Kalman filters

Kalman filters are very good at finding the state x of a system. But what kinds of Kalman filters are there? How do they work? And what are their properties? That's what this chapter is about.

1 The time-varying Kalman filter

1.1 The Kalman filtering problem

Let's suppose that we are observing some stochastic process y. Based on this process, we should estimate another stochastic process z. In other words, we need to determine $z(t)$ given F_{t-1}^y. A special case of this problem is the **Kalman filtering problem**. Now z (or x) equals the state of some system with output y. In other words, given the system

$$x(t+1) = A(t)x(t) + M(t)v(t), \quad x(t_0) = x_0 \in G(m_0, Q_0)m \tag{1.1}$$

$$y(t) = C(t)x(t) + N(t)v(t), \quad v(t) \in G(0, Q_v(t)), \tag{1.2}$$

we need to determine the conditional distribution of $x(t)$ given F_{t-1}^y.

1.2 The time-varying Kalman filter

To solve this problem, we can use a **time-varying Kalman filter**. However, this only works if $N(t)Q_v(t)N(t)^T > 0$. (If this is not the case, then we can split up the system. The condition then holds for the first part of the system, while the second part is not affected by the noise.) In this case, the distribution of $x(t)|F_{t-1}^y$ is Gaussian and is specified by the characteristic function

$$E[\exp(iw^T x(t))|F_{t-1}^y] = \exp(iw^T \hat{x}(t) - \frac{1}{2} w^T Q_f(t) w), \ \forall w \in \mathbb{R}^n. \tag{1.3}$$

In the above equation, we call $\hat{x}$ the **conditional mean process** and Q_f the **conditional variance process**.

But how do we find $\hat{x}$ and Q_f? For that, we can use the recursively defined equations

$$\hat{x}(t+1) = A(t)\hat{x}(t) + K(t)[y(t) - C(t)\hat{x}(t)], \ \hat{x}(t_0) = E[x_0] = m_0, \tag{1.4}$$

$$H_{11}(t) = A(t)Q_f(t)A(t)^T + M(t)Q_v(t)M(t)^T, \tag{1.5}$$

$$H_{12}(t) = A(t)Q_f(t)C(t)^T + M(t)Q_v(t)N(t)^T, \tag{1.6}$$

$$H_{22}(t) = C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T, \tag{1.7}$$

$$K(t) = H_{12}(t)H_{22}^{-1}(t), \tag{1.8}$$

$$Q_f(t+1) = H_{11}(t) - H_{12}(t)H_{22}^{-1}(t)H_{12}(t)^T, \ Q_f(t_0) = Q_0. \tag{1.9}$$

If we use the above relations, then we have

$$\hat{x}(t) = E[x(t)|F_{t-1}^y] \qquad \text{and} \qquad Q_f(t) = E[(x(t) - \hat{x}(t))(x(t) - \hat{x}(t))^T|F_{t-1}^y]. \tag{1.10}$$

1.3 Related process

When applying the Kalman filter, we can define the **innovation process** $\bar{v}(t)$ as $\bar{v}(t) = y(t) - C(t)\hat{x}(t)$. Now $\bar{v}(t)$ is a Gaussian white noise process, with

$$\bar{v}(t) \in G(0, Q_{\bar{v}}(t)), \quad Q_{\bar{v}}(t) = H_{22}(t) = C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T. \tag{1.11}$$

Given this white noise process, the output $y(t)$ can be determined. In other words, we have $F_t^y = F_t^{\bar{v}}$. Next to this, the error process $e(t)$ is defined as $e(t) = x(t) - \hat{x}(t)$. Its recursive equation is

$$e(t+1) = (A(t) - K(t)C(t))e(t) + (M(t) - K(t)N(t))v(t). \tag{1.12}$$

2 The time-invariant Kalman filter

2.1 The Filter algebraic Ricatti equation

Previously, we considered the case where the state space matrices depended on time. Now let's assume that A, C, M, N and Q_v are constant in time. In this case, the time-varying Kalman filter may converge to a **time-invariant Kalman filter**. To see when this happens, we first make some assumptions.

- Assume that $NQ_vN^T > 0$.
- Define F and G such that

$$F = A - MQ_vN^T(NQ_vN^T)^{-1}C, \tag{2.1}$$
$$GG^T = MQ_vM^T - MQ_vN^T(NQ_vN^T)^{-1}(MQ_vN^T)^T. \tag{2.2}$$

  Assume that (F, G) is a stabilizable pair.
- Assume that (A, C) is a detectable pair.

Let's define the function $f(Q)$ as

$$f(Q_f) = AQ_fA^T + MQ_vM^T - (AQ_fC^T + MQ_vN^T)(CQ_fC^T + NQ_vN^T)^{-1}(AQ_fC^T + MQ_vN^T)^T. \tag{2.3}$$

The above assumptions now imply that there exists a unique solution $Q_f = Q_f^T \geq 0$ to the **Filter algebraic Riccati equation** $Q_f = F(Q_f)$. This solution Q_f is actually the limit $\lim_{t\to\infty} Q_f(t)$ of the time-varying Kalman filter. Also, if (F, G) is controllable as well, then we have $Q_f = Q_f^T > 0$.

2.2 The time-invariant Kalman filter

Let's define the matrices $K(Q_f)$ and $A(Q_f)$ as

$$K(Q_f) = (AQ_fC^T + MQ_vN^T)(CQ_fC^T + NQ_vN^T)^{-1}, \tag{2.4}$$
$$A(Q_f) = A - K(Q_f)C. \tag{2.5}$$

It can be shown that $\text{spec}(A(Q_f)) \subset \mathbb{D}$. The time-invariant Kalman filter is now given by

$$\bar{x}(t+1) = A\bar{x}(t) + K(y(t) - C\bar{x}(t)). \tag{2.6}$$

It is important to note that we don't denote this filter by $\hat{x}(t)$. This is because generally $\bar{x} \neq E[x(t)|F_{t-1}^y]$. But if this is the case, then why would we use the time-invariant Kalman filter, and not the time-varying Kalman filter.

The answer is that a filter with a time-varying matrix K (like the time-varying Kalman filter) is quite complicated. A filter with a constant matrix K is a lot simpler. And it can be shown that the time-invariant Kalman filter is better than any other filter with a constant matrix K. With 'better' we mean that the variance of the error $e(t) = x(t) - \bar{x}(t)$ of the Kalman filter is always better (or just as good) as any other filter.