

## 卡尔曼滤波简介说明及其算法 MATLAB实现代码

卡尔曼滤波算法实现代码 ( C, C++ 分别实现 )

### 卡尔曼滤波器简介

近来发现有些问题很多人都很感兴趣。所以在这里希望能尽自己能力跟大家讨论一些力所能及的算法。现在先讨论一下卡尔曼滤波器，如果时间和能力允许，我还希望能够写写其他的算法，例如遗传算法，傅立叶变换，数字滤波，神经网络，图像处理等等。因为这里不能写复杂的数学公式，所以也只能形象的描述。希望如果哪位是这方面的专家，欢迎讨论更正。

### 卡尔曼滤波器 – Kalman Filter

#### 1. 什么是卡尔曼滤波器 ( What is the Kalman Filter? )

在学习卡尔曼滤波器之前，首先看看为什么叫“卡尔曼”。跟其他著名的理论（例如傅立叶变换，泰勒级数等等）一样，卡尔曼也是一个人的名字，而跟他们不同的是，他是个现代人！

卡尔曼全名 Rudolf Emil Kalman，匈牙利数学家，1930 年出生于匈牙利首都布达佩斯。1953，1954 年于麻省理工学院分别获得电机工程学士及硕士学位。1957 年于哥伦比亚大学获得博士学位。我们现在要学习的卡尔曼滤波器，正是源于他的博士论文和 1960 年发表的论文《A New Approach to Linear Filtering and Prediction Problems》（线性滤波与预测问题的新方法）。如果对这篇论文有兴趣，可以到这里的地址下载：  
<http://www.cs.unc.edu/~welch/media/pdf/Kalman1960.pdf>。

简单来说，卡尔曼滤波器是一个“optimal recursive data processing algorithm（最优化自回归数据处理算法）”。对于解决很大部分的问题，他是最优，效率最高甚至是最有用的。他的广泛应用已经超过 30 年，包括机器人导航，控制，传感器数据融合甚至在军事方面的雷达系统以及导弹追踪等等。近年来更被应用于计算机图像处理，例如头脸识别，图像分割，图像边缘检测等等。

#### 2. 卡尔曼滤波器的介绍 ( Introduction to the Kalman Filter )

为了可以更加容易的理解卡尔曼滤波器，这里会应用形象的描述方法来讲解，而不是像大多数参考书那样罗列一大堆的数学公式和数学符号。但是，他的 5 条公式是其核心内容。结合现代的计算机，其实卡尔曼的程序相当的简单，只要你理解了他的那 5 条公式。

在介绍他的 5 条公式之前，先让我们来根据下面的例子一步一步的探索。

假设我们要研究的对象是一个房间的温度。根据你的经验判断，这个房间的温度是恒定的，也就

是下一分钟的温度等于现在这一分钟的温度（假设我们用一分钟来做时间单位）。假设你对你的经验不是 100% 的相信，可能会有上下偏差几度。我们把这些偏差看成是高斯白噪声（White Gaussian Noise），也就是这些偏差跟前后时间是没有关系的而且符合高斯分配（Gaussian Distribution）。另外，我们在房间里放一个温度计，但是这个温度计也不准确的，测量值会比实际值偏差。我们也把这些偏差看成是高斯白噪声。

好了，现在对于某一分钟我们有两个有关于该房间的温度值：你根据经验的预测值（系统的预测值）和温度计的值（测量值）。下面我们要用这两个值结合他们各自的噪声来估算出房间的实际温度值。

假如我们要估算  $k$  时刻的是实际温度值。首先你要根据  $k-1$  时刻的温度值，来预测  $k$  时刻的温度。因为你相信温度是恒定的，所以你会得到  $k$  时刻的温度预测值是跟  $k-1$  时刻一样的，假设是 23 度，同时该值的高斯噪声的偏差是 5 度（5 是这样得到的：如果  $k-1$  时刻估算出的最优温度值的偏差是 3，你对自己预测的不确定度是 4 度，他们平方相加再开方，就是 5）。然后，你从温度计那里得到了  $k$  时刻的温度值，假设是 25 度，同时该值的偏差是 4 度。

由于我们用于估算  $k$  时刻的实际温度有两个温度值，分别是 23 度和 25 度。究竟实际温度是多少呢？相信自己还是相信温度计呢？究竟相信谁多一点，我们可以用他们的 covariance 来判断。因为  $K_g^2 = 5^2 / (5^2 + 4^2)$ ，所以  $K_g = 0.78$ ，我们可以估算出  $k$  时刻的实际温度值是： $23 + 0.78 * (25 - 23) = 24.56$  度。可以看出，因为温度计的 covariance 比较小（比较相信温度计），所以估算出的最优温度值偏向温度计的值。

现在我们已经得到  $k$  时刻的最优温度值了，下一步就是要进入  $k+1$  时刻，进行新的最优估算。到现在为止，好像还没看到什么自回归的东西出现。对了，在进入  $k+1$  时刻之前，我们还要算出  $k$  时刻那个最优值（24.56 度）的偏差。算法如下： $((1 - K_g) * 5^2)^{0.5} = 2.35$ 。这里的 5 就是上面的  $k$  时刻你预测的那个 23 度温度值的偏差，得出的 2.35 就是进入  $k+1$  时刻以后  $k$  时刻估算出的最优温度值的偏差（对应于上面的 3）。

就是这样，卡尔曼滤波器就不断的把 covariance 递归，从而估算出最优的温度值。他运行的很快，而且它只保留了上一时刻的 covariance。上面的  $K_g$ ，就是卡尔曼增益（Kalman Gain）。他可以随不同的时刻而改变他自己的值，是不是很神奇！

下面就要言归正传，讨论真正工程系统上的卡尔曼。

### 3. 卡尔曼滤波器算法 (The Kalman Filter Algorithm)

在这一部分，我们就来描述源于 Dr Kalman 的卡尔曼滤波器。下面的描述，会涉及一些基本的概念知识，包括概率（Probability），随即变量（Random Variable），高斯或正态分配（Gaussian Distribution）还有 State-space Model 等等。但对于卡尔曼滤波器的详细证明，这里不能一一描述。

首先，我们先要引入一个离散控制过程的系统。该系统可用一个线性随机微分方程（Linear Stochastic Difference equation）来描述：

$$X(k)=A X(k-1)+B U(k)+W(k)$$

再加上系统的测量值：

$$Z(k)=H X(k)+V(k)$$

上两式子中， $X(k)$  是  $k$  时刻的系统状态， $U(k)$  是  $k$  时刻对系统的控制量。 $A$  和  $B$  是系统参数，对于多模型系统，他们为矩阵。 $Z(k)$  是  $k$  时刻的测量值， $H$  是测量系统的参数，对于多测量系统， $H$  为矩阵。 $W(k)$  和  $V(k)$  分别表示过程和测量的噪声。他们被假设成高斯白噪声 (White Gaussian Noise)，他们的 covariance 分别是  $Q, R$  (这里我们假设他们不随系统状态变化而变化)。

对于满足上面的条件 (线性随机微分系统，过程和测量都是高斯白噪声)，卡尔曼滤波器是最优的信息处理器。下面我们用他们结合他们的 covariances 来估算系统的最优化输出 (类似上一节那个温度的例子)。

首先我们要利用系统的过程模型，来预测下一状态的系统。假设现在的系统状态是  $X(k-1|k-1)$ ，根据系统的模型，可以基于系统的上一状态而预测出现在状态：

$$X(k|k-1)=A X(k-1|k-1)+B U(k) \dots (1)$$

式(1) 中， $X(k|k-1)$  是利用上一状态预测的结果， $X(k-1|k-1)$  是上一状态最优的结果， $U(k)$  为现在状态的控制量，如果没有控制量，它可以为 0。

到现在为止，我们的系统结果已经更新了，可是，对应于  $X(k|k-1)$  的 covariance 还没更新。我们用  $P$  表示 covariance：

$$P(k|k-1)=A P(k-1|k-1) A' +Q \dots (2)$$

式(2) 中， $P(k|k-1)$  是  $X(k|k-1)$  对应的 covariance， $P(k-1|k-1)$  是  $X(k-1|k-1)$  对应的 covariance， $A'$  表示  $A$  的转置矩阵， $Q$  是系统过程的 covariance。式子 1, 2 就是卡尔曼滤波器 5 个公式当中的前两个，也就是对系统的预测。

现在我们有现在状态的预测结果，然后我们再收集现在状态的测量值。结合预测值和测量值，我们可以得到现在状态  $(k)$  的最优化估算值  $X(k|k)$ ：

$$X(k|k)=X(k|k-1)+K_g(k)(Z(k)-H X(k|k-1)) \dots (3)$$

其中  $K_g$  为卡尔曼增益 (Kalman Gain)：

$$K_g(k)=P(k|k-1) H' / (H P(k|k-1) H' + R) \dots (4)$$

到现在为止，我们已经得到了  $k$  状态下最优的估算值  $X(k|k)$ 。但是为了要另卡尔曼滤波器不断的运行下去直到系统过程结束，我们还要更新  $k$  状态下  $X(k|k)$  的 covariance：

$$P(k|k)=(I-K_g(k) H) P(k|k-1) \dots (5)$$

其中  $I$  为 1 的矩阵，对于单模型单测量， $I=1$ 。当系统进入  $k+1$  状态时， $P(k|k)$  就是式子 (2) 的  $P(k-1|k-1)$ 。这样，算法就可以自回归的运算下去。

卡尔曼滤波器的原理基本描述了，式子 1, 2, 3, 4 和 5 就是他的 5 个基本公式。根据这 5 个公式，可以很容易的实现计算机的程序。

下面，我会用程序举一个实际运行的例子。。。

#### 4. 简单例子

( A Simple Example )

这里我们结合第二第三节，举一个非常简单的例子来说明卡尔曼滤波器的工作过程。所举的例子是进一步描述第二节的例子，而且还会配以程序模拟结果。

根据第二节的描述，把房间看成一个系统，然后对这个系统建模。当然，我们见的模型不需要非常地精确。我们所知道的这个房间的温度是跟前一时刻的温度相同的，所以  $A=1$ 。没有控制量，所以  $U(k)=0$ 。因此得出：

$$X(k|k-1)=X(k-1|k-1) \dots (6)$$

式子 (2) 可以改成：

$$P(k|k-1)=P(k-1|k-1)+Q \dots (7)$$

因为测量的值是温度计的，跟温度直接对应，所以  $H=1$ 。式子 3, 4, 5 可以改成以下：

$$X(k|k)=X(k|k-1)+Kg(k)(Z(k)-X(k|k-1)) \dots (8)$$

$$Kg(k)=P(k|k-1)/(P(k|k-1)+R) \dots (9)$$

$$P(k|k)=(1-Kg(k))P(k|k-1) \dots (10)$$

现在我们模拟一组测量值作为输入。假设房间的真实温度为 25 度，我模拟了 200 个测量值，这些测量值的平均值为 25 度，但是加入了标准偏差为几度的高斯白噪声（在图中为蓝线）。

为了令卡尔曼滤波器开始工作，我们需要告诉卡尔曼两个零时刻的初始值，是  $X(0|0)$  和  $P(0|0)$ 。他们的值不用太在意，随便给一个就可以了，因为随着卡尔曼的工作， $X$  会逐渐的收敛。但是对于  $P$ ，一般不要取 0，因为这样可能会令卡尔曼完全相信你给定的  $X(0|0)$  是系统最优的，从而使算法不能收敛。我选了  $X(0|0)=1$  度， $P(0|0)=10$ 。

该系统的真实温度为 25 度，图中用黑线表示。图中红线是卡尔曼滤波器输出的最优化结果（该结果在算法中设置了  $Q=1e-6$ ， $R=1e-1$ ）。

最佳线性滤波理论起源于 40 年代美国科学家 Wiener 和前苏联科学家 K 等人的研究工作，后人统称为维纳滤波理论。从理论上说，维纳滤波的最大缺点是必须用到无限过去的数  
据，不适用于实时处理。为了克服这一缺点，60 年代 Kalman 把状态空间模型引入滤波理论，并导出了一套递推估计算法，后人称之为卡尔曼滤波理论。卡尔曼滤波是以最小均方误差为估计的最佳准则，来寻求一套递推估计的算法，其基本思想是：采用信号与噪声的状态空间模型，利用前一时刻地估计值和现时刻的观测值来更新对状态变量的估计，求出现时刻的估计值。它适合于实时处理和计算机运算。

现设线性时变系统的离散状态方程和观测方程为：

$$X(k) = F(k,k-1) \cdot X(k-1) + T(k,k-1) \cdot U(k-1)$$

$$Y(k) = H(k) \cdot X(k) + N(k)$$

其中

$X(k)$  和  $Y(k)$  分别是  $k$  时刻的状态矢量和观测矢量

$F(k,k-1)$  为状态转移矩阵

$U(k)$  为  $k$  时刻动态噪声

$T(k,k-1)$  为系统控制矩阵

$H(k)$  为  $k$  时刻观测矩阵

$N(k)$  为  $k$  时刻观测噪声

则卡尔曼滤波的算法流程为：

预估计  $\hat{X}(k) = F(k,k-1) \cdot X(k-1)$

1. 计算预估计协方差矩阵

$$C(k) = F(k,k-1) \times C(k-1) \times F(k,k-1)' + T(k,k-1) \times Q(k) \times T(k,k-1)'$$

$$Q(k) = U(k) \times U(k)'$$

2. 计算卡尔曼增益矩阵

$$K(k) = C(k) \times H(k)' \times [H(k) \times C(k) \times H(k)' + R(k)]^{-1}$$

$$R(k) = N(k) \times N(k)'$$

3. 更新估计

$$\hat{X}(k) = \hat{X}(k) + K(k) \times [Y(k) - H(k) \times \hat{X}(k)]$$

4. 计算更新后估计协方差矩阵

$$C(k) =$$

$$[I - K(k) \times H(k)] \times C(k) \times [I - K(k) \times H(k)]' + K(k) \times R(k) \times K(k)'$$

5.  $\hat{X}(k+1) = \hat{X}(k)$

$$C(k+1) = C(k)$$

重复以上步骤

## Kalman Filter

科技 2010-05-29 21:13:49 阅读 90 评论 0 字号：大中小 订阅

Kalman Filter 是一个高效的递归滤波器，它可以实现从一系列的噪声测量中，估计动态系统的状态。广泛应用于包含 Radar、计算机视觉在内的等工程应用领域，在控制理论和控制系统工程中也是一个非常重要的课题。连同线性均方规划，卡尔曼滤波器可以用于解决 LQG(Linear-quadratic-Gaussian control) 问题。卡尔曼滤波器，线性均方归化及线性均方高斯控制器，是大部分控制领域基础难题的主要解决途径。

### 目录

- 1 应用实例
- 2 命名和发展历史
- 3 基本动态系统模型
- 4 卡尔曼滤波器
  - 4.1 预测
  - 4.2 更新
  - 4.3 不变量
- 5 实例
- 6 推导
  - 6.1 后验估计协方差矩阵推导
  - 6.2 Kalman 增益推导
  - 6.3 后验误差协方差矩阵简化
- 7 信息滤波
- 8 非线性滤波器
  - 8.1 扩展 Kalman 滤波
  - 8.2 Unscented Kalman filter
- 9 Kalman-Bucy 滤波
- 10 应用
- 11 参见
- 12 参考文献
- 13 外部链接
- 1 应用实例

一个简单的应用是估计物体的位置和速度； 简要描述如下： 假设我们可以获取一个物体的包含噪声的一系列位置观测数据，我们可以获得此物体的精确速度和位置连续更新信息。

例如，对于雷达来说，我们关心的是跟踪目标，而目标的位置，速度，加速度的测量值是时刻含有误差的， 卡尔曼滤波器利用目标的动态信息， 去掉噪声影响， 获取目标此刻好的位置估计(滤波)，将来位置估计(预测)，也可以是过去位置估计的(插值或平滑)

### 2 命名和发展历史

这个滤波器以它的发明者 Rudolf.E.Kalman 而命名，但是在 Kalman 之前，Thorvald Nicolai Thiele 和 Peter Swerling 已经提出了类似的算法。 Stanley Schmidt 首次实现了 Kalman 滤波器。在一次对 NASA Ames Research Center 访问中，卡尔曼发现他的方法对于解决阿波罗计划的轨迹预测很有用，后来阿波罗飞船导航 电脑就使用了这种滤波器。这个滤波器可以追溯到 Swerling(1958),Kalman(1960),Kalman 和 Bucy(1961)发表的论文。

这个滤波器有时叫做 Stratonovich-Kalman-Bucy 滤波器。因为更为一般的非线性滤波器最初由 Ruslan L.Stratonovich 发明，而 Stratonovich-Kalman-Bucy 滤波器只是非线性滤波器的一个特例。事实上，1960 年夏季，Kalman 和 Stratonovich 在一个 Moscow 召开的会议中相遇，

而作为非线性特例的线性滤波方程，早已经由 Stratonovich 在此以前发表了。

在控制领域，Kalman 滤波被称为线性二次型估计，目前，卡尔曼滤波已经有很多不同的实现，有施密特扩展滤波器、信息滤波器以及一系列的 Bierman 和 Thornton 发明的平方根滤波器等，而卡尔曼最初提出的形式现在称为简单卡尔曼滤波器。也许最常见的卡尔曼滤波器应用是锁相环，它在收音机、计算机和几乎全部视频或通讯设备中广泛存在。

### 3 基本动态系统模型

Kalman 滤波基于时域描述的线性动态系统，它的模型是 Markov Chain，而 Markov Chain 建立在一个被高斯噪声干扰的线性算子之上。系统的状态可以用一个元素为实数的向量表示。

随着离散时间的增加，这个线性算子就会作用到当前状态之上，产生一个新的状态，并且会带入一定的噪声，同时一些已知的控制信息也会加入。同时另外一个受噪声干扰的线性算子将产生这些隐含状态的可见输出。Kalman 滤波可以被看作为类似隐马尔科夫模型，它们的显著不同点在于：隐状态变量的取值空间是一个连续的空间，而离散状态空间则不是；另为，隐马尔科夫模型可以描述下一个状态的一个任意分布，这也与应用于 Kalman 滤波器中的高斯噪声模型相反。Kalman 滤波器方程和隐马尔科夫方程之间有很大的二重性，关于 Kalman 滤波方程和隐马尔科夫方程之间二重性参看 Roweis and Ghahramani(1999)[4]。

为了从一系列的噪声观测中，应用 Kalman 滤波估计观测过程的内部状态。我们必须把这个过程在 Kalman 滤波器的框架下建立模型，这就意味着，对于

每一步  $k$  我们要定义矩阵  $F_k$ 、 $H_k$ 、 $Q_k$ 、 $R_k$ 、 $B_k$  如下：

Kalman Filter 假设  $k$  时刻的真实状态是从  $k-1$  时刻演化而来，符合下式

$$X_k = F_k x_{k-1} + B_k u_k + w_k$$

这里

$F_k$  是作用在前一状态的状态转移模型（状态转移矩阵）

$B_k$  是作用在控制向量上的控制输入模型（输入输出矩阵）

$w_k$  是过程噪声，假设是均值为 0 的白噪声，协方差为  $Q_k$  则： $w_k \sim N(0, Q_k)$

在  $k$  时刻，假设真实状态  $x_k$  的观测， $Z_k$  满足如下公式：

$$Z_k = H_k x_k + v_k$$

其中  $H_k$  是观测模型（观测矩阵），它把真实状态映射到观测空间， $v_k$  是观测噪声，假设它

是均值是 0，方差是  $R_k$  的高斯白噪声： $v_k \sim N(0, R_k)$

Kalman Filter 基本动态系统模型如图 (1) 所示，圆圈代表向量，方块代表矩阵，星号代表高斯噪声，其协方差在右下方标出。

初始状态以及每一时刻的噪声向量  $\{x_0, w_1, \dots, w_k, v_1 \dots v_k\}$  都为认为是互相独立的。实际中，真实世界中动态系统并不是严格的符合此模型。但是 Kalman 模型是设计在噪声过程工作的，一个近似的符合已经可以使这个滤波器非常有用了，更多复杂模型关于 Kalman Filter 模型的变种，将在下述中讨论：

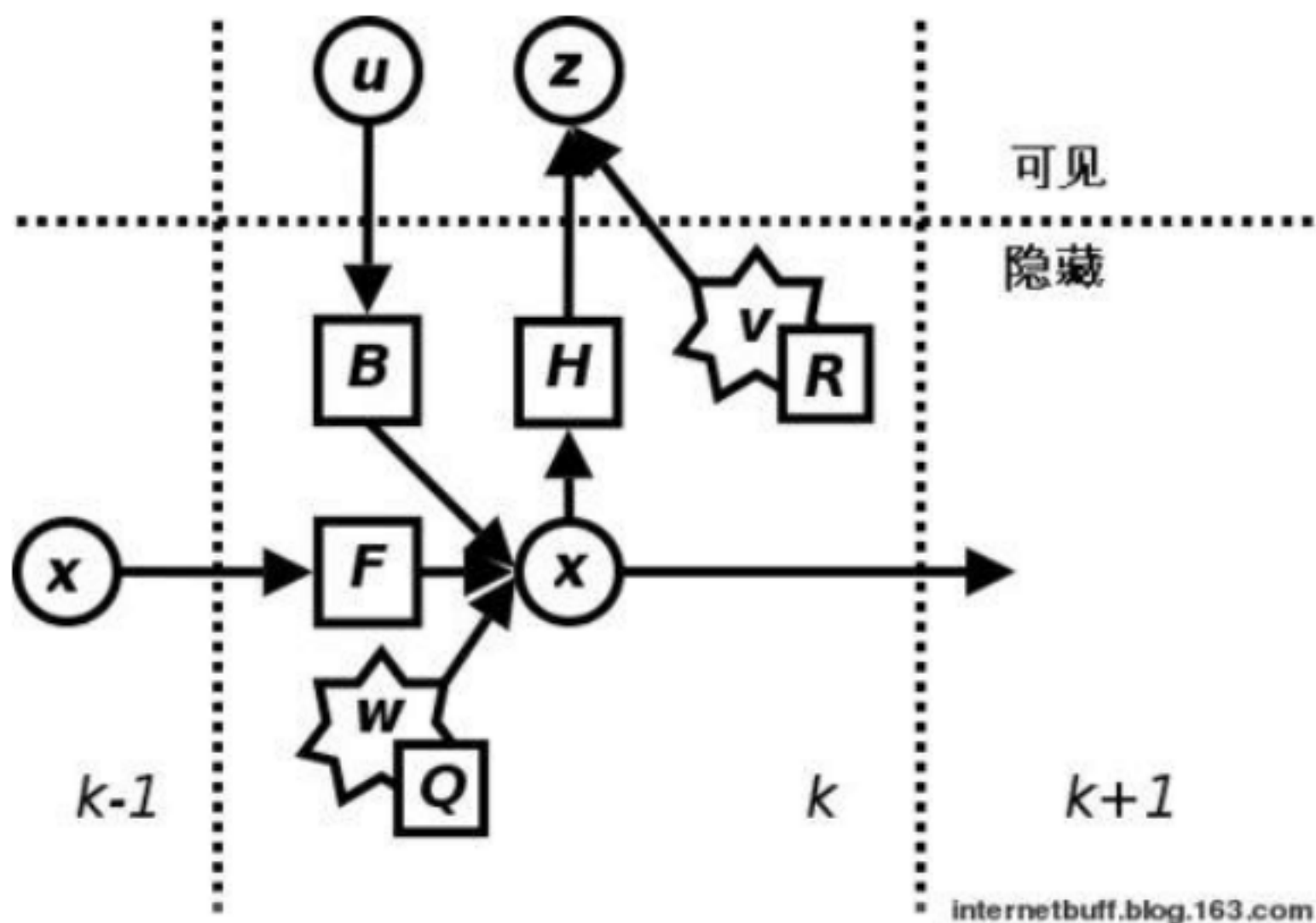


图 (1)

#### 4 卡尔曼滤波器

Kalman Filter 是一个递归的估计，即只要获知上一时刻的状态估计和当前状态的观测就可以计算出当前状态的估计，不同于其他的估计技术，Kalman 滤波器不需要观测或 / 和估计的历史记录，Kalman Filter 是一个纯粹的时域滤波器，而不像低通滤波器等频域滤波器那样，需要在频域中设计，然后转换到时域中应用。

下面， $\hat{x}_{n|m}$  代表已知从 m 到 n-1 包括 m 时刻的观测在 n 时刻的估计值

卡尔曼滤波器的状态由以下两个变量表示：

$\hat{x}_{k|k}$  已知 k 时刻以前时刻观测值，k 时刻的状态估计值

$P_{k|k}$  误差协方差矩阵，度量状态估计的精度程度

Kalman 滤波包括两个阶段：预测和更新；在估计阶段，滤波器应用上一状态的估计做出对当前状态的估计。在更新阶段，滤波器利用在当前状态的观测值优化预测阶段的预测值，以获的一个更精确的当前状态的估计。

##### 4.1 预测

状态预测：

$$\hat{x}_{k|k-1} = F_k \hat{x}_{k-1|k-1} + B_{k-1} u_{k-1}$$

估计协方差预测：

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^T + Q_{k-1}$$

##### 4.2 更新

新息或测量余量

$$\tilde{y}_k = Z_k - H_k \hat{x}_{k|k-1}$$



新息协方差

$$S_k = H_k P_{k|k-1} H_k^T + R_k$$

Kalman 增益

$$K_k = P_{k|k-1} H_k^T S_k^{-1}$$

状态估计更新

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + K_k \tilde{y}_k$$

状态协方差更新

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}$$

使用上述公式计算  $P_{k|k}$  仅在最优卡尔曼增益的时候有效。使用其他增益公式要复杂一些，  
看见推导

#### 4.3 不变量

如果模型准确， $x_{0|0}$  和  $P_{0|0}$  值将准确反映最初状态的分布，那么下面所有不变量保持不变，  
所有估计的误差均值为 0：

$$E(x_k - x_{k|k}) = E(x_k - x_{k|k-1}) = 0$$

$$E(\tilde{y}_k) = 0$$

这里  $E[\xi]$  表示  $\xi$  的期望，而协方差矩阵则反映的估计的协方差

$$P_{k|k} = \text{cov}(x_k - \hat{x}_{k|k})$$

$$P_{k|k-1} = \text{cov}(x_k - \hat{x}_{k|k-1})$$

$$S_k = \text{cov}(\tilde{y}_k)$$

#### 5 实例

考虑在一个无摩擦、无限长的直轨道上的一辆小车，它的初始位置在 0 点，但是它会随机的  
受到冲击作用，我们每隔  $\Delta t$  测量一次小车的位置，但是这些测量数据不是很精确。我们想  
建立一个关于小车位置和速度的模型，这里我们描述如何建立这个模型，以及从这个模型出  
发如何推导出 Kalman 滤波器。

因为小车没有控制输入，我们可以忽略  $B_k$  和  $u_k$ 。由于  $F$ ， $H$ ， $R$  和  $Q$  全是恒值，我们可以  
忽略时间下标。

小车的位置和速度用线性空间可以描述如下：

$$x_k = \begin{bmatrix} x \\ \dot{x} \end{bmatrix}$$

这里  $\dot{x}$  表示速度，也就是位置对时间的微分。

我们假设在时间间隔  $k-1$  和  $k$  之间，小车受到一个恒定的冲击  $a_k$ ， $a_k$  服从均值为 0，方差为  $\delta_a$  的正态分布，根据 Newton 动力学方程，可得到：

$$x_k = Fx_{k-1} + Ga_k$$

其中

$$F = \begin{bmatrix} 1 & \Delta t \\ 0 & 1 \end{bmatrix}, \quad G = \begin{bmatrix} \Delta t^2 \\ \Delta t \end{bmatrix}$$

我们发现：
$$Q = \text{cov}(Ga) = E[(Ga)(Ga)^T] = GE[a^2]G^T = \delta_a^2 GG^T$$

在每一时刻，我们获取真实位置的  $z_k$ 。我们假设噪声服从噪声干扰测量，假设测量噪声服从均值为 0，标准差为  $\delta_z$  正态分布。

$$z_k = Hx_k + v_k \quad \text{其中 } H = [1 \quad 0], \quad R = E[v_k v_k^T] = [\delta_z^2]$$

我们可以得到足够精度的初始状态数据，所以我们可以初始化  $\hat{x}_{00} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ， $P_{00} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ ，如果初始位置和速度不是精确的知道，那么协方差矩阵应该初始化为一个对角线元素  $B$  为适当大小的矩阵如下：

$$P_{00} = \begin{bmatrix} B & 0 \\ 0 & B \end{bmatrix}$$

这样与模型中已有信息相比，滤波器更趋向于使用首次的测量数据信息。

## 6 推导

### 6.1 后验估计协方差矩阵推导

首先开始不变量后验估计协方差矩阵  $P_{k|k}$  的推导： $P_{k|k} = \text{cov}(x_k - \hat{x}_{k|k})$  带入  $\hat{x}_{k|k}$  定义，可得

$$P_{k|k} = \text{cov}(x_k - (\hat{x}_{k|k-1} + K_k \tilde{y}_k))$$

，代入  $\tilde{y}_k$  可得

$$P_{k|k} = \text{cov}(x_k - (\hat{x}_{k|k-1} + K_k (z_k - H_k \hat{x}_{k|k-1})))$$

代入  $z_k$  可得

$$P_{k|k} = \text{cov}(x_k - (\hat{x}_{k|k-1} + K_k (H_k x_k + v_k - H_k \hat{x}_{k|k-1})))$$

整理误差向量可得，

$$P_{k|k} = \text{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1}) - K_k v_k)$$

由于误差向量  $v_k$  与其他不相关，所以

$$P_{k|k} = \text{cov}((I - K_k H_k)(x_k - \hat{x}_{k|k-1})) + \text{cov}(K_k v_k)$$

由协方差矩阵性质则

$$P_{k|k} = (I - K_k H_k) \text{cov}(x_k - \hat{x}_{k|k-1}) (I - K_k H_k)^T + K_k \text{cov}(v_k) K_k^T$$

：使用不变量  $P_{k|k-1}$  以及

$R_k$  的定义这一项可以写作  $P_{k|k} = (I - K_k H_k) P_{k|k-1} (I - K_k H_k)^T + K_k R_k K_k^T$  , 此公式 (Joseph form) 对任意增益  $K_k$  的都成立, 如果  $K_k$  最优卡尔曼增益, 则可以进一步简化, 见下文。

## 6.2 Kalman 增益推导

Kalman 滤波器是一个最小均方误差估计器, 先验状态误差估计可表示为  $x_k - \hat{x}_{k|k}$  我们最小

化这个矢量幅度平方的期望值  $E[|x_k - \hat{x}_{k|k}|^2]$ , 这等价于最小化后验估计协方差矩阵  $P_{k|k}$  的

迹, 通过展开合并  $P_{k|k}$  公式, 可得

$$\begin{aligned} P_{k|k} &= P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^T K_k^T + K_k (H_k P_{k|k-1} H_k^T + R_k) K_k^T \\ &= P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^T K_k^T + K_k S_k K_k^T \end{aligned}$$

$$\frac{\partial \text{tr}(P_{k|k})}{\partial K_k} = -2(H_k P_{k|k-1})^T + 2K_k S_k = 0$$

当矩阵导数为 0 时, 矩阵的迹取最小值,

从这个式子解出 Kalman 增益  $K_k$

$$K_k S_k = (H_k P_{k|k-1})^T = P_{k|k-1} H_k^T$$

$$K_k = P_{k|k-1} H_k^T S_k^{-1}$$

这个增益就是最优 Kalman 增益, 应用它可以得到最小均方误差。

## 6.3 后验误差协方差矩阵简化

当应用上述最优 Kalman 增益时, 后验误差协方差可以得到简化, 在最优 Kalman 增益两边

同时乘以  $S_k K_k^T$ , 可得  $K_k S_k K_k^T = P_{k|k-1} H_k^T K_k^T$ , 参见后验误差协方差公式展开

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} - P_{k|k-1} H_k^T K_k^T + K_k S_k K_k^T$$

, 带入上式, 可得:

$$P_{k|k} = P_{k|k-1} - K_k H_k P_{k|k-1} = (I - K_k H_k) P_{k|k-1}$$

这个公式的计算比较简单, 所以实际中总是使

用这个公式, 但是需注意这公式仅在最优卡尔曼增益时它才成立。如果算术精度总是很低而导致数值稳定性出现问题, 或者特意使用非最优卡尔曼增益, 那么就不能使用这个简化; 必须使用上面导出的后验误差协方差公式。

## 7 信息滤波

在信息滤波器 (逆方差滤波器) 中, 协方差估计和状态估计将会被信息矩阵和信息向量所取代, 它们的定义如下:

$$Y_{k|k} = P_{k|k}^{-1}$$

$$\hat{y}_{k|k} = P_{k|k}^{-1} \hat{x}_{k|k}$$

类似的预测协方差和预测状态也有等价的信息形式, 定义如下:

$$Y_{k|k-1} = P_{k|k-1}^{-1}$$

$$\hat{y}_{k|k-1} = P_{k|k-1}^{-1} \hat{x}_{k|k-1}$$

同样测量协方差和测量向量定义为：

$$I_k = H_k^T R_k^{-1} H_k$$

$$i_k = H_k^T R_k^{-1} Z_k$$

信息更新现在变成一个加和形式：

$$Y_{k|k} = Y_{k|k-1} + \sum_{j=1}^N I_{k,j}$$

$$\hat{y}_{k|k} = \hat{y}_{k|k-1} + \sum_{j=1}^N i_{k,j}$$

信息滤波器的主要优点在于  $N$  和测量数据都可以用于滤波，简单的通过信息矩阵和信息向量的加和。

为了预测信息滤波器，信息矩阵和信息向量必须变换到它们的等价状态空间，或者应用下述信息空间更新：

$$M_k = [F_k^{-1}]^T Y_{k-1|k-1} F_k^{-1}$$

$$C_k = M_k [M_k + Q_k^{-1}]^{-1}$$

$$L_k = I - C_k$$

$$Y_{k|k-1} = L_k M_k L_k^T + C_k Q_k^{-1} C_k^T$$

$$\hat{y}_{k|k-1} = L_k [F_k^{-1}]^T \hat{y}_{k-1|k-1}$$

这里  $F$  和  $Q$  必须可逆。

## 8 非线性滤波器

### 8.1 扩展 Kalman 滤波

#### 估计过程

如以上所述，卡尔曼滤波器估计一个线性随机差分方程描述的离散时间过程的状态变量

$x \in \mathbb{R}^n$ ，但是如果被估计的过程和（或）观测变量与过程的关系不是线性关系。那该如何处理

呢？一些很有趣和成功的 Kalman 滤波器应用就是处理这些情况的。将期望和方差线性化的卡尔曼滤波器称作扩展卡尔曼滤波器（Extended Kalman Filter），简称 EKF。

同泰勒级数类似，面对非线性关系时，我们可以通过求过程和量测方程的偏导来线性化并计算当前估计，为了实现这个目的，我们必须修改上面的一些描述，我们假设过程仍具有状态

向量  $x \in \mathbb{R}^n$ ，但其状态方程已变为非线性随机差分方程的形式。

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$$

观测变量  $Z \in \mathbb{R}^m$  为：

$$Z_k = h(x_k, v_k)$$

这里随机变量  $w_k$  和  $v_k$  分别为过程噪声和观测噪声。差分方程式 (1.1) 中的非线性函数  $f$  将

过去  $k-1$  时刻状态与现在  $k$  时刻状态联系起来。在测量方程 (2.2) 中，输入函数  $u_k$  和零均值过程噪声  $w_k$  是它的参数。非线性函数  $h$  反映了状态变量  $x_k$  和观测变量  $z_k$  的关系。实际中我们并不知道每一时刻噪声  $w_k$  和  $v_k$  各自真实值，但是我们可以在假设他们不存在的前提下，近似估计状态向量和测量向量：

$$\tilde{x} = f(\hat{x}_{k-1}, u_k, 0)$$

$$\tilde{z}_k = h(\tilde{x}_k, 0)$$

这里  $\hat{x}_k$  是相对于前一时刻  $k$  的后验状态估计。

有一点非常重要，那就是扩展卡尔曼滤波器的一个基本缺陷：离散随机变量的分布（或连续随机变量的密度）在经过非线性系统转化后不再是正态的了。扩展卡尔曼滤波器其实就是一个通过线性化而达到渐进最优贝叶斯决策的特殊状态估计器。[Julier96] 中描述了一项有趣的研究，Julier 设计了扩展卡尔曼滤波器的一种变体，使得通过非线性转换后的随机变量仍具有正态分布特性。

滤波器的计算原型

为了估计一个具有非线性差分 and 量测关系的过程，我们先给出式 1.3 和式 1.4 的一个新的线性化表示：

$$x_k \approx \tilde{x}_k + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}$$

$$z_k \approx \tilde{z}_k + H(x_k - \tilde{x}) + Vv_k$$

其中：

$x_k$  和  $z_k$  是状态向量和观测向量的真值；

$\tilde{x}_k$  和  $\tilde{z}_k$  来自 1.3 式和 1.4 式，是状态向量和观测向量的近似值；

$\hat{x}_k$  是  $k$  时刻状态向量的后验估计；

随机变量  $w_k$  和  $v_k$  表示过程激励噪声和观测噪声。

$A$  是  $f$  对  $x$  的偏导的雅可比矩阵：

$$A_{[i,j]} = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0)$$

$W$  是  $f$  对  $w$  的偏导的雅可比矩阵：

$$W_{[i,j]} = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\hat{x}_{k-1}, u_{k-1}, 0)$$

$H$  是  $h$  对  $x$  的偏导的雅可比矩阵：

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_k, 0)$$

$V$  是  $h$  对  $v$  的偏导的雅可比矩阵：

$$V_{[i,j]}=\frac{\partial h_{[i]}}{\partial v_{[j]}}(\tilde{x}_k,0)$$