# OPTIMAL FILTERING

**LECTURE 4**

1. The innovations process

2. Derivation of the Kalman filter, innovations approach

**Reading instructions:** Kailath, Sect. 9.1-9.4



# THE INNOVATIONS PROCESS

Let $\{y_k\}$, be a sequence of Gaussian random variables. The innovations process $\{\tilde{y}_{k|k-1}\}$ can be regarded as the *new information* brought by $y_k$ that cannot be determined from the past.

$$\tilde{y}_{k|k-1} = y_k - \hat{y}_{k|k-1} \ .$$

Sometimes we denote the innovations

$$e_k = \tilde{y}_{k|k-1} = y_k - \hat{y}_{k|k-1} \ .$$

- Note that $e_k$ is a linear function of $y_0, y_1, \ldots, y_k$ since $\hat{y}_{k|k-1}$ is the l.l.s.e.

- $e_k$ is orthogonal to $y_0, y_1, \ldots, y_{k-1}$ by the projection principle

This implies that $\{e_k\}$ is a **white process**.

To see this:

$$e_k = \underbrace{\mathbf{w}}_{1 \times (k+1) \text{ vector}} \mathbf{y}_k \quad \text{where} \quad \mathbf{y}_k = \begin{bmatrix} y_0 \\ \vdots \\ y_k \end{bmatrix}$$

linear transformation of the observations.

Then, we have

$$E\{e_k e_l^*\} = E\{e_k \mathbf{y}_l^* \mathbf{w}^*\} = 0 \ \text{ if } \ k > l$$
$$E\{e_k e_l^*\} = E\{\mathbf{w} \mathbf{y}_k e_l^*\} = 0 \ \text{ if } \ l > k$$



# Innovations Process, Other Properties

- In HW#2, you showed that $\{e_k\}$ can be obtained by a causal linear transformation of $\{y_k\}$. This transformation is *causally* invertible. Can be obtained, for example, by Gram-Schmidt orthogonalization.
- The same information is contained in $\{y_k\}$ and $\{e_k\}$!

$$\Longrightarrow$$

$$\begin{aligned} \hat{x}_{\cdot|k} &= \text{the l.l.s.e. of } x \text{ given } \{y_0, \ldots, y_k\} \\ &= \text{the l.l.s.e. of } x \text{ given } \{e_0, \ldots, e_k\} \end{aligned}$$

$$\begin{aligned} \hat{x}_{\cdot|k} &= E\{x|y_0, \ldots, y_k\} = E\{x|e_0, \ldots, e_k\} = \backslash e_k \text{ white}\backslash \\ &= E\{x|e_0\} + \cdots + E\{x|e_{k-1}\} + E\{x|e_k\} \\ &= \hat{x}_{\cdot|k-1} + E\{x|e_k\} \end{aligned}$$

# To Summarize

$$\hat{x}_{\cdot|k} = \hat{x}_{\cdot|k-1} + E\{x|e_k\}$$

$$e_k = y_k - \hat{y}_{k|k-1} = y_k - \sum_{j=0}^{k-1} E\{y_k e_j^*\} E\{e_j e_j^*\}^{-1} e_j$$

These simple formulas are the key to many results in l.l.s.e. theory.



# The Kalman Filter, Innovations Approach

**Assumptions:**

$$x_{k+1} = F_k x_k + G_k w_k \qquad k \geq 0$$
$$y_k = H_k x_k + v_k$$

$$E \begin{bmatrix} w_k \\ v_k \end{bmatrix} \begin{bmatrix} w_l^* & v_l^* \end{bmatrix} = \begin{bmatrix} Q_k & S_k \\ S_k^* & R_k \end{bmatrix} \delta_{k,l}$$

$$Ex_0 = 0, \qquad Ew_k = 0, \qquad Ev_k = 0$$

$$Ex_0 x_0^* = \Pi_0, \qquad Ex_0 v_k^* = 0, \qquad Ex_0 w_k^* = 0$$

**Innovations:**

$$e_k = y_k - \hat{y}_{k|k-1} \tag{1}$$

# Main results from the derivations

$$\hat{y}_{k|k-1} = H_k \hat{x}_{k|k-1} \tag{2}$$

$$\hat{x}_{k+1|k-1} = F_k \hat{x}_{k|k-1} \tag{4}$$

$$\begin{aligned} e_k &= y_k - H_k \hat{x}_{k|k-1} \qquad k > 0 \\ \hat{x}_{k+1|k} &= F_k \hat{x}_{k|k-1} + K_k e_k \qquad k \geq 0 \\ K_k &= E\{x_{k+1} e_k^*\} R_{e_k}^{-1} \\ e_0 &= y_0, \qquad \hat{x}_{0|-1} = 0 \end{aligned} \tag{5}$$

Same without innovations:

$$\hat{x}_{k+1|k} = (F_k - K_k H_k)\hat{x}_{k|k-1} + K_k y_k \tag{5'}$$



# Finding $K_k$ and $R_{e_k}^{-1}$

**Note:** These quantities are non-random and determined by our model and assumptions, *not* on the actual observations.

**Notation:**

$$\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$$

$$\underbrace{P_k}_{\text{previously called } P_{k|k-1}} = E\{\tilde{x}_{k|k-1} \tilde{x}_{k|k-1}^*\}$$

**Main steps in the derivations:**

$$R_{e_k} = H_k P_k H_k^* + R_k \tag{6}$$

$$K_k = (F_k P_k H_k^* + G_k S_k) R_{e_k}^{-1} \tag{7}$$

**Discrete-time Riccati equation:**

$$P_{k+1} = F_k P_k F_k^* + G_k Q_k G_k^* - K_k R_{e_k} K_k^* \qquad k \geq 0 \tag{8}$$

# Resulting Kalman Filter Equations

$$
\begin{aligned}
e_k &= y_k - H_k \hat{x}_{k|k-1} \qquad k > 0 \\
\hat{x}_{k+1|k} &= F_k \hat{x}_{k|k-1} + K_k e_k \qquad k \geq 0 \\
R_{e_k} &= H_k P_k H_k^* + R_k \\
K_k &= (F_k P_k H_k^* + G_k S_k) R_{e_k}^{-1} \\
P_{k+1} &= F_k P_k F_k^* + G_k Q_k G_k^* - K_k R_{e_k} K_k^* \qquad k \geq 0 \\
P_0 &= \Pi_0 \\
X_{0|-1} &= 0
\end{aligned}
$$

So-called *predicted estimates* form.



# Kalman Gain

$K_k$ is called the *Kalman gain*.

It tells us how much we should adjust our estimate $\hat{x}_{k+1|k}$ based on the measurements $y_k$.

$K_k$ small $\Longrightarrow$ trust the model

$K_k$ large $\Longrightarrow$ trust the measurements