Part1递归算法

2023年3月30日 15:56

Filter ---> optimal recursive(递归) data processing algorithm

不确定性：

1）不存在完美的数学模型
2）系统的扰动不可控，也很难建模
3）测量的传感器存在着误差

递归的思想：

假设有k个测量的结果$(Z_1 \cdots Z_k)$，那么我们估计真实的数据如下所示：
我们令根据k次测量结果得到的均值为$\hat{X}_k$，那么可以得到下面这样的公式：

$$\hat{X}_k = \frac{1}{K}(Z_1 + Z_2 + \cdots Z_K)$$
$$= \frac{1}{K}(Z_1 + Z_2 + \cdots Z_{K-1}) + \frac{1}{K}Z_K$$
$$= \frac{1}{K}\frac{K-1}{K-1}(Z_1 + Z_2 + \cdots Z_{K-1}) + \frac{1}{K}Z_K$$
$$= \frac{K-1}{K}\hat{X}_{k-1} + \frac{1}{K}Z_K$$
$$= \hat{X}_{k-1} - \frac{1}{K}\hat{X}_{k-1} + \frac{1}{K}Z_K = \hat{X}_{k-1} + \frac{1}{K}(Z_K - \hat{X}_{k-1})$$

When $k \dashrightarrow \infty$, $\frac{1}{K} \dashrightarrow 0$, *then* $\hat{X}_k \dashrightarrow \hat{X}_{k-1}$
说明随着测量的增加，测量的结果也就不再重要了。
而当k值比较小的时候，Z_K的影响就比较大了。
此时我们可以提炼一下系数。令$K_k = \frac{1}{K}$(大K指的是Kalman)
则$\hat{X}_k = \hat{X}_{k-1} + K_k(Z_K - \hat{X}_{k-1})$
含义为
当前的估计值 = 上一次的估计值 +系数 *（当前的测量值-上一次的估计值）
K_k又被成为kalman gain
Kalman filter 的一个重要的优势，就在于它并不用追逐很早之前的估计值，仅需要最近一次的估计值即可

下面介绍几个常用的参数
估计误差：$e_{est}(estimate)$
测量误差：$e_{MEA}(measurement)$

则此时kalman gain 可以写为：
$$K_k = \frac{e_{est_{k-1}}}{e_{est_{k-1}} + e_{MEA_k}} \text{(后续会解释)}$$
讨论，
当k时刻，$e_{est_{k-1}} \gg e_{MEA_k}$时，$K_k \dashrightarrow 1$，$\hat{X}_k = \hat{X}_{k-1} + K_k(Z_K - \hat{X}_{k-1}) \approx \hat{X}_k = \hat{X}_{k-1} + (Z_K - \hat{X}_{k-1}) = Z_K$，则说明估计的误差大，要更加信任测量值
当k时刻，$e_{est_{k-1}} \ll e_{MEA_k}$时，$K_k \dashrightarrow 0$，$\hat{X}_k = \hat{X}_{k-1} + K_k(Z_K - \hat{X}_{k-1}) \approx \hat{X}_{k-1}$，则说明估计的误差比较小，要更加信任估计值。

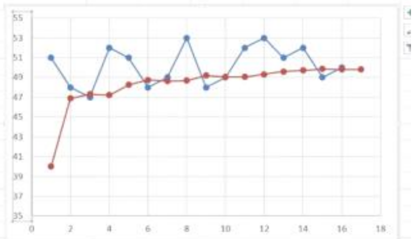
因此利用kalman filter去解决相关问题的最基本的方法为：
Step1: 计算kalman gain，$K_k = \frac{e_{est_{k-1}}}{e_{est_{k-1}} + e_{MEA_k}}$
Step2: 计算$\hat{X}_k = \hat{X}_{k-1} + K_k(Z_K - \hat{X}_{k-1})$
Step3: 更新估计误差$e_{est_k} = (1 - K_k)e_{est_{k-1}}$------(后续会解释)

例题：
假设一物体的长度为50mm
第一次测量$\hat{X}_0 = 40mm$；$e_{est_0} = 5mm$；$Z_1 = 51mm$；$e_{MEA_0} = 3mm$；
则

k	Z_k	e_{MEA_k}	$\hat{x}_k$	K_k	e_{EST_k}
0			40		5
1	51	3	46.875	0.625	1.875
2					
3					

$k=1$：
$K_k = \frac{5}{5+3} = 0.625$
$\hat{x}_k = 40 + 0.625(51 - 40) = 46.875$
$e_{EST} = (1 - 0.625)5 = 1.875$

k	Zk	e(MEAk)	hat(x)k	Kk	e(ESTk)
0			40		5
1	51	3	46.88	0.625	1.875
2	48	3	47.31	0.385	1.154
3	47	3	47.22	0.278	0.833
4	52	3	48.26	0.217	0.652
5	51	3	48.75	0.179	0.536
6	48	3	48.64	0.152	0.455
7	49	3	48.68	0.132	0.395
8	53	3	49.19	0.116	0.349
9	48	3	49.06	0.104	0.313
10	49	3	49.06	0.094	0.283
11	52	3	49.31	0.086	0.259
12	53	3	49.6	0.079	0.238
13	51	3	49.71	0.074	0.221
14	52	3	49.86	0.068	0.205
15	49	3	49.81	0.064	0.192
16	50	3	49.82	0.06	0.181

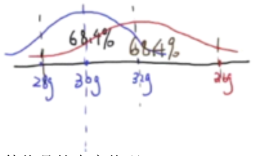


最终递归收敛到真实值附近

Part2 数学基础_数据融合_协方差矩阵_状态空间方程_观测器问题

2023年3月30日　　17:02

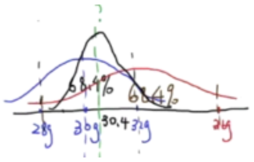
Data Fusion:
例子开始：用两个不同的测量工具测量同一件物品，$Z_1 = 30, Z_2 = 32$，我们已知这两个的测量工具的测量标准差为：$\sigma_1 = 2, \sigma_2 = 4$，并且都符合正态分布。



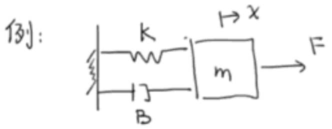
那么如何去估计这件物品的真实值呢？

此时我们想要充分利用这两个数据，那么我们就需要把他们结合起来。我们设估计值为$\hat{Z}$

$\hat{Z} = Z_1 + K(Z_2 - Z_1)$
那么目前就是求解一个合适的K使得测量的误差最小--->var($\hat{Z}$)最小
var($\hat{Z}$)=var$(Z_1 + K(Z_2 - Z_1))$=var($(1-k)Z_1 + KZ_2$)
可以看到前后两项是相互独立的，因此
var($\hat{Z}$)= var($(1-k)Z_1$)+var(KZ_2)=$(1-k)^2 var(Z_1) + K^2 var(Z_2)$=$((1-k)\sigma_1)^2 + (K\sigma_2)^2$
然后我们就可以对k进行求导：
$-2(1-k)\sigma_1{}^2 + 2K\sigma_2{}^2 = 0$，则可以求出：
$K = \dfrac{\sigma_1{}^2}{\sigma_1{}^2 + \sigma_2{}^2}$



State space observation:



弹簧系数为k，　阻尼系数为B，移动距离为x，　施加的力为F，写出这个木块的状态方程

根据F=ma，显然我们可以得到
$m\ddot{x} + B\dot{x} + kx = F$
令$F = u$，也就是系统的输入。
令$x_1 = x, x_2 = \dot{x}$，
则$\dot{x_1} = x_2$，
$\dot{x_2} = \frac{1}{m}u - \frac{B}{m}\dot{x} - \frac{K}{m}x = \frac{1}{m}u - \frac{B}{m}x_2 - \frac{K}{m}x_1$，
假设此时还具有测量量：
$Z_1 = x_1\ Z_2 = x_2$
则：

$$\begin{bmatrix}\dot{x_1}\\ \dot{x_2}\end{bmatrix} = \begin{bmatrix}0 & 1\\ -\frac{K}{m} & -\frac{B}{m}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ 1\end{bmatrix}u$$

$$\begin{bmatrix}Z_1\\ Z_2\end{bmatrix} = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}x_1\\ x_2\end{bmatrix}$$

因此状态空间方程可以写为：
$\dot{x}(t) = Ax(t) + Bu(t)$
$Z(t) = Hx(t)$

$\dot{x}(t)$表示的其实就是对于时间的变化

假如是离散的时刻，那么状态空间方程可以写为：
$X_k = AX_{k-1} + BU_{k-1}$
$Z_k = HX_k$，
其中下标表示为时间单位
当然并不是照抄上面的形式下来的，还是有具体的离散推导得到的

而由于我们提到过，kalman filter 解决的过程中存在着不确定性，那么
$X_k = AX_{k-1} + BU_{k-1} + \omega_{k-1}$，
ω_{k-1}被称为过程噪音，
而同样，观测设备也同样存在着不确定性，那么
$Z_k = HX_k + \upsilon_k$，
υ_k被成为测量噪音。
那么在模型也不准确，测量也不准确的情况下，如何去估计得到一个精确的$\hat{X_k}$？
这就是kalman filter要解决的问题。

Covariance Matrix:
将方差和协方差在一个矩阵中表现出来，表明了变量间的一个联动的关系。
例子：

	x	y	z
球员	身高	体重	年龄
瓦尔迪	179	74	33
奥巴梅扬	187	80	31
萨拉赫	175	71	28

方差 $\sigma_x^2 = \frac{1}{3}((179-180.3)^2 + (187-180.3)^2 + (175-180.3)^2) = 24.89$
$\sigma_y^2 = 14$
$\sigma_z^2 = 4.22$

协方差 $\sigma_x\sigma_y = \frac{1}{3}((\underbrace{179-180.3}_{<0})(\underbrace{74-75}_{<0}) + (\underbrace{187-180.3}_{>0})(\underbrace{80-75}_{>0}) + (175-180.3)(71-75)) = 18.7 = \sigma_y\sigma_x$

得到的协方差矩阵为：

$$P = \begin{bmatrix}\sigma_x^2 & \sigma_x\sigma_y & \sigma_x\sigma_z\\ \sigma_y\sigma_x & \sigma_y^2 & \sigma_y\sigma_z\\ \sigma_z\sigma_x & \sigma_z\sigma_y & \sigma_z^2\end{bmatrix}$$

G	H	I	J
	身高	体重	年龄
身高	32.69	29.75	1.4
体重	29.75	38.06	3.98
年龄	1.4	3.978	19.4

协方差值越大，说明相关程度越高，例如身高和体重，而年龄和身高、体重的值比较小，说明相关程度较小。

Part3 Kalman gain 推导

2023年3月30日　19:25

上一节我们知道了状态方程的公式，
$X_k = AX_{k-1} + BU_{k-1} + \omega_{k-1}$，
$Z_k = HX_k + \upsilon_k$，

其中ω_{k-1}是过程噪声，υ_k是测量噪声，这一部分是不可测的，
但是一般在自然界中可以将其假设成服从正态分布，
也就是说
$P(\omega) \sim (0, Q)$，
其中0是均值，而Q是协方差矩阵
$Q = E(\omega\omega^T)$

Q这个公式这一步是怎么来的呢
假设X_k是一个二维的向量，包含$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$
那么自然而然ω_{k-1}也是一个二维的向量，即$\begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$
那么$\omega\omega^T = \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} [\omega_1\ \omega_2] = \begin{bmatrix} \omega_1\omega_1 & \omega_1\omega_2 \\ \omega_2\omega_1 & \omega_2\ \omega_2 \end{bmatrix}$，
那么$E(\omega\omega^T) = \begin{bmatrix} E\omega_1\omega_1 & E\ \omega_1\omega_2 \\ E\ \omega_2\omega_1 & E\ \omega_2\ \omega_2 \end{bmatrix}$，
由于$\text{var}(x) = E(x^2) - E^2(x)$，
由于E(x)=0,
因此$\text{var}(x) = E(x^2)$，
故$Q = E(\omega\omega^T)$就是协方差矩阵了

同上，我们也认为υ_k符合正态分布
$P(\upsilon) \sim (0, R)$，
R就是测量噪声的协方差矩阵。

我们建模中不知道的便是ω和υ，其余的矩阵是已知的，而我们通过
$X_k = AX_{k-1} + BU_{k-1}$这个公式得到的$\hat{X}_k$其实是不正确的，
在这里在上面增加一个小小的负号，表明它是先验估计值。
$\hat{X}_k^- = A\hat{X}_{k-1} + BU_{k-1}$，(这里的$\hat{X}_{k-1}$也是估计值)
由于我们是知道测试结果Z_k，
那么在不考虑测量噪音的情况下，可以大体得到一个估计的$\hat{X}_k$，
那么我们令这个估计的$\hat{X}_k$命名为$\hat{X}_{k_{mea}}$,表示这是根据测试结果估计得到的$\hat{X}_k$
那么，根据公式$Z_k = HX_k$，可以得到
$\hat{X}_{k_{mea}} = H^{-1}Z_k$,(由于是可以观测的系统，因此H一定是可逆的)
因此我们得到了两个不同的关于X_k的结果，一个是先验结果$\hat{X}_k^-$*（算出来的）*,另一个是测出来的结果$\hat{X}_{k_{mea}}$*（测出来的）*,
但是着两个其实都是不准确的，因为其没有考虑噪声的影响
我们真正要求的是$\hat{X}_k$（表示后验）,包含了噪声的影响。
那么这里就要用到数据融合（Data Fusion）的概念了。
我们知道$\hat{X}_k^-$,$\hat{X}_{k_{mea}}$
那么$\hat{X}_k = \hat{X}_k^- + G(\hat{X}_{k_{mea}} - \hat{X}_k^-)$，(G是方阵)
$\hat{X}_{k_{mea}} = H^{-1}Z_k$，
那么$\hat{X}_k = \hat{X}_k^- + G(H^{-1}Z_k - \hat{X}_k^-)$，
那么令$K_k = GH^{-1}$,便可以得到$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)$，
那么在此时$K_k \notin (0,1)$, 而是$K_k \notin (0, H^{-1})$，
因此我们的目标就变得十分清晰了，就是求K_k使得考虑了噪声之后的估计值$\hat{X}_k$趋近于真实值X_k，
(那怎么取呢，很显然这里与误差息息相关)
那么此时引入误差:
$e_k = X_k - \hat{X}_k$,同理$P(e_k) \sim (0, P)$,*P也是那个协方差矩阵*
而当在不同的维度上的方差越小，那么说明这个e_k越接近0,因此估计值和真实值也就是最相近的。
所以，要选择合适的K_k使得tr(P)(矩阵对角线相加)最小，那么优化问题就变成了下面这个公式。
$P = E(e_ke_k^T) = E\left((X_k - \hat{X}_k)(X_k - \hat{X}_k)^T\right)$，
将$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)$代入，
可以得到$P = E\left((X_k - (\hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)))(X_k - (\hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)))^T\right)$，

那么此时将$Z_k = HX_k + \upsilon_k$,进行替换。
那么$X_k - \hat{X}_k = X_k - \hat{X}_k^- - K_k(HX_k + \upsilon_k) + K_kH\hat{X}_k^- = X_k - \hat{X}_k^- - K_kH(X_k - \hat{X}_k^-) - K_k\upsilon_k = (I - K_kH)(X_k - \hat{X}_k^-) - K_k\upsilon_k$，
由于$e_k = X_k - \hat{X}_k$,因此也可以命名$X_k - \hat{X}_k^-$为e_k的先验, 记为e_k^-，
因此此时的P可以改写为$P = E(e_ke_k^T) = E\left(\left((I - K_kH)e_k^- - K_k\upsilon_k\right)\left((I - K_kH)e_k^- - K_k\upsilon_k\right)^T\right)$，
根据矩阵运算 $(AB)^T = B^TA^T$，$(A+B)^T = A^T + B^T$，
则$P = E\left(\left((I - K_kH)e_k^- - K_k\upsilon_k\right)\left(e_k^{-T}(I - K_kH)^T - \upsilon_k^TK_k^T\right)\right) = E((I - K_kH)e_k^-e_k^{-T}(I - K_kH)^T - (I - K_kH)e_k^-\upsilon_k^TK_k^T - K_k\upsilon_ke_k^-e_k^{-T}(I - K_kH)^T + K_k\upsilon_k\upsilon_k^TK_k^T)$，

求上述的期望，等价于求上述每一项的期望之和，

即$\mathrm{P}=\mathrm{E}\left((I-K_kH)e_k^-e_k^{-T}(I-K_kH)^T\right)+\mathrm{E}\left(-(I-K_kH)e_k^-\upsilon_k^TK_k^T\right)+\mathrm{E}\left(-K_k\upsilon_ke_k^{-T}(I-K_kH)^T\right)+\mathrm{E}\left(K_k\upsilon_k\upsilon_k^TK_k^T\right)$,

由于$(I-K_kH)$和K_k^T都是常数，因此中间的两项可以将常数提取出来，

$\mathrm{P}=\mathrm{E}\left((I-K_kH)e_k^-e_k^{-T}(I-K_kH)^T\right)-(I-K_kH)\mathrm{E}(e_k^-\upsilon_k^T)K_k^T-K_k\mathrm{E}(\upsilon_ke_k^{-T})(I-K_kH)^T+\mathrm{E}(K_k\upsilon_k\upsilon_k^TK_k^T)$,

而因为e_k^-是算出来的,υ_k是测量误差,故这两个肯定是相互独立的，

故$\mathrm{E}(e_k^-\upsilon_k^T)=E(e_k^-)E(\upsilon_k^T)=0$, 同理第三项也是等于0(高斯分布),

故$\mathrm{P}=\mathrm{E}\left((I-K_kH)e_k^-e_k^{-T}(I-K_kH)^T\right)+\mathrm{E}(K_k\upsilon_k\upsilon_k^TK_k^T)$,

故$\mathrm{P}=(I-K_kH)\mathrm{E}(e_k^-e_k^{-T})(I-K_kH)^T+K_k\mathrm{E}(\upsilon_k\upsilon_k^T)K_k^T$,

那么此时便令$P^-=\mathrm{E}(e_k^-e_k^{-T})$,先验误差的协方差矩阵,

故$\mathrm{P}=(I-K_kH)P^-(I-K_kH)^T+K_k\mathrm{E}(\upsilon_k\upsilon_k^T)K_k^T=(P^--K_kHP^-)(I-H^TK_k^T)+K_kRK_k^T=P^--K_kHP^--P^-H^TK_k^T+K_kHP^-H^TK_k^T+K_kRK_k^T$

求tr(P), 此时我们发现K_kHP^-的转置就是$P^-H^TK_k^T$,如果一个矩阵是另一个矩阵的转置的话，那么两个矩阵的迹tr是一样的，故

$\mathrm{tr(P)}=\mathrm{tr}(P^-)-2\mathrm{tr}(K_kHP^-)+\mathrm{tr}(K_kHP^-H^TK_k^T)+\mathrm{tr}(K_kRK_k^T)$,

此时求得最小值，那么便是对K_k求导,

$\frac{d(\mathrm{tr(P)})}{d(K_k)}=0$,

因为第一项没有K_k,故为0，第二项根据雅可比矩阵

假设有两个矩阵A,B, 那么$\mathrm{A}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$, $\mathrm{B}=\begin{bmatrix}b_{11} & b_{12}\\ b_{21} & b_{22}\end{bmatrix}$,

那么相乘可得到下面的形式:

$\begin{bmatrix}a_{11}b_{11}+a_{12}b_{21} & a_{11}b_{12}+a_{12}b_{22}\\ a_{21}b_{11}+a_{22}b_{21} & a_{21}b_{12}+a_{22}b_{22}\end{bmatrix}$,

那么$\mathrm{tr(AB)}=a_{11}b_{11}+a_{12}b_{21}+a_{21}b_{12}+a_{22}b_{22}$,

根据雅可比矩阵

$$\frac{d(\mathrm{tr(AB)})}{d(A)}=\begin{bmatrix}\frac{d(\mathrm{tr(AB)})}{d(a_{11})} & \frac{d(\mathrm{tr(AB)})}{d(a_{12})}\\ \frac{d(\mathrm{tr(AB)})}{d(a_{21})} & \frac{d(\mathrm{tr(AB)})}{d(a_{22})}\end{bmatrix}=\begin{bmatrix}b_{11} & b_{21}\\ b_{12} & b_{22}\end{bmatrix}=B^T$$

所以第二项求导后为$-2(HP^-)^T$,同理第三项是$\mathrm{AB}A^T$的形式,按照上面的推导,可以得到结果为2AB,

那么第三项求导后为$2K_kHP^-H^T$,同理可以得到第四项。

因此$\frac{d(\mathrm{tr(P)})}{d(K_k)}=0-2(HP^-)^T+2K_kHP^-H^T+2K_kR=0$,而协方差矩阵$P^-$的转置有等于他本身

故可以得到$-P^-H^T+K_kHP^-H^T+K_kR=0\longrightarrow K_k(HP^-H^T+R)=P^-H^T$

故$K_k=P^-H^T(HP^-H^T+R)^{-1}$,为方便理解，我这里先写成分子分母的形式，但实际矩阵不应该这么写,

$K_k=\frac{P^-H^T}{HP^-H^T+R}$,

可以看到当R特别大的时候，也就是测量的误差特别大的时候，$K_k\longrightarrow 0$, 而由上面我们可以得知, 当$K_k\longrightarrow 0$时，$\hat{X}_k\longrightarrow\hat{X}_k^-$,确实是满足噪声特别大的时候

而当R比较大的时候，$K_k\longrightarrow H^{-1}$,而由上面我们可以得知, 当$K_k\longrightarrow H^{-1}$时，$\hat{X}_k\longrightarrow\hat{X}_{k_{mea}}$,确实是满足噪声特别小的时候,要听从测量值的结论。

Part4 误差协方差矩阵

2023年3月30日 21:37

这是之前所推出来的内容

$X_k = AX_{k-1} + BU_{k-1} + W_{k-1}$ $\omega \sim p(0, Q)$

$Z_k = HX_k + V_k$ $V \sim p(0, R)$

先验估计

$\hat{X}_k^- = A\hat{X}_{k-1} + BU_{k-1}$

后验估计

$\hat{X}_k = \hat{X}_k^- + K_k(Z_k - H\hat{X}_k^-)$

卡尔曼增益

$K_k = \frac{P_k^- H^T}{HP_k^- H^T + R}$

而目前，我们虽然求得了K_k的公式，但是P^-也就是图中$P^-{}_k$目前还是未知的，所以我们需要的也就是求先验误差的协方差矩阵。

$P^- = \mathrm{E}(e_k{}^- e_k{}^{-T})$.$e_k{}^- = X_k - \hat{X}_k{}^-$*(就是真实值减去先验估计值)*

真实值$X_k = AX_{k-1} + BU_{k-1} + \omega_{k-1}$

先验估计值$\hat{X}_k{}^- = A\hat{X}_{k-1} + BU_{k-1}$

那么便可以得到.$e_k{}^- = AX_{k-1} + BU_{k-1} + \omega_{k-1} - (A\hat{X}_{k-1} + BU_{k-1}) = AX_{k-1} + \omega_{k-1} - A\hat{X}_{k-1} = A(X_{k-1} - \hat{X}_{k-1}) + \omega_{k-1}$

而$X_{k-1} - \hat{X}_{k-1} = e_{k-1}$

故$e_k{}^- = Ae_{k-1} + \omega_{k-1}$,

所以$P^- = \mathrm{E}(e_k{}^- e_k{}^{-T}) = E\left((Ae_{k-1} + \omega_{k-1})(Ae_{k-1} + \omega_{k-1})^T\right) = E\left((Ae_{k-1} + \omega_{k-1})(e_{k-1}{}^T A^T + \omega_{k-1}{}^T)\right) = E(Ae_{k-1}e_{k-1}{}^T A^T + Ae_{k-1}\omega_{k-1}{}^T + \omega_{k-1}e_{k-1}{}^T A^T + \omega_{k-1}\omega_{k-1}{}^T)$

$P^- = E(Ae_{k-1}e_{k-1}{}^T A^T) + E(Ae_{k-1}\omega_{k-1}{}^T) + E(\omega_{k-1}e_{k-1}{}^T A^T) + E(\omega_{k-1}\omega_{k-1}{}^T)$,

而由于ω_{k-1}作用在X_k上,跟$e_{k-1} = X_{k-1} - \hat{X}_{k-1}$,并没有关系，故两个是独立的，又都是高斯分布，所以中间那两个便是0,所以

$P^- = E(Ae_{k-1}e_{k-1}{}^T A^T) + E(\omega_{k-1}\omega_{k-1}{}^T) = AE(e_{k-1}e_{k-1}{}^T)A^T + E(\omega_{k-1}\omega_{k-1}{}^T)$,

而$E(e_{k-1}e_{k-1}{}^T) = P^-{}_{k-1}$,$E(\omega_{k-1}\omega_{k-1}{}^T) = Q$

所以$P^-{}_k = AP^-{}_{k-1}A^T + Q$

有了这些式子之后，就可以用kalman filter 去估计状态变量的值了。

有两步，首先是预测，然后是校正，在预测中也有两步

预测中：

先求先验估计值$\hat{X}_k{}^- = A\hat{X}_{k-1} + BU_{k-1}$

然后求先验误差协方差矩阵：$P^-{}_k = AP^-{}_{k-1}A^T + Q$

在校正中：

首先计算kalman gain：$\mathrm{K_k} = P^- H^T (HP^- H^T + R)^{-1}$

然后再求后验估计：$\hat{X}_k = \hat{X}_k{}^- + K_k(Z_k - H\hat{X}_k{}^-)$

先求先验估计值$\hat{X}_k{}^- = A\hat{X}_{k-1} + BU_{k-1}$

先验误差协方差矩阵：$P^-{}_k = AP^-{}_{k-1}A^T + Q$

Kalman gain：$\mathrm{K_k} = P^- H^T (HP^- H^T + R)^{-1}$

后验估计：$\hat{X}_k = \hat{X}_k{}^- + K_k(Z_k - H\hat{X}_k{}^-)$

更新先验误差协方差矩阵$\mathrm{P} = (I - K_k H)P^-$

在这里我们{求先验误差协方差矩阵}的时候会用到上一次的先验误差协方差矩阵

因此在用完之后还需要再更新先验误差协方差矩阵

在上一部分中我们计算得到的

$\mathrm{P} = P^- - K_k HP^- - P^- H^T K_k{}^T + K_k HP^- H^T K_k{}^T + K_k RK_k{}^T = P^- - K_k HP^- - P^- H^T K_k{}^T + K_k(HP^- H^T + R)K_k{}^T$,

代入$\mathrm{K_k} = P^- H^T (HP^- H^T + R)^{-1}$,

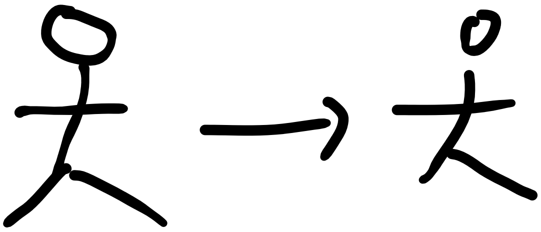
$\mathrm{P} = P^- - K_k HP^- - P^- H^T K_k{}^T + P^- H^T K_k{}^T = P^- - K_k HP^- = (I - K_k H)P^-$

这就是完整的kalman filter的五个公式了

Part5 Kalman Filter 的具体应用

2023年3月30日 23:17

2D例子



一个人在走路，状态可以由两个变量进行表示，x_1表示人的位置，x_2表示人的速度，假设人在匀速直线行驶(也就是$x_{2_k}=x_{2_{k-1}}$),采样时刻ΔT为K时刻和K-1时刻的时间差

位置:$x_{1,k}=x_{1,k-1}+\Delta T x_{2,k-1}$, 令$\Delta T=1$

接下来，我们就要考虑过程中的噪声了。

$x_{1,k}=x_{1,k-1}+\Delta T x_{2,k-1}+\omega_{1,k-1}$

速度也要考虑噪声：

$x_{2_k}=x_{2_{k-1}}+\omega_{2,k-1}$

加上天上的卫星进行实时的观测。

$Z_{1,k}=x_{1,k}+v_{1,k}$ $Z_{2,k}=x_{2,k}+v_{2,k}$

写成矩阵的形式：

$$\begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1,k-1} \\ x_{2,k-1} \end{bmatrix} + \begin{bmatrix} w_{1,k-1} \\ w_{2,k-1} \end{bmatrix}$$

$$\begin{bmatrix} z_{1,k} \\ z_{2,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_{1,k} \\ x_{2,k} \end{bmatrix} + \begin{bmatrix} v_{1,k} \\ v_{2,k} \end{bmatrix}$$

Extend Kalman filter(EKF)

2023年3月31日 17:10

线性系统的状态方程：

$x_k = Ax_{k-1} + Bu_{k-1} + \omega_{k-1}$
$Z_k = H_{k-1} + v_k$
$\mathrm{P}(\omega) \sim (0, Q)$,
$\mathrm{P}(v) \sim (0, R)$

有了这些式子之后，就可以用kalman filter 去估计状态变量的值了。

有两步，首先是预测，然后是校正，在预测中也有两步

预测中：

先求先验估计值$\hat{X}_k^{\ -} = A\hat{X}_{k-1} + BU_{k-1}$

然后求先验误差协方差矩阵：$P^-{}_k = AP^-{}_{k-1}A^T + Q$

在校正中：

首先计算kalman gain：$\mathrm{K_k} = P^- H^T (HP^- H^T + R)^{-1}$

然后再求后验估计：$\hat{X}_k = \hat{X}_k^{\ -} + K_k(Z_k - H\hat{X}_k^{\ -})$

在这里我们(求先验误差协方差矩阵)的时候会用到上一次的先验误差协方差矩阵

因此在用完之后还需要再更新先验误差协方差矩阵

在上一部分中我们计算得到的

$\mathrm{P} = P^- - K_k HP^- - P^- H^T K_k^{\ T} + K_k HP^- H^T K_k^{\ T} + K_k RK_k^{\ T} = P^- - K_k HP^- - P^- H^T K_k^{\ T} +$
$K_k\ (HP^- H^T + R) K_k^{\ T}$,
代入$\mathrm{K_k} = P^- H^T (HP^- H^T + R)^{-1}$,
$\mathrm{P} = P^- - K_k HP^- - P^- H^T K_k^{\ T} + P^- H^T K_k^{\ T} = P^- - K_k HP^- = (I - K_k H) P^-$

这就是完整的kalman filter的五个公式了

考虑高斯白噪声的非线性系统:

$$\begin{cases} x_{k+1} = f(x_k) + w_k \\ z_k = h(x_k) + v_k \end{cases}$$

x_k为状态向量，z_k量测向量，$f(\cdot)$和$h(\cdot)$分别为系统非线性状态函数和量测函数
w_k和v_k分别是零均值，协方差为Q_k和R_k的不相关高斯白噪声。

假设我们已知 k时刻状态估计值 $\hat{x}_{k|k}$和估计方差 $P_{k|k}$，

我们将非线性函数 $f(x_k)$在$\hat{x}_{k|k}$处进行一阶泰勒展开可 得：

$$f(x_k) = f(\hat{x}_{k|k}) + \frac{\partial f}{\partial x_k}\Big|_{x_k = \hat{x}_{k|k}} (x_k - \hat{x}_{k|k}) + o(x_k - \hat{x}_{k|k})$$

其中 $o(x_k - \hat{x}_{k|k})$为高阶项，我们定义 $\frac{\partial f}{\partial x_k}\Big|_{x_k = \hat{x}_{k|k}} = F_k$,忽略高阶项，状态方程 可以化简为：

$x_{k+1} = f(\hat{x}_{k|k}) + F_k(x_k - \hat{x}_{k|k}) + w_k$

一步状态预测： $\hat{x}_{k+1|k} = \mathrm{E}[f(\hat{x}_{k|k}) + F_k(x_k - \hat{x}_{k|k}) + w_k] = f(\hat{x}_{k|k})$

一步预测协方差：
$$\begin{aligned} P_{k+1|k} &= \mathrm{E}[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^{\mathrm{T}}] \\ &= \mathrm{E}\{[F_k(x_k - \hat{x}_{k|k}) + w_k][F_k(x_k - \hat{x}_{k|k}) + w_k]^{\mathrm{T}}\} \\ &= F_k P_{k|k} F_k^{\mathrm{T}} + \mathrm{Q_k} \end{aligned}$$

将非线性函数$h(\cdot)$在一步状态预测$\hat{x}_{k+1|k}$处一阶泰勒展开得：

$$h(x_{k+1}) = h(\hat{x}_{k+1|k}) + \frac{\partial h}{\partial x_{k+1}}\Big|_{x_{k+1} = \hat{x}_{k+1|k}} (x_{k+1} - \hat{x}_{k+1|k}) + o(x_{k+1} - \hat{x}_{k+1|k})$$

其中$o(x_{k+1} - \hat{x}_{k+1|k})$为高阶项，我们令$\frac{\partial h}{\partial x_{k+1}}\Big|_{x_{k+1} = \hat{x}_{k+1|k}} = H_{k+1}$,所以量测方程可表示为:

$z_{k+1} = h(\hat{x}_{k+1|k}) + H_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}$

所以量测一步预测：$\hat{z}_{k+1} = \mathrm{E}[h(\hat{x}_{k+1|k}) + H_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}] = h(\hat{x}_{k+1|k})$

量测预测误差协方差阵:
$$\begin{aligned} P_{zz,k+1|k} &= \mathrm{E}[(z_{k+1} - \hat{z}_{k+1})(z_{k+1} - \hat{z}_{k+1})^{\mathrm{T}}] \\ &= \mathrm{E}\{[H_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}][H_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}]^{\mathrm{T}}\} \\ &= H_{k+1} P_{k+1|k} H_{k+1}^{\mathrm{T}} + R_{k+1} \end{aligned}$$

状态与量测间的互协方差矩阵：
$$\begin{aligned} P_{xz,k+1|k} &= \mathrm{E}[(x_{k+1} - \hat{x}_{k+1})(z_{k+1} - \hat{z}_{k+1})^{\mathrm{T}}] \\ &= \mathrm{E}\{(x_{k+1} - \hat{x}_{k+1})[H_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}]^{\mathrm{T}}\} \\ &= P_{k+1|k} H_{k+1}^{\mathrm{T}} \end{aligned}$$

状态增益矩阵:

而对于一个非线性(Nonlinear)的系统来说

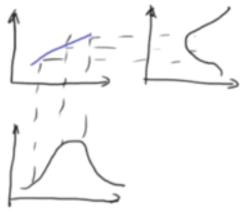
只能写成下面的这种形式：

$x_k = f(x_{k-1}, u_{k-1}, \omega_{k-1})$
$z_k = h(x_k, v_k)$
$\mathrm{P}(\omega) \sim (0, Q)$,
$\mathrm{P}(\upsilon) \sim (0, \mathrm{R})$

正态分布的随机变量通过非线性系统之后就不再是正态的了

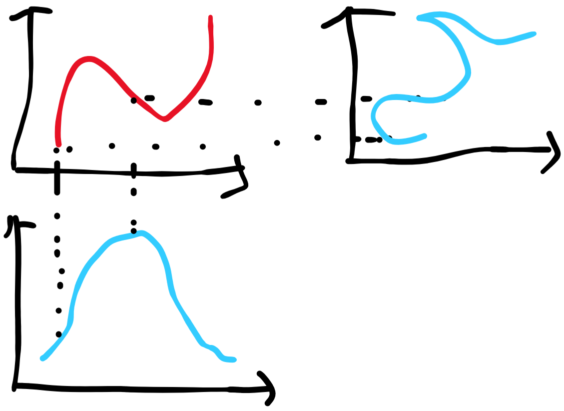
这其实很好懂：

假如是一个线性系统的话，那么我们可以由下面的图可以看到



正态分布经过映射之后还是正态分布。

那如果是非线性系统的话，那么显然他将不会是处于正态分布的情况了



所以说如果还是想要对这个系统进行卡尔曼滤波的话，就要对这个系统进行线性化（linearization）

在这里会用到泰勒级数展开。

$\mathrm{f(x)} = \mathrm{f}(x_0) + \frac{\partial f(x)}{\partial x}(x - x_0) + \frac{\partial^2 f(x)}{2!\,\partial x^2}(x - x_0)^2 + \cdots + \Delta(o)$

就是在非线性函数的具体某一点上进行泰勒级数展开，把他变成一个线性的系统，然后把一个完成的非线性函数离散成无数条离散的线性函数。

但是系统是存在误差的，我们永远不知道函数中的真实点是多少，因此也就无法在该点进行泰勒级数展开。

因此退而求其次，我们将$\mathrm{f}(x_k)$在$\hat{x}_{k-1}$处进行线性化。也就是上一时刻的后验估计。

因此在上面的公式可以写为：

$x_k = f(x_{k-1}, u_{k-1}, \omega_{k-1}) = f(\hat{x}_{k-1}, u_{k-1}, \omega_{k-1}) + A_k(x_k - \hat{x}_{k-1}) + B_K\ \omega_{k-1}$
$A_k = \frac{\partial f}{\partial x}|\hat{x}_{k-1}, u_{k-1}, \omega_{k-1}$
$B_k = \frac{\partial f}{\partial \omega}|\hat{x}_{k-1}, u_{k-1}, \omega_{k-1}$

因为我们还不了解 ω_{k-1}的具体情况，因此为了简化，我们将其忽略掉

$x_k = f(x_{k-1}, u_{k-1}, \omega_{k-1}) = f(\hat{x}_{k-1}, u_{k-1}, \omega_{k-1}) + A_k(x_k - \hat{x}_{k-1}) + B_K\ \omega_{k-1}$

A是雅可比矩阵

$A_k = \frac{\partial f}{\partial x}|\hat{x}_{k-1}, u_{k-1}$
$B_k = \frac{\partial f}{\partial \omega}|\hat{x}_{k-1}, u_{k-1}$

下面举一个例子

例:

$X_1 = X_1 + \sin X_2 = f_1$
$X_2 = X_1^2 = f_2$

$$A = \frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} \end{bmatrix} = \begin{bmatrix} 1 & \cos x_2 \\ 2x_1 & 0 \end{bmatrix}\Bigg|_{\hat{x}_{1k-1}, \hat{x}_{2k-1}}$$

$$\begin{bmatrix} 1 & \cos \hat{x}_{2k-1} \end{bmatrix}$$

$= P_{k+1|k}H_{k+1}$

状态增益矩阵：

$$K_{k+1} = P_{xz,k+1|k}(P_{zz,k+1|k})^{-1} = P_{k+1|k}H_{k+1}^{\mathrm{T}}(H_{k+1}P_{k+1|k}H_{k+1}^{\mathrm{T}} + R_{k+1})^{-1}$$

$k+1$时刻状态估计值： $\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}(z_{k+1} - \hat{z}_{k+1|k})$

状态估计误差协方差矩阵：

$$\begin{aligned} P_{k+1|k+1} &= \mathrm{E}[(x_{k+1} - \hat{x}_{k+1|k+1})(x_{k+1} - \hat{x}_{k+1|k+1})^{\mathrm{T}}] \\ &= \mathrm{E}\{[x_{k+1} - \hat{x}_{k+1|k} - K_{k+1}(z_{k+1} - \hat{z}_{k+1|k})][x_{k+1} - \hat{x}_{k+1|k} - K_{k+1}(z_{k+1} - \hat{z}_{k+1|k})]^{\mathrm{T}}\} \\ &= (I - K_{k+1}H_{k+1})P_{k+1|k}(I - K_{k+1}H_{k+1})^{T} + K_{k+1}R_{k+1}K_{k+1}^{T} \\ &= (I - K_{k+1}H_{k+1})P_{k+1|k} \end{aligned}$$

$$\begin{bmatrix} 2x_1 & 0 \end{bmatrix}\Big|_{\hat{x}_{1,k-1},\ \hat{x}_{2,k-1}}$$

$$= \begin{bmatrix} 1 & \omega_3 \hat{x}_{2,k-1} \\ 2\hat{x}_{1,k-1} & 0 \end{bmatrix}$$

可以看到A_k在随着k的变化而不断地变化

我们将$f(\hat{x}_{k-1}, u_{k-1}, 0)$*定义为*$\hat{x}_k^-$

对于测量方程而言：那么就让z_k在$\hat{x}_k$附近进行线性化。

$z_k = h(x_k, v_k) = h(\hat{x}_k, v_k) + H_K(x_k - \hat{x}_k) + C_k v_k$

同样是假设$v_k = 0$

$H_k = \frac{\partial f}{\partial x}|\hat{x}_k,$

$C_k = \frac{\partial f}{\partial v}|\hat{x}_k,$

我们将$h(\hat{x}_k, 0)$*定义为*$\hat{z}_k$

那么式子便可以改写为：

$x_k = f(x_{k-1}, u_{k-1}, \omega_{k-1}) = \hat{x}_k + A_k(x_k - \hat{x}_{k-1}) + B_K\, \omega_{k-1}$

$z_k = h(x_k, v_k) = \hat{z}_k + H_K(x_k - \hat{x}_k) + C_k v_k$

而$B_K\, \omega_{k-1}$和$C_k v_k$仍然是满足正态分布的，不过变成了$(0, B_K\, Q{B_K}^T)$和$(0, C_k R{C_k}^T)$

左乘加上右乘转置，这样得到的结果就类似于平方

那么那五个式子便可以写成

先验： $\hat{x}_k{}^- = f(\hat{x}_{k-1}, u_{k-1}, 0)$

先验误差的协方差矩阵为：$P^-{}_k = AP^-{}_{k-1}A^T + B_K\, Q{B_K}^T$

Kalman fiter的增益改为：$\mathrm{K_k} = P^- H^T(HP^- H^T + C_k R{C_k}^T)^{-1}$

那么再求后验估计：$\hat{X}_k = \hat{X}_k{}^- + K_k\left(Z_k - h(\hat{x}_k, 0)\right)$

$\mathrm{P} = P^- - K_k HP^- - P^- H^T{K_k}^T + P^- H^T{K_k}^T = P^- - K_k HP^- = (I - K_k H)P^-$ *不变*