

# Lecture #4

Gustaf Hendeby  
gustaf.hendeby@liu.se

Version: 2023-12-04

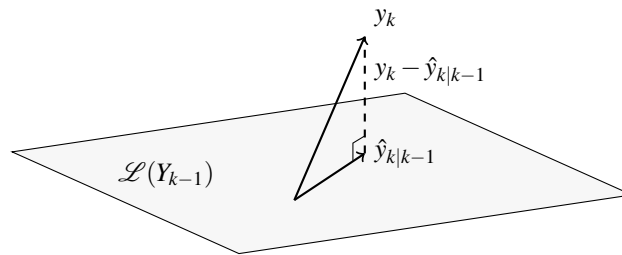
## 1 The Innovation Process

Let  $\{y_k\}$  be a sequence of random variables, and let  $\hat{y}_{k|k-1}$  be the lse of  $y_k$  given  $y_0, y_1, \dots, y_{k-1}$ . The “new information” brought by  $y_k$ , that could not be determined from past observations, are

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

which is denoted the innovation.

Note that  $e_k$  is a linear function of  $\{y_i\}_{i=0}^k$ .  $e_k \perp \mathcal{L}(\{y_i\}_{i=0}^{k-1})$  by the projection principle, thus  $e_k$  is a **white sequence (process)**.



**To see this:** Let

$$e_k = AY_k, \quad Y_k = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix},$$

then

$$E[e_k e_l^*] = E[e_k (AY_l^*)] = E[e_k Y_l^*] A^* = \langle e_k, Y_l \rangle A^* = 0, \quad k > l$$

$$E[e_k e_l^*] = A E[Y_k e_l^*] = A \langle Y_k, e_l \rangle = 0, \quad l > k$$

**Note:** In Homework 2 you showed that  $E_k^T = (e_0^T \dots e_k^T)$  can be obtained by a **causal** linear transformation of  $Y_k^T = (y_0^T \dots y_k^T)$ . This transformation is causally invertible.

### Matrix form

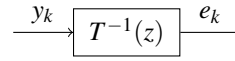
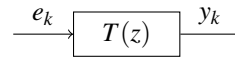
$$Y_k = \begin{pmatrix} & & & \mathbf{0} \\ & & \times & \\ & & & \\ \times & & & \end{pmatrix} E_k, \quad E_k = \begin{pmatrix} e_0 \\ e_1 \\ \vdots \\ e_k \end{pmatrix}$$

$$E_k = \begin{pmatrix} & & & \mathbf{0} \\ & & \times & \\ & & & \\ \times & & & \end{pmatrix} Y_k, \quad Y_k = \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_k \end{pmatrix}$$

$$R_{yy} = TR_{ee}T^*,$$

$T$ — lower triangular, with unit diagonal

### Transfer function



$$\Phi_{yy}(z) = \sigma_e^2 T(z)T(z^{-1}),$$

$$0 < \Phi_{yy}(z) \Big|_{z=e^{j\omega}} < \infty, \forall \omega$$

The “matrix” transformation can be obtained by, e.g., Gram-Schmidt orthogonalization, *i.e.*,

$$e_k = y_k - \sum_{j=0}^{k-1} \langle y_k, e_j \rangle \|e_j\|^{-2} e_j$$

Hence, the same information is contained in  $Y_k$  and  $E_k$ , which implies that

$$\begin{aligned} \hat{x}_{l|k} &= \text{llse. of } x_l \text{ given } Y_k = \text{llse. of } x_l \text{ given } E_k \\ &= \text{Proj}(x_l | \mathcal{L}(E_k)) = \text{orthogonal} / = \sum_{i=0}^k \langle x_l, e_i \rangle \|e_i\|^{-2} e_i \\ &= \hat{x}_{l|k-1} + \underbrace{\langle x_l, e_k \rangle}_{E[x_l e_k^*]} \underbrace{\|e_k\|^{-2}}_{R_{e,k}^{-1}} e_k \end{aligned}$$

### Summary

$$\hat{x}_{l|k} = \hat{x}_{l|k-1} + \langle x_l, e_k \rangle \|e_k\|^{-2} e_k \quad (\text{Recursive update of llse.})$$

$$e_k = y_k - \hat{y}_{k|k-1} = y_k - \sum_{j=0}^{k-1} \langle y_k, e_j \rangle \|e_j\|^{-2} e_j \quad (\text{Innovations sequence from observations})$$

See how the Wiener-Hopf equations were solved in Lecture 2 using the whitening filter.

### Computational issues

- Generally:  $\mathcal{O}(k^3)$  operation to calculate  $T$  (factorize  $R_{yy}$ ) or equivalently the innovations.
- If  $\{y_k\}$  stationary process  $\rightarrow \mathcal{O}(k^2)$  operations.
- State-space model, where  $\dim(x_k) = n \ll k \rightarrow \mathcal{O}(kn^3)$  operations.

## Kalman Filter — an innovation approach

State-space model ( $k \geq 0$ ):

$$\begin{aligned} x_{k+1} &= F_k x_k + G_k w_k \\ y_k &= H_k x_k + v_k \end{aligned}$$

$$\mathbb{E} \begin{bmatrix} \begin{pmatrix} w_k \\ v_k \\ x_0 \end{pmatrix} \begin{pmatrix} w_l \\ v_l \\ x_0 \\ 1 \end{pmatrix}^* \end{bmatrix} = \begin{pmatrix} Q_k \delta_{l-k} & S_k \delta_{l-k} & 0 & 0 \\ S_k^* \delta_{l-k} & R_k \delta_{l-k} & 0 & 0 \\ 0 & 0 & \Pi_0 & 0 \end{pmatrix}$$

(The last column implies the involved stochastic variables are zero-mean.) Let  $e_k = y_k - \hat{y}_{k|k-1}$ , where

$$\hat{y}_{k|k-1} = \text{Proj}(y_k | Y_{k-1}) = \text{Proj}(H_k x_k + v_k | Y_{k-1}) = \left/ v_k \perp Y_{k-1} \right/ = H_k \hat{x}_{k|k-1}$$

Thus, finding the innovations is equivalent to find the one-step predictor of the state  $x_k$ . Since  $\{e_k\}$  is a white sequence, then

$$\begin{aligned} \hat{x}_{k+1|k} &= \sum_{i=0}^k \langle x_{k+1}, e_i \rangle \|e_i\|^{-2} e_i \\ &= \hat{x}_{k+1|k-1} + \langle x_{k+1}, e_k \rangle \|e_k\|^{-2} e_k \\ &= \hat{x}_{k+1|k-1} + K_{p,k} e_k, \end{aligned} \quad K_{p,k} = \langle x_{k+1}, e_k \rangle \|e_k\|^{-2} = \mathbb{E}[x_{k+1} e_k^*]$$

Note, we would like an expression in terms of  $\hat{x}_{k|k-1}$  instead of  $\hat{x}_{k+1|k-1}$ .

$$\begin{aligned} \hat{x}_{k+1|k-1} &= \text{Proj}(x_{k+1} | Y_{k-1}) = \text{Proj}(F x_k + G w_k | Y_{k-1}) \\ &= F_k \text{Proj}(x_k | Y_{k-1}) = F_k \hat{x}_{k|k-1} \end{aligned}$$

Bringing it all to together we get the recursion

$$\hat{x}_{k+1|k} = F_k \hat{x}_{k|k-1} + K_{p,k} (y_k - H_k \hat{x}_{k|k-1}) = (F_k - K_{p,k} H_k) \hat{x}_{k|k-1} + K_{p,k} y_k.$$

What remains is to find a recursive way to calculate the Kalman predictive gain  $K_{p,k}$ .

Let  $\tilde{x}_{k|k-1} = x_k - \hat{x}_{k|k-1}$  and  $P_{k|k-1} = \mathbb{E}[\tilde{x}_{k|k-1} \tilde{x}_{k|k-1}^*]$  then

$$\begin{aligned} e_k &= y_k - H_k \hat{x}_{k|k-1} = H_k \tilde{x}_{k|k-1} + v_k \\ R_{e,k} &= \mathbb{E}[e_k e_k^*] = \left/ v_k \perp \tilde{x}_{k|k-1} \right/ = H_k P_{k|k-1} H_k^* + R_k \end{aligned}$$

Further we have that

$$\mathbb{E}[x_{k+1} e_k^*] = F_k \mathbb{E}[x_k e_k^*] + G_k \mathbb{E}[w_k e_k^*]$$

where

$$\begin{aligned} \mathbb{E}[x_k e_k^*] &= \mathbb{E}[x_k (H_k \tilde{x}_{k|k-1} + v_k)^*] \\ &= \mathbb{E}[x_k \tilde{x}_{k|k-1}^*] H_k^* + \mathbb{E}[x_k v_k^*] = \left/ \hat{x}_{k|k-1} \perp \tilde{x}_{k|k-1} \right/ \Rightarrow \mathbb{E}[\hat{x}_{k|k-1} \tilde{x}_{k|k-1}^*] = 0 \left/ \right. \\ &= \mathbb{E}[(x_k - \hat{x}_{k|k-1}) \tilde{x}_{k|k-1}^*] H_k^* = P_{k|k-1} H_k^* \\ \mathbb{E}[w_k e_k^*] &= \mathbb{E}[w_k (H_k \tilde{x}_{k|k-1} + v_k)^*] = \mathbb{E}[w_k v_k^*] = S_k \end{aligned}$$

yielding

$$K_{p,k} = (F_k P_{k|k-1} H_k^* + G_k S_k) (H_k P_{k|k-1} H_k^* + R_k)^{-1}$$

Now we need a recursion for  $P_{k|k-1}$ .

$$\begin{aligned} \tilde{x}_{k+1|k} &= x_{k+1} - \hat{x}_{k+1|k} = F_k x_k + G_k w_k - (F_k \hat{x}_{k|k-1} + K_{p,k} e_k) \\ &= F_k x_k + G_k w_k - F_k \hat{x}_{k|k-1} - K_{p,k} (H_k \tilde{x}_{k|k-1} + v_k) \\ &= (F_k - K_{p,k} H_k) \tilde{x}_{k|k-1} + \begin{pmatrix} G_k & K_{p,k} \end{pmatrix} \begin{pmatrix} w_k \\ -v_k \end{pmatrix} \end{aligned}$$

yielding ( $\tilde{x}_{k|k-1} \perp w_k, v_k$ )

$$\begin{aligned} P_{k+1|k} &= (F_k - K_{p,k} H_k) P_{k|k-1} (F_k - K_{p,k} H_k)^* + \begin{pmatrix} G_k & K_{p,k} \end{pmatrix} \begin{pmatrix} Q_k & -S_k \\ -S_k^* & R_k \end{pmatrix} \begin{pmatrix} G_k & K_{p,k} \end{pmatrix}^* \\ &= \left/ \dots \right/ = F_k P_{k|k-1} F_k^* + G_k Q_k G_k^* - K_{p,k} R_{e,k} K_{p,k}^* \end{aligned}$$

This is the *discrete time algebraic Riccati equation* (DARE)

## Kalman Filter (prediction form)

$P_0 = \Pi_0, \quad \hat{x}_{0 -1} = 0$ (assuming zero-mean)	Initial values
$e_k = y_k - H_k \hat{x}_{k k-1}$	Innovation
$R_{e,k} = H_k P_{k k-1} H_k^* + R_k$	Innovation covariance
$K_{p,k} = (F_k P_{k k-1} H_k^* + G_k S_k) R_{e,k}^{-1}$	Kalman prediction gain
$\hat{x}_{k+1 k} = F_k \hat{x}_{k k-1} + K_{p,k} e_k$	Prediction
$P_{k+1 k} = F_k P_{k k-1} F_k^* + G_k Q_k G_k^* - K_{p,k} R_{e,k} K_{p,k}^*$	State covariance

$K_{p,k}$  tells us how much we should adjust our estimate  $\hat{x}_{k|k-1}$  given observations  $y_k$ :

- $K_{p,k}$  — small: Trust the model ( $\|Q\|/\|R\|$  small)
- $K_{p,k}$  — large: Trust the measurements/observations ( $\|Q\|/\|R\|$  large)