \\ & & -\mathbf{A} & \ddots & \\ & & & \ddots & 1 \\ & & & & -\mathbf{A} & 1 \end{bmatrix}$$

solution:

$$\begin{bmatrix} 1 & & & & \\ \mathbf{A} & 1 & & & \\ \mathbf{A}^2 & \mathbf{A} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \mathbf{A}^{K-3} & \dots & 1 \\ \mathbf{A}^K & \mathbf{A}^{K-1} & \mathbf{A}^{K-2} & \dots & \mathbf{A} & 1 \end{bmatrix} \begin{bmatrix} 1 & & & & \\ -\mathbf{A} & 1 & & & \\ & -\mathbf{A} & 1 & & \\ & & -\mathbf{A} & \ddots & \\ & & & \ddots & 1 \\ & & & & -\mathbf{A} & 1 \end{bmatrix} = E$$

此处  $E$  为单位矩阵, 即证

## 3.5.7

我们已经介绍了在批量最小二乘解中, 后验协方差:

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

同时我们也知道, Cholesky 分解:

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

的计算代价  $O(N(K+1))$ , 这是由于系统具备稀疏性, 反之, 我们有:

$$\hat{P} = L^{-T} L^{-1}$$

请说明这种计算方法计算  $\hat{P}$  的复杂度。

solution:

假设 $L$ 逆矩阵的已经计算得到，由于 $LL^T$ 矩阵相乘得到的特殊稀疏结构，

$\hat{P}$ 为 $(K+1) \times (K+1)$ 的二维矩阵

计算 $\hat{P}$ 的复杂度：

$$\begin{aligned} \text{计算频数} &= N([K+1] + [(K+1) + (K)] + \dots + [(K+1) + (K) + (K-1) + \dots + 1]) + \\ &+ ([K] + [(K) + (K-1)] + \dots + [(K) + (K-1) + \dots + 1]) \\ &= N(\sum_{i=1}^{K+1} i^2 + \sum_{i=1}^K i^2) \\ &= N(\frac{(K+1)(K+2)(2K+3)}{6} + \frac{K(K+1)(2K+1)}{6}) \end{aligned}$$

因此计算复杂度为 $O(NK^3)$

## 4.非线性非高斯系统的状态估计

### 4.6.1

考虑如下离散时间系统：

$$\begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} x_{k-1} \\ y_{k-1} \\ \theta_{k-1} \end{bmatrix} + T \begin{bmatrix} \cos\theta_{k-1} & 0 \\ \sin\theta_{k-1} & 0 \\ 0 & 1 \end{bmatrix} \left( \begin{bmatrix} v_k \\ \omega_k \end{bmatrix} + \boldsymbol{\omega}_k \right)$$

其中 $\boldsymbol{\omega}_k \sim N(0, Q)$

$$\begin{bmatrix} r_k \\ \phi_k \end{bmatrix} = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} \\ \text{atan2}(-y_k, -x_k) - \theta_k \end{bmatrix} + \boldsymbol{n}_k$$

$\boldsymbol{n}_k \sim N(0, R)$

该系统可以看做移动机器人在xy平面上移动，测量值为移动机器人距离原点的距离和方位，请建立EFK方程来估计移动机器人的位姿，并写出雅可比 $F_{k-1}$ ， $G_k$ 和协方差 $Q'_k$ ， $R'_k$ 的表达式。

**solution:**

状态方程：

$$\begin{bmatrix} x_k \\ y_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} x_{k-1} + Tv_k \cos\theta_{k-1} + \omega_{xk} \\ y_{k-1} + T\omega_k \sin\theta_{k-1} + \omega_{yk} \\ \theta_{k-1} + T\omega_k + \omega_{\theta k} \end{bmatrix}$$

这里(注意这里加粗符号与普通符号的区别，尤其在控制部分的 $\omega$ 容易跟噪声混淆)：

$$\boldsymbol{\omega}_k = \begin{bmatrix} \omega_{xk} \\ \omega_{yk} \\ \omega_{\theta k} \end{bmatrix}$$



根据4-25:

$$F_{k-1} = \begin{bmatrix} 1 & 0 & -Tv_k \sin \theta_{k-1} \\ 0 & 1 & T\omega_k \cos \theta_{k-1} \\ 0 & 0 & 1 \end{bmatrix}$$

$$\omega'_k = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_{xk} \\ \omega_{yk} \\ \omega_{\theta k} \end{bmatrix} = \omega_k$$

测量方程:

$$\begin{bmatrix} r_k \\ \phi_k \end{bmatrix} = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} + n_{rk} \\ \text{atan2}(-y_k, -x_k) - \theta_k + n_{\phi k} \end{bmatrix}$$

其中:

$$n_k = \begin{bmatrix} n_{rk} \\ n_{\phi k} \end{bmatrix}$$

$$G_k = \begin{bmatrix} \frac{x_k}{\sqrt{x_k^2 + y_k^2}} & \frac{y_k}{\sqrt{x_k^2 + y_k^2}} & 0 \\ \frac{-y_k}{\sqrt{x_k^2 + y_k^2}} & \frac{x_k}{\sqrt{x_k^2 + y_k^2}} & -1 \end{bmatrix}$$

$$n'_k = n_k$$

$$Q'_k = E(\omega'_k \omega'^T_k) = Q$$

$$R'_k = E(n'_k n'^T_k) = R$$

EFK计算公式将上面计算结果带入式(4.32)即可

## 4.6.2

考虑将高斯先验 $N(\mu_x, \sigma_x^2)$ 传递进非线性函数 $f(x) = x^3$ 中, 请使用蒙特卡洛、线性化和sigmapoint变换方法来确定变换后的均值和协方差, 并对结果进行评价, 提示: 使用Isserlis定理计算高阶矩。

**solution:**

蒙特卡洛方法:

$$x_i = \mu_x + \delta x_i$$

$$\delta x_i \sim N(0, \sigma_x^2)$$

$$y_i = f(x_i) = (\mu_x + \delta x_i)^3 = \mu_x^3 + 3\mu_x^2 \delta x_i + 3\mu_x (\delta x_i)^2 + (\delta x_i^3)$$

根据Isserlis定理:

$$E(\delta x_i) = 0$$

$$E((\delta x_i)^2) = \sigma_x^2$$

$$E((\delta x_i)^3) = 0$$

$$E((\delta x_i)^4) = 3\sigma_x^4$$

$$E((\delta x_i)^5) = 0$$

$$(\delta x_i)^6 = 15\sigma_x^6$$

$$\mu_y = E(y_i) = 4\mu_x^3 + 3\mu_x\sigma_x^2 + 0 = 4\mu_x^3 + 3\mu_x\sigma_x^2$$

$$\sigma_y^2 = E((y_i - \mu_y)^2) = E((3\mu_x^2\delta x_i + (\delta x_i)^3)^2)$$

$$= 9\mu_x^4\sigma_x^2 + 15\sigma_x^6 + 18\mu_x^2\sigma_x^4$$

线性化方法:

$$y_i = f(\mu_x + \delta x_i) \approx \mu_x^3 + 3\mu_x^2\delta x_i$$

$$\mu_y = E(y_i) = \mu_x^3$$

$$\sigma_y^2 = E((y_i - \mu_y)^2) = 9\mu_x^4\sigma_x^2$$

*sigmapoint*方法:

$$L=1$$

$$x_0 = \mu_x$$

$$x_1 = \mu_x + \sqrt{(1 + \kappa)}\sigma_x$$

$$x_2 = \mu_x - \sqrt{(1 + \kappa)}\sigma_x$$

$$y_0 = f(x_0) = \mu_x^3$$

$$y_1 = (\mu_x + \sqrt{(1 + \kappa)}\sigma_x)^3$$

$$y_2 = (\mu_x - \sqrt{(1 + \kappa)}\sigma_x)^3$$

$$\mu_y = \frac{1}{\kappa+1}(\kappa y_0 + 0.5 \sum_{i=1}^2 y_i) = \mu_x^3 + 3\mu_x\sigma_x^2$$

$$\sigma_y^2 = \frac{1}{1+\kappa}(\kappa(y_0 - \mu_y)^2 + 0.5 \sum_{i=1}^2 (y_i - \mu_y)^2) = 9\kappa\mu_x^2\sigma_x^4 + (3\mu_x^2\sigma_x + (1 + \kappa)\sigma_x^3)^2$$

## 4.6.3

略

## 4.6.4

在sigmapoint卡尔曼滤波部分，当观测模型和测量噪声线性相关时，测量协方差为：

$$\sum_{yy,k} = \sum_{j=0}^{2N} \beta_j (\check{y}_{k,j} - \mu_{y,k})(\check{y}_{k,j} - \mu_{y,k})^T + R_k$$

请验证这个方程也可以写成：

$$\sum_{yy,k} = Z_k Z_k^T + R_k$$

其中

$$col_j Z_k = \sqrt{\beta_j} (\check{y}_{k,j} - \mu_{y,k})$$

**solution:**

令：

$$v_j = \check{y}_{k,j} - \mu_{y,k} = (x_{j0}, x_{j1}, \dots, x_{jM})^T$$

则：

$$Z = (\sqrt{\beta_0} v_0, \sqrt{\beta_1} v_1, \dots, \sqrt{\beta_{2N}} v_{2N})$$

因此：

$$(Z Z^T)_{p,q} = \sum_{j=0}^{2N} \beta_j x_{j,p} x_{j,q}$$

$$\sum_{j=0}^{2N} \beta_j (\check{y}_{k,j} - \mu_{y,k})(\check{y}_{k,j} - \mu_{y,k})^T = Y$$

$$Y_{p,q} = \sum_{j=0}^{2N} \beta_j x_{j,p} x_{j,q}$$

即证

## 4.6.5

不会

## 参考文献：

1.Matrix Cookbook---Kaare Brandt Petersen