

---

# 目 录

1.	卡尔曼滤波器简介 .....	2
2.	由状态观测器建立卡尔曼滤波器状态方程 .....	3
2.1	状态观测器简介 .....	3
2.2	离散情况下状态观测器表示 .....	4
2.3	卡尔曼滤波器系统框图 .....	5
3.	卡尔曼滤波器参数推导 .....	7
3.1	先验估计协方差 $P_k^-$ .....	7
3.2	后验估计协方差 $P_k$ .....	7
3.3	卡尔曼增益 $K_k$ .....	8
3.4	滤波器公式总结 .....	10
4.	卡尔曼滤波器应用示例 .....	10
4.1	惯性测量单元(IMU)卡尔曼滤波 .....	10
4.1.1	卡尔曼滤波主要公式回顾 .....	10
4.1.2	IMU 状态方程与测量方程建模 .....	11
4.1.3	IMU 卡尔曼滤波器方程 .....	13
4.1.4	算法实现 .....	14
4.1.5	MATLAB 程序 .....	16
5.	扩展卡尔曼滤波器 .....	17
5.1	扩展卡尔曼滤波介绍 .....	17
5.2	非线性卡尔曼滤波线性化推导（状态方程建立） .....	18
5.3	扩展卡尔曼滤波器系统框图 .....	21
5.4	扩展卡尔曼滤波器与卡尔曼滤波器对比 .....	22
6.	扩展卡尔曼滤波器参数推导 .....	23
6.1	先验估计协方差 .....	23
6.2	后验估计协方差 .....	23
6.3	扩展卡尔曼增益 .....	24
6.4	滤波器公式总结 .....	25
7.	扩展卡尔曼滤波器应用示例 .....	26
7.1	惯性测量单元(IMU)扩展卡尔曼滤波 .....	26
7.1.1	四元数简介及相关公式 .....	26
7.1.2	扩展卡尔曼滤波主要公式回顾 .....	26
8.	参考文献 .....	26
9.	附录 .....	27
9.1	附录一：加速度计计算姿态角推导 .....	27

## 1. 卡尔曼滤波器简介

卡尔曼滤波器用于估计离散~~线性~~时间过程的状态变量  $x$ ，这个离散事件过程由以下离散随机差分方程描述：

**状态变量方程：**  $x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$

**测量方程：**  $z_k = Hx_k + v_k$

其中  $w_k, v_k$  分别表示过程噪声(process noise)和观测噪声(measurement noise)。假设他们为相互独立，正太分布的白噪声，即遵循高斯分布：

$$p(w) \sim N(0, Q)$$

$$p(v) \sim N(0, R)$$

实际过程中  $Q, R$  可能随着每次迭代过程而变化，这里假设他们是常数。

定义  $\hat{x}_k^- \in R^n$  ( $-$ 代表先验， $\hat{\cdot}$ 代表估计)为在已知第  $k$  步以前状态情况下第  $k$  步的~~先验~~

~~状态估计~~。 $\hat{x}_k \in R^n$  为在已知第  $k$  步以前状态情况下第  $k$  步的~~后验状态估计~~。因此得到~~先验估计误差~~和~~后验估计误差~~：

$$e_k^- \equiv x_k - \hat{x}_k^-$$

$$e_k \equiv x_k - \hat{x}_k$$

其协方差为：

先验估计协方差： $P_k^- = E(e_k^- e_k^{-T}) = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T]$ ， $E$  代表取期望

后验估计协方差： $P_k = E(e_k e_k^T) = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$

此处先给出线性离散卡尔曼滤波器的时间更新方程和测量更新方程，推导过程详见第二章和第三章：

**时间更新方程（预测方程）：**

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$$

$$P_k^- = AP_{k-1}A^T + Q$$

**测量更新方程（修正方程）：**

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H) P_k^-$$

## 2. 由状态观测器建立卡尔曼滤波器状态方程

### 2.1 状态观测器简介

状态观测器 (state observer) 是控制理论中配合实际系统的一种系统, 在已知实际系统架构的情形下, 可以根据一实际系统的输入及输出而估测其内部的状态。

对状态能控的线性定常系统, 可以通过线性状态反馈来进行任意极点配置, 以使闭环系统具有所期望的极点及性能品质指标。但是, 由于描述内部运动特性的状态变量有时并不是能直接测量的, 更甚者有时并没有实际物理量与之直接相对应而成为一种抽象的数学变量。在这些情况下, 以状态变量作为反馈变量来构成状态反馈系统带来了具体工程实现上的困难。

因此提出了状态变量的重构或观测估计问题。那么什么是状态变量的重构或观测估计问题?

设法另外构造一个物理或算法可实现的动态系统, 它以原系统的输入和输出作为它的输入, 而它的状态变量的值能渐近逼近原系统的状态变量的值或者其某种线性组合, 则这种渐近逼近的状态变量的值即为原系统的状态变量的估计值, 并可用于状态反馈闭环系统中代替原状态变量作为反馈量来构成状态反馈律。这种重构或估计系统状态变量值的装置称为**状态观测器**。它可以是由电子、电气等装置构成的物理系统, 也可以是由计算机和计算模型及软件来实现的软系统。

观测器可分为: 全维观测器、降维观测器。本文中**以全维观测器**举例。

简单的来说, 有一可观线性定常连续系统, 其状态空间模型为:

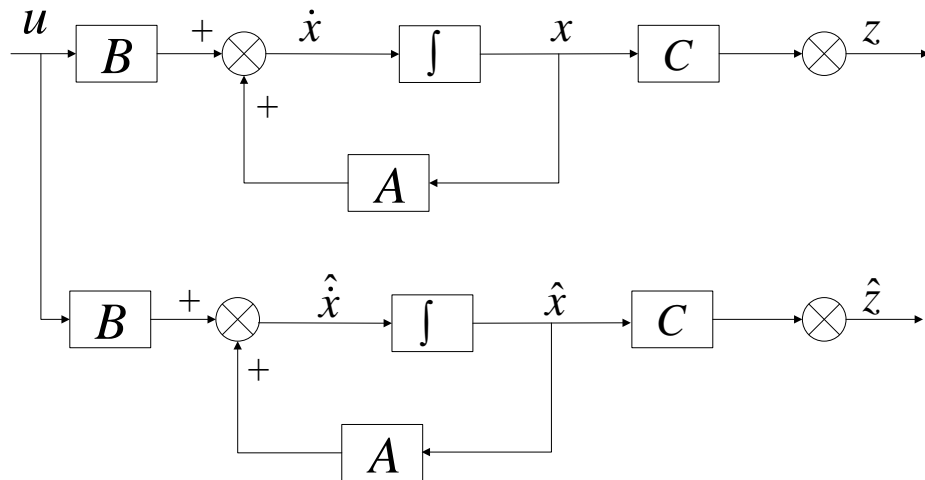
$$\begin{cases} \dot{x} = Ax + Bu \\ z = Cx \end{cases}$$

若状态变量  $x(t)$  不能完全直接测量到, 则希望利用仿真技术来构造一个和被控系统有同样动力学性质(即有同样的系数矩阵  $A$ 、 $B$  和  $C$ )的如下系统来重构被控系统的状态变量:

$$\begin{cases} \dot{\hat{x}} = A\hat{x} + Bu \\ \hat{z} = C\hat{x} \end{cases}$$

上标代表**被观测**。

其基本的系统框图如下:



比较被观测系统和观测系统状态方程, 有:

$$\dot{x}(t) - \dot{\hat{x}}(t) = A(x(t) - \hat{x}(t)),$$

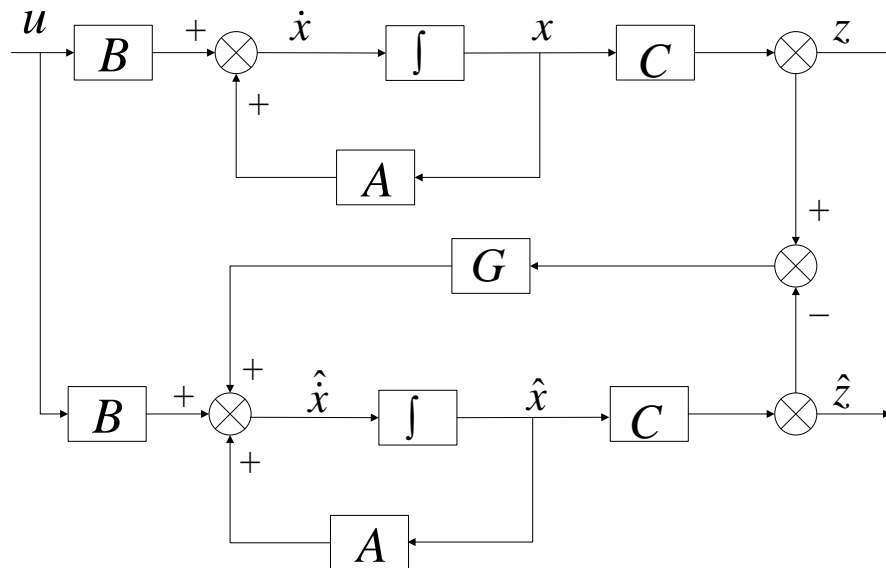
解之状态估计误差的解为：

$$x(t) - \hat{x}(t) = e^{At} [x(0) - \hat{x}(0)]$$

易得，要使得观测值接近原始值，即  $x(t) - \hat{x}(t) = 0$ ，则  $x(0) - \hat{x}(0) = 0$ 。

然而，有些系统难以得到初始状态变量  $x(0)$ ，即无法保证  $x(0) - \hat{x}(0) = 0$ 。

为了解决这个问题，首先对原系统进行分析，发现其处于开环状态，只利用了被控系统输入  $u(t)$ ，而未利用输出  $y(t)$ ，即观测器得到的  $\hat{x}(t)$  只是  $x(t)$  的一种开环估计值，一旦观测值出现偏差，无法予以纠偏。故此处引入输出插值反馈，思路为：根据反馈控制原理， $z(t) - \hat{z}(t)$  的差值可用来调节  $\hat{x}(t)$ ，使  $x(t) - \hat{x}(t) \rightarrow 0$ ，并同时引入输出误差反馈矩阵  $G$ ，如下图所示：



引入输出反馈后的状态观测器方程为：

$$\begin{aligned} \dot{x} - \dot{\hat{x}} &= Ax + Bu - (A\hat{x} + Bu + G(z - \hat{z})) \\ &= A(x - \hat{x}) - GC(x - \hat{x}) \\ &= (A - GC)(x - \hat{x}) \end{aligned}$$

$A - GC$  称为状态观测器的系统矩阵。

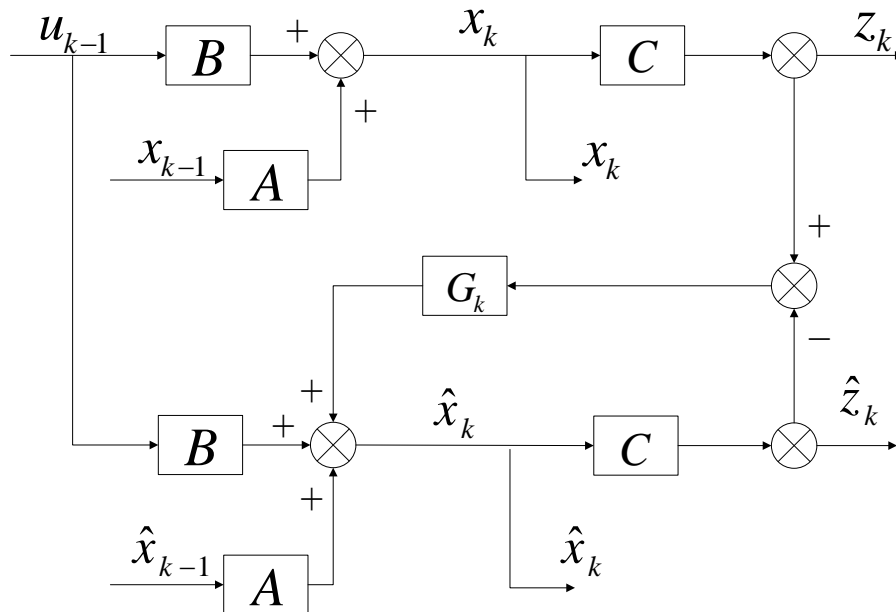
上述误差方程的解为：

$$x(t) - \hat{x}(t) = e^{(A-GC)t} [x(0) - \hat{x}(0)]$$

此处的意义在于通过输出反馈引入矩阵  $G$ ，使得通过  $G$  来调整使  $A - GC$  的特征值均具有负实部，则  $x(t) - \hat{x}(t)$  的值渐近衰减为零，即观测值趋向于实际值。 $G$  的设计方法此处不再赘述，参见现代控制理论相关课本。

## 2.2 离散情况下状态观测器表示

离散情况下带输出反馈的状态观测器表示如下图：



此时原系统状态方程表达式为：

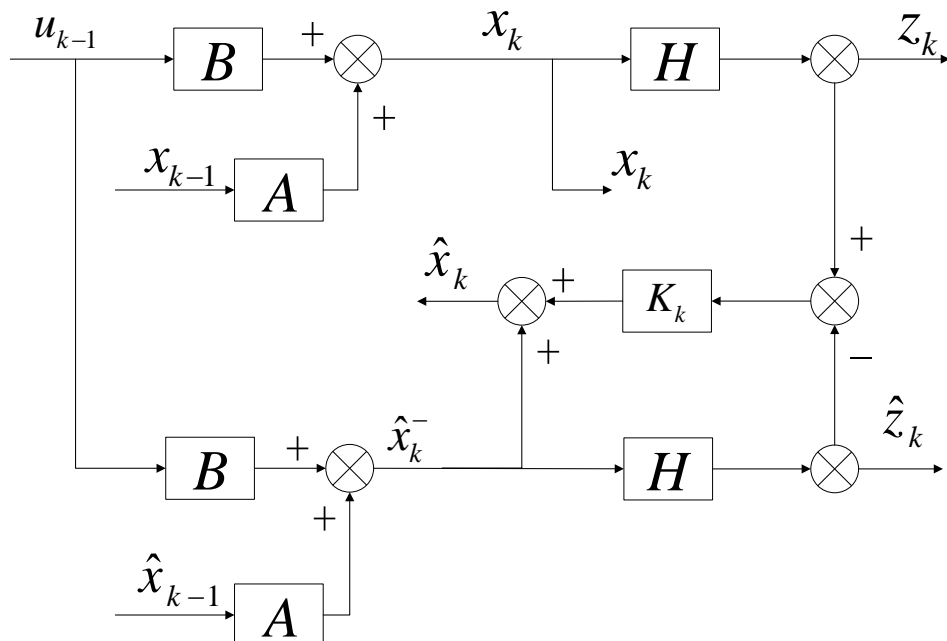
$$\begin{cases} x_k = Ax_{k-1} + Bu \\ z_k = Cx_k \end{cases}$$

状态观测器表达式为

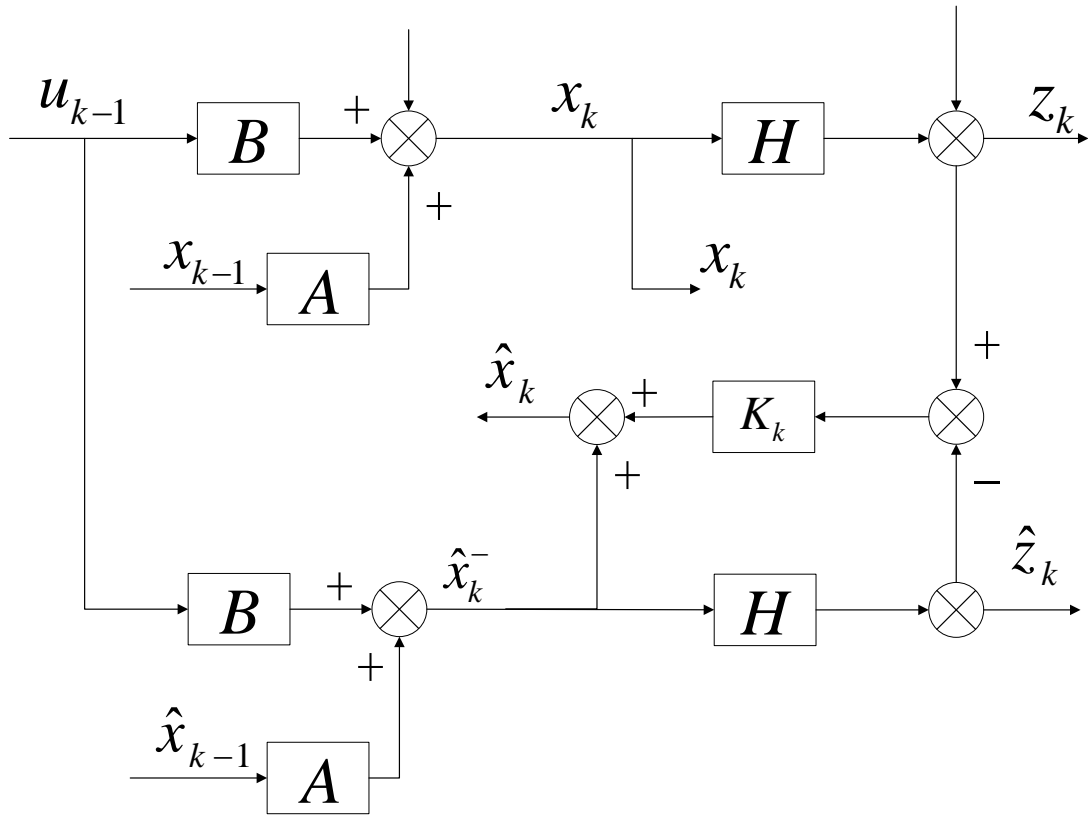
$$\begin{cases} \hat{x}_k = A\hat{x}_{k-1} + Bu + G_k(z_k - \hat{z}_k) \\ \hat{z}_k = C\hat{x}_k \end{cases}$$

### 2.3 卡尔曼滤波器系统框图

将上述式子中的  $C$  改成  $H$ ， $G$  改成  $K$ ，则可得到用来表述卡尔曼滤波器的状态变量框图，同时我们将利用该框图来推导卡尔曼滤波器的状态方程（此处为便于表示先验与后验以及输出反馈的关系，将估计过程中的第一个运算符拆成两个）：



补上过程噪声  $w_k$  和观测噪声  $v_k$  可得如下框图



记增加**输出反馈前**（即在观测器的  $\hat{x}_k$  求取的过程中没有引入  $K_k(z_k - H\hat{x}_k^-)$  项）的对  $x_k$  状

态估计为**先验状态估计**，即只根据  $x_{k-1}$  来估计  $x_k$ ，不对系统做校正，记作  $\hat{x}_k^-$ （ $-$  代表先

验， $\wedge$  代表估计），增加**输出反馈后**的状态估计为**后验状态估计**，记作  $\hat{x}_k$ ，则有

先验估计的状态表达式为：

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$$

后验估计的状态表达式为：

$$\begin{aligned}\hat{x}_k &= A\hat{x}_{k-1} + Bu_{k-1} + K_k(z_k - H\hat{x}_k^-) \\ &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)\end{aligned}$$

可以看到此处的两个状态表达式即为线性离散卡尔曼滤波器测量更新方程和时间更新方程的状态表达式。同时为形象理解，可以把  $\hat{x}_k^-$  看成预测项，则  $K_k(z_k - H\hat{x}_k^-)$  为修正项。

注：不同的地方对先验，后验概率等有不同的标记，例如有的博客中记作  $\hat{x}_{k-1|k-1}$ ， $\hat{x}_{k|k-1}$ ，

$\hat{x}_{k|k}$ 。也可以通过我们的模型来理解。 $\hat{x}_{k|k-1}$  为利用前一状态即  $k-1$  的状态来给出  $\hat{x}_k$  的估计

值，可以看到其实是上述框图中未添加输出反馈时对状态的估计值，即先验估计； $\hat{x}_{k|k}$  则

为添加了输出反馈后的估计值，即利用  $k$  的状态来给出  $\hat{x}_k$  的估计值，称为后验估计。

$\hat{x}_{k-1|k-1}$  即为利用  $k-1$  的状态来给出  $\hat{x}_{k-1}$  的估计值，即  $k-1$  时刻的后验估计。

### 3. 卡尔曼滤波器参数推导

#### 3.1 先验估计协方差 $P_k^-$

由上得到先验估计协方差表达式为：

$$P_k^- = E(e_k^- e_k^{-T}) = E\left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\right]$$

把其中  $x_k, \hat{x}_k^-$  的表达式代入：

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$$

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$$

代入则有：

$$\begin{aligned} x_k - \hat{x}_k^- &= Ax_{k-1} + Bu_{k-1} + w_{k-1} - A\hat{x}_{k-1} - Bu_{k-1} \\ &= A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1} \end{aligned}$$

则：

$$\begin{aligned} P_k^- &= E(e_k^- e_k^{-T}) \\ &= E\left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\right] \\ &= E\left[A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}\right] \left[A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}\right]^T \\ &= AE\left[(x_{k-1} - \hat{x}_{k-1})(x_{k-1} - \hat{x}_{k-1})^T\right]A^T + E[w_{k-1}w_{k-1}^T] \\ &= AP_{k-1}A^T + Q \end{aligned}$$

则得到时间更新方程中的第二个式子

#### 3.2 后验估计协方差 $P_k$

由上得到后验估计协方差表达式为：

$$P_k = E(e_k e_k^T) = E\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right]$$

把其中  $x_k, \hat{x}_k$  的表达式代入：

$$\begin{aligned} x_k &= Ax_{k-1} + Bu_{k-1} + w_{k-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \\ &= A\hat{x}_{k-1} + Bu_{k-1} + K_k(z_k - H\hat{x}_k^-) \\ &= A\hat{x}_{k-1} + Bu_{k-1} + K_k H(x_k - \hat{x}_k^-) + K_k v_k \end{aligned}$$

(上式计算过程中把最开始的测量方程  $z_k = Hx_k + v_k$  代入)

代入则有：

$$\begin{aligned}
 x_k - \hat{x}_k &= x_k - \left( \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-) \right) \\
 &= x_k - \left( \hat{x}_k^- + K_k (Hx_k + v_k - H\hat{x}_k^-) \right) \\
 &= x_k - \left( \hat{x}_k^- + K_k H (x_k - \hat{x}_k^-) + K_k v_k \right) \\
 &= x_k - \hat{x}_k^- - K_k H (x_k - \hat{x}_k^-) - K_k v_k \\
 &= (I - K_k H) (x_k - \hat{x}_k^-) - K_k v_k
 \end{aligned}$$

或者用

$$\begin{aligned}
 x_k - \hat{x}_k &= Ax_{k-1} + Bu_{k-1} + w_{k-1} - A\hat{x}_{k-1} - Bu_{k-1} - K_k H (x_k - \hat{x}_k^-) - K_k v_k \\
 &= A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1} - K_k H (x_k - \hat{x}_k^-) - K_k v_k \\
 &= A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1} - K_k H (A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}) - K_k v_k \\
 &= (I - K_k H) (A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}) - K_k v_k
 \end{aligned}$$

在 3.1 中已经求得：

$$\begin{aligned}
 x_k - \hat{x}_k^- &= Ax_{k-1} + Bu_{k-1} + w_{k-1} - A\hat{x}_{k-1} - Bu_{k-1} \\
 &= A(x_{k-1} - \hat{x}_{k-1}) + w_{k-1}
 \end{aligned}$$

即上述两种方法求得的先验估计误差  $x_k - \hat{x}_k$  相等。此处用第一种方法，即：

$$x_k - \hat{x}_k = (I - K_k H) (x_k - \hat{x}_k^-) - K_k v_k$$

故有

$$\begin{aligned}
 P_k &= E(e_k e_k^T) \\
 &= E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \\
 &= E\left[(I - K_k H)(x_{k-1} - \hat{x}_{k-1}^-)(x_{k-1} - \hat{x}_{k-1}^-)^T (I - K_k H)^T + K_k v_k v_k^T K_k^T\right] \\
 &= (I - K_k H) P_k^- (I - K_k H)^T + K_k R K_k^T
 \end{aligned}$$

可以看到，此处的后验估计误差与前面给出的卡尔曼滤波测量更新方程仍有差距，待求出卡尔曼滤波器增益后代入可化简，请先看 3.3。

化简后得到：

$$P_k = (I - K_k H) P_k^-$$

### 3.3 卡尔曼增益 $K_k$

滤波器增益  $K_k$  含义较为抽象，结合上面状态观测器的理论来说，在此处的意义在于通过输出反馈引入矩阵  $K_k$ ，使得通过调整  $K_k$  可以使  $x_k - \hat{x}_k$  的方差渐近衰减为最小，即观测值趋向于实际值。

使后验估计  $x_k - \hat{x}_k$  的方差  $P_k$  达到最小，也是使协方差矩阵  $E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$  的



对角线元素的和（也就是线性代数中的迹）最小化的过程。此处补充说明，协方差的对角线元素即为各个对应元素的方差。

故要求  $K_k$ ，使得  $P_k$  的迹取到一个最小值。即  $P_k$  关于  $K_k$  的导数为零。

首先将  $P_k$  展开

$$\begin{aligned} P_k &= (I - K_k H) P_k^- (I - K_k H)^T + K_k R K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + K_k H P_k^- (K_k H)^T + K_k R K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + K_k (H P_k^- H^T + R) K_k^T \end{aligned}$$

此时得到的为协方差，先对其取对角线元素之和，然后对其迹关于  $K_k$  求导。取迹

$tr(P_k)$  有：

$$tr(P_k) = tr(P_k^-) - 2tr(K_k H P_k^-) + tr(K_k (H P_k^- H^T + R) K_k^T)$$

\*此处说明，上式中直接  $tr(K_k H P_k^-) = tr(P_k^- (K_k H)^T)$ ，因为协方差矩阵  $P_k^-$  是对称阵，且根据矩阵的迹等于其转置矩阵的迹，故有：

$$tr(P_k^- (K_k H)^T) = tr(P_k^{-T} (K_k H)^T) = tr((K_k H P_k^-)^T) = tr(K_k H P_k^-)$$

对  $tr(P_k)$  关于  $K_k$  求导：

$$\frac{\partial T(P_k)}{\partial K_k} = \frac{\partial (tr(P_k^-) - 2tr(K_k H P_k^-) + tr(K_k (H P_k^- H^T + R) K_k^T))}{\partial K_k}$$

当导数等于零时，即  $\frac{\partial tr(P_k)}{\partial K_k} = 0$ ，有：

$$\begin{aligned} & \frac{\partial (tr(P_k^-) - 2tr(K_k H P_k^-) + tr(K_k (H P_k^- H^T + R) K_k^T))}{\partial K_k} \\ &= \frac{\partial (tr(P_k^-))}{\partial K_k} - 2 \frac{\partial (tr(K_k H P_k^-))}{\partial K_k} + \frac{\partial (tr(K_k (H P_k^- H^T + R) K_k^T))}{\partial K_k} \\ &= 0 - 2(H P_k^-)^T + K_k \left[ (H P_k^- H^T + R) + (H P_k^- H^T + R)^T \right] \\ &= 0 - 2(H P_k^-)^T + 2K_k \left[ (H P_k^- H^T + R) \right] \\ &= 0 \end{aligned}$$

即：

$$\begin{aligned} 2K_k \left[ (HP_k^- H^T + R) \right] &= 2(HP_k^-)^T \\ K_k &= (P_k^-)^T (H)^T (HP_k^- H^T + R)^{-1} \\ K_k &= P_k^- (H)^T (HP_k^- H^T + R)^{-1} \end{aligned}$$

此时将  $K_k = P_k^- (H)^T (HP_k^- H^T + R)^{-1}$

代入后验估计误差  $P_k = (I - K_k H) P_k^- (I - K_k H)^T + K_k R K_k^T$

有：

$$\begin{aligned} P_k &= (I - K_k H) P_k^- (I - K_k H)^T + K_k R K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + K_k (H P_k^- H^T + R) K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + P_k^- (H)^T (HP_k^- H^T + R)^{-1} (HP_k^- H^T + R) K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + P_k^- (H)^T K_k^T \\ &= P_k^- - K_k H P_k^- + 0 \\ &= (I - K_k H) P_k^- \end{aligned}$$

### 3.4 滤波器公式总结

此时对卡尔曼滤波器五个方程进行总结，两个状态方程  $\hat{x}_k, \hat{x}_k^-$  可由离散线性状态观测器的系统框图来表示。而先验估计协方差，后验估计协方差，卡尔曼增益分别给出了推导。其中卡尔曼增益为保证后验估计的方差取到最小值时的  $K_k$  的解。五个方程如下：

时间更新方程：

$$\begin{aligned} \hat{x}_k^- &= A \hat{x}_{k-1} + B u_{k-1} \\ P_k^- &= A P_{k-1} A^T + Q \end{aligned}$$

测量更新方程：

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k (z_k - H \hat{x}_k^-) \\ P_k &= (I - K_k H) P_k^- \end{aligned}$$

## 4. 卡尔曼滤波器应用示例

### 4.1 惯性测量单元(IMU)卡尔曼滤波

此例主要参照 Lauszus 的博客文章 A practical approach to Kalman filter and how to implement it。主要内容为使用六轴 IMU(加速度与陀螺仪)提供的加速度数据和角速度数据进行 IMU 欧拉角的估计。

#### 4.1.1 卡尔曼滤波主要公式回顾

首先我们对 IMU 的状态方程进行建模。在这之前请先回顾第一章中对于卡尔曼滤

波的简介。主要包括如下方程：

状态变量方程：  $x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$

测量方程：  $z_k = Hx_k + v_k$

过程噪声  $w_k$  和观测噪声  $v_k$ ：

$$p(w) \sim N(0, Q)$$

$$p(v) \sim N(0, R)$$

先验估计协方差：  $P_k^- = E(e_k^- e_k^{-T}) = E[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T]$

后验估计协方差：  $P_k = E(e_k e_k^T) = E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$

时间更新方程：

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$$

$$P_k^- = AP_{k-1}A^T + Q$$

测量更新方程：

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1}$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-)$$

$$P_k = (I - K_k H) P_k^-$$

#### 4.1.2 IMU 状态方程与测量方程建模

讨论状态方程

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}。$$

首先对其中的状态变量  $x_k$  进行定义：

$$x_k = \begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_k$$

其中  $\theta$  代表角度（加速度计测得，具体折算参见附录 1）， $\dot{\theta}_b$  陀螺仪测出的角速度的偏差，这个偏差说明了陀螺仪产生的漂移的程度。

系统矩阵  $A$  定义为：

$$A = \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix}$$

输入  $u_{k-1}$  在这里是陀螺仪得到的角速度(单位 $^\circ/\text{s}$ )，定义为：

$$u_{k-1} = \dot{\theta}_{k-1}$$

输入系数矩阵  $B$  定义为:

$$B = \begin{bmatrix} \Delta t \\ 0 \end{bmatrix}$$

$B$  的定义很容易理解, 因为当你把角速度  $\dot{\theta}$  与时间间隔  $\Delta t$  相乘时, 就能得到  $\Delta t$  时间内转动的角度。因为我们无法直接计算陀螺仪漂移与角速度的关系, 我们把矩阵第二项设为零。

$w_k$  遵循高斯分布的规律, 均值为 0, 方差为  $Q_k$ 。

$$w_k \sim N(0, Q_k)$$

$Q_k$  作为过程噪声协方差矩阵, 用来模拟加速度计测得的角度和加速度偏差的噪声。

$$Q_k = \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t$$

这也很容易理解, 过程噪声会随着时间的增加而放大, 因为它取决于自身上一个状态的量, 就像陀螺仪会随着时间产生漂移一样。注意, 如果你把某个噪声的值设的越大, 代表你对这个状态变量不够信赖, 即这个状态值的估计存在更多的噪声。举个例子, 如果估计的角度开始漂移, 说明偏差过大不值得信赖, 即你应该增加  $Q_{\dot{\theta}_b}$  的值了。反之, 若估计的角度趋于平缓, 则角度的值值得信赖, 则应降低  $Q_\theta$ 。

Lauszus 在他的博客中给出了两个  $Q$  的推荐值分别为  $Q_\theta = 0.001$ ,  $Q_{\dot{\theta}_b} = 0.003$ 。他指出这两个  $Q$  对大多数 IMU 都适用。

那么此处, 可以将状态变量方程总结为:

$$\begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_k = \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_{k-1} + \begin{bmatrix} \Delta t \\ 0 \end{bmatrix} \dot{\theta}_{k-1} + w_{k-1}$$

$$w_{k-1} \sim N\left(0, \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t\right)$$

讨论测量方程

$$z_k = Hx_k + v_k$$

矩阵  $H$  是观测矩阵, 用来描述测量的变量与状态变量之间的关系。因为最终测量得到的值是 IMU 的角度值, 故  $H$  可定义为:

$$H = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

噪声  $v_k$  遵循高斯分布:

$$v_k \sim N(0, R)$$

因为  $v_k$  实际上是一个  $1 \times 1$  的向量，即一个常数值，故  $R$  的值就是  $v_k$  的方差。此处我们假设测量噪声与时间无关，是一个常量：

$$\text{var}(v_k) = \text{var}(v)$$

注意，如果方差  $\text{var}(v)$  设置得太高，则滤波器响应会变得非常的慢，对新的测量值信赖的更少，但是如果方差太小，滤波器可能产生超调，并且伴随一定的噪声，因为我们过于信赖加速度计了。

Lauszus 在他的博客中给出了  $R$  的推荐值为  $R = 0.03$ ，他指出这个数值对于大多数 IMU 都适用。

所以有

$$\begin{aligned} z_k &= Hx_k + v_k \\ &= [1 \quad 0] \begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_k + v_k \\ &= \theta_k + v_k \end{aligned}$$

#### 4.1.3 IMU 卡尔曼滤波器方程

首先是时间更新方程（预测方程），即利用  $k-1$  时候的变量去得到  $k$  时刻的 **先验估计**：

$$\begin{aligned} \hat{x}_k^- &= A\hat{x}_{k-1} + Bu_{k-1} \\ &= \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_{k-1} + \begin{bmatrix} \Delta t \\ 0 \end{bmatrix} \dot{\theta}_{k-1} \end{aligned}$$

先验估计协方差矩阵  $P_k^-$  为：

$$\begin{aligned} P_k^- &= AP_{k-1}A^T + Q_k \\ &= \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix} P_{k-1} \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix}^T + \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t \end{aligned}$$

由此看到协方差矩阵  $P_k^-$  是一个  $2 \times 2$  的矩阵：

$$P_k^- = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

注意，如果你不知道  $P$  矩阵在一开始的状态，你可以把它设置为：

$$P_0 = \begin{bmatrix} L & 0 \\ 0 & L \end{bmatrix}$$

其中  $L$  代表一个很大的数字，在作者的平衡车博客中，他知道起始的角度并且通过标定知道了陀螺仪的偏差，所以他假设姿态一开始是知道的，即  $x_0$  与它的估计值  $\hat{x}_0$  相等，所

以他把一开始的误差矩阵  $P_0$  设为 0。

然后是状态更新方程，即利用  $k-1$  时刻以及  $k$  时刻的变量去得到  $k$  时刻的**后验估计**。

对卡尔曼增益  $K_k$  来说

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} \left( \begin{bmatrix} 1 & 0 \end{bmatrix} P_k^- \begin{bmatrix} 1 \\ 0 \end{bmatrix} + R \right)^{-1} \end{aligned}$$

可得卡尔曼增益  $K_k$  是一个  $2 \times 1$  的二维列向量。

对后验估计  $\hat{x}_k$ ，即对先验估计增加输出反馈后的估计来说，

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - H \hat{x}_k^-)$$

$\hat{x}_k^-$ ,  $K_k$ ,  $H$ ,  $z_k$  均可由之前给出的式子得到并且代入，此处公式太长，不再写出。

同理后验估计协方差矩阵亦可给出结果：

$$P_k = (I - K_k H) P_k^-$$

#### 4.1.4 算法实现

##### Step 1 先验估计

$$\begin{aligned} \hat{x}_k^- &= A \hat{x}_{k-1} + B u_{k-1} \\ &= \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \theta \\ \dot{\theta}_b \end{bmatrix}_{k-1} + \begin{bmatrix} \Delta t \\ 0 \end{bmatrix} \dot{\theta}_{k-1} \\ &= \begin{bmatrix} \theta_{k-1} + \Delta t (\dot{\theta} - \dot{\theta}_b)_{k-1} \\ (\dot{\theta}_b)_{k-1} \end{bmatrix} \end{aligned}$$

此处可以看到，状态变量  $x_k$  的先验估计等于  $k-1$  时刻的后验估计，以及纠偏后角速度

$(\dot{\theta} - \dot{\theta}_b)_{k-1}$  与时间  $\Delta t$  的乘积。因为我们无法直接测量偏差，故先验偏差的估计保持不变。

Step 2 先验估计误差

$$\begin{aligned}
 P_k^- &= AP_{k-1}A^T + Q_k \\
 &= \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}_{k-1} \begin{bmatrix} 1 & 0 \\ -\Delta t & 1 \end{bmatrix} + \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t \\
 &= \begin{bmatrix} P_{11} - \Delta t P_{21} & P_{12} - \Delta t P_{22} \\ P_{21} & P_{22} \end{bmatrix}_{k-1} \begin{bmatrix} 1 & 0 \\ -\Delta t & 1 \end{bmatrix} + \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t \\
 &= \begin{bmatrix} P_{11} - \Delta t P_{21} - \Delta t (P_{12} - \Delta t P_{22}) & P_{12} - \Delta t P_{22} \\ P_{21} - \Delta t P_{22} & P_{22} \end{bmatrix}_{k-1} + \begin{bmatrix} Q_\theta & 0 \\ 0 & Q_{\dot{\theta}_b} \end{bmatrix} \Delta t \\
 &= \begin{bmatrix} P_{11} + \Delta t (\Delta t P_{22} + Q_\theta - P_{21} - P_{12}) & P_{12} - \Delta t P_{22} \\ P_{21} - \Delta t P_{22} & P_{22} + Q_{\dot{\theta}_b} \Delta t \end{bmatrix}_{k-1}
 \end{aligned}$$

再将 Lauszus 在他的博客中给出了方差的推荐值：  $Q_\theta = 0.001$ ，  $Q_{\dot{\theta}_b} = 0.003$  代入。

Step 3 卡尔曼增益

首先求取括号里的一项：

$$\begin{aligned}
 \text{设 } P_k^- &= \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \\
 HP_k^- H^T + R &= \begin{bmatrix} 1 & 0 \end{bmatrix} P_k^- \begin{bmatrix} 1 \\ 0 \end{bmatrix} + R \\
 &= P_{11} + \text{var}(v)
 \end{aligned}$$

然后求解卡尔曼增益：

$$\begin{aligned}
 K_k &= P_k^- H^T (HP_k^- H^T + R)^{-1} \\
 \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}_k &= \frac{\begin{bmatrix} P_{11} \\ P_{21} \end{bmatrix}}{P_{11} + \text{var}(v)}
 \end{aligned}$$

再将 Lauszus 在他的博客中给出方差的推荐值：  $R = 0.03$  代入。

Step 4 后验估计

首先求输出反馈中的输出的差值：

$$\begin{aligned}
 \hat{y}_k &= z_k - H\hat{x}_k^- \\
 &= \theta_k + v_k - \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} \theta_{k-1} + \Delta t (\dot{\theta} - \dot{\theta}_b)_{k-1} \\ (\dot{\theta}_b)_{k-1} \end{bmatrix} \\
 &= \theta_k + v_k - \left( \theta_{k-1} + \Delta t (\dot{\theta} - \dot{\theta}_b)_{k-1} \right)
 \end{aligned}$$

然后对后验估计进行求取:

$$\begin{aligned}\hat{x}_k &= \hat{x}_k^- + K_k (z_k - H\hat{x}_k^-) \\ &= \hat{x}_k^- + K_k \hat{y}_k \\ &= \begin{bmatrix} \theta_{k-1} + \Delta t (\dot{\theta} - \dot{\theta}_b)_{k-1} \\ (\dot{\theta}_b)_{k-1} \end{bmatrix} + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}_k \hat{y}_k\end{aligned}$$

Step 5 后验协方差

$$\text{设 } P_k^- = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix},$$

$$\begin{aligned}P_k &= (I - K_k H) P_k^- \\ &= \left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}_k \begin{bmatrix} 1 & 0 \end{bmatrix} \right) P_k^- \\ &= P_k^- - \begin{bmatrix} K_1 & 0 \\ K_2 & 0 \end{bmatrix}_k P_k^- \\ &= \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} - \begin{bmatrix} K_1 & 0 \\ K_2 & 0 \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \\ &= \begin{bmatrix} P_{11} - K_1 P_{11} & P_{12} - K_1 P_{12} \\ P_{21} - K_2 P_{11} & P_{22} - K_2 P_{12} \end{bmatrix}\end{aligned}$$

#### 4.1.5 MATLAB 程序

见附件。



## 5. 扩展卡尔曼滤波器

### 5.1 扩展卡尔曼滤波介绍

卡尔曼滤波器用于估计离散线性时间过程的状态变量  $x$ ，对于离散非线性时间状态过程，可以使用扩展卡尔曼滤波。

首先，为了表示这个非线性离散过程，给出非线性状态下的两个方程：

**状态变量方程：**  $x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$

**测量方程：**  $z_k = h(x_k, v_k)$

其中  $w_k, v_k$  分别表示过程噪声(process noise)和观测噪声(measurement noise)，虽然  $w_k, v_k$  仍服从高斯分布，如下：

$$p(w) \sim N(0, Q)$$

$$p(v) \sim N(0, R)$$

但是由于非线性函数的关系，实际噪声已经不再遵循高斯分布。

定义  $\hat{x}_k^- \in R^n$  ( $-$  代表先验， $\hat{\cdot}$  代表估计)为在已知第  $k$  步以前（不包括  $k$  步）状态情况

下第  $k$  步的估计，即**先验状态估计**。 $\hat{x}_k \in R^n$  为在已知第  $k$  步及以前（包括  $k$  步）状态情

况下第  $k$  步的估计，即**后验状态估计**。

因为假定估计过程不存在过程噪声和观测噪声，有

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

**先验测量值：**  $\hat{z}_k^- = h(\hat{x}_k^-, 0)$

此处的图解可参照 5.3。

扩展卡尔曼滤波器中，为了对非线性方程进行线性化，通过一阶泰勒展开，我们就得到了非线性卡尔曼滤波器线性化之后的**状态方程**和**测量方程**。具体推导过程见 5.2。

**状态变量方程：**  $x_k \approx \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}$

**测量方程：**  $z_k \approx \hat{z}_k^- + H(x_k - \hat{x}_k^-) + Vv_k$

记  $\varepsilon_k = Ww_{k-1}$ ， $\eta_k = Vv_k$

则  $\varepsilon_k$ ， $\eta_k$  分别代表具有零均值和协方差矩阵  $WQ_kW^T$ ， $VR_kV^T$  的过程噪声和测量噪声，可写作如下：

$$p(\varepsilon_k) = p(Ww_{k-1}) \sim N(0, WQ_kW^T)$$

$$p(\eta_k) = p(Vv_k) \sim N(0, VR_kV^T)$$

定义  $\hat{x}_k^- \in R^n$  ( $-$  代表先验,  $\hat{\cdot}$  代表估计) 为在已知第  $k$  步以前状态情况下第  $k$  步的 **先验状态估计**。 $\hat{x}_k \in R^n$  为在已知第  $k$  步以前状态情况下第  $k$  步的 **后验状态估计**。因此得到 **先验估计误差** 和 **后验估计误差**:

$$\begin{aligned} e_k^- &\equiv x_k - \hat{x}_k^- \\ e_k &\equiv x_k - \hat{x}_k \end{aligned}$$

其协方差为:

$$\text{先验估计协方差: } P_k^- = E(e_k^- e_k^{-T}) = E\left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\right], \text{ E 代表取期望}$$

$$\text{后验估计协方差: } P_k = E(e_k e_k^T) = E\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right]$$

此处先给出扩展卡尔曼滤波器的时间更新方程和测量更新方程, 推导过程详见第六章:

**时间更新方程 (预测方程):**

$$\begin{aligned} \hat{x}_k^- &= f(\hat{x}_{k-1}, u_{k-1}, 0) \\ P_k^- &= A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T \end{aligned}$$

**测量更新方程 (修正方程):**

$$\begin{aligned} K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0)) \\ P_k &= (I - K_k H_k) P_k^- \end{aligned}$$

此处应注意: 使用一阶泰勒展开来对非线性系统进行线性化的过程中, 一阶线性化的过程忽略了高阶项, 只有被线性化的那个点附近的线性化模型和真实的模型相近时, 才能达到好的线性化效果。远的话会导致误差过大, 那么这个时候卡尔曼滤波器的效果就不好, 甚至导致滤波器发散。比方说, 在对状态变量方程  $x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$  在点  $(\hat{x}_{k-1}, u_{k-1}, 0)$  线性化时, 若  $(x_{k-1}, u_{k-1}, w_{k-1})$  与  $(\hat{x}_{k-1}, u_{k-1}, 0)$  差距过大, 会导致滤波器发散。

## 5.2 非线性卡尔曼滤波线性化推导 (状态方程建立)

首先对 **状态变量方程**  $x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$  进行线性化

为对其进行线性化, 要对原来的多变量方程进行一阶泰勒展开, 方便表示, 首先构造两个变量, 这样是原来式子中的三个值变成一个值:

$$\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}, \begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}$$

因此有

$$x_k = f(x_{k-1}, u_{k-1}, w_{k-1}) = f\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}\right), \quad \hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}, 0) = f\left(\begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right)$$

然后对  $f\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}\right)$  在点  $\begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}$  处做一阶泰勒展开, 有

$$\begin{aligned} f\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}\right) &= f\left(\begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right) + \frac{\partial f(\hat{x}_{k-1}, u_{k-1}, 0)_{m^*1}}{\partial (x_{k-1}, u_{k-1}, w_{k-1})_{m^*1}} \left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix} - \begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right) + \sigma\left(\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix} - \begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right)^2\right) \\ f\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}\right) &\approx f\left(\begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right) + \frac{\partial f(\hat{x}_{k-1}, u_{k-1}, 0)_{m^*1}}{\partial (x_{k-1}, u_{k-1}, w_{k-1})_{m^*1}} \begin{bmatrix} x_{k-1} - \hat{x}_{k-1} \\ 0 \\ w_{k-1} \end{bmatrix} \\ f\left(\begin{bmatrix} x_{k-1} \\ u_{k-1} \\ w_{k-1} \end{bmatrix}\right) &\approx f\left(\begin{bmatrix} \hat{x}_{k-1} \\ u_{k-1} \\ 0 \end{bmatrix}\right) + \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_{m_x}} & \frac{\partial f_1}{\partial u_1} & \dots & \frac{\partial f_1}{\partial u_{m_u}} & \frac{\partial f_1}{\partial w_1} & \dots & \frac{\partial f_1}{\partial w_{m_w}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{m-1}}{\partial x_1} & \dots & \frac{\partial f_{m-1}}{\partial x_{m_x}} & \frac{\partial f_{m-1}}{\partial u_1} & \dots & \frac{\partial f_{m-1}}{\partial u_{m_u}} & \frac{\partial f_{m-1}}{\partial w_1} & \dots & \frac{\partial f_{m-1}}{\partial w_{m_w}} \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_{m_x}} & \frac{\partial f_m}{\partial u_1} & \dots & \frac{\partial f_m}{\partial u_{m_u}} & \frac{\partial f_m}{\partial w_1} & \dots & \frac{\partial f_m}{\partial w_{m_w}} \end{bmatrix} \begin{bmatrix} x_{k-1} - \hat{x}_{k-1} \\ 0 \\ w_{k-1} \end{bmatrix} \\ x_k &\approx \hat{x}_k^- + \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_{m_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{m-1}}{\partial x_1} & \dots & \frac{\partial f_{m-1}}{\partial x_{m_x}} \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_{m_x}} \end{bmatrix} (x_{k-1} - \hat{x}_{k-1}) + 0 + \begin{bmatrix} \frac{\partial f_1}{\partial w_1} & \dots & \frac{\partial f_1}{\partial w_{m_w}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{m-1}}{\partial w_1} & \dots & \frac{\partial f_{m-1}}{\partial w_{m_w}} \\ \frac{\partial f_m}{\partial w_1} & \dots & \frac{\partial f_m}{\partial w_{m_w}} \end{bmatrix} w_{k-1} \\ x_k &\approx \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1} \end{aligned}$$

其中:

$$A = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_{m_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{m-1}}{\partial x_1} & \dots & \frac{\partial f_{m-1}}{\partial x_{m_x}} \\ \frac{\partial f_m}{\partial x_1} & \dots & \frac{\partial f_m}{\partial x_{m_x}} \end{bmatrix} = \frac{\partial f(\hat{x}_{k-1}, u_{k-1}, 0)}{\partial x}, \quad W = \begin{bmatrix} \frac{\partial f_1}{\partial w_1} & \dots & \frac{\partial f_1}{\partial w_{m_w}} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_{m-1}}{\partial w_1} & \dots & \frac{\partial f_{m-1}}{\partial w_{m_w}} \\ \frac{\partial f_m}{\partial w_1} & \dots & \frac{\partial f_m}{\partial w_{m_w}} \end{bmatrix} = \frac{\partial f(\hat{x}_{k-1}, u_{k-1}, 0)}{\partial w}$$

然后对测量方程  $z_k = h(x_k, v_k)$  进行线性化，

同理，对  $h\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}\right)$  在点  $\begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}$  处做一阶泰勒展开，有

$$h\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}\right) = h\left(\begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}\right) + \frac{\partial h(\hat{x}_k^-, 0)}{\partial (x_k, v_k)_{n*1}} \left(\begin{bmatrix} x_k \\ v_k \end{bmatrix} - \begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}\right) + \sigma\left(\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix} - \begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}\right)^2\right)$$

$$h\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}\right) \approx h\left(\begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}\right) + \frac{\partial h(\hat{x}_k^-, 0)}{\partial (x_k, v_k)_{n*1}} \begin{bmatrix} x_k - \hat{x}_k^- \\ v_k \end{bmatrix}$$

$$h\left(\begin{bmatrix} x_k \\ v_k \end{bmatrix}\right) \approx h\left(\begin{bmatrix} \hat{x}_k^- \\ 0 \end{bmatrix}\right) + \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_{n_x}} & \frac{\partial h_1}{\partial v_1} & \dots & \frac{\partial h_1}{\partial v_{n_v}} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial x_1} & \dots & \frac{\partial h_{n-1}}{\partial x_{n_x}} & \frac{\partial h_{n-1}}{\partial v_1} & \dots & \frac{\partial h_{n-1}}{\partial v_{n_v}} \\ \frac{\partial h_n}{\partial x_1} & \dots & \frac{\partial h_n}{\partial x_{n_x}} & \frac{\partial h_n}{\partial v_1} & \dots & \frac{\partial h_n}{\partial v_{n_v}} \end{bmatrix} \begin{bmatrix} x_k - \hat{x}_k^- \\ v_k \end{bmatrix}$$

$$z_k \approx \hat{z}_k^- + \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_{n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial x_1} & \dots & \frac{\partial h_{n-1}}{\partial x_{n_x}} \\ \frac{\partial h_n}{\partial x_1} & \dots & \frac{\partial h_n}{\partial x_{n_x}} \end{bmatrix} (x_k - \hat{x}_k^-) + \begin{bmatrix} \frac{\partial h_1}{\partial v_1} & \dots & \frac{\partial h_1}{\partial v_{n_v}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial v_1} & \dots & \frac{\partial h_{n-1}}{\partial v_{n_v}} \\ \frac{\partial h_n}{\partial v_1} & \dots & \frac{\partial h_n}{\partial v_{n_v}} \end{bmatrix} v_k$$

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

其中：

$$H = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_{n_x}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial x_1} & \dots & \frac{\partial h_{n-1}}{\partial x_{n_x}} \\ \frac{\partial h_n}{\partial x_1} & \dots & \frac{\partial h_n}{\partial x_{n_x}} \end{bmatrix} = \frac{\partial h(\hat{x}_k^-, 0)}{\partial x}, \quad V = \begin{bmatrix} \frac{\partial h_1}{\partial v_1} & \dots & \frac{\partial h_1}{\partial v_{n_v}} \\ \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial v_1} & \dots & \frac{\partial h_{n-1}}{\partial v_{n_v}} \\ \frac{\partial h_n}{\partial v_1} & \dots & \frac{\partial h_n}{\partial v_{n_v}} \end{bmatrix} = \frac{\partial h(\hat{x}_k^-, 0)}{\partial v}$$

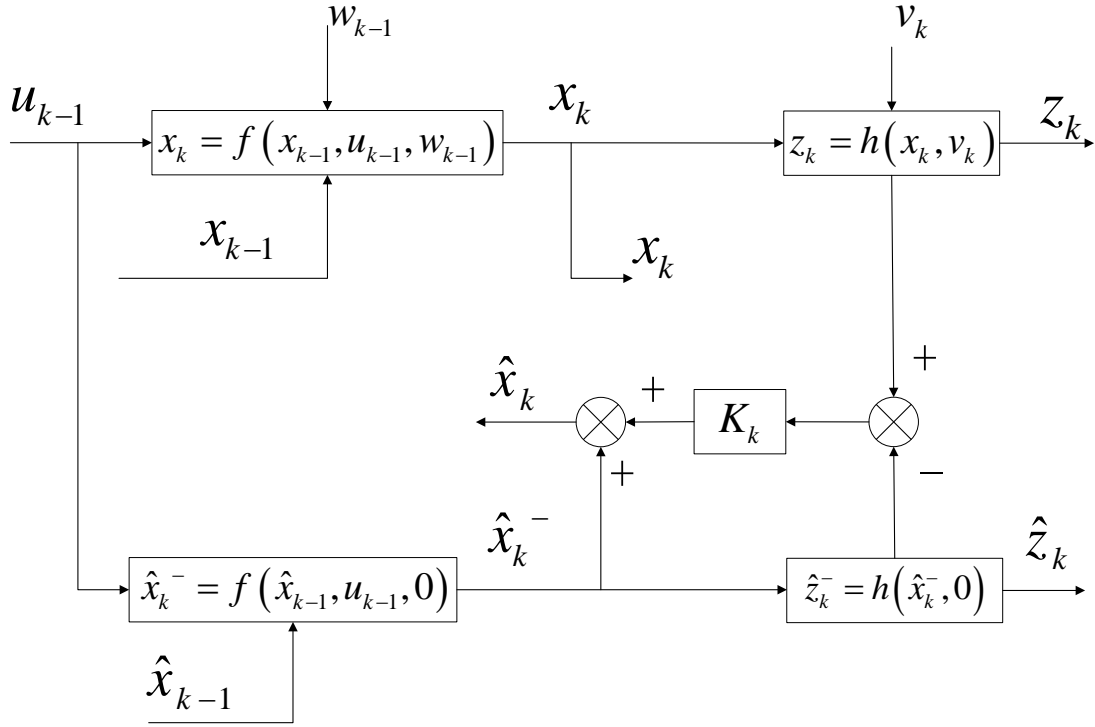
那么这里我们就得到了非线性卡尔曼滤波器线性化之后的状态方程和测量方程。

**状态方程：**  $x_k \approx \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}$

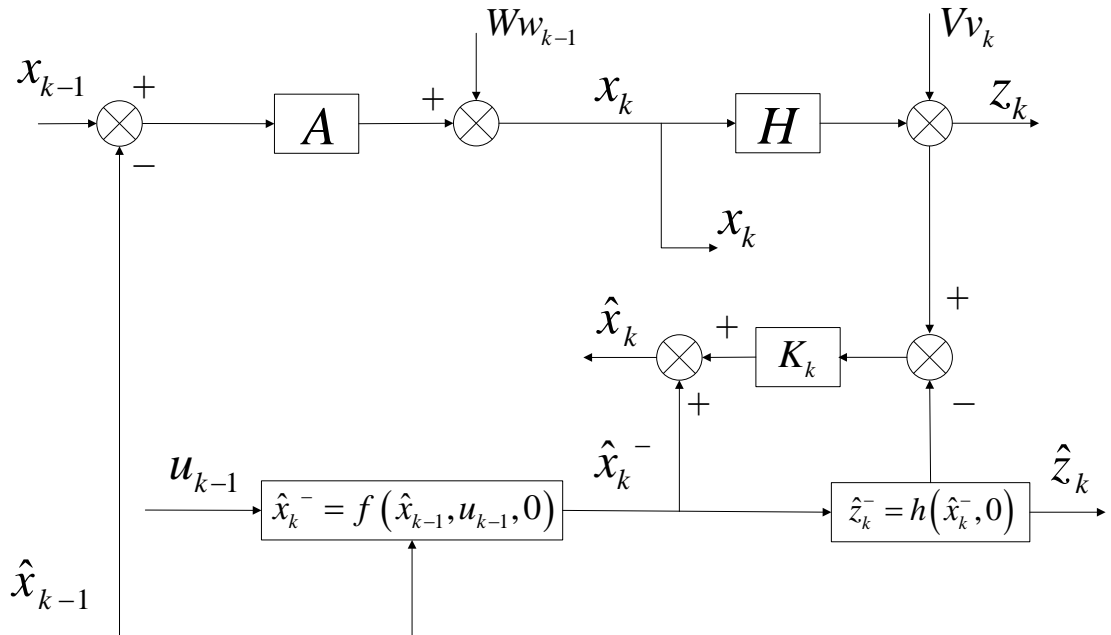
**测量方程：**  $z_k \approx \hat{z}_k^- + H(x_k - \hat{x}_k^-) + Vv_k$

### 5.3 扩展卡尔曼滤波器系统框图

非线性系统框图如下：



线性化后系统框图如下：



## 5.4 扩展卡尔曼滤波器与卡尔曼滤波器对比

这里给出扩展卡尔曼滤波和卡尔曼滤波器对比图：

	卡尔曼滤波	扩展卡尔曼滤波
是否线性	线性	非线性
是否连续	离散	离散
系统框图	见 2.3	见 5.3
状态变量方程	$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1}$	$x_k = f(x_{k-1}, u_{k-1}, w_{k-1})$ 线性化后： $x_k \approx \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}$
测量方程	$z_k = Hx_k + v_k$	$z_k = h(x_k, v_k)$ 线性化后： $z_k \approx \hat{z}_k^- + H(x_k - \hat{x}_k^-) + Vv_k$
噪声	$p(w) \sim N(0, Q)$ $p(v) \sim N(0, R)$	$p(\varepsilon_k) = p(Ww_{k-1}) \sim N(0, WQ_kW^T)$ $p(\eta_k) = p(Vv_k) \sim N(0, VR_kV^T)$
预测方程	$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1}$ $P_k^- = AP_{k-1}A^T + Q$	$\hat{x}_k^- = f(\hat{x}_{k-1}, u_{k-1}, 0)$ $P_k^- = A_kP_{k-1}A_k^T + W_kQ_{k-1}W_k^T$
修正方程	$K_k = P_k^-H^T(HP_k^-H^T + R)^{-1}$ $\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-)$ $P_k = (I - K_kH)P_k^-$	$K_k = P_k^-H_k^T(H_kP_k^-H_k^T + V_kR_kV_k^T)^{-1}$ $\hat{x}_k = \hat{x}_k^- + K_k(z_k - h(\hat{x}_k^-, 0))$ $P_k = (I - K_kH_k)P_k^-$
收敛性	未定	线性化不到位，容易发散

## 6. 扩展卡尔曼滤波器参数推导

### 6.1 先验估计协方差

由上得到先验估计协方差表达式为：

$$P_k^- = E(e_k^- e_k^{-T}) = E\left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\right]$$

把其中  $x_k, \hat{x}_k^-$  的表达式代入：

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

代入则有：

$$\begin{aligned} x_k - \hat{x}_k^- &= \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1} - \hat{x}_k^- \\ &= A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1} \end{aligned}$$

则：

$$\begin{aligned} P_k^- &= E(e_k^- e_k^{-T}) \\ &= E\left[(x_k - \hat{x}_k^-)(x_k - \hat{x}_k^-)^T\right] \\ &= E\left[A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}\right]\left[A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1}\right]^T \\ &= AE\left[(x_{k-1} - \hat{x}_{k-1})(x_{k-1} - \hat{x}_{k-1})^T\right]A^T + E[Ww_{k-1}w_{k-1}^T W^T] \\ &= AP_{k-1}A^T + WQW^T \end{aligned}$$

则得到时间更新方程中的第二个式子。其中  $A, W, Q$  均会随着时间而改变。

### 6.2 后验估计协方差

由上得到后验估计协方差表达式为：

$$P_k = E(e_k e_k^T) = E\left[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T\right]$$

把其中  $x_k, \hat{x}_k$  的表达式代入：

$$\begin{aligned} x_k &\approx \hat{x}_k^- + A(x_{k-1} - \hat{x}_{k-1}) + Ww_{k-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k(z_k - \hat{z}_k^-) \end{aligned}$$

（上式计算过程中把最开始的测量方程  $z_k \approx \hat{z}_k^- + H(x_k - \hat{x}_k^-) + Vv_k$  代入）

代入则有：

$$\begin{aligned} x_k - \hat{x}_k &= x_k - \left(\hat{x}_k^- + K_k(z_k - \hat{z}_k^-)\right) \\ &= x_k - \hat{x}_k^- - K_k\left(\hat{z}_k^- + H(x_k - \hat{x}_k^-) + Vv_k - \hat{z}_k^-\right) \\ &= x_k - \hat{x}_k^- - K_k\left(H(x_k - \hat{x}_k^-) + Vv_k\right) \\ &= (I - K_k H)(x_k - \hat{x}_k^-) - K_k Vv_k \end{aligned}$$

即：

$$x_k - \hat{x}_k = (I - K_k H)(x_k - \hat{x}_k^-) - K_k V v_k$$

故有

$$\begin{aligned} P_k &= E(e_k e_k^T) \\ &= E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \\ &= E\left[(I - K_k H)(x_{k-1} - \hat{x}_{k-1}^-)(x_{k-1} - \hat{x}_{k-1}^-)^T (I - K_k H)^T + K_k V v_k v_k^T V^T K_k^T\right] \\ &= (I - K_k H) P_k^- (I - K_k H)^T + K_k V R V^T K_k^T \end{aligned}$$

可以看到，此处的后验估计误差与前面给出的卡尔曼滤波测量更新方程仍有差距，待求出卡尔曼滤波器增益后代入可化简，请先看 6.3。

化简后得到：

$$P_k = (I - K_k H) P_k^-$$

### 6.3 扩展卡尔曼增益

滤波器增益  $K_k$  含义较为抽象，结合上面状态观测器的理论来说，在此处的意义在于通过输出反馈引入矩阵  $K_k$ ，使得通过调整  $K_k$  可以使  $x_k - \hat{x}_k$  的方差渐近衰减为最小，即观测值趋向于实际值。

使后验估计  $x_k - \hat{x}_k$  的方差  $P_k$  达到最小，也是使协方差矩阵  $E[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$  的对角线元素的和（也就是线性代数中的迹）最小化的过程。此处补充说明，协方差的 diagonal 元素即为各个对应元素的方差。

故要求  $K_k$ ，使得  $P_k$  的迹取到一个最小值。即  $P_k$  关于  $K_k$  的导数为零。

首先将  $P_k$  展开

$$\begin{aligned} P_k &= (I - K_k H) P_k^- (I - K_k H)^T + K_k V R V^T K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + K_k H P_k^- (K_k H)^T + K_k V R V^T K_k^T \\ &= P_k^- - K_k H P_k^- - P_k^- (K_k H)^T + K_k (H P_k^- H^T + V R V^T) K_k^T \end{aligned}$$

此时得到的为协方差，先对其取对角线元素之和，然后对其迹关于  $K_k$  求导。取迹

$tr(P_k)$  有：

$$tr(P_k) = tr(P_k^-) - 2tr(K_k H P_k^-) + tr(K_k (H P_k^- H^T + V R V^T) K_k^T)$$

\*此处说明，上式中直接  $tr(K_k H P_k^-) = tr(P_k^- (K_k H)^T)$ ，因为协方差矩阵  $P_k^-$  是对称阵，且根据矩阵的迹等于其转置矩阵的迹，故有：

$$tr(P_k^- (K_k H)^T) = tr(P_k^{-T} (K_k H)^T) = tr((K_k H P_k^-)^T) = tr(K_k H P_k^-)$$



对  $tr(P_k)$  关于  $K_k$  求导：

$$\frac{\partial T(P_k)}{\partial K_k} = \frac{\partial \left( tr(P_k^-) - 2tr(K_k HP_k^-) + tr(K_k (HP_k^- H^T + VRV^T) K_k^T) \right)}{\partial K_k}$$

当导数等于零时，即  $\frac{\partial tr(P_k)}{\partial K_k} = 0$ ，有：

$$\begin{aligned} & \frac{\partial \left( tr(P_k^-) - 2tr(K_k HP_k^-) + tr(K_k (HP_k^- H^T + VRV^T) K_k^T) \right)}{\partial K_k} \\ &= \frac{\partial \left( tr(P_k^-) \right)}{\partial K_k} - 2 \frac{\partial \left( tr(K_k HP_k^-) \right)}{\partial K_k} + \frac{\partial \left( tr(K_k (HP_k^- H^T + VRV^T) K_k^T) \right)}{\partial K_k} \\ &= 0 - 2 \left( HP_k^- \right)^T + K_k \left[ \left( HP_k^- H^T + VRV^T \right) + \left( HP_k^- H^T + VRV^T \right)^T \right] \\ &= 0 - 2 \left( HP_k^- \right)^T + 2K_k \left[ \left( HP_k^- H^T + VRV^T \right) \right] \\ &= 0 \end{aligned}$$

即：

$$\begin{aligned} 2K_k \left[ \left( HP_k^- H^T + VRV^T \right) \right] &= 2 \left( HP_k^- \right)^T \\ K_k &= \left( P_k^- \right)^T (H)^T \left( HP_k^- H^T + VRV^T \right)^{-1} \\ K_k &= P_k^- (H)^T \left( HP_k^- H^T + VRV^T \right)^{-1} \end{aligned}$$

此时将  $K_k = P_k^- (H)^T \left( HP_k^- H^T + VRV^T \right)^{-1}$

代入后验估计误差  $P_k = (I - K_k H) P_k^- (I - K_k H)^T + K_k VRV^T K_k^T$

有：

$$\begin{aligned} P_k &= (I - K_k H) P_k^- (I - K_k H)^T + K_k VRV^T K_k^T \\ &= P_k^- - K_k HP_k^- - P_k^- (K_k H)^T + K_k (HP_k^- H^T + VRV^T) K_k^T \\ &= P_k^- - K_k HP_k^- - P_k^- (K_k H)^T + P_k^- (H)^T \left( HP_k^- H^T + VRV^T \right)^{-1} \left( HP_k^- H^T + VRV^T \right) K_k^T \\ &= P_k^- - K_k HP_k^- - P_k^- (K_k H)^T + P_k^- (H)^T K_k^T \\ &= P_k^- - K_k HP_k^- + 0 \\ &= (I - K_k H) P_k^- \end{aligned}$$

## 6.4 滤波器公式总结

此时对扩展卡尔曼滤波器五个方程进行总结，两个状态方程  $\hat{x}_k, \hat{x}_k^-$  可由系统框图来表示。而先验估计协方差，后验估计协方差，扩展卡尔曼增益分别给出了推导。其中扩展卡

尔曼增益为保证后验估计的方差取到最小值时的  $K_k$  的解。五个方程如下：

时间更新方程（预测方程）：

$$\begin{aligned}\hat{x}_k^- &= f(\hat{x}_{k-1}, u_{k-1}, 0) \\ P_k^- &= A_k P_{k-1} A_k^T + W_k Q_{k-1} W_k^T\end{aligned}$$

测量更新方程（修正方程）：

$$\begin{aligned}K_k &= P_k^- H_k^T (H_k P_k^- H_k^T + V_k R_k V_k^T)^{-1} \\ \hat{x}_k &= \hat{x}_k^- + K_k (z_k - h(\hat{x}_k^-, 0)) \\ P_k &= (I - K_k H_k) P_k^-\end{aligned}$$

## 7. 扩展卡尔曼滤波器应用示例

### 7.1 惯性测量单元(IMU)扩展卡尔曼滤波

#### 7.1.1 四元数简介及相关公式

#### 7.1.2 扩展卡尔曼滤波主要公式回顾

## 8. 参考文献

- [1] 《现代控制理论》第六版，刘豹
- [2] An-Introduction-to-the-Kalman-Filter(中文).pdf
- [3] <https://blog.csdn.net/heyijia0327/article/details/17487467>
- [4] [https://en.wikipedia.org/wiki/Matrix\\_calculus#Scalar-by-matrix](https://en.wikipedia.org/wiki/Matrix_calculus#Scalar-by-matrix)
- [5] A practical approach to Kalman filter and how to implement it, Lauszus

## 9. 附录

### 9.1 附录一：加速度计计算姿态角推导

系统选取世界（world）坐标系（ $X_w, Y_w, Z_w$ ）和载体（target）坐标系（ $X_t, Y_t, Z_t$ ）。已知世界坐标系到载体坐标系旋转矩阵为：

$$\begin{bmatrix} X_t \\ Y_t \\ Z_t \end{bmatrix} = R_w^t \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$$

其中变换矩阵  $R_w^t$  为向左乘，即绕变换前的初始静坐标系旋转。参见机器人学“左静右动”。

变换矩阵  $R_w^t$  一般由  $R_x(\phi)$ ， $R_y(\theta)$ ， $R_z(\psi)$  三种值自由组合而成，如（z-y-x, z-x-y, y-z-x, x-z-y, y-x-z, x-y-z）。有：

$$R_x(\phi) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix}$$

$$R_y(\theta) = \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix}$$

$$R_z(\psi) = \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

因为向左乘，代表绕初始的静坐标系的三轴转动，如 x-y-z，则代表先绕初始的 x 轴转动  $\phi$ ，再绕初始的 y 轴转动  $\theta$ ，再绕初始的 z 轴转动  $\psi$ ，但应注意，此时的符号记为

$$R_{zyx} = R_z(\psi)R_y(\theta)R_x(\phi)。$$

第 4 章中的 MATLAB 程序选用的是  $R_{xyz}$  的计算顺序，即先绕初始的 z 轴转动  $\psi$ ，再绕初始的 y 轴转动  $\theta$ ，再绕初始的 x 轴转动  $\phi$ 。

$$R_w^t = R_x(\phi)R_y(\theta)R_z(\psi)$$

若载体静止，物体在世界（world）坐标系中测得加速度为[0,0,g]，则通过变换后可知，在载体坐标系中测得加速度为：

$$\begin{aligned}
 \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} &= R_{xyz} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \\
 &= R_x(\phi) R_y(\theta) R_z(\psi) \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \\
 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\phi) & \sin(\phi) \\ 0 & -\sin(\phi) & \cos(\phi) \end{pmatrix} \begin{pmatrix} \cos(\theta) & 0 & -\sin(\theta) \\ 0 & 1 & 0 \\ \sin(\theta) & 0 & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\psi) & -\sin(\psi) & 0 \\ \sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \\
 &= \begin{pmatrix} \cos(\theta)\cos(\psi) & \cos(\theta)\sin(\psi) & -\sin(\theta) \\ \sin(\phi)\cos(\psi)\sin(\theta) - \cos(\phi)\sin(\psi) & \sin(\phi)\sin(\theta)\sin(\psi) + \cos(\phi)\cos(\psi) & \cos(\theta)\sin(\phi) \\ \cos(\phi)\cos(\psi)\sin(\theta) + \sin(\phi)\sin(\psi) & \cos(\phi)\sin(\theta)\sin(\psi) - \sin(\phi)\cos(\psi) & \cos(\theta)\cos(\phi) \end{pmatrix} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \\
 &= \begin{pmatrix} -\sin(\theta) \\ \cos(\theta)\sin(\phi) \\ \cos(\theta)\cos(\phi) \end{pmatrix} g
 \end{aligned}$$

所以求解得：

$$\begin{aligned}
 \tan \phi &= \left( \frac{a_y}{a_z} \right) \\
 \tan \theta &= \left( \frac{-a_x}{\sqrt{a_y^2 + a_z^2}} \right)
 \end{aligned}$$

可用来求绕 x, y 轴的转角, z 轴的角度无法直接求得。