卡尔曼滤波理论推导

作者 **电子科技大学 信息与通信工程学院 小明同学**

初识卡尔曼滤波

第一次看到卡尔曼滤波是在b站的一个自动化up主的视频里，卡尔曼滤波在自动控制领域有着广泛的应用，刚好随机信号分析课程对卡尔曼滤波算法有要求，老师课件上满满的纯公式推导我实在是啃不动，也算是课程设计的需要，于是就萌生了写这篇文档的想法。我这篇文档的总结也是我学习卡尔曼滤波的过程，浏览了网上能搜到的十多篇相关博客、论文，集百家之大成，绕开了很多晦涩难懂的公式推导，至少在我本人的角度是很好理解的。下面我们进入正题。

什么是卡尔曼滤波

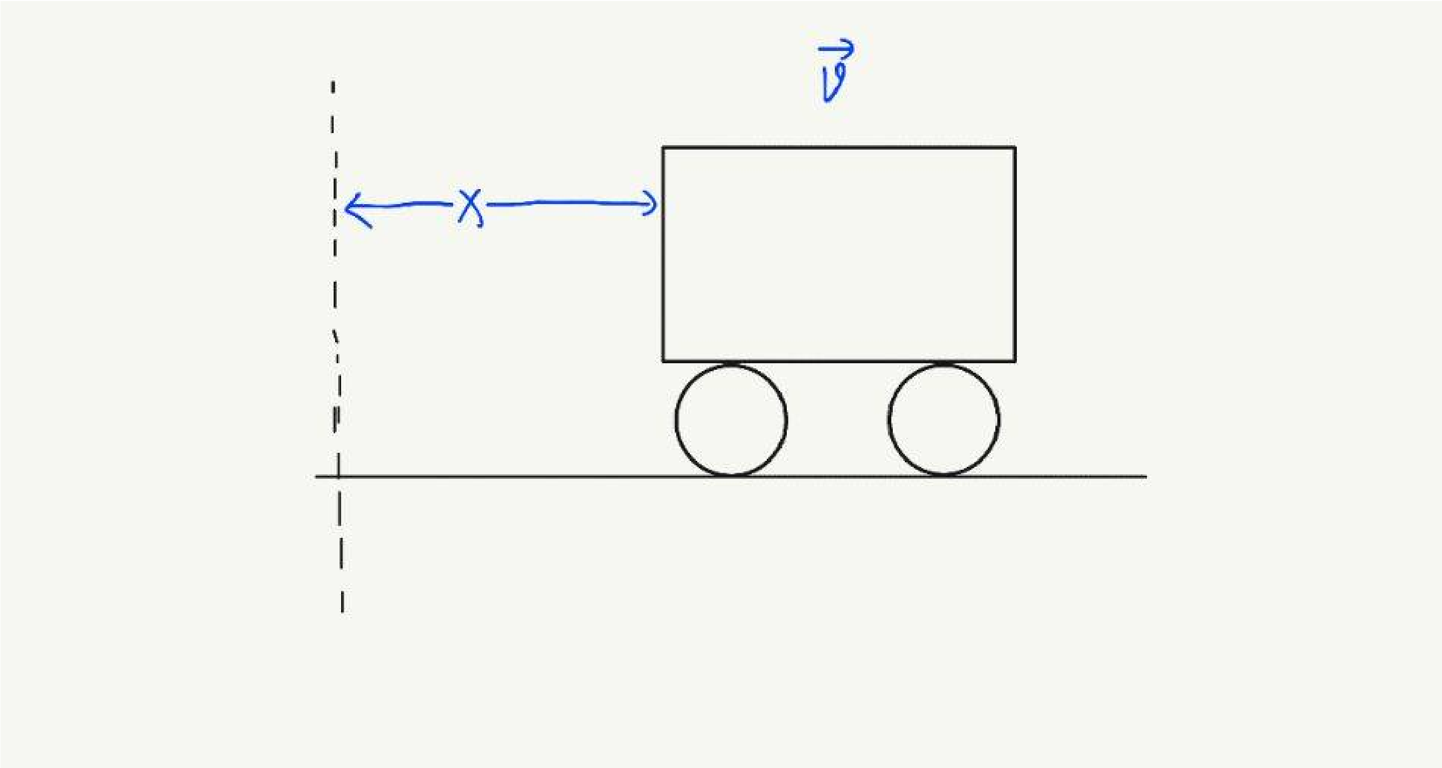
卡尔曼滤波（Kalman filter）是一种高效率的递归滤波器（自回归滤波器），它能够从一系列的不完全及包含噪声的测量中，估计动态系统的状态。
但相比于维基百科的说法，我更喜欢这篇知乎文章链接的解释
对于这个滤波器，我们几乎可以下这么一个定论：只要是存在不确定信息的动态系统，卡尔曼滤波就可以对系统下一步要做什么做出有根据的推测。即便有噪声信息干扰，卡尔曼滤波通常也能很好的弄清楚究竟发生了什么，找出现象间不易察觉的相关性。

问题引入

我相信，所有算法的学习都要建立在理解的基础上，绝不是枯燥的公式推导。各种矩阵变换让人没有阅读的欲望，我相信这肯定不是我一个人的感受（狗头。但是要理解高维空间的变换又属实不是一件容易的事情。所以，我学习新方法的步骤一般是一维理解，二维推导，再去尝试性理解高维空间。当然，我们需要一些基本的线性代数和概率论的知识储备。

小车问题

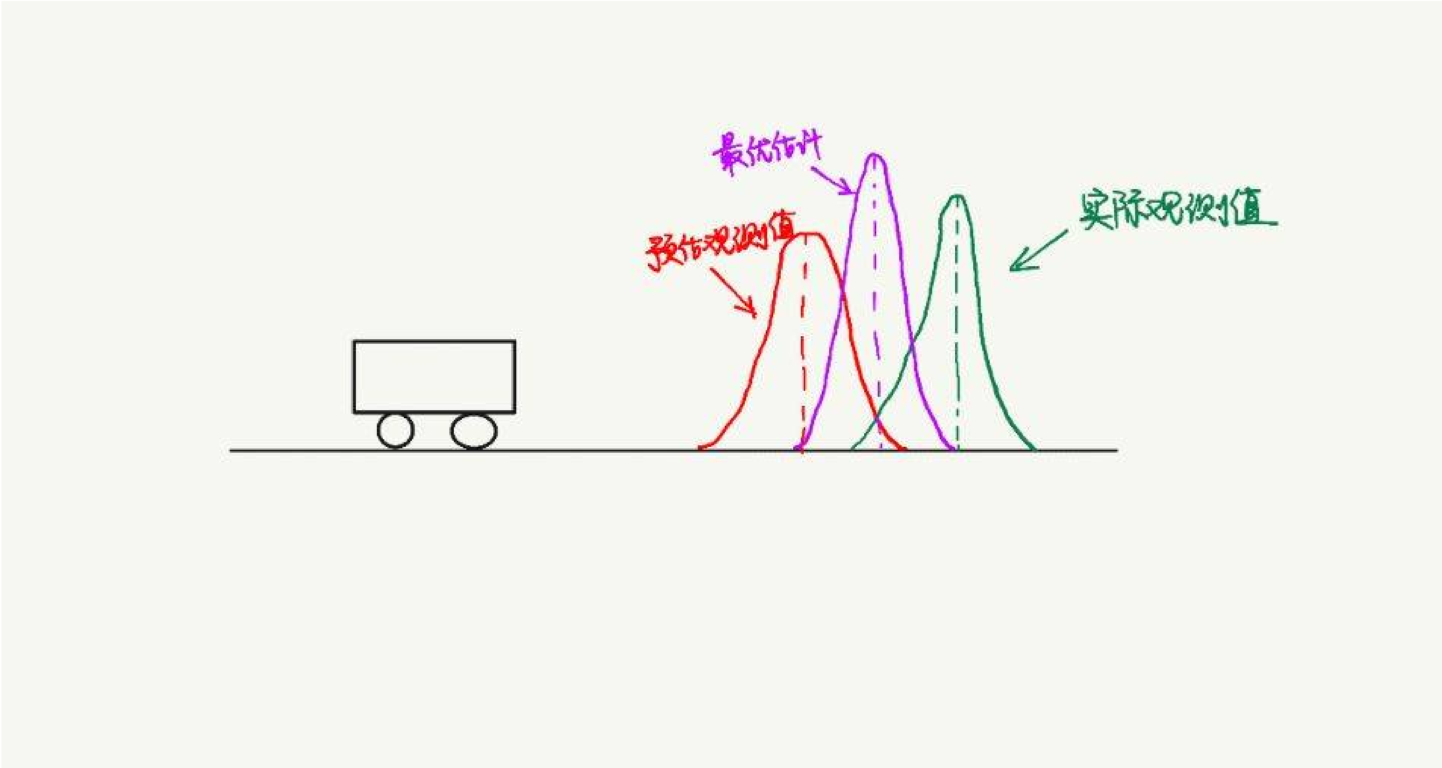
既然卡尔曼滤波在测距上有广泛的应用，而测距相较于滤波更容易理解，我们不妨就以测距为起点开启卡尔曼滤波的学习旅程。
我们首先来看这样一个模型



我们有一个小车，用矩阵$X = \begin{pmatrix} x \\ v \end{pmatrix}$来表示小车的状态，其中x是小车的位移，v是小车当前的速度。我们现在想要知道小车在下一时刻的状态，有以下两种方法：

1. 通过小车当前状态（x,v）(加速度a,后面会讨论，这里先默认为匀速运动a=0)计算下一状态。理论上这种方法可以完美计算，但是由于实际问题的影响（比如遇到雨雪天轮胎打滑的情况），我们不可能用这种方法准确估计下一状态。
2. 在小车上安装距离传感器和速度传感器用来观测当前状态。显然这种方法可以应对复杂的实际情况，可是这同样不能给出准确的估计（在同一个地点多次使用GPS定位，每次的结果也不会都一样）

于是，综合以上两种方法，我们有这样一种处理方式



将两种方法的结果用某种加权方法（后面会讲）进行融合，得到最优估计。再以这个最优估计为基准，去得到下一次的最优估计，如此迭代下去，就是卡尔曼滤波的基本思想。

理论推导

经过了小车问题的分析，相信卡尔曼滤波的基本思想不难理解，下面就开始正经地推公式啦！

还是以小车运动为例用矩阵$X_k = \begin{bmatrix} x_k \\ v_k \end{bmatrix}$来表示小车的状态，用协方差矩阵$P_k = \begin{bmatrix} \Sigma_{xx} & \Sigma_{xv} \\ \Sigma_{vx} & \Sigma_{vv} \end{bmatrix}$来表示状态参数之间的相关性，再加上我们幼儿园大班学习的匀变速直线运动的位移方程和速度方程

$$\begin{cases} x_k = x_{k-1} + \Delta t \cdot v_{k-1} + \frac{1}{2}\Delta t^2 \cdot a \\ v_k = \qquad\qquad\quad v_{k-1} + \quad \Delta t \cdot a \end{cases}$$

用矩阵的形式，可以把这两个方程写成下面这种形式：

$$\hat{X}_k = F_k \hat{X}_{k-1} + B_k u_k + W_k$$

在这个问题中$F_k = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}$称为状态转移矩阵，$B_k = \begin{bmatrix} \frac{1}{2}\Delta t^2 \\ \Delta t \end{bmatrix}$称为状态控制矩阵$u_k = [a]$称为状态控制向量，$W_k$是实际情况引入的外部噪声。矩阵形式是一种广义的写法，在这个小车问题模型中可以用以上内容来表示

下面我们要计算预测值与真实值之间误差的协方差矩阵

在这里我们要用到协方差矩阵的性质

$$Cov\left(x\right)=\Sigma$$

$$Cov\left(Ax\right)=A\Sigma A^T$$

所以经过了上面状态转移的变换后，预测值与真实值的协方差矩阵为

$$P_k=F_kP_{k-1}F_k^T+Q_k$$

其中Q为转换过程中引入的外部噪声。

下面我们要进行预测值向预估观测值的转换

$$\mu_{expected}=H_k\hat{X}_k$$

$$\Sigma_{expected}=H_kP_kH_k^T$$

这里H_k是定义的观测矩阵，在小车运动的模型里可以把它看作1，但是在其他模型中，我们需要的观测值跟预测值很有可能连维度都不一样，所以需要引入一个观测矩阵用来转换。

在进行下一步之前，我们还需要补充一个数学知识点：融合高斯分布

两个高斯函数相乘，再进行归一化后：

$$\mu'=\mu_0+\frac{\sigma_0^2\left(\mu_1-\mu_0\right)}{\sigma_0^2+\sigma_1^2}$$

$$\sigma'^2=\sigma_0^2-\frac{\sigma_0^4}{\sigma_0^2+\sigma_1^2}$$

具体的推导过程本文不再赘述，可以参考这篇知乎链接

进一步，如果我们令$K=\frac{\sigma_0^2}{\sigma_0^2+\sigma_1^2}$,则可以将上面两式简化为：

$$\mu'=\mu_0+K(\mu_1-\mu_0)$$

$$\sigma'^2=\sigma_0^2-K\sigma_0^2$$

更进一步，将化简后的结果写成矩阵形式

$$K=\Sigma_0(\Sigma_0+\Sigma_1)^{-1}$$

$$\vec{\mu'}=\vec{\mu_0}+K(\vec{\mu_1}-\vec{\mu_0})$$

$$\Sigma'=\Sigma_0-K\Sigma_0$$

这时候我们已有预测部分$(\mu_0,\Sigma_0)=(H_k\hat{X}_k,H_kP_kH_k^T)$,实际观测部分记为$(\mu_1,\Sigma_1)=(z_k,R_k)$，其中$z_k$为观测到的值，$R_k$为观测误差，也可以理解为噪声。这里运用我们开头提出的卡尔曼滤波的基本思想：将预测部份和观测部分融合，得到最优估计，再进行迭代

融合方法就是刚刚提到的融合高斯分布，将两组数据带入最后得到的矩阵形式的表达式中，就可以推出卡尔曼滤波对于当前状态的最优估计

$$K=H_kP_kH_k^T(H_kP_kH_k^T+R_k)^{-1}$$

$$H_k\hat{X}_k{}'=H_k\hat{X}_k+K(z_k-H_k\hat{X}_k)$$

$$H_kP_k'H_k^T=H_kP_kH_k^T-KH_kP_kH_k^T$$

由于K中含有Hk项，可以对后两个等式进行化简：

$$\hat{X}_k{}' = \hat{X}_k + K'(z_k - H_k\hat{X}_k)$$

$$P_k' = P_k - K'H_kP_k$$

这里的$K' = P_kH_k^T(H_kP_kH_k^T + R_k)^{-1}$就是传说中的卡尔曼增益

至此，卡尔曼滤波的理论推导就已经完成了，不知道大家学会了没有呢^_^

卡尔曼滤波的应用

卡尔曼滤波的推导过程似乎有点虐，不过我看来尚可接受，在应用时，只需要掌握卡尔曼滤波的五大核心，既可以理顺思路，又可以理解精髓！

1. 根据估计值计算下一次的预测值

$$\hat{X}_k = F_k\hat{X_{k-1}} + B_ku_k + W_k$$

2. 计算预测值的协方差矩阵

$$P_k = F_kP_{k-1}F_k^T + Q_k$$

3. 计算卡尔曼增益

$$K' = P_kH_k^T(H_kP_kH_k^T + R_k)^{-1}$$

4. 计算估计值

$$\hat{X}_k{}' = \hat{X}_k + K'(z_k - H_k\hat{X}_k)$$

5. 计算估计值的协方差矩阵

$$P_k' = P_k - K'H_kP_k$$

在应用时，我们只需要将这五行代码迭代，卡尔曼滤波还有一个好处就是需要的存储空间小，每次迭代只需要用到前一次的值即可！