

# 目 录

|                   |    |
|-------------------|----|
| 卡尔曼滤波.....        | 2  |
| 前言.....           | 2  |
| 一 天方夜谭.....       | 2  |
| 二 贝叶斯滤波.....      | 3  |
| 2.1 贝叶斯理论.....    | 3  |
| 2.2 白噪声和有色噪声..... | 4  |
| 2.3 贝叶斯框架.....    | 5  |
| 2.4 一些重要结论.....   | 7  |
| 三 卡尔曼滤波.....      | 8  |
| 3.1 问题阐述.....     | 8  |
| 3.2 优化模型.....     | 10 |
| 3.3 线性系统.....     | 10 |
| 3.4 初值问题.....     | 15 |
| 3.5 新息残差.....     | 15 |
| 四 工程实践.....       | 17 |
| 参考文献.....         | 20 |
| 声明.....           | 21 |

# 卡尔曼滤波

作者：追风少年

邮箱：ouyangacas@gmail.com

时间：2017/01/18

本文为作者原创文章，未经作者允许不得转载，谢谢合作！

## 前言

卡尔曼滤波器(Kalman Filter)发展已经接近 50 多年了，理论上可谓比较成熟，其最常见的应用有惯性导航系统、四旋翼、GPS-IMU 组合导航系统等。卡尔曼滤波器，是一种基于高斯白噪声的线性滤波系统，原理看似简单，但在实际应用过程中，常常出现似懂非懂、一知半解的尴尬，毕竟卡尔曼滤波建立的模型在数学上如此完美，因此工程应用上需要灵活变动，设计一个合理有效的滤波器成为本文想阐述的最重要的问题之一。本文将从卡尔曼滤波的概念、推导、性质和应用等多方面介绍什么是卡尔曼滤波器，希望能够为想学习卡尔曼滤波的读者解决疑惑。

## 一 天方夜谭

首先，我们回顾历史，翻开历史的扉页，控制理论从古典控制学派到现代控制学派推进。何为古典控制学派？那就是我们的万能助手 PID 控制器。PID 控制<sup>1</sup>的思想在于，模型常常不需要对应用系统有太多先验概率知识，类似于驾驶员控制油门：离目标车速越远，司机给油量呈线性增长，至于多大，给油即可，当观测车速逼近目标速度时给油减少，控制方法可谓完美。很明显，这是一种模糊控制方法，实际应用中简单粗暴有效果，已至于发展这么多年，依然应用于各种控制算法中。

尽管 PID 控制器效果很好，但老这样暴力控制对汽车硬件，舒适度不好，尤其是控制精度达不到要求，理论的发展总是一环扣一环，状态估计理论恰好解决这一弊端。卡尔曼滤波研究系统的噪声模型，系统基于误差的噪声模型对系统进行最优估计(近似)，然后做出最佳决策。卡尔曼滤波是一种状态估计器，实际应用中，当某些物理量不可测量时，就要寻找一些容易观测的变量，寻找这两者之间的联系(如函数映射关系  $\hat{y} = f(\hat{x})$ )，卡尔曼滤波主要负责估计这些噪声的大小，达到最优控制，是一种精确控制手段。

---

<sup>1</sup> 这里的控制是指一般意义的控制，即改变状态变量的值。

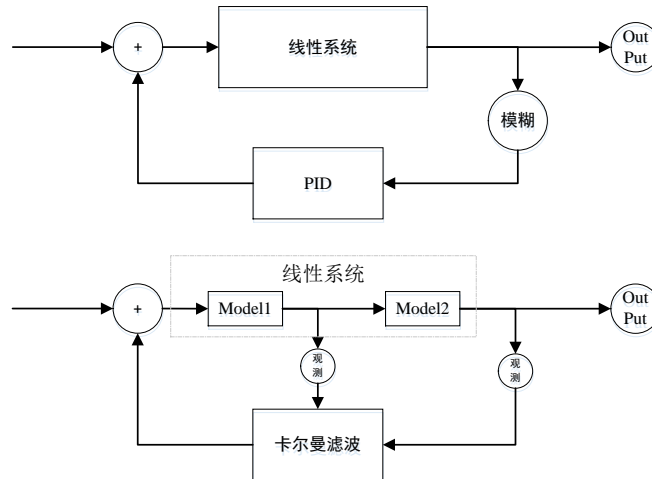


图1 PID控制与卡尔曼滤波

实际应用过程中，很多时候无法获取模型中噪声的概率统计特性，只能将噪声做出近似，假设其为高斯噪声。所以说，卡尔曼滤波是一个严重依赖模型(模型假设噪声为高斯白噪声)的滤波器，一旦假设不成立(比如间歇性抖动，噪声不为高斯)，滤波器会发散或错误动作。在实际应用过程中，若遇到滤波器不收敛或者效果不好，就需要分析设计的滤波器出错到底是什么原因造成的。对比 PID 和卡尔曼滤波的设计思想，可以总结出一条工程经验，当系统建立的模型不够精确(如 PID、反馈控制，无噪声模型)，控制2稳定性好；而当系统建立的模型精准(误差为高斯噪声)，控制精度高但易受干扰。在实际设计滤波器过程中，过于严格地假设和不那么严格的假设，一个追求控制精度，一个追求稳定性，是卡尔曼滤波器设计的准则之一，应折衷选择。

## 二 贝叶斯滤波

### 2.1 贝叶斯理论

在滤波算法中，因为无法获取精确的状态值，因此只能对状态量进行近似，因此也衍生出了许多不同的估计算法，也称为参数统计估计理论。按照代价函数的选取差异，可以大致分为最小均方误差估计(MMSE)、条件中位值估计和最大后验估计(MAP)，以及其变体最大似然估计(ML)等方法。

所有概率模型研究中最重要结论莫过于贝叶斯理论。假设有两个随机变量  $x \in X$  和  $z \in Z^3$ ，其联合概率密度函数为  $f(x, z)$ 。根据条件概率密度函数的链式法则，可以将其扩展为

<sup>2</sup> 这里的控制是指一般意义的控制，即改变状态变量的值。

<sup>3</sup> 一般地，选取符号  $x$  表示状态量，选取符号  $z$  表示观测量

$$f(x, z) = f(x | z)f(z) = f(z | x)f(x) \quad (1)$$

将(1)式进行变形，得出贝叶斯理论：

$$f(x | z) = \frac{f(z | x)f(x)}{f(z)} \quad (2)$$

其中， $f(x | z)$ 表示**后验概率**密度函数，它由 $f(z)$ 、 $f(z | x)$ 、 $f(x)$ 唯一决定， $f(x)$ 为状态量 $x$ （未知量）的先验概率密度分布函数，是状态量的先验信息； $f(z | x)$ 为观测量 $z$ 与 $x$ 之间的关系表达式，通常取决于选择的观测量(如传感器特性)； $f(z)$ 为常数。

最优估计的问题可以表述为：已知后验概率密度函数 $f(x | z)$ ，则最大后验估计是选取 $f(x | z)$ 为最大值时对应的 $x^*$ 为参数估计值，可以表示为

$$x^* = \underset{x}{\operatorname{argmax}} f(x | z) \quad (3)$$

如果 $f(x | z)$ 存在最大值，且具有连续的一阶导数，则获取最大值的必要条件是

$$\frac{df(x | z)}{dx} = 0 \text{ 或 } \frac{d \ln f(x | z)}{dx} = 0 \quad (4)$$

## 2.2 白噪声和有色噪声

白噪声，通俗一点来讲，是一种高斯概率模型，其满足均值为0，方差为 $\Sigma$ 的高斯模型，即 $x \sim N(0, \Sigma)$ 。噪声因其时间不确定性，无法计算傅里叶变换，一般地，工程上选取噪声的自相关函数进行傅里叶变换，称为功率谱。

从功率谱<sup>4</sup>定义上，白噪声是指噪声功率谱密度在整个实数范围内为常值的随机变量。我们知道冲击函数 $\delta(t)$ 和常值函数为一对傅里叶变换对，而白噪声功率谱密度函数为常值函数，这正好说明了白噪声的时间相关常数为零(自相关函数 $S(t)$ 为冲激函数 $\delta(t)$ )，即噪声的随机性与上一时刻的值无关。

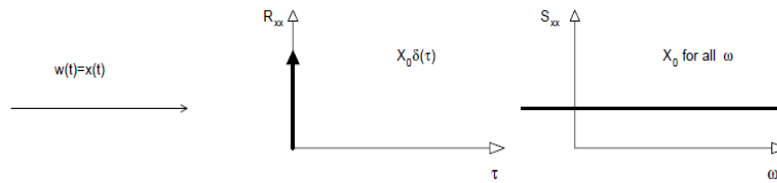


图2 白噪声自相关函数与功率谱

同时，白噪声在计算上有很多好的性质：假设有随机变量 $X$ 和 $Y$ ，其满足 $x \sim N(\mu_x, \Sigma_x)$ ，

<sup>4</sup> 功率谱是对噪声时间相关特性的傅里叶描述，研究噪声的功率特性。

$y \sim N(\mu_y, \Sigma_y)$ ，则有

1)  $E(X + Y) = EX + EY$

2)  $E(XY) = EX * EY$

3)  $E(h(X, Y)) \sim N(\mu, \Sigma)$ ，其中  $h(x, y)$  为线性函数

条件 3) 说明白噪声  $X, Y$  经过线性函数映射后输出的噪声仍然是白噪声，简化了系统模型的复杂度。

区别于白噪声，**有色噪声**当然就是功率谱密度不为常值的随机变量，不同功率谱形状决定了噪声的颜色，因其和光谱类似而得名。显然，有色噪声的自相关函数不是冲激函数，噪声值在时间域存在相关性，在实际应用过程中，常见的有色噪声有一阶马尔可夫过程(Markov Process)和随机游走噪声(Random Walk)。

**一阶马尔可夫过程**，是指噪声值仅与前一时刻的值存在相关性，可以用数学表示为

$$x_t = e^{-\beta \Delta t} x_{t-1} + w_t$$

其中  $\beta^{-1}$  为自相关时间， $\Delta t$  为时间采样周期。 $w$  为白噪声， $w \sim N(0, \sigma_x^2)$  且  $\sigma_{x_t}^2 = \sigma_{w_t}^2 / (1 - e^{-2\beta \Delta t})$ 。其相关函数为

$$R_x(\tau) = \sigma_{x_t}^2 e^{-\beta|\tau|}$$

**随机游走噪声**，是指噪声由白噪声积分而得的噪声，可以用数学表示为

$$x_t = x_{t-1} + w_t$$

在惯性传感器中，常常使用这三种模型对加速度计和陀螺仪进行建模。噪声由白噪声和随机游走噪声加性混合而成，若需要对噪声进行分离，可以考虑阿兰方差进行计算。

## 2.3 贝叶斯框架

什么是状态估计  $p(x|Z, U)$  ? Estimate the state  $x$  of a system given observations  $z$  and controls  $u$  . 下面将从数学推导的方式，给出贝叶斯框架的表达式：

### 1. 滤波目标

$$bel(x_t) = p(x_t | z_{1:t}, u_{1:t})$$

Definition of the belief

### 2. t 时刻分离 $x_t$ 和 $z_t$

$$\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta \underline{p(z_t \mid x_t, z_{1:t-1}, u_{1:t})} p(x_t \mid z_{1:t-1}, u_{1:t})
\end{aligned}$$

Bayes' rule

3. 传感器后验概率  $p(z_t \mid x_t)$

$$\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta \underline{p(z_t \mid x_t)} p(x_t \mid z_{1:t-1}, u_{1:t})
\end{aligned}$$

Markov assumption

4. t 时刻分离  $x_t$  与  $x_{t-1}$

$$\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} \underline{p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t})} \\
&\quad \underline{p(x_{t-1} \mid z_{1:t-1}, u_{1:t})} dx_{t-1}
\end{aligned}$$

Law of total probability

5. 状态推移传递概率  $p(x_t \mid x_{t-1}, u_t)$

$$\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\quad p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) dx_{t-1} \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} \underline{p(x_t \mid x_{t-1}, u_t)} p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) dx_{t-1}
\end{aligned}$$

Markov assumption

6. t-1 时刻  $p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1})$

$$\begin{aligned}
bel(x_t) &= p(x_t \mid z_{1:t}, u_{1:t}) \\
&= \eta p(z_t \mid x_t, z_{1:t-1}, u_{1:t}) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) p(x_t \mid z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} p(x_t \mid x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\quad p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) dx_{t-1} \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} p(x_t \mid x_{t-1}, u_t) p(x_{t-1} \mid z_{1:t-1}, u_{1:t}) dx_{t-1} \\
&= \eta p(z_t \mid x_t) \int_{x_{t-1}} p(x_t \mid x_{t-1}, u_t) p(x_{t-1} \mid z_{1:t-1}, u_{1:t-1}) \underline{p(x_{t-1} \mid u_{1:t-1})} dx_{t-1}
\end{aligned}$$

Markov assumption

7. 递归传递表达式

$$\begin{aligned}
bel(x_t) &= p(x_t | z_{1:t}, u_{1:t}) \\
&= \eta p(z_t | x_t, z_{1:t-1}, u_{1:t}) p(x_t | z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t | x_t) p(x_t | z_{1:t-1}, u_{1:t}) \\
&= \eta p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}, z_{1:t-1}, u_{1:t}) \\
&\quad p(x_{t-1} | z_{1:t-1}, u_{1:t}) dx_{t-1} \\
&= \eta p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}, u_t) p(x_{t-1} | z_{1:t-1}, u_{1:t}) dx_{t-1} \\
&= \eta p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}, u_t) p(x_{t-1} | z_{1:t-1}, u_{1:t-1}) dx_{t-1} \\
&= \eta p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}, u_t) \underline{bel(x_{t-1})} dx_{t-1}
\end{aligned}$$

**Recursive term**

根据以上推导,可以得出贝叶斯滤波的一般框架,滤波的过程总结为:状态预测+量测更新。  
 即 t-1 时刻状态  $x_{t-1}$  在控制量作用下预测  $x_k = f(x_{k-1}, u_k, w_k)$ , 然后预测量的不确定性依靠观测  
 量  $z_t$  对其校正  $p(z_t | x_t)$ 。

$$bel(x_t) = \eta p(z_t | x_t) \int_{x_{t-1}} p(x_t | x_{t-1}, u_t) bel(x_{t-1}) dx_{t-1}$$

常见的滤波器如卡尔曼滤波、粒子滤波等均采用贝叶斯滤波框架,滤波器在此充当传递函数特性,不断迭代更新状态量的条件概率密度函数,如何获取最佳递归特性是不同滤波器的区别。从这里就可以看出,白噪声因其独特的性质,在线性系统中噪声传递总是保证状态量的条件概率密度函数仍为高斯型,极大地减少了递推计算的复杂度。

## 2.4 一些重要结论

下面引用几条结论作为后续卡尔曼推导过程的先导条件,详细证明和论述请参照参考文献 [1], 描述如下:

**Result 3.2.1** : *The estimator that minimizes the conditional mean-square error is the conditional mean, [5], [10],*

$$\hat{\theta}(k) = E[\theta | Y_1^k]. \quad (3.10)$$

**Result 3.3.1** *If  $\theta$  and  $Y_1^k$  are jointly Gaussian random vectors, then,*

$$E[\theta | Y_1^k] = E[\theta] + R_{\theta Y_1^k} R_{Y_1^k}^{-1} [Y_1^k - E[Y_1^k]],$$

where

$$\begin{aligned}
R_{\theta Y_1^k} &= E[(\theta - E(\theta))(Y_1^k - E(Y_1^k))^T], \\
R_{Y_1^k Y_1^k} &= E[(Y_1^k - E(Y_1^k))(Y_1^k - E(Y_1^k))^T].
\end{aligned}$$

上面结论十分重要,它陈述, **如果  $\theta$  和  $Y_1^k$  为联合高斯概率密度变量,则  $\theta$  的最小条件均方**

误差估计可以表示为一组观测量  $Y_1^k$  的线性组合。结论能够被重写成如下形式

$$E[\theta|Y_1^k] = f(E(\theta), E(Y_1^k)) + \sum_{i=1}^k W_i Y_i,$$

如果  $\theta \in Y_1^k$  不是联合高斯概率密度变量，则  $E[\theta|Y_1^k]$  为一组观测量  $Y_1^k$  的非线性函数。

**Result 3.3.2** *In the situation considered in Result 3.3.1,  $\hat{\theta}(k)$  is an unbiased estimate of  $\theta$ , i.e.,*

$$E[\hat{\theta}(k)] = E[\theta]. \quad (3.23)$$

**Result 3.3.5** *Consider that  $\theta \in Y_1^k$  are not jointly Gaussian, but  $E[\theta]$ ,  $E[Y_1^k]$ ,  $R_{Y_1^k Y_1^k}$  and  $R_{\theta Y_1^k}$  are known. Then, the linear estimator that minimizes the mean square error is (still) given by*

$$\hat{\theta}(k) = E[\theta] + R_{\theta Y_1^k} R_{Y_1^k}^{-1} (Y_1^k - E[Y_1^k]). \quad (3.24)$$

### 三 卡尔曼滤波

#### 3.1 问题阐述

一般的非线性时变系统可以用如下数学关系式表达，其状态方程和量测方程分别为

$$\begin{cases} x_{k+1} = f(x_k, u_k, w_k) \\ z_{k+1} = h(x_{k+1}, v_{k+1}) \end{cases}$$

**Given:**

- f, h, 噪声特性，初始状态
- 控制量  $u_0, u_1, \dots, u_{k-1}$
- 观测值  $z_1, \dots, z_{k-1}, z_k$

**Obtain:**

k 时刻  $x_k$  的最优估计

任何一个滤波器的目标都是根据现有被噪声污染的传感器数据求解状态量的最优估计。从贝叶斯估计来看，滤波器会根据量测数据和系统噪声特性，递推传递并更新状态量的条件概率密度函数。例如滤波器随时间不断估计和更新状态量的条件概率密度函数

$$p(x_k | z_1, \dots, z_k, u_0, \dots, u_{k-1})$$

概率密度分布函数给出了状态估计的可信度，概率越大，说明状态估计量的不确定性越小。

鉴于此，为了表述方便，将历史观测量和控制量表示为



$$U_0^{k-1} = \{u_0, \dots, u_{k-1}\}$$

$$Z_1^k = \{z_1, \dots, z_k\}$$

系统的滤波过程可以表述为如下：

- **Given**  $x_0$ 
  - Nature apply  $\omega_0$ ,
  - We apply  $u_0$ ,
  - The system moves to state  $x_1$ ,
  - We make a measurement  $z_1$ .
- **Question:** which is the best estimate of  $x_1$  ?
 

**Answer:** is obtained from  $p(x_1 | Z_1^1, U_0^0)$

  - Nature apply  $\omega_1$ ,
  - We apply  $u_1$ ,
  - The system moves to state  $x_2$ ,
  - We make a measurement  $z_2$ .
- **Question:** which is the best estimate of  $x_2$  ?
 

**Answer:** is obtained from  $p(x_2 | Z_1^2, U_0^1)$

  - ...
  - ...
  - ...
  - ...
- **Question:** which is the best estimate of  $x_{k-1}$  ?
 

**Answer:** is obtained from  $p(x_{k-1} | Z_1^{k-1}, U_0^{k-2})$

  - Nature apply  $\omega_{k-1}$ ,
  - We apply  $u_{k-1}$ ,
  - The system moves to state  $x_k$ ,
  - We make a measurement  $z_k$ .
- **Question:** which is the best estimate of  $x_k$  ?
 

**Answer:** is obtained from  $p(x_k | Z_1^k, U_0^{k-1})$

  - ...
  - ...
  - ...
  - ...

为了获取最优估计，滤波器不断融合新的量测和控制量并以此求出当前状态量的条件概率密度函数。从时间递推的角度，因量测包含随机噪声，滤波器起到了传递噪声的作用，将量测和控制量的噪声传入到状态量上，从系统模型上得出状态量的概率特性。最后，概率特性已知，利用不同的优化策略便可获取最优估计值。

$$\begin{array}{ll}
p(x_0) & \\
p(x_1|Y_1^1, U_0^0) & \longrightarrow \hat{x}_1 \\
p(x_2|Y_1^2, U_0^1) & \longrightarrow \hat{x}_2 \\
\vdots & \longrightarrow \vdots \\
p(x_{k-1}|Y_1^{k-1}, U_0^{k-2}) & \longrightarrow \hat{x}_{k-1} \\
p(x_k|Y_1^k, U_0^{k-1}) & \longrightarrow \hat{x}_k
\end{array}$$

### 3.2 优化模型

由前面定义知，滤波问题即为最优化问题，通过建立估计量的代价函数从而求取代价函数最小化对应的状态值。因此卡尔曼滤波中状态量的估计问题，本质上可以描述为：若给出一系列观测信息  $z_1, \dots, z_k$ ，根据优化算法(或优化目标)，估计出最优的参数向量  $\hat{\theta}_k$ ，例如

$$\hat{\theta}_k = g(z_1, \dots, z_k)$$

建立均方误差代价函数  $J(\theta_k) = E[\tilde{\theta}_k^T \tilde{\theta}_k]$ ，其中  $\tilde{\theta}_k = \theta_k - \hat{\theta}_k$ ，则状态量的最优估计即使得均方误差最小时的估计值  $\hat{\theta}_k$ ，即

$$\hat{\theta}_k = \arg \min E[(\theta_k - \hat{\theta}_k)^T (\theta_k - \hat{\theta}_k)]$$

对均方误差代价函数进行展开，可得期望  $E$  的条件概率积分表达式：

$$E[\tilde{\theta}(k)^T \tilde{\theta}(k)] = \int_{-\infty}^{\infty} E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_1^k] p(Y_1^k) dY_1^k$$

因观测量  $Y$  为常数，于是对原均方误差代价函数求最小值的过程可以等价于对其条件均方误差代价函数求最小值。因此滤波问题可以等价为

$$\hat{\theta}(k) = \operatorname{argmin} E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_1^k]$$

### 3.3 线性系统

#### 3.3.1 线性时变系统

在前面 3.1 部分描述了一般的非线性时变系统的滤波问题，现在仅仅考虑某些指定的线性时变问题，其描述如下：

$$\begin{aligned}
x_k &= A_{k-1} x_{k-1} + B_{k-1} u_{k-1} + w_{k-1} \\
z_k &= C_k x_k + v_k
\end{aligned}$$

其中  $\{w_k\}$  称为过程噪声， $\{v_k\}$  为量测噪声，均为系统的零均值的高斯白噪声，其均值

$$E[w_k] = E[v_k] = 0$$

和联合协方差矩阵为

$$E \left[ \begin{pmatrix} w_k \\ v_k \end{pmatrix} (w_k^T v_k^T) \right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix}$$

### 3.3.2 滤波过程

令

$$p(x_k | Y_1^k, U_0^{k-1}) \sim N(\hat{x}(k|k), P(k|k))$$

表示一个高斯型条件概率密度函数，则状态估计  $\hat{x}(k|k)$  为该条件概率函数的条件均值，协方差均值  $P(k|k)$  量化状态估计的不确定性

$$\begin{aligned} \hat{x}(k|k) &= E[x(k) | Y_1^k, U_0^{k-1}] \\ P(k|k) &= E[(x(k) - \hat{x}(k|k))(x(k) - \hat{x}(k|k))^T | Y_1^k, U_0^{k-1}]. \end{aligned}$$

比起前面所述滤波器传递整个条件概率密度函数，卡尔曼滤波仅仅传递上一个时刻和当前时刻的条件概率密度。卡尔曼滤波是一个递推过程，为了估计  $x(k+1|k+1)$ ，仅仅要求上一次的状态估计  $x(k|k)$  和当前时刻的观测量  $y(k+1)$ 。

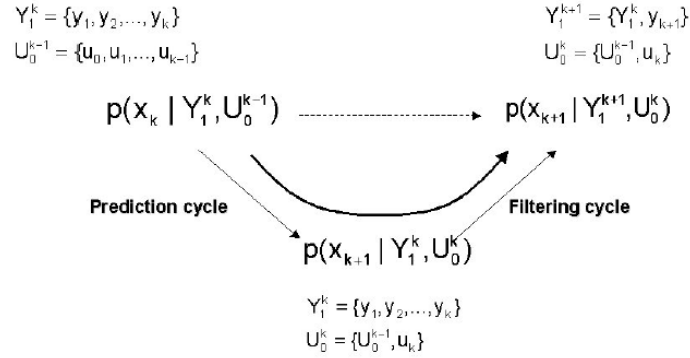


图3 卡尔曼滤波预测过程和滤波过程

卡尔曼滤波过程可以分为两个过程：预测部分和滤波部分。首先，根据状态方程对  $k+1$  时刻进行状态预测，预测使得状态量的不确定度增强(如协方差  $P$  增大)，然后基于预测量和观测量进行滤波修正，得出最优状态估计值。如此依次进行递推，我们可以将卡尔曼滤波过程归纳成如下流程：

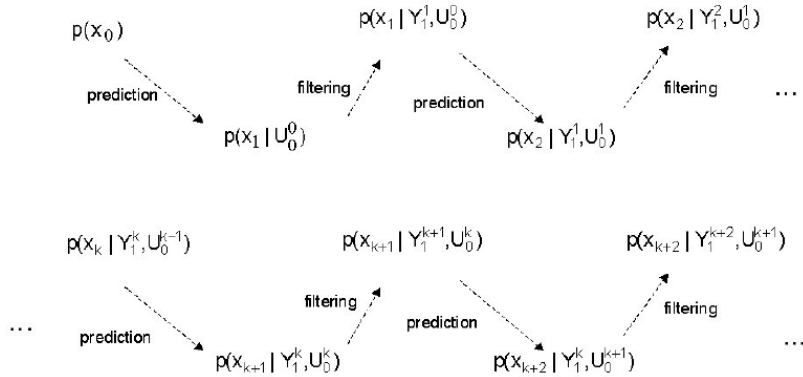


图 4 卡尔曼滤波迭代过程

总之，卡尔曼滤波是一个动态递推应用过程，滤波器循环地进行预测和校正两个环节，如上图 4 所示。

令

$$p(x_k | Y_1^k, U_0^{k-1}) \sim N(\hat{x}(k|k), P(k|k))$$

$$p(x_{k+1} | Y_1^k, U_0^k) \sim N(\hat{x}(k+1|k), P(k+1|k))$$

其中  $\hat{x}(k|k)$  和  $\hat{x}(k+1|k)$  满足

$$\hat{x}(k|k) = E[x(k) | Y_1^k, U_0^{k-1}]$$

$$\hat{x}(k+1|k) = E[x(k+1) | Y_1^k, U_0^k]$$

$$P(k|k) = E[(x_k - \hat{x}(k|k))(x_k - \hat{x}(k|k))^T | Y_1^k, U_0^{k-1}]$$

$$P(k+1|k) = E[(x_{k+1} - \hat{x}(k+1|k))(x_{k+1} - \hat{x}(k+1|k))^T | Y_1^k, U_0^k]$$

### 3.3.3 卡尔曼滤波推导

下面我们将对卡尔曼滤波的具体数学表达进行推导，为方便描述，我们假设  $k$  时刻  $p(x_k | Y_1^k, U_0^{k-1})$  已知，即  $\hat{x}(k|k), P(k|k)$  给出。

**STEP 1:**  $p(x_{k+1} | Y_1^k, U_0^k)$  估计

[状态预测](#)

高斯型随机变量可以唯一由均值和方差表示，对状态方程等式两边分别求均值，可以得出如下表达式

$$E[x_{k+1} | Y_1^k, U_0^k] = A_k E[x_k | Y_1^k, U_0^{k-1}] + B_k E[u_k | Y_1^k, U_0^{k-1}] + G E[w_k | Y_1^k, U_0^{k-1}]$$

考虑  $w_k$  和  $Y_1^k$  为相互独立的随机变量且  $w_k$  为零均值的白噪声，则

$$\hat{x}(k+1|k) = A_k \hat{x}(k|k) + B_k u_k$$

定义该过程的**预测误差**为

$$\tilde{x}(k+1|k) = x(k+1) - \hat{x}(k+1|k)$$

将  $x(k+1)$  和  $\hat{x}(k+1|k)$  表达式带入上式得

$$\tilde{x}(k+1|k) = A_k x_k + B_k u_k + G_k w_k - A_k \hat{x}(k|k) - B_k u_k = A_k \tilde{x}(k|k) + G_k w_k$$

同理，可以将滤波过程的误差定义为

$$\tilde{x}(k|k) = x(k) - \hat{x}(k|k)$$

因  $\tilde{x}(k|k)$  和  $w_k$  为相互独立的随机变量，可以得出

$$E[\tilde{x}(k+1|k)\tilde{x}(k+1|k)^T | Y_1^k, U_0^k] = A_k E[\tilde{x}(k|k)\tilde{x}(k|k)^T | Y_1^k, U_0^{k-1}] A_k^T + G_k Q_k G_k^T$$

即

$$P(k+1|k) = A_k P(k|k) A_k^T + G_k Q_k G_k^T$$

**STEP 2:**  $p(y_{k+1} | Y_1^k, U_0^k)$  估计

[量测预测](#)

首先，从量测方程可以得出

$$p(y_{k+1} | Y_1^k, U_0^k) = p(C_{k+1} x_{k+1} + v_{k+1} | Y_1^k, U_0^k)$$

这是一个高斯型随机变量， $v_{k+1}$  和  $Y_1^k$  相互独立，则

$$\hat{y}(k+1|k) = C_{k+1} \hat{x}(k+1|k)$$

定义**观测量预测误差**为

$$\tilde{y}(k+1|k) = y(k+1) - \hat{y}(k+1|k)$$

将  $y(k+1)$  和  $\hat{y}(k+1|k)$  表达式带入上式得

$$\tilde{y}(k+1|k) = C_{k+1} \tilde{x}(k+1|k) + v_{k+1}$$

类似地，其协方差矩阵可以表示为

$$P_{\tilde{y}}(k+1|k) = C_{k+1} P(k+1|k) C_{k+1}^T + R_{k+1}$$

在  $\tilde{y}(k+1|k)^T$  左边乘以  $x_{k+1}$ ，可得

$$x_{k+1} \tilde{y}(k+1|k)^T = x_{k+1} \tilde{x}(k+1|k)^T C_{k+1}^T + x_{k+1} v_{k+1}^T$$

于是

$$E[x_{k+1} \tilde{y}(k+1|k)^T] = P(k+1|k) C_{k+1}^T \quad (3.3.3)$$

**STEP 3:**  $p(x_{k+1} | Y_1^{k+1}, U_0^k)$  估计

[滤波](#)

为了估计  $x_{k+1}$  的条件概率密度函数，可以做如下等效，

$$Y_1^{k+1} \in \{Y_1^k, \tilde{y}(k+1|k)\}$$

其包含的概率信息等价

$$E[x_{k+1} | Y_1^{k+1}, U_0^k] = E[x_{k+1} | Y_1^k, \tilde{y}(k+1|k), U_0^k]$$

事实上,  $Y_1^k \in \tilde{y}(k+1|k)$  为相互独立的变量, 由 2.4 节中 **Result 3.3.1** 可知

$$\hat{x}(k+1|k+1) = E[x_{k+1} | Y_1^k] + E[x_{k+1} \tilde{y}(k+1|k)] P_{\tilde{y}(k+1|k)}^{-1} \tilde{y}(k+1|k)$$

将式(3.3.3)带入上式, 可得滤波最优估计为

$$\begin{aligned} \hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + K(k+1)(y(k+1) - C_{k+1}\hat{x}(k+1|k)) \\ &+ P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}]^{-1}(y(k+1) - C_{k+1}\hat{x}(k+1|k)) \end{aligned}$$

定义卡尔曼滤波增益

$$K(k+1) = P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}]^{-1}$$

于是

$$\hat{x}(k+1|k+1) = \hat{x}(k+1|k) + K(k+1)(y(k+1) - C_{k+1}\hat{x}(k+1|k))$$

定义滤波误差为

$$\begin{aligned} \tilde{x}(k+1|k+1) &= x(k+1) - \hat{x}(k+1|k+1) \\ \tilde{x}(k+1|k+1) &= \tilde{x}(k+1|k) - \\ &P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}]^{-1}(C_{k+1}\hat{x}(k+1|k) - y(k+1)) \\ P(k+1|k+1) &= P(k+1|k) - \\ &P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}]^{-1}C_{k+1}P(k+1|k) \end{aligned}$$

$$\text{滤波输出} = \text{预测状态估计} + \textit{Gain} * \textit{Error}$$

**Summary:**

预测

$$\begin{aligned} \hat{x}(k+1|k) &= A_k \hat{x}(k|k) + B_k u_k \\ P(k+1|k) &= A_k P(k|k) A_k^T + G_k Q_k G_k^T \end{aligned}$$

滤波

$$\begin{aligned} \hat{x}(k+1|k+1) &= \hat{x}(k+1|k) + K(k+1)(y(k+1) - C_{k+1}\hat{x}(k+1|k)) \\ K(k+1) &= P(k+1|k)C_{k+1}^T[C_{k+1}P(k+1|k)C_{k+1}^T + R_{k+1}]^{-1} \\ P(k+1|k+1) &= P(k+1|k)[I - K(k+1)C_{k+1}] \end{aligned}$$

初始状态

$$\begin{aligned}\hat{x}(0|-1) &= \bar{x}_0 \\ P(0|-1) &= \Sigma_0\end{aligned}$$

### 3.4 初值问题

卡尔曼滤波在开始状态需要进行初始化，即需要选择一个起始估计  $\hat{x}(0|-1)$  和协方差矩阵  $P(0|-1)$ 。一般地，初始值对滤波器状态估计的影响会随着时间推移而消散( $x_k = A^k x_0$ )，他们并不会影响到滤波器收敛的稳定性。通常在线性滤波器中，不管初始状态值和其状态方差是否变换，滤波器总会收敛到相同的稳定状态上。当然一个接近真实值的初始值可以提高滤波器的收敛速度。

但是在扩展卡尔曼滤波过程，因为非线性处理需要对状态量进行一阶近似，偏离真实值将导致近似误差过大，有时会导致滤波器无法收敛。故初始值的选择应尽量靠近真实值  $\hat{x}(0|-1) = \bar{x}_0$  和  $P(0|-1) = \Sigma_0$ 。一般地，因无法获取初始值，往往需要依靠其他算法估计出近似值，然后再带入滤波器算法中，该过程在导航系统中称之为初始对齐。

### 3.5 新息残差

在  $k-1$  时刻，滤波器对  $k$  时刻的测量值  $\hat{z}(k|k-1)$  进行预测，即

$$\hat{z}(k|k-1) = C_k \hat{x}(k|k-1)$$

则观测量  $z_k$  与预测值  $\hat{z}(k|k-1)$  之差叫做新息  $v_k$  (Innovation)，有时也被叫做残差(residual)。

$$v_k = z_k - C_k \hat{x}(k|k-1)$$

新息是一种十分重要的测量方式，它取滤波器估计值与观测序列之差。通常，**比起状态估计值，“真实”的状态值很难知晓**，新息成为评估估计器性能表现是否好的唯一测量手段。新息可以被用来调谐估计器，也可以帮助开发者建立正确的观测值和过程噪声源系统模型。同时，新息还可以用来验证外部观测量是否为有效值(野值检验)，以此判断是否将当前观测量加入滤波器进行更新迭代。总之，新息是一种非常重要的方法，在滤波器设计过程中起到至关重要的作用，是我们进行实际工程项目中不可或缺的一个数学概念。

新息最重要的性质就是，它是一个正交、时间不相关、白噪声序列。

$$E[v_k | Z_1^k] = 0, \quad E[v(i)v^T(i)] = S(i)\delta_{ij}$$

新息的方差可以表示为

$$\begin{aligned}
S(k) &= E[v(k)v^T(k)] \\
&= E[(z_k - C_k \hat{x}(k|k-1))(z_k - C_k \hat{x}(k|k-1))^T] \\
&= E[(C_k(x_k - \hat{x}(k|k-1)) + v_k)(C_k(x_k - \hat{x}(k|k-1)) + v_k)^T] \\
&= R_k + C_k P(k|k-1) C_k^T
\end{aligned}$$

因为新息是一个高斯型白噪声，则

$$q(i) = v^T(i) S(i) v(i)$$

随机变量  $q(i)$  被称为归一化新息方差，服从  $\chi^2$  分布(卡方分布)，其自由度  $m = \dim(z(k))$ ，均值为  $m$ ，方差为  $2m$ 。卡方检验提供了一个置信边界帮助验证观测量和滤波器性能，滤波器的一致性检验就是依据这个性质。若存在**假设 H: 滤波器对系统适应性很好**，则为接受该假设 H 必须要对新息序列进行测试。

$$P\{H \text{ accepted} | q_i \in (2 - \sigma \text{ gate})\} = 0.95$$

根据  $\chi^2$  分布的性质，我们可以找到新息序列对应的  $\pm\sigma$  和  $\pm 2\sigma$  边界，后者界限给出了比例大约为 95% 的新息需要落入  $2-\sigma$  门限范围内。

如果新息存在非随机周期性变化或者非随机性漂移，新息序列极有可能是时间相关的(非白噪声)。为了测试这一特点，需要计算新息的自相关函数。正如前面描述的一样，当对时间相关做一致性检验时，若自相关性超出它的变化边界并远离零值时，则噪声为白的假设不成立。白噪声验证失败，一般会存在两个原因：**过程噪声有着不可忽视的高阶量没有被模型所考虑或者观测噪声与过程噪声值大小之比(噪声参数不对)过大**。同时，我们也可以依据自相关函数的形状提取一些有用的信息，分析出新息为何是时间相关的。



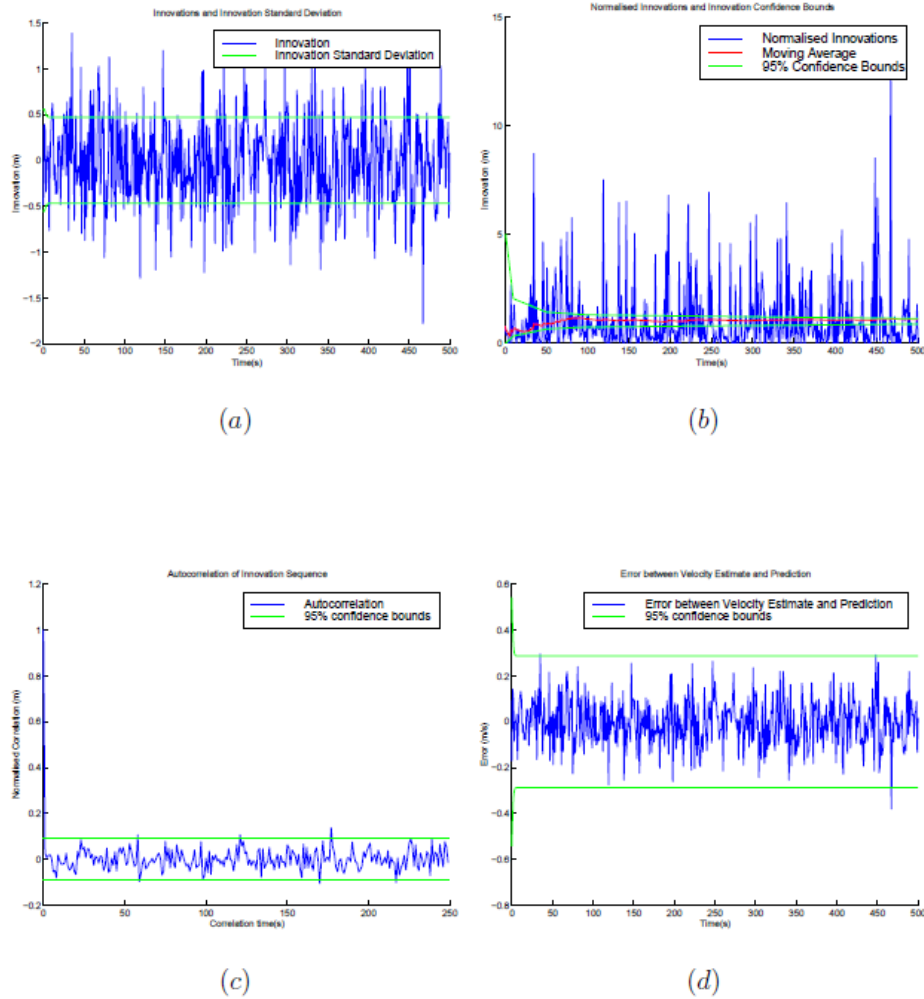


Figure 15: Exactly Matched Constant-Velocity Filter with  $\hat{\sigma}_q = \sigma_q = 0.01$  and  $\hat{\sigma}_r = \sigma_r = 0.1$ . (a) Innovation Sequence and Estimated Innovation Standard Deviation; (b) Normalized Innovations Squared, Moving Average and 95% Confidence Bounds; (c) Normalized, Biased Autocorrelation of the Innovation Sequence and 95% Confidence Bounds; (d) Error between Velocity Estimate and Velocity Prediction, With 95% Confidence Bounds.

参考文献[2]列举了几个实验，分析当过程噪声过大、过小，观测噪声过大、过小时，与之对应的新息序列自相关性、卡方分布如何变化，希望读者能够好好参阅，这对于检验**滤波器模型是否有效**(假设 H)是一个很好的手段，这里不再赘述，仅仅给出正常情况下新息序列图形，如上图所示。

## 四 工程实践

卡尔曼滤波是一个数学上十分完美的滤波模型。在工程实践应用中，如果模型建立不够准确，将会遭遇滤波器无法收敛，算法发散等灾难。本节给出设计卡尔曼滤波器的一些工程实现建议：

**Rule 1: 模型准确则滤波器性能佳。**在卡尔曼滤波设计之前，务必对系统过程(状态)、传感器(量测)有清晰的了解，可以考虑采集系统数据，并对其进行概率特性分析比如均值、方差、概率密度曲线、噪声功率谱等等。一般地，工程上不存在完全呈现高斯白噪声特性的随机噪声，常常需要将某些噪声近似成高斯白噪声处理。所以，对滤波器一致性(或者相容性)进行概率检验与说明模型的合理性，成为至关重要的步骤之一。

**Rule 2: 过程噪声  $R$  和量测噪声  $Q$ 。**在卡尔曼滤波在计算过程中涉及到噪声参数  $R$  和  $Q$ 。实际上大多噪声无法直接测量，鉴于此，工程上仅仅考虑噪声方差幅值的数量级，选择过程噪声  $Q$  和量测噪声  $R$  与其真实值在一个数量级上，已基本满足滤波的要求。幸运地是，有些应用可以在静态环境测试噪声方差，但实际应用中动态环境会导致噪声增大，这种情况下静止测量的参数无法适用与动态环境，则可以考虑以静态参数为基数放大 10~100 倍，带入滤波器中计算。

**Rule 3: 滤波器收敛性。**回顾卡尔曼滤波推导过程， $\hat{x}(k+1|k)$  的计算是基于前一时刻  $k$  估计值  $\hat{x}(k|k)$  在真实值附近这一假设  $H_1$ ，即  $p(x_k | Y_1^k, U_0^{k-1}) \sim N(\hat{x}(k|k), P(k|k))$ 。若系统提供观测信息不足(假设  $H_1$  不再成立)时，就会出现估计值  $\hat{x}(k|k)$  不在真实值附近的情况，滤波器可能发散。如在定位跟踪领域，某些应用无法根据当前观测获取状态量的足够信息，必须根据历史轨迹内推进行状态估计，这时卡尔曼滤波可能无法适用或者需要对其状态增广。故卡尔曼滤波，比起粒子滤波、贝叶斯滤波等，在观测量冗余时，具有计算量小、模型简单易于实现等优点，优势更为明显。

#### **Rule 4: 工程小技巧。**

第一，对积分的协方差矩阵  $P$  进行限幅，这是为了保证给出的上一时刻的估计值置信度不至于太差，导致滤波器发散。

第二，根据新息序列曲线，适当调节方差矩阵  $R$  和  $Q$ ，以使滤波器性能最佳。也可以考虑在线估计噪声的方差，实现自适应滤波。

第三，当传感器数据出现某些高频量(如野值，突变)时，应在卡尔曼滤波之前，先对其进行频域滤波，消除抖动对滤波器的影响。

第四，对滤波器是否发散进行周期性检测，一旦发现滤波器有发散的趋势，应对滤波器进行特殊处理，将估计值尽可能纠正到真实值附近，帮助其收敛。

**Rule 5: 理论终究是理论。**在工程应用中，应灵活运用卡尔曼滤波器，更多地要去实践，在调试过程中对卡尔曼滤波进行更深层次的理解。博主在不断试错中，发现卡尔曼滤波对模型要求十分高，大量的时间都是在如何建立合理的近似模型，使得滤波器性能达到最优。所以掌握滤波器的原理对工程应用十分重要。

下面给出一个卡尔曼滤波 Demo，旨在解决室内定位中 UWB 定位和 MEMS-IMU 进行组合导航的算法。本算法利用 MATLAB 实现，数据均为仿真产生，希望能够帮助读者更深理解卡尔曼滤波器。

网址：[https://github.com/DahaiHigh/demo\\_for\\_kalmanFilter.git](https://github.com/DahaiHigh/demo_for_kalmanFilter.git)

## 参考文献

- [1] Ribeiro M I. Kalman and extended kalman filters: Concept, derivation and properties[J]. Institute for Systems and Robotics, 2004, 43.
- [2] Durrant-Whyte H. Introduction to estimation and the Kalman filter[J]. Australian Centre for Field Robotics, 2001, 28(3): 65-94.
- [3] Zhao Y, Horemuz M, Sjöberg L E. Stochastic modelling and analysis of IMU sensor errors[J]. Archiwum Fotogrametrii, Kartografii i Teledetekcji, 2011, 22.

## 声明

1、本文为作者原创文章，未经作者允许不得转载或用于商业目的，谢谢合作。

2、如有侵犯您的知识产权和版权问题，请通知本人，本人会即时做出处理删除文章。

3、如果本文章在知识点上有错误，欢迎指出错误所在，欢迎多多交流。谢谢！

4、本百度文库的文章、图片以及转载的文章引起的任何意外、疏忽、合约毁坏、诽谤、版权或知识产权侵犯及其所造成的损失（包括因下载而感染电脑病毒），本人概不负责，亦不承担任何法律责任。

5、因本百度文库的文章、图片以及转载的文章所产生的任何法律（包括宪法，民法，刑法，书法，公检法，基本法，劳动法，婚姻法，输入法，没办法，国际法，今日说法，吸星大法，与台湾关系法及文中涉及或可能涉及、以及未涉及之法，各地治安管理条例）纠纷或责任本人概不负责。

6、一切网民在进入本百度文库的各个页面时，默认表示遵守本条款并完全同意。

7、此声明最终解释权归本人所有。