

# 滤波

Liam

2023 年 9 月 10 日

## 1 一般滤波器

### 1.1 限幅滤波法

- 优点：克服偶然因素引起的脉冲干扰
- 缺点：无法抑制周期性干扰，平滑度差
- 方法：

如果本次值与上一次值的绝对值差  $\leq A$ ，则本次值有效。

反之则无效

```
1 int filter(int newvalue)
2 {
3     if(((NewValue - Value) > FILTER_A) || ((Value -
4         NewValue) > FILTER_A))
5         return Value;
6     else
7         return NewValue;
8 }
```

## 1.2 中位值滤波法

- 连续采样奇数  $N$  次，将样本按照大小排序，取中间值为有效值。
- 优点：能够克服偶然因素波动干扰，对变化缓慢的观察量有好的滤波效果
- 缺点：不适用快速变化的量。

```
1 int filter(buffertype *to_filter_buf)
2 {
3     // 采样值从小到大排列（冒泡法）
4     for(j = 0; j < FILTER_N - 1; j++) {
5         for(i = 0; i < FILTER_N - 1 - j; i++) {
6             if(filter_buf[i] > filter_buf[i + 1]) {
7                 filter_temp = filter_buf[i];
8                 filter_buf[i] = filter_buf[i + 1];
9                 filter_buf[i + 1] = filter_temp;
10            }
11        }
12    }
13    return filter_buf[(FILTER_N - 1) / 2];
14 }
```

## 1.3 算术平均滤波法

- 优点：适用于具有随机干扰信号的滤波，适用于信号一般在某一值上下波动。
- 缺点：测量数据过慢或者过快的实时控制不好。

```
1  int filter(sometype data)
2  {
3      static sometype filter_sum = 0;
4      for(int i=0; i<FILER_N,i++)
5      {
6          filter_sum += data;
7      }
8      return (int)(filter_sum / FILTER_N);
9  }
```

## 1.4 消抖滤波法

- 优点：对变化缓慢的被测量具有较好的滤波效果
- 作用：避免在临界值附近反复跳动 or 抖动。
- 缺点：对于快速变化的量不适用
- 方法：
  1. 设置一个滤波计数器，将每次采样的值和当前有效值进行对比
  2. 如果采样值 = 当前有效值，则计数器清零；如果采样值 <> 当前有效值，则计数器 +1，并判断计数器是否 >= 上限 N（溢出）；
  3. 如果计数器溢出，则将本次值替换当前有效值，并清计数器。

```
1
2  int Filter_Value;
3  int Value;
4
5  #define FILTER_N 12
6  int i = 0;
```

```
7 int Filter() {  
8     int new_value;  
9     new_value = Get_AD();  
10    if(Value != new_value) {  
11        i++;  
12        if(i > FILTER_N) {  
13            i = 0;  
14            Value = new_value;  
15        }  
16    }  
17    else  
18        i = 0;  
19    return Value;  
20 }
```

## 2 贝叶斯滤波

### 2.1 概率往事

#### 2.1.1 随机变量

定义在样本空间  $\Omega$  上的实值函数：

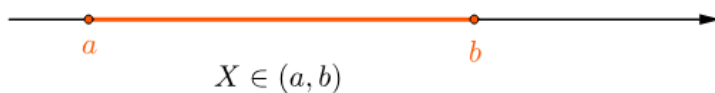
$$X = X(\omega) \quad (1)$$

被称为**随机变量** (random variable)。 $\omega$  表示样本空间  $\Omega$  里的随机事件，可能与数量有关（如传感器随机测得的前方障碍物的距离值，随机掷骰子得到的点数），也可能与数量无关（如随机掷硬币的结果可能为正面或反面）。随机变量表示随机试验各种结果的实值单值函数。随机事件不论与数量是否直接有关，都可以数量化，即都能用数量化的方式表达。

**离散型随机变量：**如果随机变量的函数值是实数轴上孤立的点



**连续型随机变量：**随机变量的函数值是实数轴上某个区间上所有的值



**累积分布函数：**

设  $X$  是一个随机变量，如果  $\exists x$ ，则：

$$F(x) = P(X \leq x) \quad (2)$$

为随机变量的**累计分布函数**（Cumulative Distribution Function, CDF），又称概率分布函数，简称分布函数。

若  $X$  为离散型随机变量，存在函数  $f_X(x)$ ，使其分布函数满足：

$$F(x) = P(X \leq x) = \sum_{u \leq x} f_X(u) \quad (3)$$

若  $X$  为连续型随机变量，存在函数  $f_X(x)$ ，使其分布函数满足：

$$F(x) = P(X \leq x) = \int_{-\infty}^x f_X(u) \quad (4)$$

并且函数满足：

- 单调性
- 有界性

- 右连续性

**概率质量函数**假设  $X$  为离散型随机变量, 其全部可能取值为  $x_1, x_2, \dots$ , 其概率分布列可表述为:

则称:

$$f_X(x) = \begin{cases} p(x) = P\{X = x_i\}, i = 1, 2, 3, \\ 0 \end{cases} \quad (5)$$

概率质量函数表征了离散型随机变量在各特定取值上的概率, 满足:

- 非负性
- 规范性
- 可加性

离散型随机变量概率质量函数的叠加得到了离散型随机变量的累积分布函数。

### 概率密度函数

设  $x$  为连续型随机变量, 存在函数  $f_X(x)$ , 满足:

- 非负性
- 规范和可加性

$$\int_{-\infty}^{\infty} f_X(x) dx = 1$$

则称函数为连续型随机变量  $X$  的概率密度函数 (Probability Density Function, PDF), 连续型随机变量概率密度函数的积分得到了连续型随机变量的累积分布函数。

概率密度函数表征了连续型随机变量的输出值在某个确定的取值点附近的可能性。连续型随机变量的输出值在某个确定的取值点上的概率为 0, 但并不代表该事件不会发生。

### 2.1.2 联合概率

**二维离散型随机变量的联合概率质量函数** 若二维离散型随机变量  $(X, Y)$  所有可能的取值：

$$(x_i, y_j), i, j = 1, 2, 3, \dots \quad (6)$$

且存在函数  $f_{x,y}(x, y)$ :

$$f_{x,y}(x, y) = p(x, y) = P(X = x, Y = y), (x, y) \in \{x_i, y_j\}, i, j = 1, 2, 3, \dots \quad (7)$$

则函数满足以下性质：

- 非负性
- 规范和可加性

则称函数为  $(x, y)$  的**联合概率质量函数**，又称联合分布列，此定义可以推广到多为离散随机变量。

二维离散型随机变量联合概率质量函数的叠加得到了二维离散型随机变量的联合累积分布函数：

$$F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \sum_{u \leq x} \sum_{v \leq y} f_{X,Y}(u, v) \quad (8)$$

**二维连续型随机变量的联合概率密度函数** 若对于随机变量  $(X, Y)$ ,  $\exists f_{X,Y}(x, y)$  满足：满足：规范和可加性  $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx dy = 1$ ，非负性则称函数为联合概率密度函数（Joint Probability Density Function），此定义可以推广到多维连续型随机变量。

二维连续型随机变量联合概率密度函数的积分得到了二维连续型随机变量的联合累积分布函数

### 2.1.3 边缘概率

**二维离散型随机变量的边缘概率质量函数** 若对于二维离散型随机变量  $(X, Y)$  所有可能的取值为  $(x_i, y_j), i, j = 1, 2, 3, \dots$  且联合概率质量函数

$f_{X,Y}(x, y)$  为:

$$f_{x,y}(x, y) = p(x, y) = P(X = x, Y = y), (x, y) \in \{x_i, y_i\}, i, j = 1, 2, 3, \dots \quad (9)$$

对  $j$  求和所得的函数:

$$f_Y(y) = \sum_{i=1}^{\infty} f_{X,Y}(x_i, y) = \sum_{i=1}^{\infty} P(X = x_i, Y = y) = P(Y = y), y \in y_i, j = 1, 2, 3, \dots \quad (10)$$

称为  $Y$  的边缘概率质量函数 (Marginal Probability Mass Function), 又称边缘分布列。

二维离散型随机变量边缘概率质量函数的叠加得到了二维离散型随机变量的边缘累积分布函数 (Marginal Cumulative Distribution Function):

$$\begin{aligned} F_X(x) &= P(X \leq x) = \sum_{u \leq x} f_x(u) = \sum_{u \leq x} \sum_{j=1}^{\infty} P(X = u, Y = y_j) = \\ &\sum_{u < x} P(X = u) \\ F_Y(y) &= P(Y \leq y) = \sum_{v \leq y} f_y(v) = \sum_{v \leq y} \sum_{i=1}^{\infty} P(X = x_i, Y = v) = \\ &\sum_{u < x} P(Y = v) \end{aligned} \quad (11)$$

#### 2.1.4 条件概率

**二维离散型随机变量的条件概率质量函数** 若  $(x, y)$  是二维离散型随机变量, 对于固定的  $j$ , 若  $P(Y=y_j) > 0$ , 则称为:

$$f_{X|Y}(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{f_{x,y}(x, y)}{f_Y(y)}, (x, y) \in \{x_i, y_j\}, i, j = 1, 2, 3, \dots \quad (12)$$

为  $X = x$  条件下离散型随机变量  $X$  的条件概率质量函数。

简言之, 对于二维离散型随机变量: 二维离散型随机变量条件概率质量函数的叠加得到了二维离散型随机变量的条件累积分布函数 (Conditional Cumulative Distribution Function):

$$f_{X|Y}(x|y) = P(X \leq x|Y = y) = \sum_{u \leq x} f_{X|Y}(u|y) = \sum_{u \leq x} \frac{f_{x,y}(u, y)}{f_Y(y)} \quad (13)$$



$$\text{条件概率质量函数} = \frac{\text{联合概率质量函数}}{\text{边缘概率质量函数}}$$

### 二维连续型随机变量的条件概率密度函数

假设二维连续型随机变量  $(x, y)$  的联合概率密度函数为  $f_{x,y}(x, y)$ ，若对于固定的  $Y = y$ ，有边缘概率密度函数  $f_Y(y) > 0$ ，则称函数：

$$f_{X|Y}(x, y) = \frac{f_{x,y}(x, y)}{f_Y(y)} \quad (14)$$

为随机变量  $X$  在条件下的条件概率密度函数（Conditional Probability Density Function）。二维连续型随机变量条件概率密度函数的积分得到了二维

$$\text{条件概率密度函数} = \frac{\text{联合概率密度函数}}{\text{边缘概率密度函数}}$$

连续型随机变量的条件累积分布函数：

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \int_{-\infty}^x f_{X|Y}(u|y) du = \int_{-\infty}^x \frac{f_{X|Y}(u, y)}{f_Y(y)} du \quad (15)$$

## 2.1.5 全概率

### 二维离散型随机变量的全概率公式

二维连续型随机变量的全概率公式 已知条件概率质量函数：

$$f_{X|Y}(x|y) = \frac{f_{x,y}(x, y)}{f_Y(y)} f_Y(y) = \frac{f_{x,y}(x, y)}{f_X(x)} \quad (16)$$

容易得到二维连续型随机变量联合概率密度函数的等价形式：

$$f_{X,Y}(x, y) = f_{X|Y}(x|y)f_Y(y) = f_{Y|X}(y|x)f_X(x) \quad (17)$$

由边缘概率密度函数与联合概率密度函数间的数量关系，得到二维连续型随机变量的全概率公式：

$$f_X(x) = \int_{-\infty}^{\infty} f_{X|Y}(x|y)f_Y(y)dy \quad (18)$$

### 2.1.6 随机过程

**随机过程**是定义在  $\Omega \times T$  上的二元函数  $X(\omega, t)$ , 简单定义为一组随机变量的集合，即指定一参数集  $T$ （又称指标集，通常为时间集），其中：

- 对于固定的时间  $t$ ,  $X(w, t)$  为随机变量，简记为  $X(t)$  或  $X_t$
- 对于固定的  $w$ （即固定每一时刻对应的随机变量的取值）， $X(\omega, t)$  为时间  $t$  的一般函数，称为样本函数（Sample Function）或样本轨道（Sample Path），简记为  $x(t)$ 。随机过程也可定义为一组样本函数的集合。

注意：随机变量  $X(t)$  或  $X_t$  并不是时间  $t$  的函数，它只表示所有样本函数图片在  $t$  时刻的取值。

依据随机变量和参数集的连续离散情况，随机过程有以下四种类型：

- 连续型随机过程：随机变量连续，参数连续；
- 离散型随机过程：随机变量离散，参数连续；
- 连续型随机序列：随机变量连续，参数离散；
- 离散型随机序列：随机变量离散，参数离散。

不同于确定过程，随机过程中的随机变量彼此间不独立。应该理解，贝叶斯滤波处理的是一个随机过程，而且往往是一个连续型随机过程。

## 2.2 贝叶斯公式

**二维离散型随机变量的贝叶斯公式** 对于二维离散型随机变量  $(x, y)$ ，由其条件概率质量函数与全概率公式，容易得到其贝叶斯公式：

$$f_{X|Y}(x|y) = \frac{f_{X,y}(x, y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\sum_{i=1}^{\infty} f_{Y|X}(y|x_i)f_X(x_i)} \quad (19)$$

**二维连续型随机变量的贝叶斯公式**

$$f_{X|Y}(x|y) = \frac{f_{X,y}(x, y)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{\infty} f_{Y|X}(y|x_i)f_X(x_i)} \quad (20)$$

二维连续型随机变量的条件累积分布函数推导：

$$\begin{aligned} F_{X|Y}(x|y) &= P(X \leq x | Y = y) \\ &= \sum_{u=-\infty}^x P(X = u | Y = y) \Rightarrow \text{化连续为离散无穷小的累加} \\ &= \sum_{u=-\infty}^x \frac{P(Y = y | X = u)P(X = u)}{P(Y = y)} \Rightarrow \text{二维离散型随机变量的贝叶斯公式} \\ &= \lim_{\epsilon \rightarrow 0} \sum_{u=-\infty}^x \frac{P(y \leq Y \leq y + \epsilon | X = u)P(u \leq X \leq u + \epsilon)}{P(y \leq Y \leq y + \epsilon)} \Rightarrow \text{化无穷小为极限形式} \\ &= \lim_{\epsilon \rightarrow 0} \sum_{u=-\infty}^x \frac{[F_{Y|X}(y + \epsilon | u) - F_{Y|X}(y | u)][F_X(u + \epsilon) - F_X(u)]}{F_Y(y + \epsilon) - F_Y(y)} \Rightarrow \text{分布函数性质} \\ &= \lim_{\epsilon \rightarrow 0} \sum_{u=-\infty}^x \frac{[f_{Y|X}(\xi_1 | u) \cdot \epsilon][f_X(\xi_2) \cdot \epsilon]}{f_Y(\xi_3) \cdot \epsilon} \Rightarrow \text{拉格朗日中值定理, } \xi_1, \xi_3 \in (y, y + \epsilon), \xi_2 \in (u, u + \epsilon) \\ &= \lim_{\epsilon \rightarrow 0} \sum_{u=-\infty}^x \frac{f_{Y|X}(y | u)f_X(u)}{f_Y(y)} \cdot \epsilon \Rightarrow \epsilon \rightarrow 0 \text{ 时, } \xi_1 \rightarrow y, \xi_2 \rightarrow u, \xi_3 \rightarrow y \\ &= \int_{-\infty}^x \frac{f_{Y|X}(y | u)f_X(u)}{f_Y(y)} du \Rightarrow \text{积分定义} \\ &= \int_{-\infty}^x \frac{f_{Y|X}(y | x)f_X(x)}{f_Y(y)} dx \Rightarrow \text{替换自变量符号 } u \text{ 为 } x \end{aligned}$$

### 2.2.1 先验概率、似然概率与后验概率

在二维连续型随机变量的贝叶斯公式中，有如下定义：

- $f_X(x)$  被称为先验概率密度 (Prior Probability Density)，表示根据以往的经验和分析，在本次试验或采样前便可获得的随机变量  $X$  的概率密度；
- $f_{Y|X}(y|x)$  被称为似然概率密度 (Likelihood Probability Density)，表示在状态随机变量  $X$  取值为  $x$  的条件下，观测随机变量  $Y$  取值为  $y$  的概率密度，状态为因，观测为果，即由因推果；

- $f_{X|Y}(x|y)$  被称为后验概率密度 (Posterior Probability Density), 表示在观测随机变量  $Y$  取值为  $y$  的条件下, 状态随机变量  $X$  取值为  $x$  的概率密度, 状态为因, 观测为果, 即由果推因。
- 当  $y$  为定值的时候,  $\mu = \left[ \int_{-\infty}^{+\infty} f_{Y|X}(y|x)f(x)dx \right]^{-1}$ , 为常数, 常被称为贝叶斯公式的归一化常数。

### 2.2.2 似然概率

上文中提到, 似然概率密度函数  $f_{Y|X}(y|X)$  表示在状态随机变量  $X$  取值为  $X$  的条件下, 观测随机变量  $Y$  取值为  $y$  的概率密度。似然概率密度函数表征了传感器检测精度, 对于给定的状态条件  $X = x$ , 观测结果  $Y = y$  的概率分布通常有三种模型:

#### (1) 等可能型

观测值在状态量真值附近呈均匀分布, 此时的似然概率密度函数为常数。

#### (2) 阶梯型

观测值在状态量真值附近呈阶梯分布, 此时的似然概率密度函数为分段常数。

#### (3) 正态分布型

观测值在状态量真值附近呈高斯分布, 此时的似然概率密度函数为高斯函数:  $f_{Y|X}(y|x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-x)^2}{2\sigma^2}}$

若假定似然概率密度函数为高斯函数, 此时, 似然概率密度函数的均值  $x$  代表状态量真值,  $\sigma$  代表传感器检测精度范围。若同时假定先验概率密度函数为高斯函数, 即

$$f_X(x) \sim N(\mu_1, \sigma_1^2), f_{Y|X}(y|x) \sim N(\mu_2, \sigma_2^2) \quad (21)$$

后验概率密度函数方差既小于先验概率密度函数方差, 也小于似然概率密度函数方差, 系统不确定度降低。

### 2.2.3 贝叶斯滤波推导

#### 问题建模 (1) 问题描述

对于某状态量随机变量  $X$ ，从初始时刻 0 开始，对其进行观测，得到 0  $k$  时刻的观测值：

$$y_0, y_1, y_2, \dots, y_k \quad (22)$$

求解  $k$  时刻状态量随机变量  $X_k$  的最优估计  $\hat{x}_k$ 。

#### (2) 求解思路

以贝叶斯公式为求解方向，将问题转化为求解状态量随机变量  $X_k$  后验概率密度函数的期望：

$$\hat{x}_k = E[f_{X_k}^+(x)] \quad (23)$$

进而需要求解状态量随机变量  $X_k$  的先验概率密度函数与似然概率密度函数。我们认为， $k$  时刻的状态量随机变量  $X_k$  与且仅与上一时刻的状态量随机变量  $X_{k-1}$  有关， $k$  时刻的观测量随机变量  $Y_k$  与且仅与  $k$  时刻的状态量随机变量  $X_k$  有关，其中的数量关系我们分别称之为状态方程与观测方程：

$$\begin{cases} X_k = f(X_{k-1}) + Q_k & \Rightarrow \\ Y_k = h(X_k) + R_k & \Rightarrow \end{cases} \quad (24)$$

$f(x)$  被称为状态转移函数， $h(x)$  被称为观测函数。

对于 0 时刻的初始状态量随机变量  $X_0$ ，认为观测值  $y_0$  即为其真值，其后验概率密度函数即为其先验概率密度函数。我们可以根据经验知识（建模精度和传感器精度）写出 0 时刻的初始状态量随机变量  $X_0$  的后验概率密度函数  $f_{X_0}^+(x)$ 、 $k$  时刻过程噪声随机变量  $Q_k$  的概率密度函数  $f_{Q_k}(x)$  和  $k$  时刻观测噪声随机变量  $R_k$  的概率密度函数  $f_{R_k}(x)$ 。

#### 符号定义

1. 各时刻的状态量随机变量  $X_0, X_1, X_2, \dots, X_k$
2. 各时刻的观测量随机变量  $Y_0, Y_1, Y_2, \dots, Y_k$

3. 各时刻的观测值  $y_0, y_1, y_2, \dots, y_k$
4. 各时刻的过程噪声随机变量  $Q_1, Q_2, \dots, Q_k$
5. 各时刻的观测噪声随机变量  $R_1, R_2, \dots, R_k$
6. 各时刻的过程噪声随机变量概率密度函数  $f_{Q_1}(x), f_{Q_2}(x), \dots, f_{Q_k}(x)$
7. 各时刻的观测噪声随机变量概率密度函数  $f_{R_1}(x), f_{R_2}(x), \dots, f_{R_k}(x)$
8. 各时刻的状态量随机变量先验概率密度函数  $f_{X_0}^-(x), f_{X_1}^-(x), f_{X_2}^-(x), \dots, f_{X_k}^-(x)$
9. 各时刻的状态量随机变量后验概率密度函数  $f_{X_0}^+(x), f_{X_1}^+(x), f_{X_2}^+(x), \dots, f_{X_k}^+(x)$
10. 各时刻状态量随机变量与观测量随机变量的似然概率密度函数

$$f_{Y_1|X_1}(y_1 | x), f_{Y_2|X_2}(y_2 | x), \dots, f_{Y_k|X_k}(y_k | x)$$

### 重要假设

$X_0$  分别与  $Q_1, Q_2, \dots, Q_k$  相互独立;  $X_0$  分别与  $R_1, R_2, \dots, R_k$  相互独立;  $X_{k-1}$  与  $Q_k$  相互独立;  $X_k$  与  $R_k$  相互独立。(5) 重要定理

条件概率里的条件可以作逻辑推导。例如:

$$P(X = 1 | Y = 2, Z = 3) = P(X+Y = 3 | Y = 2, Z = 3) = P(X+Y = 3 | Y = 2, Z-Y = 1) \quad (25)$$

预测步推导已知 0 时刻状态量随机变量  $X_0$  的后验概率密度函数  $f_{X_0}^+(x)$ , 状态转移函数  $f(x)$ , 1 时刻过程噪声随机变量  $Q_1$  的概率密度函数  $f_{Q_1}(x)$ , 求解 1 时刻状态量随机变量  $X_1$  的先验概率密度函数  $f_{X_1}^-(x)$ 。

类似二维连续型随机变量贝叶斯公式的推导过程, 我们从求解  $X_1$  的先验累积分布函数  $F_{X_1}^-$  入手。

$$\begin{aligned}
F_{X_1}^-(x) &= P(X_1 \leq x) \\
&= \sum_{u=-\infty}^x P(X_1 = u) \Rightarrow \\
&= \sum_{u=-\infty}^x \sum_{v=-\infty}^{+\infty} P(X_1 = u \mid X_0 = v) P(X_0 = v) \Rightarrow \\
&= \sum_{u=-\infty}^x \sum_{v=-\infty}^{+\infty} P[X_1 - f(X_0) = u - f(v) \mid X_0 = v] P(X_0 = v) \Rightarrow \\
&= \sum_{u=-\infty}^x \sum_{v=-\infty}^{+\infty} P[Q_1 = u - f(v) \mid X_0 = v] P(X_0 = v) \Rightarrow \\
&= \sum_{u=-\infty}^x \sum_{v=-\infty}^{+\infty} P[Q_1 = u - f(v)] P(X_0 = v) \Rightarrow \textcolor{red}{X_{k-1} Q_k} \\
&= \sum_{u=-\infty}^x \left\{ \lim_{\epsilon \rightarrow 0} \sum_{v=-\infty}^{+\infty} f_{Q_1}[u - f(v)] \text{üü} f_{X_0}^+(v) \text{üü} \right\} \Rightarrow \\
&= \sum_{u=-\infty}^x \left\{ \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{+\infty} f_{Q_1}[u - f(v)] f_{X_0}^+(v) dv \text{üü} \right\} \Rightarrow \\
&= \int_{-\infty}^x \int_{-\infty}^{+\infty} f_{Q_1}[u - f(v)] f_{X_0}^+(v) dv du \Rightarrow \\
&= \int_{-\infty}^x \int_{-\infty}^{+\infty} f_{Q_1}[x - f(v)] f_{X_0}^+(v) dv dx \Rightarrow \textcolor{red}{ux}
\end{aligned} \tag{26}$$

故，1 时刻状态量随机变量  $X_1$  的先验概率密度函数为：

$$f_{X_1}^-(x) = \frac{dF_{X_1}^-(x)}{dx} = \int_{-\infty}^{+\infty} f_{Q_1}[x - f(v)] f_{X_0}^+(v) dv \tag{27}$$

推导完毕。可以发现，先验概率密度函数本质来源于状态方程。

2.4.3 更新步推导已知 1 时刻观测量随机变量  $Y_1$  的取值  $y_1$ ，求解 1 时刻状态量随机变量与观测量随机变量的似然概率密度函数  $f_{Y_1|X_1}(y_1 \mid x)$ ，并联合预测步得到的 1 时刻状态量随机变量  $X_1$  的先验概率密度函数  $f_{X_1}^-(x)$ ，求解 1 时刻状态量随机变量  $X_1$  的后验概率密度函数  $f_{X_1}^+(x)$ 。

首先，求解似然概率密度函数  $f_{Y_1|X_1}(y_1 \mid x)$ ：

$$\begin{aligned}
f_{Y_1|X_1}(y_1|x) &= \lim_{\epsilon \rightarrow 0} \frac{F_{Y_1|X_1}(y_1 + \epsilon|x) - F_{Y_1|X_1}(y_1|x)}{\epsilon} \Rightarrow \text{导数的定义} \\
&= \lim_{\epsilon \rightarrow 0} \frac{P(y_1 \leq Y_1 \leq y_1 + \epsilon | X_1 = x)}{\epsilon} \Rightarrow \text{累积分布函数的性质} \\
&= \lim_{\epsilon \rightarrow 0} \frac{P[y_1 - h(x) \leq Y_1 - h(X_1) \leq y_1 - h(x) + \epsilon | X_1 = x]}{\epsilon} \Rightarrow \text{条件概率里的条件可以} \\
&= \lim_{\epsilon \rightarrow 0} \frac{P[y_1 - h(x) \leq R_1 \leq y_1 - h(x) + \epsilon | X_1 = x]}{\epsilon} \Rightarrow \text{观测方程} \\
&= \lim_{\epsilon \rightarrow 0} \frac{P[y_1 - h(x) \leq R_1 \leq y_1 - h(x) + \epsilon]}{\epsilon} \Rightarrow X_k \text{与} R_k \text{相互独立} \\
&= \lim_{\epsilon \rightarrow 0} \frac{F_{R_1}[y_1 - h(x) + \epsilon] - F_{R_1}[y_1 - h(x)]}{\epsilon} \Rightarrow \text{累积分布函数的性质} \\
&= f_{R_1}[y_1 - h(x)] \Rightarrow \text{导数的定义}
\end{aligned} \tag{28}$$

可以发现，似然概率密度函数本质来源于观测方程。

然后，联合预测步得到的 1 时刻状态量随机变量  $X_1$  的先验概率密度函数  $f_{X_1}^-(x)$ ，求解 1 时刻状态量随机变量  $X_1$  的后验概率密度函数  $f_{X_1}^+(x)$ ：

$$f_{X_1}^+(x) = \eta_1 \mathring{u} f_{Y_1|X_1}(y_1|x) \mathring{u} f_{X_1}^-(x) = \eta_1 \mathring{u} f_{R_1}[y_1 - h(x)] \mathring{u} f_{X_1}^-(x) \tag{29}$$

其中，归一化常数  $\eta_1$  为：

$$\eta_1 = \left[ \int_{-\infty}^{+\infty} f_{Y_1|X_1}(y_1|x) f_{X_1}^-(x) dx \right]^{-1} = \left\{ \int_{-\infty}^{+\infty} f_{R_1}[y_1 - h(x)] f_{X_1}^-(x) dx \right\}^{-1} \tag{30}$$

推导完毕。

2.4.4 递推流程由预测步和更新步的推导结果，可得到由 0 时刻状态量随机变量  $X_0$  的后验概率密度函数  $f_{X_0}^+(x)$  到  $k$  时刻状态量随机变量  $X_k$  的后验概率密度函数  $f_{X_k}^+(x)$  的递推流程：



$$\begin{aligned}
f_{X_0}^+(x) &\xrightarrow{\text{预测}} f_{X_1}^-(x) = \int_{-\infty}^{+\infty} f_{Q_1}[x - f(v)]f_{X_0}^+(v)dv \xrightarrow{\text{观测更新}} f_{X_1}^+(x) = \eta_1 \mathfrak{U} f_{R_1}[y_1 - h(x)] \mathfrak{U} f_{X_1}^-(x) \\
&\xrightarrow{\text{预测}} f_{X_2}^-(x) = \int_{-\infty}^{+\infty} f_{Q_2}[x - f(v)]f_{X_1}^+(v)dv \xrightarrow{\text{观测更新}} f_{X_2}^+(x) = \eta_2 \mathfrak{U} f_{R_2}[y_2 - h(x)] \mathfrak{U} f_{X_2}^-(x) \\
&\dots \\
&\xrightarrow{\text{预测}} f_{X_k}^-(x) = \int_{-\infty}^{+\infty} f_{Q_k}[x - f(v)]f_{X_{k-1}}^+(v)dv \xrightarrow{\text{观测更新}} f_{X_k}^+(x) = \eta_k \mathfrak{U} f_{R_k}[y_k - h(x)] \mathfrak{U} f_{X_k}^-(x)
\end{aligned} \tag{31}$$

其中，归一化常数  $\eta_k$  为：

$$\eta_k = \left\{ \int_{-\infty}^{+\infty} f_{R_k}[y_k - h(x)]f_{X_k}^-(x)dx \right\}^{-1} \tag{32}$$

最终，可得到  $k$  时刻状态量随机变量  $X_k$  的最优估计  $\hat{x}_k$ ：

$$\hat{x}_k = E[f_{X_k}^+(x)] = \int_{-\infty}^{+\infty} x f_{X_k}^+(x)dx \tag{33}$$

### 完整算法框架

(1) 设初值

初始 0 时刻状态量随机变量  $X_0$  的后验概率密度函数：

$$f_{X_0}^+(x)$$

(2) 预测步

$k$  时刻状态量随机变量  $X_k$  的先验概率密度函数：

$$f_{X_k}^-(x) = \int_{-\infty}^{+\infty} f_{Q_k}[x - f(v)]f_{X_{k-1}}^+(v)dv$$

(3) 更新步

$k$  时刻状态量随机变量  $X_k$  的后验概率密度函数：

$$f_{X_k}^+(x) = \eta_k \mathfrak{U} f_{R_k}[y_k - h(x)] \mathfrak{U} f_{X_k}^-(x)$$

归一化常数  $\eta_k$ :

$$\eta_k = \left\{ \int_{-\infty}^{+\infty} f_{R_k}[y_k - h(x)] f_{X_k}^-(x) dx \right\}^{-1}$$

(4) 求解状态量后验估计

$k$  时刻状态量随机变量  $X_k$  的后验估计:

$$\hat{x}_k^+ = E[f_{X_k}^+(x)] = \int_{-\infty}^{+\infty} x f_{X_k}^+(x) dx$$

2.5 贝叶斯滤波的缺点及解决方法 2.5.1 缺点从上文的推导及结论中可以发现, 求解预测步中的先验概率密度函数  $f_{X_k}^-(x)$ 、更新步中的归一化常数  $\eta_k$ 、最终的最优估计  $\hat{x}_k$  时均涉及到无穷积分, 而大多数情况无法得到解析解, 使得贝叶斯滤波算法的直接应用十分困难。

2.5.2 解决办法为了解决贝叶斯滤波中的无穷积分问题, 通常从两个角度出发:

(1) 作理想假设

假设状态转移函数  $f(x)$  和观测函数  $h(x)$  均为线性函数, 过程噪声随机变量  $Q_k$  和观测噪声随机变量  $R_k$  均服从均值为 0 的正态分布——卡尔曼滤波 (Kalman Filter) 假设状态转移函数  $f(x)$  和 (或) 观测函数  $h(x)$  为非线性函数, 过程噪声随机变量  $Q_k$  和观测噪声随机变量  $R_k$  均服从均值为 0 的正态分布——扩展卡尔曼滤波 (Extended Kalman Filter) 和无迹卡尔曼滤波 (Unscented Kalman Filter) (2) 化连续为离散

将无穷积分转化为数值积分, 一般有以下方法:

- 高斯积分 (不常用)
- 蒙特卡罗积分 (粒子滤波, Particle Filter)
- 直方图滤波

## 3 高斯滤波

高斯滤波, 称为一个重要的递归状态估计器家族

### 3.1 卡尔曼滤波

**卡尔曼滤波**是以贝叶斯滤波为理论基础, 并通过假设状态量随机变量(以下简称状态量)、观测量均服从**正态分布**, 假设过程噪声、观测噪声均服从均值为 0 的正态分布, 以及假设**状态转移函数**和**观测函数**均为**线性函数**, 实现对连续型随机过程的递推状态估计。简言之, 卡尔曼滤波是在贝叶斯滤波框架下求解线性高斯问题。

简单来讲, 卡尔曼滤波器就是根据上一时刻的状态, 预测当前时刻的状态, 将预测的状态量与当前时刻的观测量进行加权, 加权后的结果才认为是当前实际的状态量, 而非仅听信当前的观测量。

#### 3.1.1 卡尔曼滤波的假设

卡尔曼滤波以贝叶斯滤波为理论基础, 并作了六个前提假设: // (1)  
假设一: 状态量服从正态分布

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$$

(2) 假设二: 观测量服从正态分布

$$Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$

(3) 假设三: 过程噪声服从均值为 0 的正态分布

$$Q \sim \mathcal{N}(0, \sigma_Q^2)$$

(4) 假设四: 观测噪声服从均值为 0 的正态分布

$$Q \sim \mathcal{N}(0, \sigma_R^2)$$

(5) 假设五：状态转移函数为线性函数

$$f(X_k) = F * X_{k-1} + B * u_k$$

其中,  $F$  为状态转移比例项, 对于单一状态量的卡尔曼滤波中,  $F$  为一常数;  $B$  为控制比例项,  $u_k$  为控制量,  $B$  和  $u_k$  的乘积可视为线性状态转移函数中的截距项。在简单的系统中, 常常没有控制项  $B$  和  $u_k$ 。

### 3.1.2 矩阵形式的卡尔曼滤波

当状态量和观测量不再是单一的随机变量而是由多个随机变量组成的序列时, 卡尔曼滤波中各个量的维数也将随之改变:

1. 状态量  $X$  由随机变量演变为随机向量, 随机向量中的每一个分量为一个状态量随机变量。维数为  $n_X \times 1$
2. 状态转移比例项  $F$  演变为矩阵, 维数为  $n_X \times n_X$
3. 控制量  $u_k$  演变为矩阵, 维数为  $n_u \times 1$
4. 控制比例项  $B$  演变为矩阵, 维数为  $n_X \times n_u$
5. 状态量概率密度函数均值  $\mu$  演变为矩阵, 维数为  $n_X \times 1$
6. 状态量概率密度函数方差  $\sigma^2$  演变为协方差矩阵, 用  $\Sigma$  表示, 维数为  $n_X \times n_X$
7. 过程噪声方差  $\sigma_Q^2$  演变为协方差矩阵, 用  $\Sigma_Q$  表示, 维数为  $n_X \times n_X$
8. 观测量  $Y$  由随机变量演变为随机向量, 随机向量中的每一个分量为一个观测量随机变量。维数为  $n_Y \times 1$
9. 观测值  $y_k$  由单一值演变为由单一值组成的值矩阵, 维数为  $n_Y \times 1$
10. 观测比例项  $H$  演变为矩阵, 维数为  $n_Y \times n_X$
11. 观测噪声方差  $\sigma_R^2$  演变为协方差矩阵, 用  $\Sigma_R$  表示, 维数为  $n_Y \times n_Y$
12. 卡尔曼增益系数  $K$  演变为矩阵, 维数为  $n_X \times n_Y$

对应的五个公式演变为:

$$\mu_k^- = F * \mu_{k-1}^+ + B * u_k \quad (4.1)$$

$$\Sigma_k^- = F * \Sigma_{k-1}^+ * F^T + \Sigma_{Q_k} \quad (4.2)$$

$$\mu_k^+ = \mu_k^- + K * (y_k - H * \mu_k^-) \quad (4.3)$$

$$\Sigma_k^+ = (I - K * H) * \Sigma_k^- \quad (4.4)$$

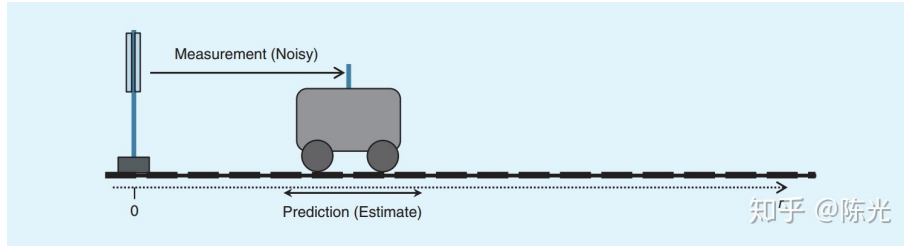
$$K = \Sigma_k^- * H^T * (H * \Sigma_k^- * H^T + \Sigma_{R_k})^{-1} \quad (4.5)$$

公式 (4.3) 中  $\mu_k^+$  即 k 时刻状态量  $X_k$  的后验估计  $\hat{x}_k^+ y_k - H * \mu_k^-$  常被称为残差 (Residual) 或新息 (Innovation); 公式 (4.4) 中的 I 代表单位矩阵, 维数为  $n_X \times n_X$ 。

从结果中还可以发现, 外部控制项  $B * u_k$  通过影响先验估计均值间接影响了后验估计均值, 但对后验估计方差没有影响。

### 3.1.3 实战写卡尔曼滤波器

**简单例子:** 假设有个小车在道路上向右侧匀速运动, 我们在左侧安装了一个测量小车距离和速度传感器, 传感器每 1 秒测一次小车的位置 s 和速度 v, 如下图所示: 我们用向量  $x_t$  来表示当前小车的状态, 该向量也是

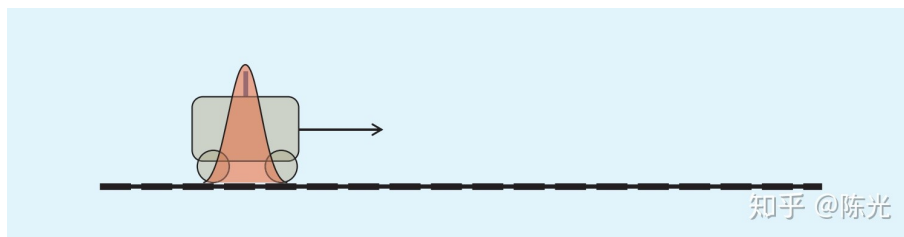


最终的输出结果, 被称作状态向量 (state vector):

$$x_t = \begin{bmatrix} s_t \\ v_t \end{bmatrix} \quad (1)$$

由于测量误差的存在, 传感器无法直接获取小车位置的真值, 只能获取在真值附近的一个近似值, 可以假设测量值在真值附近服从高斯分布。如下图

所示，测量值分布在红色区域的左侧或右侧，真值则在红色区域的波峰处。由于是第一次测量，没有小车的历史信息，我们认为小车在 1 秒时的状态



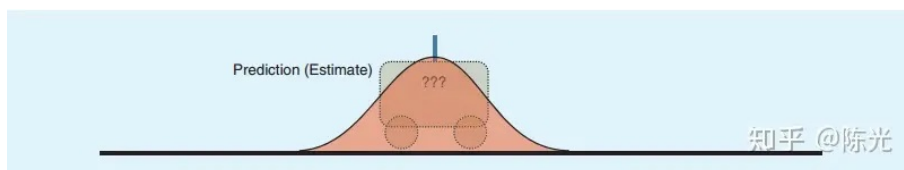
$x$  与测量值  $z$  相等，表示如下：

$$x_t = \begin{bmatrix} s_1 \\ v_1 \end{bmatrix} \quad (2)$$

公式中的 1 表示第 1 秒。

### 预测

预测是卡尔曼滤波器中很重要的一步，这一步相当于使用历史信息对未来的位置进行推测。根据第 1 秒小车的位置和速度，我们可以推测第 2 秒时，小车所在的位置应该如下图所示。会发现，图中红色区域的范围变大了，这是因为预测时加入了速度估计的噪声，是一个放大不确定性的过程。根据小车第一秒的状态进行预测，得到预测的状态  $x_{pre}$ ：



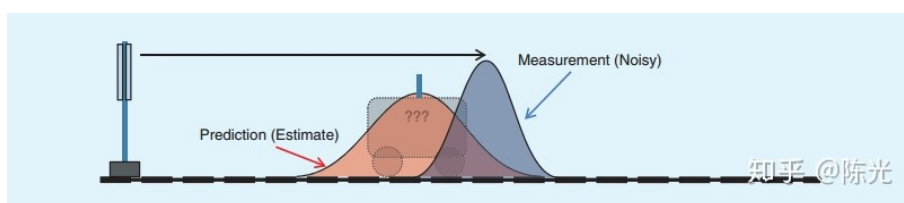
$$x_{pre} = \begin{bmatrix} s_1 + v_1 \\ v_1 \end{bmatrix} \quad (2)$$

其中，pre 是 prediction 的简称；时间间隔为 1 秒，所以预测位置为距离 + 速度 \* 1；由于小车做的是匀速运动，因此速度保持不变。

**观测** 在第 2 秒时，传感器对小车的位置做了一次观测，我们认为小车在第 2 秒时观测值为  $z_2$ ，用向量表示第 2 秒时的观测结果为：

$$z_2 = \begin{bmatrix} s_2 \\ v_2 \end{bmatrix} \quad (2)$$

很显然，第二次观测的结果也是存在误差的，我们将预测的小车位置与实际观测到的小车位置放到一个图上，即可看到：图中红色区域为预测的小车



位置，蓝色区域为第 2 秒的观测结果。

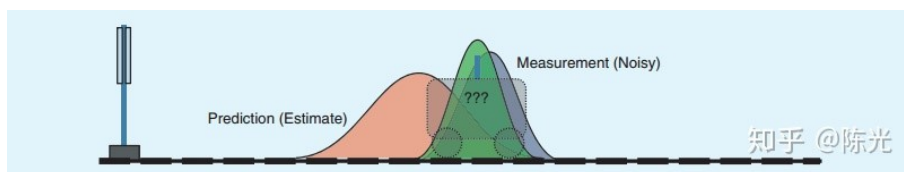
很显然，这两个结果都在真值附近。为了得到尽可能接近真值的结果，我们将这两个区域的结果进行加权，取加权后的值作为第二秒的状态向量。

为了方便理解，可以将第 2 秒的状态向量写成：

$$x_2 = w_1 \times x_{pre} + w_2 \times Z_2$$

其中， $w_1$  为预测结果的权值， $w_2$  为观测结果的权值。两个权值的计算是根据预测结果和观测结果的不确定性来的，这个不确定性就是高斯分布中的方差的大小，方差越大，波形分布越广，不确定性越高，这样一来给的权值就会越低。

加权后的状态向量的分布，可以用下图中绿色区域表示：你会发现绿



色区域的方差比红色区域和蓝色区域的小。这是因为进行加权运算时，需

要将两个高斯分布进行乘法运算，得到的新的高斯分布的方差比两个做乘法的高斯分布都小。

两个不那么确定的分布，最终得到了一个相对确定的分布，这是卡尔曼滤波的一直被推崇的原因。

### 再预测，再观测（Prediction & Measurement）

第 1 秒的初始化以及第 2 秒的预测、观测，实现卡尔曼滤波的一个周期。

同样的，我们根据第 2 秒的状态向量做第 3 秒的预测，再与第 3 秒的观测结果进行加权，就得到了第 3 秒的状态向量；再根据第 3 秒的状态向量做第 4 秒的预测，再与第 4 秒的观测结果进行加权，就得到了第 4 秒的状态向量。以此往复，就实现了一个真正意义上的卡尔曼滤波器。

以上就是卡尔曼滤波器的感性分析过程，下面我们回归理性，谈谈如何将以上过程写成代码。

#### 3.1.4 扩展卡尔曼滤波

**扩展卡尔曼滤波**（Extended Kalman Filter, EKF）是标准卡尔曼滤波在非线性情形下的一种扩展形式，它是一种高效率的递归滤波器（自回归滤波器）。

EKF 的基本思想是利用泰勒级数展开将非线性系统的状态转移函数和（或）观测函数线性化，然后采用卡尔曼滤波框架对信号进行滤波，因此它是一种次优滤波。

#### 与卡尔曼滤波相同的假设

(1) 假设一：状态量服从正态分布

$$X \sim \mathcal{N}(\mu_X, \sigma_X^2)$$

(2) 假设二：观测量服从正态分布

$$Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$$



# Kalman Filter

## Prediction

$$x' = Fx + u$$

$$P' = FPF^T + Q$$

## Measurement update

$$y = z - Hx'$$

$$S = HP'H^T + R$$

$$K = P'H^TS^{-1}$$

$$x = x' + Ky$$

$$P = (I - KH)P'$$

知乎 @陈光

图 1: 七个描述卡尔曼滤波器的公式

(3) 假设三：过程噪声服从均值为 0 的正态分布

$$Q \sim \mathcal{N}(0, \sigma_Q^2)$$

(4) 假设四：观测噪声服从均值为 0 的正态分布

$$R \sim \mathcal{N}(0, \sigma_R^2)$$

**与卡尔曼滤波不同的假设** (5) 假设五：状态转移函数和（或）观测函数为非线性函数

在卡尔曼滤波的前提假设中，认为状态方程中的状态转移函数  $f(x)$  以及观测方程中的观测函数  $h(x)$  均为线性函数。基于这种线性假设，存在常数或常矩阵  $F$ ，使得  $f(x)$  可以写成卡尔曼滤波中的线性形式，存在常数或常矩阵  $H$ ，使得  $h(x)$  也可以写成卡尔曼滤波中的线性形式。

不同于标准卡尔曼滤波，**扩展卡尔曼滤波处理的是非线性系统**，假设系统的状态转移函数和（或）观测函数为非线性函数。

**矩阵形式的扩展卡尔曼滤波** 状态转移比例项

$$F_j = \left( \begin{array}{cccc} \frac{\partial f_1}{\partial X_{k-1}^1} & \frac{\partial f_1}{\partial X_{k-1}^2} & \cdots & \frac{\partial f_1}{\partial X_{k-1}^{n_X}} \\ \frac{\partial f_2}{\partial X_{k-1}^1} & \frac{\partial f_2}{\partial X_{k-1}^2} & \cdots & \frac{\partial f_2}{\partial X_{k-1}^{n_X}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{n_X}}{\partial X_{k-1}^1} & \frac{\partial f_{n_X}}{\partial X_{k-1}^2} & \cdots & \frac{\partial f_{n_X}}{\partial X_{k-1}^{n_X}} \end{array} \right) \bigg|_{X_{k-1} = \hat{X}_{k-1}^+}$$

观测比例项

$$H_j = \left( \begin{array}{cccc} \frac{\partial h_1}{\partial X_k^1} & \frac{\partial h_1}{\partial X_k^2} & \cdots & \frac{\partial h_1}{\partial X_k^{n_X}} \\ \frac{\partial h_2}{\partial X_k^1} & \frac{\partial h_2}{\partial X_k^2} & \cdots & \frac{\partial h_2}{\partial X_k^{n_X}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{n_Y}}{\partial X_k^1} & \frac{\partial h_{n_Y}}{\partial X_k^2} & \cdots & \frac{\partial h_{n_Y}}{\partial X_k^{n_X}} \end{array} \right) \bigg|_{X_k = \hat{X}_k^-}$$

五个公式：

$$\mu_k^- = f(\mu_{k-1}^+) \quad (4.1)$$

$$\Sigma_k^- = F_j * \Sigma_{k-1}^+ * F_j^T + \Sigma_{Q_k} \quad (4.2)$$

$$\mu_k^+ = \mu_k^- + K * [y_k - h(\mu_k^-)] \quad (4.3)$$

$$\Sigma_k^+ = (I - K * H_j) * \Sigma_k^- \quad (4.4)$$

$$K = \Sigma_k^- * H_j^T * (H_j * \Sigma_k^- * H_j^T + \Sigma_{R_k})^{-1} \quad (4.5)$$

公式 (4.3) 中  $\mu_k^+$  即 k 时刻状态量  $X_k$  的后验估计  $\hat{x}_k^+$ ， $y_k - h(\mu_k^-)$  常被称为残差 (Residual) 或新息 (Innovation)；公式 (4.4) 中的 I 代表单位矩阵，维数为  $n_X \times n_X$ 。

当然，对于某个非线性系统，不一定状态转移和观测都是非线性的：

- 线性的状态转移 + 非线性的观测此时，滤波递推公式由卡尔曼滤波的预测步两公式和扩展卡尔曼滤波的更新步三公式组成。
- 非线性的状态转移 + 线性的观测此时，滤波递推公式由扩展卡尔曼滤波的预测步两公式和卡尔曼滤波的更新步三公式组成。