

**Control Engineering SC42095**

# **Kalman filtering and LQG control**

Tamás Keviczky, Azita Dabiri  
Delft Center for Systems and Control  
Delft University of Technology

# Lecture outline

## 1. Review

- LQ control
- completing the squares
- dynamic programming

## 2. Kalman filtering

## 3. Linear quadratic Gaussian control

Note: These slides are partly inspired by the slides for this course developed at the Department of Automatic Control, Lund Institute of Technology (see <http://www.control.lth.se/~kursdr>)

## Lecture outline (continued)

- Linear Quadratic (LQ) control → assumes *full state information*
- Estimating state from measurements of output = Kalman filtering
- Combination of LQ and state estimation = Linear Quadratic Gaussian (LQ) control based on *separation theorem*

# 1. Review

## 1.1 LQ control

- Minimize

$$J = \sum_{k=0}^{N-1} \left( x^T(k) Q_1 x(k) + 2x^T(k) Q_{12} u(k) + u^T(k) Q_2 u(k) \right) + x^T(N) Q_0 x(N)$$

subject to  $x(k+1) = \Phi x(k) + \Gamma u(k)$  and  $x(0) = x_0$

- Solution approach based on quadratic optimization problem and dynamic programming
- Results in state feedback controller  $u = -L(k)x(k)$  with  $L(k)$  determined by solution  $S(k)$  of Riccati recursion

## 1.2 Completing the squares

- Find  $u$  that minimizes  $x^\top Q_x x + 2x^\top Q_{xu} u + u^\top Q_u u$  with  $Q_u$  positive definite
- Let  $L$  be such that  $Q_u L = Q_{xu}^\top$ . Then

$$x^\top Q_x x + 2x^\top Q_{xu} u + u^\top Q_u u =$$

$$x^\top (Q_x - L^\top Q_u L) x + (u + Lx)^\top Q_u (u + Lx)$$

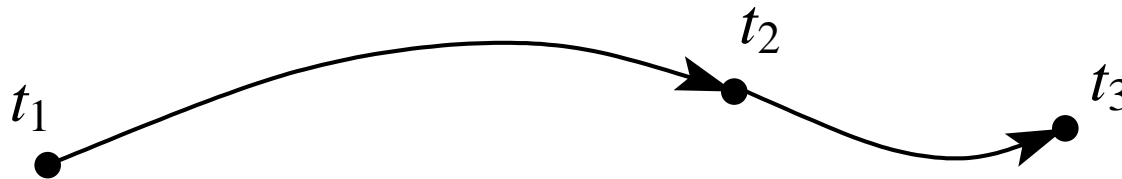
is minimized for  $u = -Lx$

and minimum value is  $x^\top (Q_x - L^\top Q_u L) x$

- If  $Q_u$  is positive definite, then  $L = Q_u^{-1} Q_{xu}^\top$

## 1.3 Dynamic programming

- **Principle of optimality:** From any point on optimal trajectory, remaining trajectory is also optimal



- allows to determine best control law over period  $[t_2, t_3]$  independent of how state at  $t_2$  was reached
- For  $N$ -step problem:
  - start from end at time  $k = N$
  - now we can determine best control law for last step independent of how state at time  $N - 1$  was reached
  - iterate backward in time to initial time  $k = 0$

## 2. Kalman filtering

- LQ control requires full state information
- In practice: only output measured  
→ how to estimate states from noisy measurements of output?
- Consider system

$$\begin{aligned}x(k+1) &= \Phi x(k) + \Gamma u(k) + v(k) \\ y(k) &= Cx(k) + e(k)\end{aligned}$$

with  $v, e$  Gaussian zero-mean white noise process with

$$\mathbb{E}[v(k)v^T(k)] = R_1, \quad \mathbb{E}[e(k)e^T(k)] = R_2, \quad \mathbb{E}[v(k)e^T(k)] = R_{12}$$

and  $x(0)$  Gaussian distributed with

$$\mathbb{E}[x(0)] = m_0$$

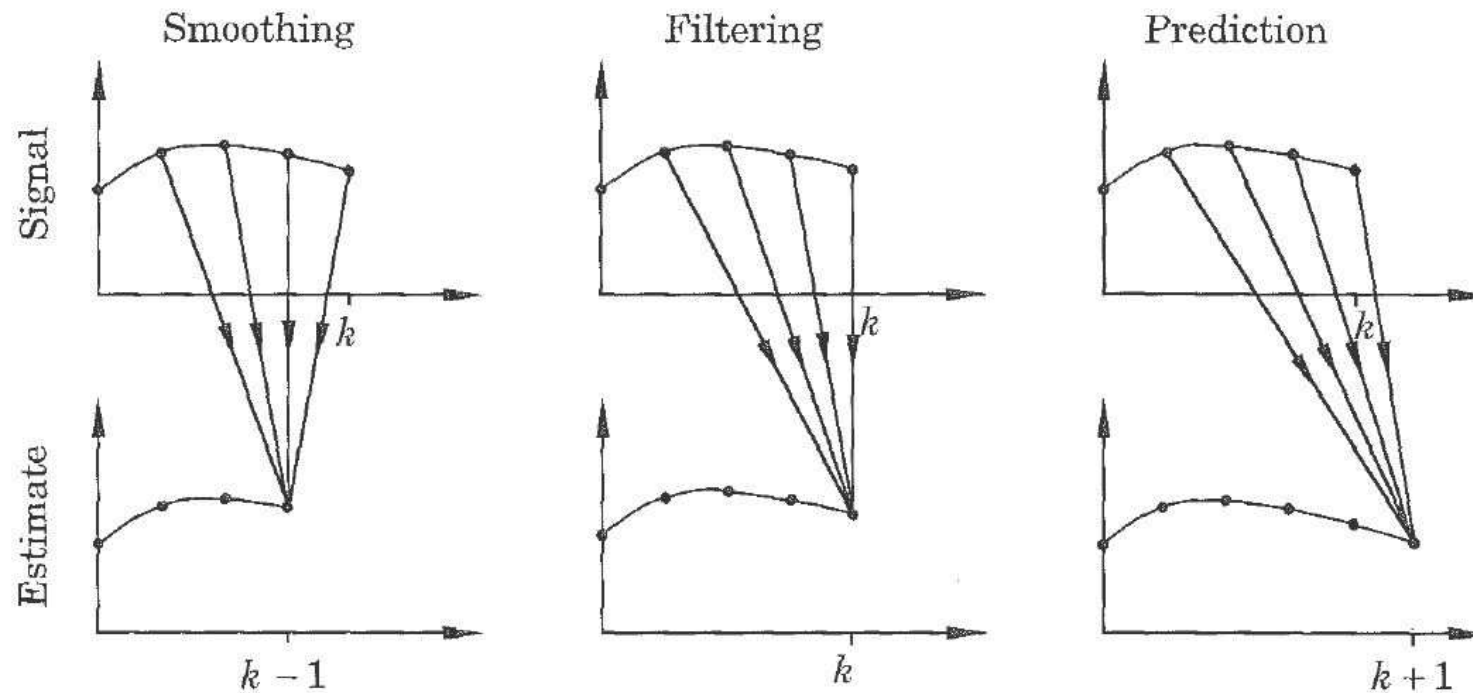
$$\text{cov}(x(0)) := \mathbb{E} \left[ (x(0) - \mathbb{E}[x(0)]) (x(0) - \mathbb{E}[x(0)])^T \right] = R_0$$

## 2.1 Problem formulation

- Given the data

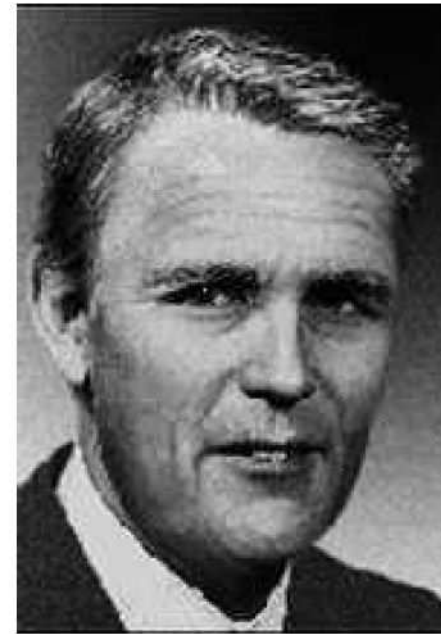
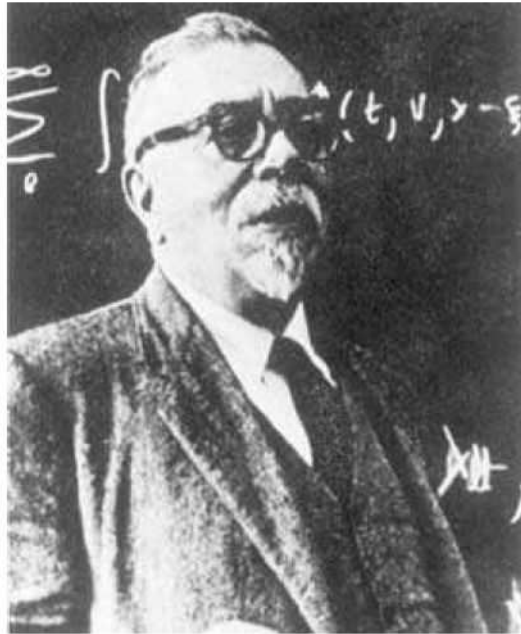
$$Y_k = \{y(i), u(i) : 0 \leq i \leq k\}$$

find the “best” (to be defined) estimate  $\hat{x}(k+m)$  of  $x(k+m)$ .  
( $m = 0$  filtering,  $m > 0$  prediction,  $m < 0$  smoothing)





## Some history



*Norbert Wiener:* Filtering, prediction, and smoothing using integral equations. Spectral factorizations.

*Rudolf E. Kalman:* Filtering, prediction, and smoothing using state-space formulas. Riccati equations.

## 2.2 Kalman filter structure

- Goal is to estimate  $x(k+1)$  by linear combination of previous inputs and outputs
- Estimator (cf. lecture on observers):

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K(k)(y(k) - C\hat{x}(k|k-1))$$

with  $\hat{x}(k+1|k)$  estimate of state  $x$  at sample step  $k+1$  using information available at step  $k$

- Error dynamics  $\tilde{x} = x - \hat{x}$  governed by

$$\begin{aligned}\tilde{x}(k+1) &= x(k+1) - \hat{x}(k+1|k) \\ &= \Phi x(k) + \Gamma u(k) + v(k) - \Phi\hat{x}(k|k-1) - \Gamma u(k) - K(k)(y(k) - C\hat{x}(k|k-1)) \\ &= \Phi x(k) + v(k) - \Phi\hat{x}(k|k-1) - K(k)(Cx(k) + e(k) - C\hat{x}(k|k-1)) \\ &= (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k)\end{aligned}$$

## 2.3 Determination of Kalman gain

- Error dynamics:  $\tilde{x}(k+1) = (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k)$
- If  $\hat{x}(0) = m_0$ , then  $\mathbf{E}[\tilde{x}(0)] = \mathbf{E}[x(0) - \hat{x}(0)] = \mathbf{E}[x(0)] - m_0 = 0$  and thus  $\mathbf{E}[\tilde{x}(k)] = 0$  for all  $k$
- How to choose  $K(k)$ ?  $\rightarrow$  **minimize covariance of  $\tilde{x}(k)$**
- We have
$$\begin{aligned}\text{cov}(\tilde{x}(k)) &= \mathbf{E} \left[ (\tilde{x}(k) - \mathbf{E}[\tilde{x}(k)]) (\tilde{x}(k) - \mathbf{E}[\tilde{x}(k)])^T \right] \\ &= \mathbf{E} [\tilde{x}(k) \tilde{x}^T(k)]\end{aligned}$$
- So if we define  $P(k) = \mathbf{E} [\tilde{x}(k) \tilde{x}^T(k)]$ , then we have to determine Kalman gain such that  $P(k)$  is minimized

## 2.3 Determination of Kalman gain (continued)

- Error dynamics:  $\tilde{x}(k+1) = (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k)$
- We have

$$\begin{aligned} P(k+1) &= \mathbf{E} [\tilde{x}(k+1)\tilde{x}^T(k+1)] \\ &= \mathbf{E} \left[ \left( (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k) \right) \left( (\Phi - K(k)C)\tilde{x}(k) + v(k) - K(k)e(k) \right)^T \right] \end{aligned}$$

- Since  $\tilde{x}(k)$  is independent of  $v(k)$  and  $e(k)$ , this results in

$$\begin{aligned} P(k+1) &= \mathbf{E} \left[ (\Phi - K(k)C)\tilde{x}(k)\tilde{x}^T(k)(\Phi - K(k)C)^T \right. \\ &\quad \left. + v(k)v^T(k) + K(k)e(k)e^T(k)K^T(k) - v(k)e^T(k)K^T(k) - K(k)e(k)v^T(k) \right] \\ &= (\Phi - K(k)C) \underbrace{\mathbf{E} [\tilde{x}(k)\tilde{x}^T(k)]}_{P(k)} (\Phi - K(k)C)^T + R_1 + K(k)R_2K^T(k) \\ &\quad - R_{12}K^T(k) - K(k)R_{12}^T \end{aligned}$$

## 2.3 Determination of Kalman gain (continued)

- So

$$\begin{aligned}
 P(k+1) &= (\Phi - K(k)C)P(k)(\Phi - K(k)C)^T + R_1 + K(k)R_2K^T(k) - R_{12}K^T(k) - K(k)R_{12}^T \\
 &= K(k)(CP(k)C^T + R_2)K^T(k) - (\Phi P(k)C^T + R_{12})K^T(k) - K(k)(CP^T(k)\Phi^T + R_{12}^T) \\
 &\quad + (\Phi P(k)\Phi^T + R_1)
 \end{aligned}$$

- Minimize  $P(k+1)$  with  $K(k)$  as decision variable
- $P(k+1)$  is quadratic function of  $K(k)$
- Use completing-of-squares solution with  $u = K^T(k)$  and  $x = I$ :

$$\begin{aligned}
 &\underbrace{K(k)}_{u^T} \underbrace{(CP(k)C^T + R_2)}_{Q_u} \underbrace{K^T(k)}_u + \underbrace{(-\Phi P(k)C^T - R_{12})}_{Q_{xu}} \underbrace{K^T(k)}_u \\
 &\quad + \underbrace{K(k)}_{u^T} \underbrace{(-CP^T(k)\Phi^T - R_{12}^T)}_{Q_{xu}^T} + \underbrace{(\Phi P(k)\Phi^T + R_1)}_{Q_x}
 \end{aligned}$$

## 2.3 Determination of Kalman gain (continued)

- Results in:

$$K^T(k) = -L(k)x = -L(k)$$

$$L(k) = Q_u^{-1}Q_{xu}^T = (CP(k)C^T + R_2)^{-1}(-\Phi P(k)C^T - R_{12})^T$$

- So

$$K(k) = -L^T(k) = (\Phi P(k)C^T + R_{12})(CP(k)C^T + R_2)^{-1}$$

- Furthermore,

$$\begin{aligned} P(k+1) &= Q_x - L^T Q_u L \\ &= \Phi P(k) \Phi^T + R_1 - K(k) (CP(k)C^T + R_2) K^T(k) \\ &= \Phi P(k) \Phi^T + R_1 - (\Phi P(k)C^T + R_{12})(CP(k)C^T + R_2)^{-1}(CP(k)\Phi^T + R_{12}^T) \end{aligned}$$

## 2.3 Determination of Kalman gain (continued)

- Initial value: 
$$\begin{aligned} P(0) &= \mathbf{E} [\tilde{x}(0)\tilde{x}^T(0)] \\ &= \mathbf{E} \left[ (x(0) - \hat{x}(0)) (x(0) - \hat{x}(0))^T \right] \\ &= \mathbf{E} \left[ (x(0) - m_0) (x(0) - m_0)^T \right] \\ &= \text{cov}(x(0)) = R_0 \end{aligned}$$

- Overall solution:**

$$\hat{x}(k+1|k) = \Phi \hat{x}(k|k-1) + \Gamma u(k) + K(k) (y(k) - C \hat{x}(k|k-1))$$

$$K(k) = (\Phi P(k) C^T + R_{12}) (C P(k) C^T + R_2)^{-1}$$

$$P(k+1) = \Phi P(k) \Phi^T + R_1 - K(k) (C P(k) C^T + R_2) K^T(k)$$

$$P(0) = R_0$$

$$\hat{x}(0|-1) = m_0$$

## 2.4 Common equivalent implementation

- **Step 0.** (Initialization)

$$P(0|-1) = P(0) = R_0$$

$$\hat{x}(0|-1) = m_0$$

Covariance and mean of initial state.



## 2.4 Common equivalent implementation (continued)

- **Step 1.** (Corrector - use the most recent measurement)

$$\hat{x}(k|k) = \hat{x}(k|k-1) + K(k) (y(k) - C\hat{x}(k|k-1))$$

$$K(k) = P(k|k-1)C^T (CP(k|k-1)C^T + R_2)^{-1}$$

$$P(k|k) = P(k|k-1) - K(k)CP(k|k-1)$$

Update estimate with  $y(k)$ , compute Kalman gain, update error covariance.

- **Step 2.** (One-step predictor)

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k) + \Gamma u(k)$$

$$P(k+1|k) = \Phi P(k|k)\Phi^T + R_1$$

Project the state and error covariance ahead.

Iterate Step 1 and 2, increase  $k$ .

## 2.5 Comments on Kalman filter solution

- From the Kalman gain  $K(k)$  update equation, we see that a large  $R_2$  (much measurement noise) leads to low influence of error on estimate.
- The  $P(k)$  estimation error covariance is a measure of the uncertainty of the estimate. It is updated as

$$P(k+1) = \Phi P(k) \Phi^T + R_1 - K(k) (C P(k) C^T + R_2) K^T(k)$$

- The first two terms on the right represent the natural evolution of the uncertainty, the last term shows how much uncertainty the Kalman filter removes.

## 2.5 Comments on Kalman filter solution (continued)

- The Kalman filter gives an unbiased estimate, i.e.,

$$\mathbf{E} [\hat{x}(k|k)] = \mathbf{E} [\hat{x}(k|k-1)] = \mathbf{E} [x(k)]$$

- If the noise is uncorrelated with  $x(0)$ , then the Kalman filter is optimal, i.e., no other *linear* filter gives a smaller variance of the estimation error.

(For non-Gaussian assumptions, nonlinear filters, particle filters or moving-horizon estimators do a much better job.)

## 2.6 Example

- Discrete-time system  $x(k+1) = x(k)$   
 $y(k) = x(k) + e(k)$   
→ constant state, to be reconstructed from noisy measurements
- Measurement noise  $e$  has standard deviation  $\sigma$  (so  $R_2 = \sigma^2$ )  
 $x(0)$  has covariance  $R_0 = 0.5$   
No process noise  $v$ , so  $R_1 = 0$  and  $R_{12} = 0$
- Kalman filter is given by

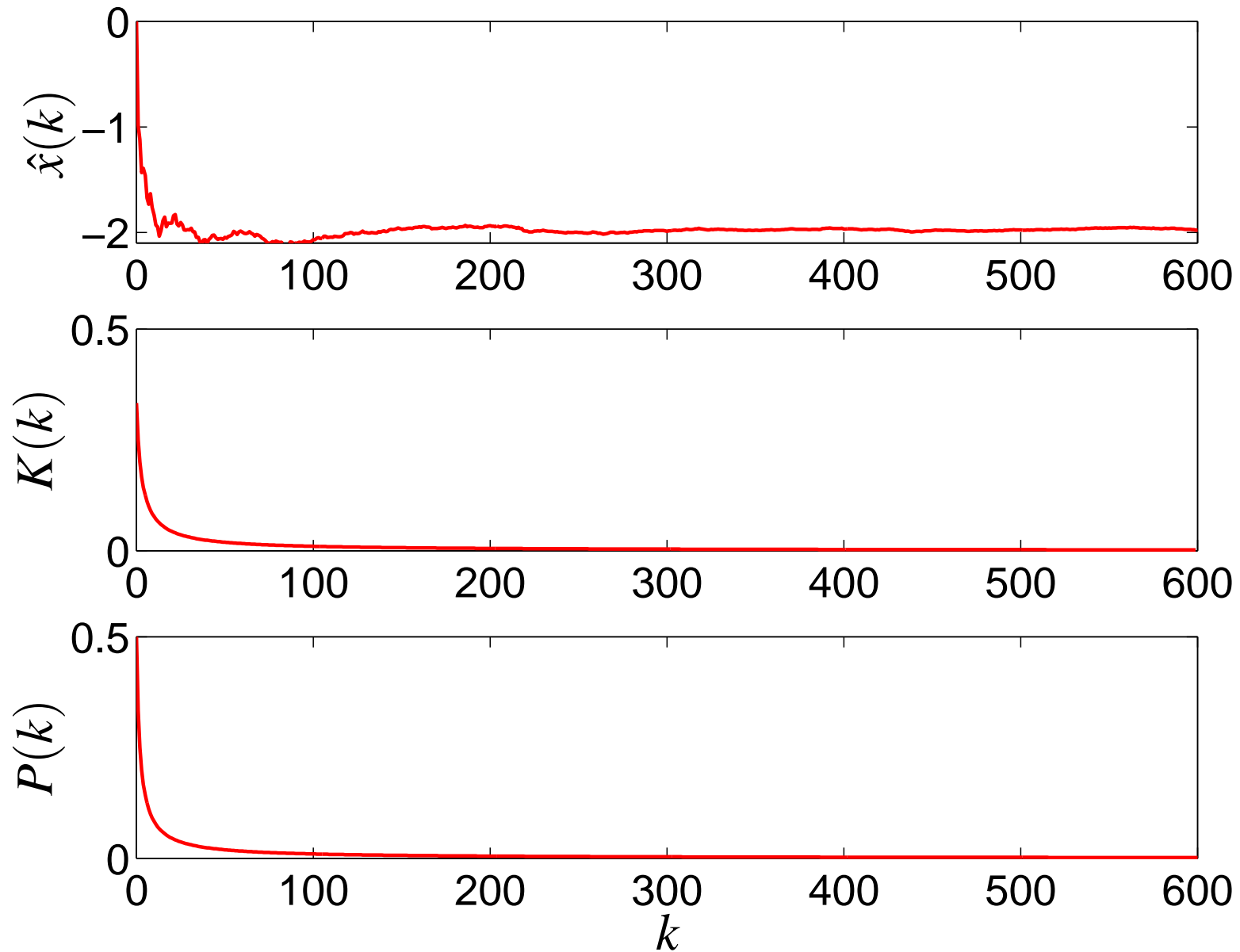
$$\hat{x}(k+1|k) = \hat{x}(k|k-1) + K(k)(y(k) - \hat{x}(k|k-1))$$

$$K(k) = \frac{P(k)}{P(k) + \sigma^2}$$

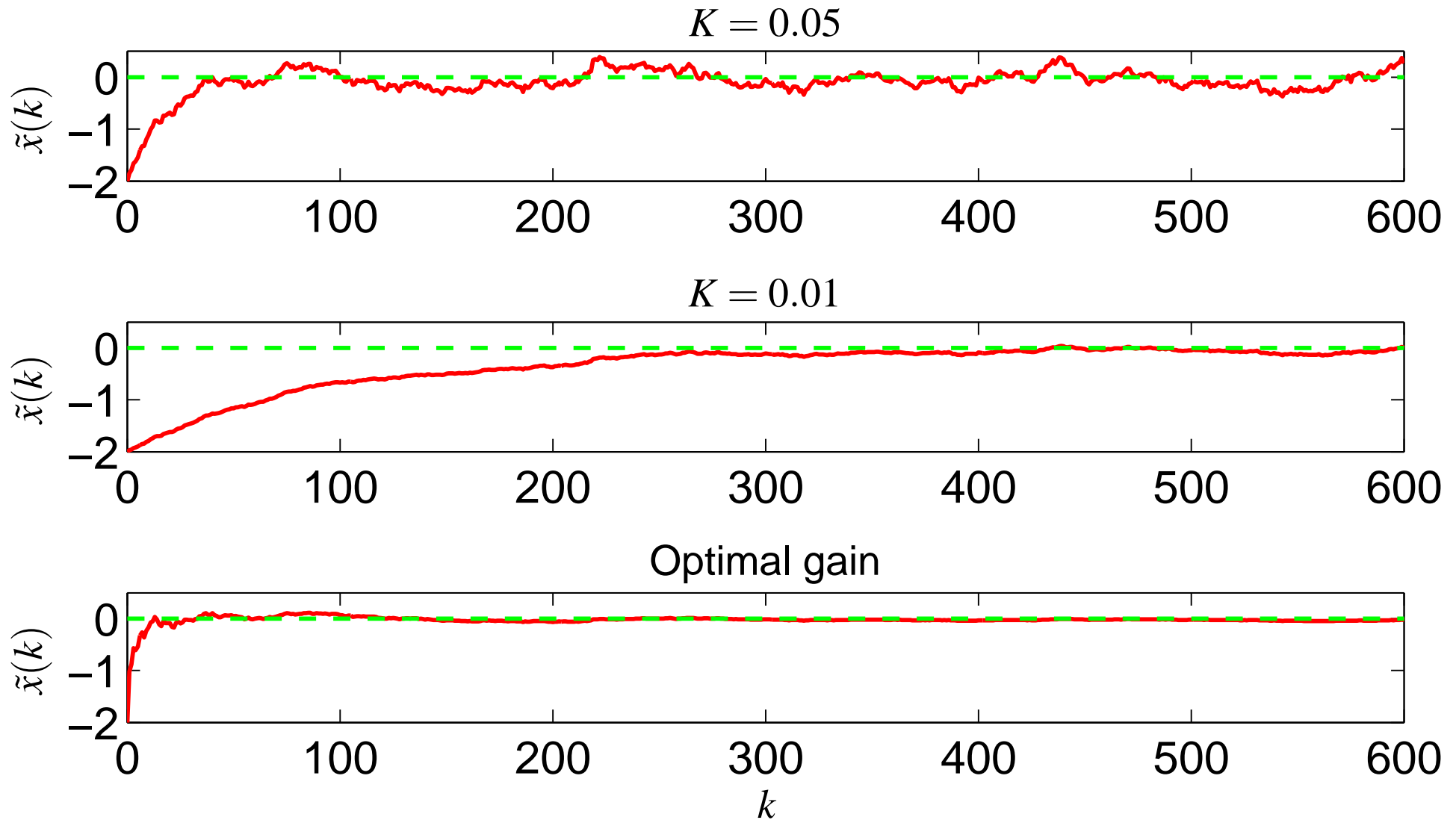
$$P(k+1) = P(k) - K(k)(P(k) + \sigma^2)K^T(k) = \frac{\sigma^2 P(k)}{P(k) + \sigma^2}$$

$$P(0) = R_0 = 0.5 \quad \hat{x}(0|-1) = m_0 = 0$$

## 2.6 Example (continued)



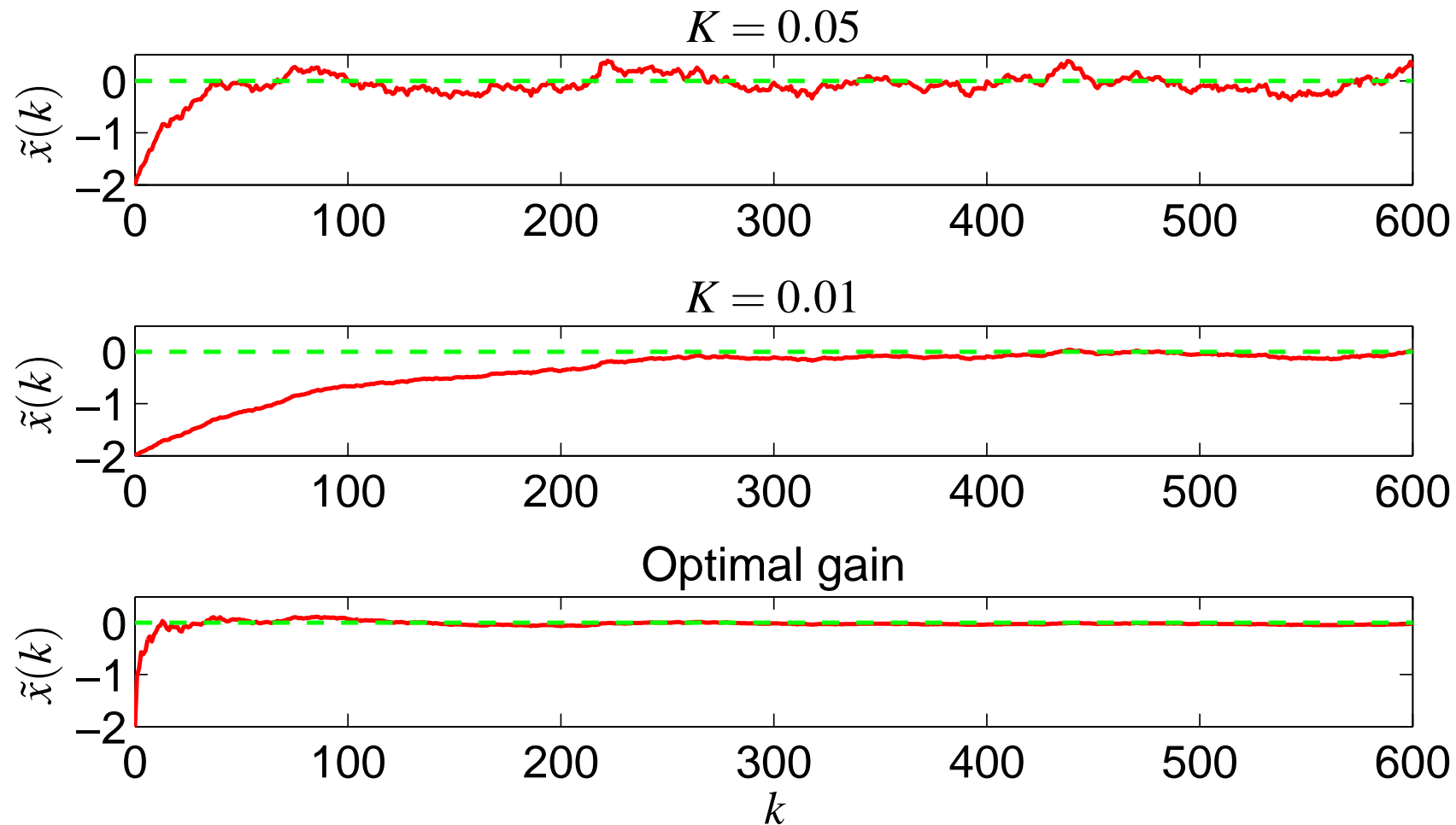
## 2.6 Example (continued)



## Example – To think about

1. What is the problem with a large (small)  $K$  in the observer?
2. What does the Kalman filter do?
3. How would you change  $P(0)$  in the Kalman filter to get a smoother (but slower) transient in  $\tilde{x}(k)$ ?
4. In practice,  $R_1$ ,  $R_2$ , and  $R_0$  are often tuning parameters. What are their influence on the estimate?

## Example – To think about (continued)



- Large fixed  $K \rightarrow$  rapid initial convergence, but large steady-state variance  
Small fixed  $K \rightarrow$  slower convergence, but better performance in steady state



## Steady-state Kalman gain

- Recall:

$$P(k+1) = \Phi P(k) \Phi^\top + R_1 - (\Phi P(k) C^\top + R_{12}) (C P(k) C^\top + R_2)^{-1} (C P(k) \Phi^\top + R_{12}^\top)$$

- Steady-state solution given by

$$\bar{P} = \Phi \bar{P} \Phi^\top + R_1 - (\Phi \bar{P} C^\top + R_{12}) (C \bar{P} C^\top + R_2)^{-1} (C \bar{P} \Phi^\top + R_{12}^\top)$$

→ is also Riccati equation!

- Note: compare with steady-state LQ controller:

$$\bar{S} = \Phi^\top \bar{S} \Phi + Q_1 - (\Phi^\top \bar{S} \Gamma + Q_{12})^\top (\Gamma^\top \bar{S} \Gamma + Q_2)^{-1} (\Gamma^\top \bar{S} \Phi + Q_{12}^\top)$$

## Duality

- Equivalence between LQ control problem and Kalman filter state estimation problem

LQ	Kalman
$k$	$N - k$
$\Phi$	$\Phi^T$
$\Gamma$	$C^T$
$Q_0$	$R_0$
$Q_1$	$R_1$
$Q_{12}$	$R_{12}$
$S$	$P$
$L$	$K^T$

## How to find Kalman filter using matlab?

- Command  $[KEST, L, P] = \text{kalman}(\text{SYS}, QN, RN, NN)$
- Calculates (full) Kalman estimator KEST for system SYS

$$x[n+1] = Ax[n] + Bu[n] + Gw[n] \quad \{\text{State equation}\}$$

$$y[n] = Cx[n] + Du[n] + Hw[n] + v[n] \quad \{\text{Measurements}\}$$

with known inputs  $u$ , process noise  $w$ , measurement noise  $v$ , and noise covariances

$$E\{ww'\} = QN, \quad E\{vv'\} = RN, \quad E\{wv'\} = NN,$$

Note: Construct SYS as  $\text{SYS} = \text{ss}(A, [B \ G], C, [D \ H], -1)$

- Also returns steady-state estimator gain  $L$  and steady-state error covariance

$$P = E\{(x - x[n|n-1])(x - x[n|n-1])'\} \quad (\text{Riccati solution})$$

### 3. Linear quadratic Gaussian control

- Given discrete-time LTI system

$$\begin{aligned}x(k+1) &= \Phi x(k) + \Gamma u(k) + v(k) \\ y(k) &= Cx(k) + e(k)\end{aligned}$$

with

$$\mathbb{E}[x(0)] = m_0, \quad \text{cov}(x(0)) = R_0, \quad \mathbb{E} \begin{bmatrix} v(k) \\ e(k) \end{bmatrix} \begin{bmatrix} v(k) \\ e(k) \end{bmatrix}^T = \begin{bmatrix} R_1 & R_{12} \\ R_{12}^T & R_2 \end{bmatrix}$$

find linear control law  $y(0), y(1), \dots, y(k-1) \mapsto u(k)$  that minimizes

$$\mathbb{E} \left[ \sum_{k=0}^{N-1} \left( x^T Q_1 x + 2x^T Q_{12} u + u^T Q_2 u \right) + x^T(N) Q_0 x(N) \right]$$

- Solution: Separation principle

## 3.1 Separation principle

- Makes it possible to use control law

$$u(k) = -L\hat{x}(k|k-1)$$

(so  $u(k) = -Lx(k) + L\tilde{x}(k)$ ) with closed-loop dynamics

$$x(k+1) = \Phi x(k) - \Gamma Lx(k) + \Gamma L\tilde{x}(k) + v(k)$$

and to view term  $\Gamma L\tilde{x}(k)$  as part of noise

→ solve LQ problem and estimation problem separately

## Proof of separation principle

- Solution of optimal observer design problem does not depend on input  $u$

So using state feedback does not influence optimality

→ Kalman filter still optimal

- Using  $u(k) = -L(k)\hat{x}(k|k-1)$  results in closed-loop system

$$\hat{x}(k+1|k) = (\Phi - \Gamma L(k))\hat{x}(k|k-1) + K(k) \underbrace{(y - C\hat{x}(k|k-1))}_{w(k)}$$

- It can be shown that for optimal  $K(k)$ ,  $w(k)$  is white noise
- So dynamics become

$$\hat{x}(k+1|k) = (\Phi - \Gamma L(k))\hat{x}(k|k-1) + K(k)w(k)$$

## Proof of separation principle (continued)

- For simplicity we assume  $Q_0 = 0$
- For  $u = -L\hat{x}$ , control design problem in terms of  $\hat{x}$  and  $\tilde{x}$  becomes

$$\begin{aligned} \min_L \mathbb{E} & \left[ \sum_{k=0}^{N-1} x^\top Q_1 x + 2x^\top Q_{12}u + u^\top Q_2 u \right] \\ &= \min_L \mathbb{E} \left[ \sum_{k=0}^{N-1} (\hat{x} + \tilde{x})^\top Q_1 (\hat{x} + \tilde{x}) + 2(\hat{x} + \tilde{x})^\top Q_{12}L\hat{x} + \hat{x}^\top L^\top Q_2 L\hat{x} \right] \end{aligned}$$

- Since it can be shown that  $\mathbb{E} [\tilde{x}^\top Q \hat{x}] = 0$ , we get

$$\min_L \mathbb{E} \left[ \sum_{k=0}^{N-1} \hat{x}^\top Q_1 \hat{x} + \tilde{x}^\top Q_1 \tilde{x} + 2\hat{x}^\top Q_{12}L\hat{x} + \hat{x}^\top L^\top Q_2 L\hat{x} \right]$$

- $\mathbb{E} [\tilde{x}^\top Q \tilde{x}]$  does not depend on  $L$  and is in fact minimal for Kalman gain  $K$

## Proof of separation principle (continued)

- Hence, we get

$$\min_L \mathbf{E} \left[ \sum_{k=0}^{N-1} \hat{x}^\top Q_1 \hat{x} + 2\hat{x}^\top Q_{12} L \hat{x} + \hat{x}^\top L^\top Q_2 L \hat{x} \right]$$

subject to

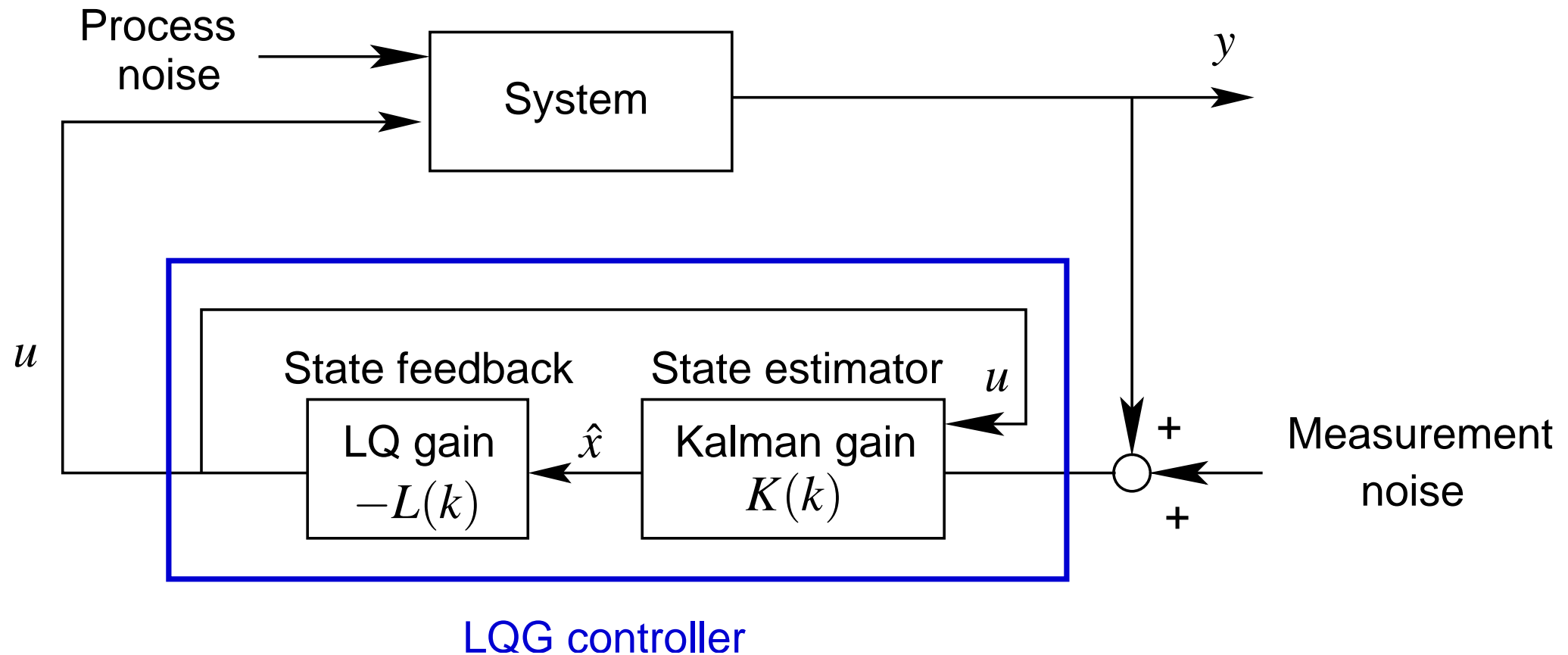
$$\hat{x}(k+1|k) = (\Phi - \Gamma L(k))\hat{x}(k|k-1) + K(k)w(k)$$

= stochastic LQ problem (but with  $\hat{x}$  instead of  $x$  and  
with  $u = -L\hat{x}$  already filled in)

→  $L(k)$  as computed before still optimal



## 3.2 LQG problem: Solution



## Stationary LQG control

- Solution:

$$u(k) = -L\hat{x}(k|k-1)$$

with

$$\hat{x}(k+1|k) = \Phi\hat{x}(k|k-1) + \Gamma u(k) + K(y(k) - C\hat{x}(k|k-1))$$

- Closed-loop dynamics (with error state  $\tilde{x}(k)$ , see slide [ce\\_kf\\_lqg.8](#)):

$$x(k+1) = (\Phi - \Gamma L)x(k) + \Gamma L\tilde{x}(k) + v(k)$$

$$\tilde{x}(k+1) = (\Phi - KC)\tilde{x}(k) + v(k) - Ke(k)$$

or

$$\begin{bmatrix} x(k+1) \\ \tilde{x}(k+1) \end{bmatrix} = \begin{bmatrix} \Phi - \Gamma L & \Gamma L \\ 0 & \Phi - KC \end{bmatrix} \begin{bmatrix} x(k) \\ \tilde{x}(k) \end{bmatrix} + \begin{bmatrix} I \\ I \end{bmatrix} v(k) + \begin{bmatrix} 0 \\ -K \end{bmatrix} e(k)$$

→ dynamics of closed-loop system determined by dynamics of LQ controller and of optimal filter

## How to construct LQG controller using matlab?

- Command `RLQG = lqgreg(KEST,K)` produces LQG controller by connecting Kalman estimator `KEST` designed with `kalman` and state-feedback gain `K` designed with `dlqr`

## Pros and cons of LQG

- + stabilizing
- + good robustness for SISO
- + works for MIMO
- robustness can be very bad for MIMO
- high-order controller
- how to choose weights?

## Summary

- Kalman filtering
  - state estimator such that error covariance is minimized
- LQG control
  - separation principle  $\rightarrow$  LQ + Kalman