

# 卡尔曼滤波增益综述报告

姓名:周峰 学号 1411082695

**摘要:** 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。本文介绍了卡尔曼滤波增益的由来，以及它在卡尔曼滤波理论中的作用，着重介绍了卡尔曼滤波增益的理论意义和它的物理意义。由卡尔曼滤波增益可以更深入的理解卡尔曼滤波，把它更好地应用于实际中。

**Abstract:** Kalman Filter is an efficient recursive filter, it can achieve the task that estimates the dynamic state of the system from a series of noise measurements. It widely be used in Radar, computer vision, include other engineering applications, is also a very important issue in control theory and control systems engineering. This paper introduces the origin of the Kalman filter gain, and it plays the important role in the Kalman filter theory, especially focuses on its the theoretical meaning and physical meaning about Kalman filter gain. We will get a deeper understanding of the Kalman filter, better applied in practice by learning of the Kalman filter gain.

**关键词:** 卡尔曼滤波 增益 误差

## 一、卡尔曼滤波器简介

### 1.1 卡尔曼滤波的由来

1960 年卡尔曼发表了用递归方法解决离散数据线性滤波问题的论文-《A New Approach to Linear Filtering and Prediction Problems》(线性滤波与预测问题的新方法), 在这篇文章里一种克服了维纳滤波缺点的新方法被提出来, 这就是我们今天称之为卡尔曼滤波的方法。卡尔曼滤波应用广泛且功能强大, 它可以估计信号的过去和当前状态甚至能估计将来的状态即使并不知道模型的确切性质。

其基本思想是以最小均方误差为最佳估计准则, 采用信号与噪声的状态空间模型利用前一时刻的估计值和当前时刻的观测值来更新对状态变量的估计, 求出当前时刻的估计值。算法根据建立的系统方程和观测方程对需要处理的信号做出满足最小均方误差的估计。

对于解决很大部分的问题, 它是最优, 效率最高甚至是最有用的。它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。

### 1.2 卡尔曼滤波原理

卡尔曼滤波器用于估计离散时间过程的状态变量  $x \in R^n$ 。这个离散时间过程由以下离散随机差分方程描述:

$$x_k = Ax_{k-1} + Bu_{k-1} + w_{k-1} \quad \text{公式 1.1}$$

定义观测变量  $z \in R^m$ , 的量测方程:

$$z_k = Hx_k + v_k \quad \text{公式 1.2}$$

随机信号  $w_k$  和  $v_k$  分别表示过程激励噪声和观测噪声。假设它们为相互独立, 正态分布的白色噪声:

$$p(w) \sim N(0, Q), \quad \text{公式 1.3}$$

$$p(v) \sim N(0, R). \quad \text{公式 1.4}$$

实际系统中, 过程激励噪声协方差矩阵  $Q$  和观测噪声协方差矩阵  $R$  可能会随每次迭代计算而变化。但在这儿我们假设它们是常数。

当控制函数  $u_{k-1}$  或过程激励噪声  $w_{k-1}$  为零时, 差分方程 1.1 中的  $n \times n$  阶增益矩阵  $A$  将上一时刻  $k-1$  的状态线性映射到当前时刻  $k$  的状态。实际中  $A$  可能随时间变化, 但在这儿假设为常数。 $n \times 1$  阶矩阵  $B$  代表可选的控制输入  $u \in R^1$  的增益。量测方程 1.2 中的  $m \times n$  阶矩阵  $H$  表示状态变量  $x_k$  对测量变量  $z_k$  的增益。实际中  $H$  可能随时间变化, 但在这儿假设为常数。

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。定义  $\hat{x}_k \in R^n$  为已知测量变量  $z_k$  时第  $k$  步的后验状态估计。由此定义先验估计误差和后验估计误差:

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

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

先验估计误差的协方差为:

$$P_k^- = E[e_k^- e_k^{-T}] \quad \text{公式 1.5}$$

后验估计的协方差:

$$P_k = E[e_k e_k^T] \quad \text{公式 1.6}$$

式 1.7 构造了卡尔曼滤波器的表达式: 先验估计  $\hat{x}_k^-$  和加权的测量变量  $z_k$  及其预测  $H\hat{x}_k^-$  之差的线性组合构成了后验状态估计  $\hat{x}_k$ 。

$$\hat{x}_k = \hat{x}_k^- + K(z_k - H\hat{x}_k^-) \quad \text{公式 1.7}$$

式 1.7 中测量变量及其预测之差  $z_k - H\hat{x}_k^-$  被称为测量过程的革新或残余。残余反映了预测值和实际值之间的不一致程度。残余为零表明二者完全吻合。式 1.7 中  $n \times m$  阶矩阵  $K$  叫做残余的增益或混合因数, 作用是使 1.6 式中的后验估计误差协方差最小。可以通过以下步骤计算  $K$ : 首先将 1.7 式代入  $e_k$  的定义式, 再将  $e_k$  代入 1.6 式中, 求得期望后, 将 1.6 式中的  $P_k$  对  $K$  求导。并使一阶导数为零从而解得  $K$  值。 $K$  的一种表示形式为:

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ &= \frac{P_k^- H^T}{H P_k^- H^T + R} \end{aligned} \quad \text{公式 1.8}$$

## 二、卡尔曼滤波增益

## 2.1 推导后验协方差矩阵 $\hat{P}_k$

按照定义，我们从误差协方差  $\hat{P}_k$  开始推导如下：

$$\hat{P}_k = \text{cov}(x_k - \hat{x}_k)$$

$$\text{代入 } \hat{x}_k = \hat{x}_{\bar{k}} + K_k(z_k - H\hat{x}_{\bar{k}})$$

$$\hat{P}_k = \text{cov}(x_k - (\hat{x}_{\bar{k}} + K_k(z_k - H\hat{x}_{\bar{k}})))$$

$$\text{再带入 } z_k = Hx_k + v_k$$

$$\hat{P}_k = \text{cov}((I - K_k H)(x_k - \hat{x}_{\bar{k}}) - K_k v_k)$$

整理测量误差向量，得：

$$\hat{P}_k = \text{cov}((I - K_k H)(x_k - \hat{x}_{\bar{k}}) - K_k v_k)$$

因为噪声项与其他项不相关，协方差=0，所以有：

$$\hat{P}_k = \text{cov}((I - K_k H)(x_k - \hat{x}_{\bar{k}})) - \text{cov}(K_k v_k)$$

利用协方差矩阵性质，提出常数矩阵，得：

$$\hat{P}_k = (I - K_k H) \text{cov}(x_k - \hat{x}_{\bar{k}}) (I - K_k H)^T - K_k \text{cov}(v_k) K_k^T$$

如果记  $\text{cov}(x_k - \hat{x}_{\bar{k}}) = \hat{P}_{\bar{k}}$ ，则有：

$$\hat{P}_k = (I - K_k H) \hat{P}_{\bar{k}} (I - K_k H)^T - K_k R K_k^T \quad 2.1$$

## 2.2 推导最优卡尔曼增益 $K_k$

最优化 K：使后验估计  $\hat{x}_k$  的协方差  $P_k$  达到最小。（换一个概念）也是使向量  $x_k - \hat{x}_k$  的二范数的数学期望值最小化的一个过程。 $E\{|x_k - \hat{x}_k|^2\}$  这等同于后验估计的协方差矩阵的迹最小化

$$P\{|x_k - \hat{x}_k|^2\} = \text{tr} \hat{P}_k$$

$$\text{记： } S_k = H \hat{P}_{\bar{k}} H^T + R$$

2.1 可以写成

$$\hat{P}_k = \hat{P}_{\bar{k}} - K_k H \hat{P}_{\bar{k}} - \hat{P}_{\bar{k}} H^T K_k^T + K_k S_k K_k^T \quad 2.2$$

把  $\text{tr}(\hat{P}_k)$  对  $K_k$  求导，并令导数=0，则可以得到  $\text{tr}(\hat{P}_k)$  取最小值时的最

优化  $K_k$  的值。

$$\frac{dtr(\hat{P}_k)}{dK_k} = -(H\hat{P}_k)^T - \hat{P}_k H^T + (S_k K_k^T)^T + (K_k S_k) = -2\hat{P}_k H^T + 2K_k S_k = 0$$

解得：

$$K_k = \hat{P}_k H^T S_k^{-1} = \hat{P}_k H^T (H\hat{P}_k H^T + R)^{-1}$$

化简后验误差协方差公式。在卡尔曼增益等于上面导出的最优值时，计算后验协方差的公式可以进行简化。对于卡尔曼增益公式

$$K_k = \hat{P}_k H^T S_k^{-1}$$

在卡尔曼增益公式两侧同时右乘  $S_k K_k^T$  得：

$$K_k S_k K_k^T = \hat{P}_k H^T K_k^T$$

把上式带入 2.2 式，可以消去后面的两项，得：

$$\hat{P}_k = \hat{P}_k - K_k H \hat{P}_k$$

整理，得：

$$\hat{P}_k = (I - K_k H) \hat{P}_k$$

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使用最优卡尔曼增益的时候它才成立。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 2.2 式表示的后验误差协方差公式。

## 2.3 卡尔曼增益的物理意义

$$K_k = \frac{\hat{P}_k H^T}{H\hat{P}_k H^T + R} \quad \text{公式 2.3}$$

其中：H 矩阵为常量； $\hat{P}_k$  与过程激励噪声的协方差矩阵 Q 有关；R 为测量噪声的协方差矩阵。 $K_k$  取值范围： $K_k \in [0, H^{-1}]$ 。

当 R 趋向于零时，有：

$$\lim_{R \rightarrow 0} K_k = H^{-1}$$

当  $\hat{P}_k$  趋向于零时，有：

$$\lim_{\hat{P}_k \rightarrow 0} K_k = 0$$

意义：决定了最优估计组成比例的“调节器”。

当 R 趋向于零时，有：

$$\lim_{R \rightarrow 0} K_k = H^{-1}$$

此时 1.7 式改为

$$\hat{x}_k = \hat{x}_{\bar{k}} + H^{-1}(z_k - H\hat{x}_{\bar{k}}) = H^{-1}z_k$$

系统表现为完全取测量值作为状态的后验估计值，而系统的先验状态估计完全被抛弃。

反之当  $\hat{P}_{\bar{k}}$  趋向于零时，可知  $Q=0$ 。

易知，此时系统完全抛弃测量值，取先验估计值。

## 2.4 卡尔曼增益的理论意义

被估计值系统的第  $K+1$  时刻的状态值  $X(K+1)$  的卡尔曼滤波值  $\hat{X}(K+1/K+1)$ ，就是  $X(K+1)$  的无偏的最小方差估计。而且，滤波误差方差阵  $P(K+1)$  是基于测量值  $Z_1, Z_2, Z_3$  等等的  $X(K+1)$  的所有线性估计中最小的均方误差阵。

对于一维的情况，测量噪声协方差矩阵增大时，增益矩阵  $K$  变小。这就表明，如果测量噪声越大，该增益取的越小，以减弱测量噪声对估计值的影响，而使预测值所占最后的结果比重加大。

从卡尔曼滤波 5 个公式的推导可以看出，当矩阵  $P(K/K-1)$ ,  $Q$ ,  $R$ ，同乘以一个常数时，增益矩阵  $K$  的值不变。

由推导过程我们还可以看出，当  $P(K-1/K-1)$  或者  $Q$  矩阵变小，或者同时变小的时候， $P(K/K-1)$  也变小， $K$  矩阵也减小。从直观上看，这是自然的，因为  $P = E(X - \hat{X})(X - \hat{X})^T$ ， $P$  变小表示估计值或者预测值比较好，又因为  $Q = E(WW^T)$ ， $Q$  变小表示状态转移随机波动减小。所以新的测量值对状态的估计值的矫正影响减弱，于是增益矩阵  $K$  应当变小。