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: $u^Tv=tr(vu^T)$

solution:

$$u = (x_1, x_2, ..., x_n)^T$$

$$v = (y_1, y_2, ..., y_n)^T$$

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

$$uv^T = egin{bmatrix} x_1y_1 & \cdots & \cdots & \cdots \ \cdots & x_2y_2 & \cdots & \cdots \ dots & dots & \ddots & dots \ \cdots & \cdots & \cdots & x_ny_n \end{bmatrix}$$

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

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:

$$egin{aligned} &= -E_{(x,y)}(ln(f(x,y))) \ &= -\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) ln(f(x,y)) dx dy \end{aligned}$$

因为x,y独立

$$\begin{split} &H(x,y)\\ &=-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x)f(y)[ln(f(x))+ln(f(y))]dxdy\\ &=-[\int_{-\infty}^{\infty}(f(x)ln(f(x))dx]*\int_{-\infty}^{\infty}f(y)dy-[\int_{-\infty}^{\infty}f(y)ln(f(y))dy]*\int_{-\infty}^{\infty}f(x)dx\\ &=-\int_{-\infty}^{\infty}(f(x)ln(f(x))dx-\int_{-\infty}^{\infty}f(y)ln(f(y))dy\\ &=H(x)+H(y) \end{split}$$

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:

$$\begin{split} & \Sigma \\ &= E[(x-\mu)(x-\mu)^T] \\ &= E(xx^T - x\mu^T - \mu x^T + \mu \mu^T) \\ &= E(xx^T) - E(x)\mu^T - \mu E(x^T) + \mu \mu^T \\ & \boxtimes \Sigma = E(xx^T) - \mu \mu^T \end{split}$$

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:

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

solution:

E(x)

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

做变换:

$$y = x - \mu$$

可得:

$$x = y + \mu$$

E(x)

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

$$=\int_{-\infty}^{\infty}rac{y}{\sqrt{(2\pi)^Ndet(\Sigma)}}exp(-rac{1}{2}y^T\Sigma^{-1}y)dy+\int_{-\infty}^{\infty}rac{\mu}{\sqrt{(2\pi)^Ndet(\Sigma)}}exp(-rac{1}{2}y^T\Sigma^{-1}y)dy$$

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:

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

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

做代换 $y = x - \mu$

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

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

下面参考文献【1】中公式(108)如下式:

$$\frac{\partial}{\partial X}(X^TBX) = BX + B^TX$$

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:

$$\frac{d}{dx}(x^TBx) = Bx + B^Tx....<1>$$

由于协方差矩阵是对称矩阵,根据等式<1>:

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

对于<2>式变换:

将<3>式带入<0>式:

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

分步积分法:

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

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:

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

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:

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

$$f_1(x) * f_2(x) * ... * f_K(x)$$

$$=rac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_{k}}\prod_{k=1}^{K}det(\Sigma_{k})}}exp(-rac{1}{2}\sum_{k=1}^{K}(x-\mu_{k})^{T}\Sigma_{k}^{-1}(x-\mu))$$

将指数部分的求和号展开:

$$f_1(x) * f_2(x) * ... * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_{k}}\prod_{k=1}^{K}\det(\Sigma_{k})}} exp(-\frac{1}{2}(x^{T}(\sum_{k=1}^{K}\Sigma_{k}^{-1})x - (\sum_{k=1}^{K}\mu_{k}^{T}\Sigma_{k}^{-1})x - x^{T}\sum_{k=1}^{K}\Sigma_{i}^{-1}\mu_{i} + \sum_{k=1}^{K}\mu_{k}^{T}\Sigma_{k}^{-1}\mu_{k}))....<0>$$

因为协方差矩阵是对称矩阵, <0>式中

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

因此:

$$f_1(x) * f_2(x) * ... * f_K(x)$$

$$= \frac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_{k}}\prod_{k=1}^{K}\det(\Sigma_{k})}} exp(-\frac{1}{2}(x^{T}(\sum_{k=1}^{K}\Sigma_{k}^{-1})x - 2x^{T}\sum_{k=1}^{K}\Sigma_{i}^{-1}\mu_{i} + \sum_{k=1}^{K}\mu_{k}^{T}\Sigma_{k}^{-1}\mu_{k})).....<1>$$

在式<1>中:

$$x^T(\sum_{k=1}^K \Sigma_k^{-1}) x$$
为二次项 $2x^T\sum_{k=1}^K \Sigma_i^{-1} \mu_i$ 为一次项

可以凑出"完全平方形式"

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

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

上式中M为一个常数:

$$f_1(x) * f_2(x) * ... * f_K(x)$$

$$=rac{1}{\sqrt{(2\pi)^{\sum_{k=1}^{K}N_{k}}\prod_{k=1}^{K}det(\Sigma_{k})}}exp(-rac{1}{2}(x-\mu)^{T}\Sigma^{-1}(x-\mu))*\ exp(M)...<2>$$

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

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

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:

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

根据统计学:

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

因此对于任意1 < i < K:

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

$$E(y) = K$$

根据Isserlis定理:

$$E\left[x_{i}x_{j}x_{k}x_{\ell}
ight]=E\left[x_{i}x_{j}
ight]E\left[x_{k}x_{\ell}
ight]+E\left[x_{i}x_{k}
ight]E\left[x_{j}x_{\ell}
ight]+E\left[x_{j}x_{k}
ight]......<0>$$

方差:

$$=E((x_1^2+x_2^2+...x_K^2)^2)$$

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

根据<0>,其中:

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

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

=3K

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

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

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:

$$x_k = x_{k-1} + v_k + \omega_k, \omega$$
服从 $N(0,Q)$ 正态分布 $y_k = x_k + n_k, n_k$ 服从 $N(0.R)$ 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 \hat{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):

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

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):

也即:

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

因此其逆矩阵:

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

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:

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

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

因此:

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

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: $rank(H^TH)=rank(H^T)=NK=5$ 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:

$$\mathbf{H}^T\mathbf{W}^{-1}\mathbf{H} = \left[egin{array}{cccccc} 2 & -1 & 0 & 0 & 0 \ -1 & 3 & -1 & 0 & 0 \ 0 & -1 & 3 & -1 & 0 \ 0 & 0 & -1 & 3 & -1 \ 0 & 0 & 0 & -1 & 2 \end{array}
ight]$$

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

solution:

证明只需将数据代入上一题目的公式中即可

这里假设:

$$L = \left[egin{array}{ccccc} L_1 & 0 & 0 & 0 & 0 \ L_{21} & L_2 & 0 & 0 & 0 \ 0 & L_{32} & L_3 & 0 & 0 \ 0 & 0 & L_{43} & L_4 & 0 \ 0 & 0 & 0 & L_{54} & L_5 \ \end{array}
ight]$$

在本问题中, L矩阵的中的每一项都是标量(0维张量), 因此:

$$L^T = \left[egin{array}{ccccc} L_1 & L_{21} & 0 & 0 & 0 \ 0 & L_2 & L_{32} & 0 & 0 \ 0 & 0 & L_3 & L_{43} & 0 \ 0 & 0 & 0 & L_4 & L_{54} \ 0 & 0 & 0 & 0 & L_5 \ \end{array}
ight]$$

将矩阵相乘:

$$LL^T = \left[egin{array}{ccccc} L_1^2 & L_1L_{21} & 0 & 0 & 0 \ L_1L_{21} & L_2^2 + L_{21}^2 & L_2L_{32} & 0 & 0 \ 0 & L_2L_{32} & L_3^2 + L_{32}^2 & L_3L_{43} & 0 \ 0 & 0 & L_3L_{43} & L_4^2 + L_{43}^2 & L_4L_{54} \ 0 & 0 & 0 & L_4L_{54} & L_5^2 + L_{54}^2 \end{array}
ight]$$

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):

$$L = \left[egin{array}{cccccc} \sqrt{2} & 0 & 0 & 0 & 0 \ -\sqrt{rac{1}{2}} & \sqrt{rac{5}{2}} & 0 & 0 & 0 \ 0 & -\sqrt{rac{5}{5}} & \sqrt{rac{13}{5}} & 0 & 0 \ 0 & 0 & -\sqrt{rac{5}{13}} & \sqrt{rac{34}{13}} & 0 \ 0 & 0 & 0 & -\sqrt{rac{13}{34}} & \sqrt{rac{55}{34}} \end{array}
ight]$$

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:

$$E\left[y_{k}y_{\ell}
ight] = \left\{egin{array}{l} R\left|k-\ell
ight| = 0 \ R/2|k-\ell| = 1 \ R/4|k-\ell| = 2 \ 0 \ ext{ otherwise} \end{array}
ight.$$

此时存在唯一的最小二乘解吗?

solution:

等价于

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

其中:

$$R^* = R \left[\begin{array}{ccccccc} 1 & 1/2 & 1/4 & 0 & 0 & 0 \\ 1/2 & 1 & 1/2 & 1/4 & 0 & 0 \\ 1/4 & 1/2 & 1 & 1/2 & 1/4 & 0 \\ 0 & 1/4 & 1/2 & 1 & 1/2 & 1/4 \\ 0 & 0 & 1/4 & 1/2 & 1 & 1/2 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1 \end{array} \right]$$

 R^* 可逆

因此 W^{-1} 存在

根据(3.37):

$$rank(H^TH) = rank(H^T) = NK = 5$$

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 $K\to\infty$ 为以下两个方程组的解:

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

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:

$$H^TW^{-1}H$$

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

$$L = \left[egin{array}{cccccc} L_0 & 0 & 0 & 0 & 0 & 0 & 0 \ L_{10} & L_1 & 0 & 0 & 0 & 0 \ 0 & L_{21} & L_2 & 0 & 0 & 0 \ 0 & 0 & L_{32} & L_3 & 0 & 0 \ 0 & 0 & 0 & L_{43} & L_4 & 0 \ 0 & 0 & 0 & 0 & L_{54} & L_5 \ \end{array}
ight]$$

$$L^T = \left[egin{array}{ccccccc} L_0^T & L_{10}^T & 0 & 0 & 0 & 0 \ 0 & L_1^T & L_{21}^T & 0 & 0 & 0 \ 0 & 0 & L_2^T & L_{32}^T & 0 & 0 \ 0 & 0 & 0 & L_3^T & L_{43}^T & 0 \ 0 & 0 & 0 & 0 & L_4^T & L_{54}^T \ 0 & 0 & 0 & 0 & 0 & L_5^T \end{array}
ight]$$

$$= \left[\begin{array}{cccccccc} L_0 & L_{10} & 0 & 0 & 0 & 0 \\ 0 & L_1 & L_{21} & 0 & 0 & 0 \\ 0 & 0 & L_2 & L_{32} & 0 & 0 \\ 0 & 0 & 0 & L_3 & L_{43} & 0 \\ 0 & 0 & 0 & 0 & L_4 & L_{54} \\ 0 & 0 & 0 & 0 & 0 & L_5 \end{array} \right]$$

对于:

$$egin{aligned} (k = 1 \dots K) \ & \mathbf{L}_{k-1} \mathbf{L}_{k-1}^T = & \mathbf{I}_{k-1} + \mathbf{A}_{k-1}^T \mathbf{Q}_k^{-1} \mathbf{A}_{k-1} \ & \mathbf{L}_{k-1} \mathbf{d}_{k-1} = & \mathbf{q}_{k-1} - \mathbf{A}_{k-1}^T \mathbf{Q}_k^{-1} \mathbf{v}_k \ & \mathbf{L}_{k,k-1} \mathbf{L}_{k-1}^T = & - \mathbf{Q}_k^{-1} \mathbf{A}_{k-1} \ & \mathbf{I}_k = & - \mathbf{L}_{k,k-1} \mathbf{L}_{k,k-1}^T + \mathbf{Q}_k^{-1} + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{C}_k \ & \mathbf{q}_k = & - \mathbf{L}_{k,k-1} \mathbf{d}_{k-1} + \mathbf{Q}_k^{-1} \mathbf{v}_k + \mathbf{C}_k^T \mathbf{R}_k^{-1} \mathbf{y}_k \end{aligned}$$

根据本系统修正:

$$egin{aligned} (k = 1 \dots K) \ & \mathbf{L}_{k-1} \mathbf{L}_{k-1} = & \mathbf{I}_{k-1} + \mathbf{Q}^{-1} \ & \mathbf{L}_{k-1} \mathbf{d}_{k-1} = & \mathbf{q}_{k-1} - \mathbf{Q}^{-1} \mathbf{v}_k \ & \mathbf{L}_{k,k-1} \mathbf{L}_{k-1} = & - \mathbf{Q}^{-1} \ & \mathbf{I}_k = & - \mathbf{L}_{k,k-1} \mathbf{L}_{k,k-1} + \mathbf{Q}^{-1} + \mathbf{R}^{-1} \ & \mathbf{q}_k = & - \mathbf{L}_{k,k-1} \mathbf{d}_{k-1} + \mathbf{Q}^{-1} \mathbf{v}_k + \mathbf{C}^T \mathbf{R}^{-1} \mathbf{y}_k \end{aligned}$$
 $\mathbf{I}_0 = \check{\mathbf{P}}_0^{-1} + \mathbf{R}^{-1}$

$$egin{aligned} \mathbf{I}_0 &= \check{\mathbf{P}}_0^{-1} + \mathbf{R}^{-1} \ \mathbf{q}_0 &= \check{\mathbf{P}}_0^{-1} \check{\mathbf{x}}_0 + \mathbf{R}^{-1} \mathbf{y}_0 \ \hat{\mathbf{x}}_K &= \mathbf{L}_K^{-1} \mathbf{d}_K \end{aligned}$$

卡尔曼滤波推导完毕

根据(3.28)

当 $K \to \infty$ 处于稳态时:

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

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

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

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):

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

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:

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

令:

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

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

$$K_{k,b,v} = -K_{k,b}K_kC_k$$

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

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~k-1时刻的信息

3.5.6

证明:

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

solution:

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

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)*(K+1)的二维矩阵

计算 \hat{P} 的复杂度:

计算频次=
$$N(([K+1]+[(K+1)+(K)]+...+[(K+1)+(K)+(K-1)+...+1])+([(K)]+[(K)+(K-1)]+...+[(K)+(K-1)+...+1]))$$

$$=N(\sum_{i=1}^{K+1}i^2+\sum_{i=1}^{K}i^2)$$

$$=N(\frac{(K+1)(K+2)(2K+3)}{6}+\frac{K(K+1)(2K+1)}{6})$$

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:

$$egin{bmatrix} x_k \ y_k \ heta_k \end{bmatrix} = egin{bmatrix} x_{k-1} \ y_{k-1} \ heta_{k-1} \end{bmatrix} + T egin{bmatrix} cos heta_{k-1} & 0 \ sin heta_{k-1} & 0 \ 0 & 1 \end{bmatrix} (egin{bmatrix} v_k \ \omega_k \end{bmatrix} + oldsymbol{\omega}_k)$$

其中 $\omega_k \sim N(0,Q)$

$$egin{bmatrix} r_k \ \phi_k \end{bmatrix} = egin{bmatrix} \sqrt{x_k^2 + y_k^2} \ atan2(-y_k, -x_k) - heta_k \end{bmatrix} + oldsymbol{n_k}$$

$$n_k \sim N(0, R)$$

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:

$$egin{bmatrix} x_k \ y_k \ heta_k \end{bmatrix} = egin{bmatrix} x_{k-1} + Tv_k cos heta_{k-1} + \omega_{k1} \ y_{k-1} + T\omega_k sin heta_{k-1} + \omega_{k1} \ heta_{k-1} + T\omega_k + \omega_{k2} \end{bmatrix}$$

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 ω 容易跟噪声混淆):

$$oldsymbol{\omega_k} = egin{bmatrix} \omega_{k1} \ \omega_{k2} \end{bmatrix}$$

根据4-25:

$$F_{k-1} = \left[egin{array}{ccc} 1 & 0 & -Tv_k sin heta_{k-1} \ 0 & 1 & Tv_k cos heta_{k-1} \ 0 & 0 & 1 \end{array}
ight]$$

$$oldsymbol{\omega_{k}^{'}} = \left[egin{array}{ccc} Tcon heta_{k-1} & 0 \ Tsin heta_{k-1} & 0 \ 0 & T \end{array}
ight] \left[egin{array}{c} \omega_{k1} \ \omega_{k2} \end{array}
ight]$$

测量方程:

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

其中:

$$egin{aligned} m{n_k} &= \left[egin{array}{c} n_{rk} \ n_{\phi k} \end{array}
ight] \ &G_k &= \left[egin{array}{ccc} rac{x_k}{x_k^2 + y_k^2} & rac{y_k}{x_k^2 + y_k^2} & 0 \ rac{-y_k}{x_k^2 + y_k^2} & rac{x_k}{x_k^2 + y_k^2} & -1 \end{array}
ight] \end{aligned}$$

$$oldsymbol{n_k'} = oldsymbol{n_k}$$

$$egin{aligned} Q_k^{'} &= E(\omega_k^{'}\omega_k^{'T}) = \left[egin{array}{ccc} Tcos heta_{k-1} & 0 \ Tsin heta_{k-1} & 0 \ 0 & T \end{array}
ight] E(\omega_k\omega_k^T) \left[egin{array}{ccc} Tcos heta_{k-1} & 0 \ 0 & T \end{array}
ight] \ \left[egin{array}{ccc} Tcos heta_{k-1} & 0 \ 0 & T \end{array}
ight] \end{aligned}$$

$$= \left[egin{array}{ccc} Tcos heta_{k-1} & 0 \ Tsin heta_{k-1} & 0 \ 0 & T \end{array}
ight]Q \left[egin{array}{ccc} Tcos heta_{k-1} & 0 \ Tsin heta_{k-1} & 0 \ 0 & T \end{array}
ight]$$

$$R_{k}^{'}=E(n_{k}^{'}n_{k}^{'T})=R$$

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:

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

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

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

sigmapoint方法:

$$x_0 = \mu_x$$

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

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

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

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

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

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

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

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:

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

请验证这个方程也可以写成:

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

其中

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

solution:

令:

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

则:

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

因此:

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

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

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

即证

4.6.5

不会

5.偏差、匹配和外点

5.5.1

考虑离散时间系统:

$$x_k = x_{k-1} + v_k + \bar{x_k}$$

$$d_k = x_k$$

其中 \bar{v}_k 是未知的输入偏差,请写出增广状态系统并确定该系统是否可观。

solution:

A=1,B=1,C=1

$$A^{'}=\left[egin{array}{cc} 1 & 1 \ 0 & 1 \end{array}
ight]$$

$$B^{'} = \begin{bmatrix} & 1 \\ & 0 \end{bmatrix}$$

$$N = 1, U = 1$$

$$C^{'}=\begin{bmatrix} & 1 & 0 \end{bmatrix}$$

$$O = C'A' = \begin{bmatrix} & 1 & 0 \\ & 1 & 1 \end{bmatrix}$$

$$rank(O) = 2 = N + U$$

增广系统可观测

5.5.2

考虑离散时间系统:

$$x_k = x_{k-1} + v_k$$

$$v_k = v_{k-1} + a_k$$

$$d_{1,k}=x_k$$

$$d_{2,k}=x_k+ar{d_k}$$

其中 $ar{d}_k$ 是未知的输入偏差(只存在于其中一个测量方程中)。请写出增广状态系统并确定该系统是否能观。

solution:

$$x_k^{'} = \left[egin{array}{ccc} x_k \ v_k \ ar{d}_k \end{array}
ight] = \left[egin{array}{ccc} 1 & 1 & 0 \ 0 & 1 & 0 \ 0 & 0 & 1 \end{array}
ight] \left[egin{array}{ccc} x_{k-1} \ v_{k-1} \ d_{k-1} \end{array}
ight] + \left[egin{array}{ccc} 1 \ 1 \ 0 \end{array}
ight] a_k$$

$$\left[egin{array}{ccc} d_{1,k}\ v_{2,k} \end{array}
ight] = \left[egin{array}{ccc} 1 & 0 & 0\ 1 & 0 & 1 \end{array}
ight] \left[egin{array}{c} x_k\ v_k\ ar{d}_k \end{array}
ight]$$

其中N = 2, U = 1

$$O = \left[egin{array}{c} C' \ C'A' \ C'A'^2 \end{array}
ight]$$

$$rank(O) = 3 = N + U$$

增广系统可观测

5.5.3

假设每个点为内点的概率为 $\omega=0.1$,如果想选择一个内点子集(n=3)的概率为p=0.999,需要多少次RANSAC迭代?

solution:

根据式(5.3.7)

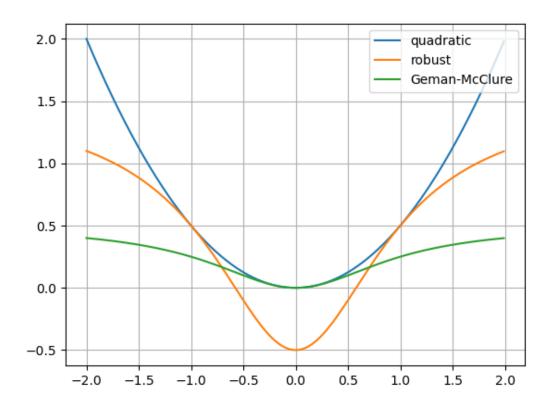
迭代次数 $k \approx 6904$

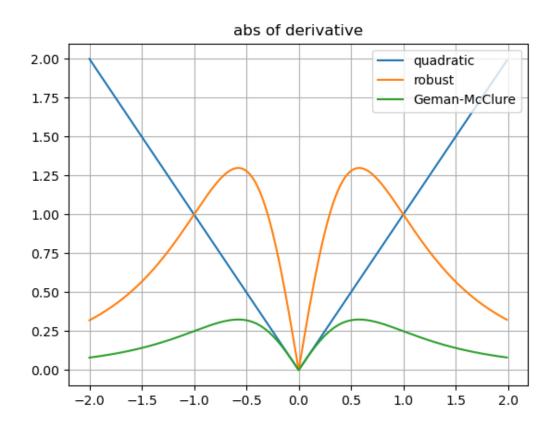
5.5.4

下面的鲁棒代价相比于German-McClure代价函数有何优势?

$$ho(u) = 0.5u^2$$
 ----- $u^2 \le 1$
 $ho(u) = \frac{2u^2}{1+u^2} - 0.5$ ----- $u^2 \ge 1$

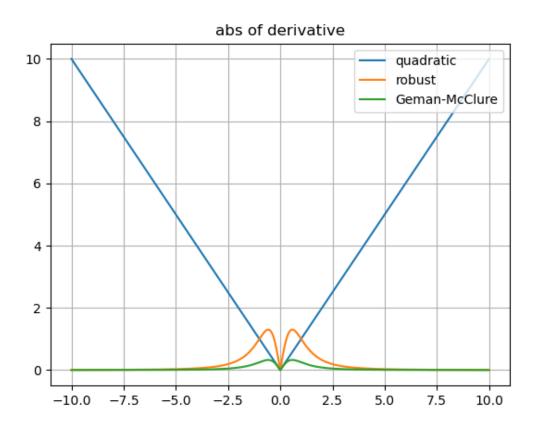
solution:





从上面第一章图可以看出,当前鲁棒函数在 $u^2 \leq 1$ 使用平方损失函数加快收敛速度 从第二章图,代价函数的导数绝对值函数可以看出,在无外点干扰的情况下,平方代价函数收敛更快

在 $u^2 \geq 1$,本鲁棒代价函数相比平方代价函数对外点的抑制作用更强,相比German-McClure代价函数在常规值的优化部分,收敛更快。当外点异常值较大时,如下图



可以看到,该算法在外点偏大区域与German-McClure代价函数具有相当的抑制作用。

总结: 在非外点区域更快迭代收敛, 在外点偏大区域具备相当的异常抑制作用, 既鲁棒又快

参考文献:

1.Matrix Cookbook---Kaare Brandt Petersen