

Problem Setup

We have an unknown target state $x_k \in \mathbb{R}^n$, we want to estimate it during each timestamp $k \in \mathbb{N} \cup \{0\}$. We create some measurement method to help us guess each x_k . The Kalman Filter's objective is to guess a theoretically perfect x_k from each measurement z_k .

TL;DR, the Kalman Filter is the process doing: get z_1 , guess x_1 ; get z_2 , guess x_2 ; get z_3 , guess x_3 , and so on.

1 Assumptions

By means of *theoretically*, we need some assumptions.

1.1 Pre-Defined Transition

Let $F_k \in \mathbb{R}^{n \times n}$ be a matrix. We assume that the next state x_{k+1} can be obtained from the current state x_k by

$$x_{k+1} = F_k x_k + u_k \quad (1)$$

where $u_k \sim \mathcal{N}(0, Q_k)$ is a white noise, $u_k \in \mathbb{R}^n$, and $Q_k \in \mathbb{R}^{n \times n}$.

This is also called *the linear model of Kalman Filter*.

1.2 Measurement to Real State

Let $H_k \in \mathbb{R}^{m \times n}$ be a matrix. We assume that the current measurement $z_k \in \mathbb{R}^m$ is obtained from the current state x_k by

$$z_k = H_k x_k + w_k \quad (2)$$

where $w_k \sim \mathcal{N}(0, R_k)$ is a white noise, $w_k \in \mathbb{R}^m$, and $R_k \in \mathbb{R}^{m \times m}$.

2 Initial Setup

We never know x_k , we can only guess it, so denote the guessed value by \hat{x}_k . Formally, we call \hat{x}_k the *estimate*.

The zero-th step of the Kalman Filter is randomly guess \hat{x}_0 since there's no z_{-1} . We further concern the correctness of each \hat{x}_k , we use the error covariance matrix $P_k = \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T]$ to formally compute it. So when we start with x_0 , we'll also have another randomly guess matrix P_0 .

TL;DR, the initial setup is one vector \hat{x}_0 and one matrix P_0 .

3 Preliminary Theorems

3.1 Minimum Variance Unbiased Estimate

Consider the measurement equation (2), we wanna seek a matrix $K_k \in \mathbb{R}^{n \times m}$ that can do the guessing $\hat{x}_k = K_k z_k$. Since w_k is a random vector, z_k is also a random vector, then so does \hat{x}_k . As a result, the error term $\hat{x}_k - x_k$ is a random vector as well. In this subsection's discussion, the timestamp subscript k won't effect any computation, so I ignore it to avoid the confusion.

We can define an inner product of two random vectors be

$$(x|y) = E(x^T y) = E\left(\sum_{i=1}^n x_i y_i\right),$$

and then induces a norm by

$$\|x\| = \sqrt{E(x^T x)} = \sqrt{E(x_1^2 + \dots + x_n^2)} = \{\text{Tr}(E(xx^T))\}^{1/2}.$$

Now, consider the l2-norm

$$\begin{aligned} \|\hat{x} - x\|^2 &= E[(\hat{x} - x)^T (\hat{x} - x)] \\ &= E[(Kz - x)^T (Kz - x)] \\ &= E[(K(Hx + w) - x)^T (K(Hx + w) - x)] \\ &= E[((KHx)^T + (Kw)^T - x^T)(K(Hx + w) - x)] \\ &= E[(KHx - x)^T (KHx - x)] + E[(Kw)^T (Kw)] \\ &= \|KHx - x\|^2 + \text{Tr}(E[Kw(Kw)^T]) \\ &= \|KHx - x\|^2 + \text{Tr}(K E[ww^T] K^T) \\ &= \|KHx - x\|^2 + \text{Tr}(K R K^T). \end{aligned}$$

The final term still depends on the actual state x , but we never know its value. To solve this, we add one more assumption $KH = I$. Observe that if $KH = I$, we have

$$E[\hat{x}] = E[Kz] = E[KHx + Kw] = E[KHx] + E[Kw] = x,$$

this implies that \hat{x} is an unbiased estimate of x .

Then the error term becomes quite simple when we assume $KH = I$

$$\|\hat{x} - x\|^2 = E[(\hat{x} - x)^T (\hat{x} - x)] = \text{Tr}(K R K^T)$$

Our mission now becomes

$$\begin{aligned} &\arg \min_K \text{Tr}(K R K^T) \\ &\text{s. t. } KH = I_n \end{aligned}$$

This is an optimization problem with $n \times n$ constraints.

Proof Needed, ref: Ch4.4 in Optimization-Vector-Space-Methods-Luenberger

4 Main Theorem

Theorem 4.1. *The optimal estimate \hat{x}_{k+1} and P_{k+1} can be generated recursively as*

$$\begin{aligned} \hat{x}_{k+1} &= F_k \hat{x}_k + F_k P_k H_k^T [H_k P_k H_k^T + R_k]^{-1} (z_k - H_k \hat{x}_k) \\ P_{k+1} &= F_k P_k \left\{ I - H_k^T [H_k P_k H_k^T + R_k]^{-1} H_k P_k \right\} F_k^T + Q_k \end{aligned}$$

Proof. Suppose we have \hat{x}_{k-1} and P_{k-1} . At k , we obtain a new measurement

$$z_k = H_k x_k + w_k$$

which gives us additional information about x_k .

□