Lecture 13

Linear quadratic Gaussian control and the extended Kalman filter

Giancarlo Ferrari Trecate[1]

[1]Dependable Control and Decision Group
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland
giancarlo.ferraritrecate@epfl.ch

Linear Quadratic Gaussian (LQG) control

Combination of LQ control with a (steady-state) KF

Problem

The system is stochastic. What "LQ" means?

- System:

$$
\begin{aligned}
x^+ &= Ax + Bu + w \qquad & w &\sim WGN(0, W) \qquad & W &\geq 0 \\
y &= Cx + v \qquad & v &\sim WGN(0, V) \qquad & V &> 0 \\
x_0 &\sim N(\bar{x}_0, \Sigma_0)
\end{aligned}
$$

- x_0, w_k, v_j uncorrelated

LQ cost

$$J = \lim_{N \to +\infty} \frac{1}{N} E\left[\Sigma_{k=0}^{N-1} x_k^T Q x_k + u_k^T R u_k\right]$$

to be minimized <u>without</u> measuring x_k

Idea

Under the assumption guaranteeing stability of IH-LQ and steady-state KF, one can use a "separation principle" for control design

Algorithm

1) Use y_k and u_k for computing $\hat{x}_{k+1|k}$ using the stationary Kalman predictor gain $\bar{L}$
2) Compute the optimal control law (gain $\bar{K}$)
3) Apply the control law

$$u_{k+1} = -\bar{K}\hat{x}_{k+1|k}$$

Main results

- The closed-loop system is AS with eigenvalues

$$\mathrm{Spec}(A - B\bar{K}) \cup \mathrm{Spec}(A - \bar{L}C)$$

- The cost J in the previous slide is minimized

Extended Kalman predictor (known as EKF-Extended Kalman Filter)

Nonlinear system

$$\Sigma_{NL} : \begin{cases} x_{k+1} = f(x_k, u_k) + w_k \\ y_k = g(x_k) + v_k \\ x_0 \sim N(\tilde{x}_0, \Sigma_0) \end{cases}$$

$x_k \in \mathbb{R}^n$

$u_k \in \mathbb{R}^m$

$y_k \in \mathbb{R}^p$

Assumptions:

- f, g are differentiable with continuity
- $w_k \sim WGN(0, W)$, $v_k \sim WGN(0, V)$, x_0, v_i, w_j independent, $\forall i, j$

Problem: how to account for the NL dynamics in KF?

First (naive) solution: linearized Kalman predictor

Let $\bar{x}_k, \bar{y}_k$ be the states/outputs in a noiseless nominal experiment

$$\bar{x}_{k+1} = f(\bar{x}_k, u_k)$$
$$\bar{y}_k = g(\bar{x}_k)$$
$$\bar{x}_0 = \tilde{x}_0$$

Let $\delta x_k = x_k - \bar{x}_k$, $\delta y_k = y_k - \bar{y}_k$ and linearize Σ_{NL} at each time k about $\bar{x}_k$ and $\bar{y}_k$

$$\overbrace{\cancel{\bar{x}_{k+1}} + \delta x_{k+1}}^{x_{k+1}} \simeq \underbrace{\cancel{f(\bar{x}_k, u_k)} + \overbrace{\left.\frac{\partial f(x,u)}{\partial x}\right|_{\bar{x}_k, u_k}}^{\bar{A}_k}}_{} \cdot \delta x_k + w_k$$

→ 1st order Taylor approximation of f about $\bar{x}_k, u_k$)

$$\overbrace{\cancel{\bar{y}_k} + \delta y_k}^{y_k} \simeq \cancel{g(\bar{x}_k)} + \underbrace{\left.\frac{\partial g(x)}{\partial x}\right|_{\bar{x}_k}}_{\bar{C}_k} \cdot \delta x_k + v_k$$

← 1st order Taylor approx. of g about $\bar{x}_k$

Then:

$$\Sigma_L : \begin{cases} \delta x_{k+1} = \bar{A}_k \delta x_k + w_k \\ \delta y_k = \bar{C}_k \delta x_k + v_k \\ \delta x_0 \sim N(0, \Sigma_0) \end{cases}$$

→ Linear Time Varying system!

Idea

- Design a KF for Σ_L

KF for the time-varying system Σ_L

Init. $\delta\hat{x}_{0|-1} = 0, \bar{\Sigma}_{0|-1} = \Sigma_0$ (statistics of δx_0)

Filtering step for δx

- Filtering step

$$\delta\hat{x}_{k|k} = \delta\hat{x}_{k|k-1} + \bar{L}_{k|k}(\delta y_k - \bar{C}_k \delta\hat{x}_{k|k-1}) \qquad (1)$$

$$\bar{L}_{k|k} = \dots, \bar{\Sigma}_{k|k} = \dots \text{ usual KF updates based on } \bar{A}_k \text{ and } \bar{C}_k$$

Remark: Since $\delta\hat{x}_{k|k} = \hat{x}_{k|k} - \bar{x}_k$ and $\delta\hat{x}_{k|k-1} = \hat{x}_{k|k-1} - \bar{x}_k$, (1) gives

$$\hat{x}_{k|k} - \cancel{\bar{x}_k} = \hat{x}_{k|k-1} - \cancel{\bar{x}_k} + \bar{L}_{k|k}\left(\overbrace{y_k}^{\delta y_k} - \underbrace{\left(g(\bar{x}_k) + \overbrace{\frac{\partial g(x)}{\partial x}}^{\bar{C}_k}\Bigg|_{\bar{x}_k} \delta\hat{x}_{k|k-1}\right)}_{\approx g(\hat{x}_{k|k-1})}\right)$$

$$\hat{x}_{k|k} \approx \hat{x}_{k|k-1} + \bar{L}_{k|k}\left(y_k - g(\hat{x}_{k|k-1})\right) \qquad (*)$$

→ Filtering step for the real state

- Prediction step

$$\delta\hat{x}_{k+1|k} = \bar{A}_k \delta\hat{x}_{k|k} \tag{2}$$

$$\bar{\Sigma}_{k+1|k} = \dots \text{ usual KF update based on } \bar{A}_k$$

Remark: (2) gives

$$\hat{x}_{k+1|k} - \underbrace{\bar{x}_{k+1}}_{f(\bar{x}_k,u_k)} = \bar{A}_k\left(\hat{x}_{k|k} - \bar{x}_k\right)$$

$$\hat{x}_{k+1|k} = \underbrace{f(\bar{x}_k, u_k) + \left.\frac{\partial f(x,u)}{\partial x}\right|_{\bar{x}_k, u_k}\left(\hat{x}_{k|k} - \bar{x}_k\right)}_{\approx f\left(\hat{x}_{k|k}, u_k\right)} \tag{**}$$

Using (∗) and (∗∗), we introduce the following filter:

Linearized KF

Init. $\hat{x}_{0|-1} \in \mathbb{R}^n, \Sigma_{0|-1} \in \mathbb{R}^{n\times n}$ (statistics of x_0)

- Filtering step

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \bar{L}_{k|k}(y_k - g(\hat{x}_{k|k-1}))$$
$$\bar{L}_{k|k} = \bar{\Sigma}_{k|k-1}\bar{C}_k^T(\bar{C}_k\bar{\Sigma}_{k|k-1}\bar{C}_k^T + V)^{-1}$$
$$\bar{\Sigma}_{k|k} = \bar{\Sigma}_{k|k-1} - \bar{\Sigma}_{k|k-1}\bar{C}_k^T(\bar{C}_k\bar{\Sigma}_{k|k-1}\bar{C}_k^T + V)^{-1}\bar{C}_k\bar{\Sigma}_{k|k-1}$$

- Prediction step

$$\hat{x}_{k+1|k} = f(\hat{x}_{k|k}, u_k)$$
$$\bar{\Sigma}_{k+1|k} = \bar{A}_k\bar{\Sigma}_{k|k}\bar{A}_k^T + W$$

Pros

$\bar{L}_{k|k}, \bar{\Sigma}_{k|k}$ can be computed in advance as usual

Cons

x_k can be very different from $\bar{x}_k$ because of noise
$\hookrightarrow$ the linearized KF utilizes the wrong model for computing $\bar{L}_{k|k}$ $\Rightarrow$ performance can be poor

Next: overcome this problem

Extended Kalman predictor (EKF)

Idea

Use the previous algorithm with $\bar{L}_k$ replaced by the gain $\hat{L}_k$ obtained by linearizing the system about $\hat{x}_k$ (the most recent estimate)

Formulae

$$\hat{A}_k = \left.\frac{\partial f(x,u)}{\partial x}\right|_{\hat{x}_{k-1|k-1},u_k} \qquad \hat{C}_k = \left.\frac{\partial g(x)}{\partial x}\right|_{\hat{x}_{k|k-1}}$$

Replace $\bar{A}_k$ with $\hat{A}_k$ and $\bar{C}_k$ with $\hat{C}_k$ in the linearized KF

Why $\hat{x}_{k|k-1}$ and not $\hat{x}_{k|k}$ when computing $\hat{C}_k$?

- For computing $\hat{x}_{k|k}$ one needs $\hat{L}_{k|k}$ and hence $\hat{C}_k$. If $\hat{C}_k$ depends upon $\hat{x}_{k|k}$ there is a circular dependence

Remark

The EKF is not optimal... but, often, it is OK in applications.

Estimation of unknown parameters with the EKF

System

$$x_{k+1} = f(x_k, u_k, \alpha) + w_k$$
$$y_k = g(x_k, \alpha) + v_k$$

$\alpha \in \mathbb{R}$: unknown parameter

Trick

1) Extend the state with the fictitious state

$$\alpha_{k+1} = \alpha_k \qquad (\square)$$

Set $\tilde{x} = \begin{bmatrix} x \\ \alpha \end{bmatrix}$ and obtain the extended system

$$\tilde{\Sigma} : \begin{cases} \tilde{x}_{k+1} = \tilde{f}(\tilde{x}_k, u_k) + \tilde{w}_k \\ y_k = \tilde{g}(\tilde{x}_k) + v_k \end{cases}$$

$\tilde{f}, \tilde{g}, \tilde{w}_k$ easily obtained from the original system and $(\square)$

2) Apply EKF to $\tilde{\Sigma}$ to estimate x, α

Remark

Even if f and g are linear in x and u, $\tilde{f}$ and $\tilde{g}$ are not

Example

$$\begin{cases} x^+ = \begin{bmatrix} 0.1 & \alpha \\ 0 & 0.2 \end{bmatrix} x = f(x, \alpha) \\ y = x_1 \end{cases}$$

Extended state equations

$$\left.\begin{aligned} x_1^+ &= 0.1x_1 + \underset{\substack{\downarrow \\ \text{nonlinear!}}}{\alpha x_2} \\ x_2^+ &= 0.2x_2 \\ \alpha^+ &= \alpha \end{aligned}\right\} \tilde{f}\left(\begin{bmatrix} x \\ \alpha \end{bmatrix}\right)$$

Example - difference between linearized and extended KF

$x_k \in \mathbb{R}$

$$x_{k+1} = x_k^2 + w_k \qquad w_k \sim WGN(0, 0.1)$$
$$y_k = x_k^3 + v_k \qquad v_k \sim WGN(0, 0.2)$$
$$x_0 \sim N(1, 1)$$

- Nominal state $x_k = \bar{x} = 1, \forall k \geq 0$
  Compute the linearized and extended KF

Linearized KF

$$\bar{A}_k = \frac{\partial x_k^2}{\partial x_k}\bigg|_{x_k=\bar{x}} = 2\, x_k \,\big|_{x_k=\bar{x}} = 2$$

$$\bar{C}_k = \frac{\partial x_k^3}{\partial x_k}\bigg|_{x_k=\bar{x}} = 3\, x_k^2 \,\big|_{x_k=\bar{x}} = 3$$

Linearized KF (check @home)

Filtering

$$
\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + \bar{L}_{k|k}(y_k - g(\hat{x}_{k|k-1})) \\
\bar{L}_{k|k} &= \bar{\Sigma}_{k|k-1} \cdot \underset{\substack{\downarrow \\ \bar{C}_k}}{3} (\underset{\substack{\downarrow \\ \bar{C}_k \bar{C}_k^T}}{9} \bar{\Sigma}_{k|k-1} + \underset{\substack{\downarrow \\ V}}{0.2})^{-1} \\
\bar{\Sigma}_{k|k} &= \bar{\Sigma}_{k|k-1} - \bar{\Sigma}^2_{k|k-1} \cdot \underset{\substack{\downarrow \\ \bar{C}_k \bar{C}_k^T}}{9} (\underset{\substack{\downarrow \\ \bar{C}_k \bar{C}_k^T}}{9} \bar{\Sigma}_{k|k-1} + 0.2)^{-1}
\end{aligned}
$$

Prediction

$$
\begin{aligned}
\hat{x}_{k+1|k} &= \hat{x}^2_{k|k} \\
\bar{\Sigma}_{k+1|k} &= \underset{\substack{\downarrow \\ \bar{A}_k \bar{A}_k^T}}{4} \bar{\Sigma}_{k|k} + \underset{\substack{\downarrow \\ W}}{0.1}
\end{aligned}
$$

EKF

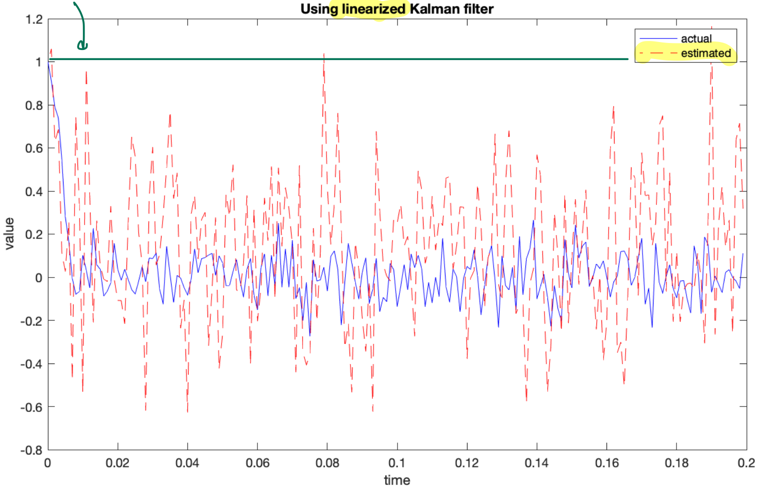
$$\hat{A}_k = \frac{\partial x_k^2}{\partial x_k}\Bigg|_{x_k=\hat{x}_{k-1|k-1}} = 2\hat{x}_{k-1|k-1}$$

$$\hat{C}_k = \frac{\partial x_k^3}{\partial x_k}\Bigg|_{x_k=\hat{x}_{k|k-1}} = 3\hat{x}_{k|k-1}^2$$

$\hookrightarrow$ replace these matrices in the previous formulae

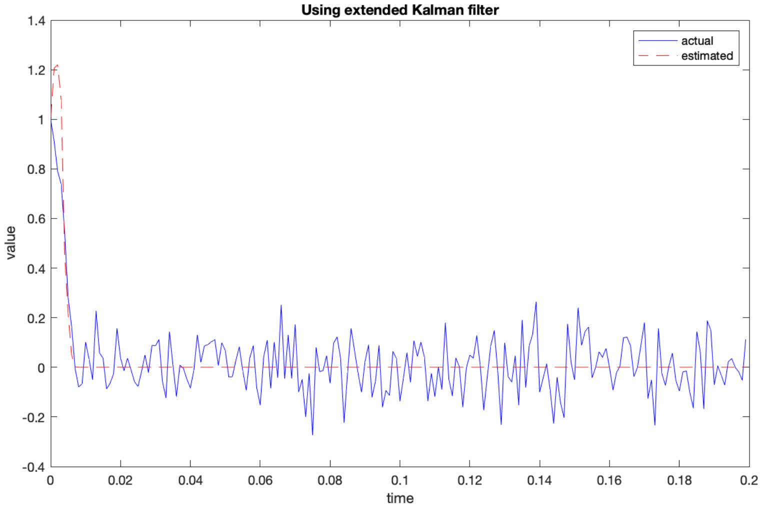
- $\hat{A}_k$ and $\hat{C}_k$ are time-varying while $\bar{A}_k$ and $\bar{C}_k$ are not!

nominal state



$$\underbrace{\frac{1}{N_s}\sum_{k=1}^{N_s}\left|\hat{x}_{k|k}-x_k\right|}_{} = 0.3051$$

sample standard deviation of the estimation error



$$\frac{1}{N_s}\sum_{k=1}^{N_s} \left|\hat{x}_{k|k} - x_k\right| = 0.0866$$

Example: Predator-prey system[1] (check @home)

Bio-system:

- Prey population: $x_{k,1}$
- Predator population: $x_{k,2}$

Model:

$$x_{k+1,1} = \left[1 + \Delta t \left(1 - \frac{x_{k,2}}{c_2}\right)\right] x_{k,1} + w_{k,1}$$

$$x_{k+1,2} = \left[1 - \Delta t \left(1 - \frac{x_{k,1}}{c_1}\right)\right] x_{k,2} + w_{k,2}$$

where Δt is the time step, c_1 and c_2 are constant model parameters and $\begin{bmatrix} w_{1,k} \\ w_{2,k} \end{bmatrix} \sim WGN(0, W)$.

Objective: estimate the predator and prey populations based on noisy measurements of the total population

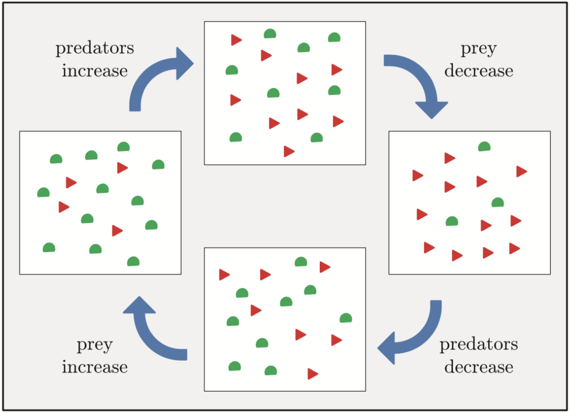
$$y_k = x_{k,1} + x_{k,2} + v_k$$

where $v_k \sim WGN(0, V)$

[1] N. Kovvali, M. Banavar, A. Spanias. *An Introduction to Kalman Filtering with MATLAB Examples.*

Example: Predator-prey system

→ There is an interaction between the two populations



→ When there is an abundance of preys, the predator population increases
→ This increase in predators drives the prey population down
→ With a scarcity of preys, the predator population is forced to decrease
→ The decrease in predators then results in an increase of the number of preys

Example: Predator-prey system

$$x_{k+1} = \begin{bmatrix} x_{k+1,1} \\ x_{k+1,2} \end{bmatrix} = f\left(\begin{bmatrix} x_{k,1} \\ x_{k,2} \end{bmatrix}\right) + \begin{bmatrix} w_{k,1} \\ w_{k,2} \end{bmatrix}$$
$$y_k = C \begin{bmatrix} x_{k,1} \\ x_{k,2} \end{bmatrix} + v_k$$

This state-space model is **non linear** and Gaussian $\rightarrow$ use EKF for state estimation

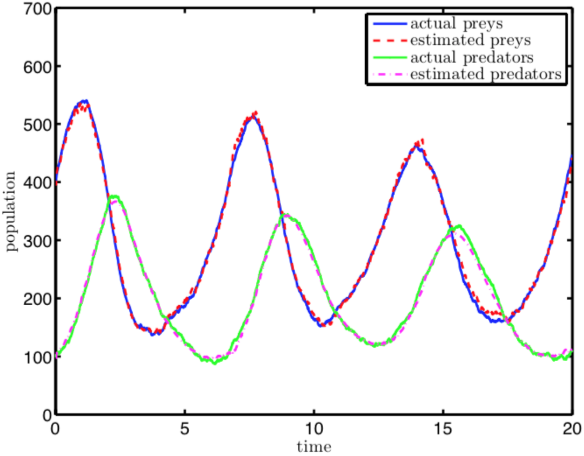
Extended KF:

$$\hat{A}_k = \begin{bmatrix} 1 + \Delta t \left(1 - \frac{x_{k-1|k-1,2}}{c_2}\right) & -\Delta t \frac{x_{k-1|k-1,1}}{c_2} \\ \Delta t \frac{x_{k-1|k-1,2}}{c_1} & 1 + \Delta t \left(1 - \frac{x_{k-1|k-1,1}}{c_1}\right) \end{bmatrix}$$

Example: Predator-Prey system

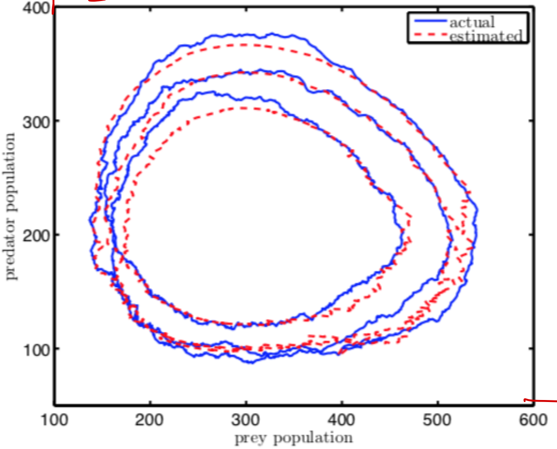
Simulations:

- $\Delta t = 0.01$, $c_1 = 300$, $c_2 = 200$, $W = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$ and $V = 100$
- Initial populations: $x_0 = \left[\begin{smallmatrix} 400 \\ 100 \end{smallmatrix}\right]$
- Initial condition for estimator: $\hat{x}_{0|0} \sim N\left(x_0, \left[\begin{smallmatrix} 100 & 0 \\ 0 & 100 \end{smallmatrix}\right]\right)$



Example: Predator-Prey system

x_2



x_1

The EKF is able to track the predator-prey dynamics with good accuracy

2h L12