Optimization-Based Control

Richard M. Murray
Control and Dynamical Systems
California Institute of Technology

Version v2.3h (20 Feb 2023)

Chapter 6

Kalman Filtering

In this chapter we derive the optimal estimator for a linear system in continuous time (also referred to as the Kalman-Bucy filter). This estimator minimizes the covariance and can be implemented as a recursive filter. We also show how to combine optimal estimation with state feedback to solve the linear quadratic Gaussian (LQG) control problem, and explore extensions of Kalman filtering for continuous time systems, such as the extended Kalman filter. Optimal estimation of discrete time systems is described in more detail in Chapter 7, in the context of sensor fusion.

Prerequisites. Readers should have basic familiarity with continuous-time stochastic systems at the level presented in Chapter 5 as well as the material in FBS2e, Chapter 8 on state space observability and estimators.

6.1 Linear Quadratic Estimators

Consider a stochastic system

$$\dot{X} = AX + Bu + FV, \qquad Y \quad = CX + W,$$

where X represents that state, u is the (deterministic) input, V represents disturbances that affect the dynamics of the system and W represents measurement noise. Assume that the disturbance V and noise W are zero-mean, Gaussian white noise (but not necessarily stationary):

$$\begin{aligned} p(w) &= \frac{1}{\sqrt{\det(2\pi R_V)}} e^{-\frac{1}{2} w^{\mathsf{T}} R_V^{-1} w} \qquad & \mathbb{E}(V(t_1)V^{\mathsf{T}}(t_2)) &= R_V(t_1)\delta(t_2 - t_1) \\ p(v) &= \frac{1}{\sqrt{\det(2\pi R_V)}} e^{-\frac{1}{2} v^{\mathsf{T}} R_W^{-1} v} \qquad & \mathbb{E}(W(t_1)W^{\mathsf{T}}(t_2)) &= R_W(t_1)\delta(t_2 - t_1) \end{aligned}$$

We also assume that the cross correlation between V and W is zero, so that the disturbances are not correlated with the noise. Note that we use multi-variable Gaussians here, with noise intensities $R_V \in \mathbb{R}^{m \times m}$ and $R_W \in \mathbb{R}^{p \times p}$. In the scalar case, $R_V = \sigma_V^2$ and $R_W = \sigma_W^2$.

We formulate the optimal estimation problem as finding the estimate $\hat{x}(t)$ that minimizes the mean square error $\mathbb{E}((x(t) - \hat{X}(t))(X(t) - \hat{x}(t))^\mathsf{T})$ given $\{y(\tau) : 0 \leq \tau \leq t\}$ where X and $\hat{X}$ satisfy the dynamics for the system and y is the measured outputs of the system. Note that our system state is not known, but we do have a description of X as a random process, and hence we can reason over the distribution of possible states of that process that are consistent with the output measurements.

The estimation problem be viewed as solving a *least squares* problem: given all previous $y(t)$, find the estimate $\hat{X}(t)$ that satisfies the dynamics and minimizes the square error between the system state and the estimated state. It can be shown that this is equivalent to finding the expected value of X subject to the "constraint" given by all of the previous measurements, so that $\hat{X}(t) = \mathbb{E}(X(t) \mid Y(\tau), \tau \leq t)$. (This was the way that Kalman originally formulated the problem, and is explored in Exercise 6.1.)

The following theorem provides the solution to the optimal estimation problem for a linear system driven by disturbances and noise that are modeled as white noise processes.

Theorem 6.1 (Kalman-Bucy, 1961)**.** *The optimal estimator has the form of a linear observer*

$$\dot{\hat{x}} = A\hat{x} + Bu - L(C\hat{x} - y)$$

where $L(t) = P(t)C^\mathsf{T} R_W^{-1}$ *and* $P(t) = \mathbb{E}((X(t) - \hat{x}(t))(X(t) - \hat{x}(t))^\mathsf{T})$ *satisfies*

$$\begin{aligned}\dot{P} &= AP + PA^\mathsf{T} - PC^\mathsf{T} R_W^{-1}(t) CP + FR_V(t)F^\mathsf{T},\\ P(0) &= \mathbb{E}(X(0)X^\mathsf{T}(0)).\end{aligned}$$

Sketch of proof. The error dynamics are given by

$$\dot{E} = (A - LC)E + \xi, \qquad \xi = FV - LW, \qquad R_\xi = FR_V F^\mathsf{T} + LR_W L^\mathsf{T}$$

The covariance matrix $P_E = P$ for this process satisfies

$$\begin{aligned}\dot{P} &= (A - LC)P + P(A - LC)^\mathsf{T} + FR_V F^\mathsf{T} + LR_W L^\mathsf{T}\\ &= AP + PA^\mathsf{T} + FR_V F^\mathsf{T} - LCP - PC^\mathsf{T} L^\mathsf{T} + LR_W L^\mathsf{T}\\ &= AP + PA^\mathsf{T} + FR_V F^\mathsf{T} + (LR_W - PC^\mathsf{T}) R_W^{-1} (LR_W - PC^\mathsf{T})^\mathsf{T}\\ &\qquad - PC^\mathsf{T} R_W^{-1} CP,\end{aligned}$$

where the last line follows by completing the square. We need to find L such that $P(t)$ is as small as possible, which can be done by choosing L so that $\dot{P}$ decreases by the maximum amount possible at each instant in time. This is accomplished by setting

$$LR_W = PC^\mathsf{T} \qquad \Longrightarrow \qquad L = PC^\mathsf{T} R_W^{-1},$$

and the final form of the update law for P follows by substitution of L. □

Note that the Kalman filter has the form of a *recursive* filter: given $P(t) = \mathbb{E}(E(t)E^\mathsf{T}(t))$ at time t, can compute how the estimate and covariance *change*. Thus we do not need to keep track of old values of the output. Furthermore, the Kalman filter gives the estimate $\hat{X}(t)$ *and* the covariance $P_E(t)$, so you can see how well the error is converging.

Another observation is that form of the covariance update can be considered to consist of a *prediction* step and a *correction* step. If we had no information about the output, then the covariance matrix would change just as in the case of the stochastic response from Chapter 5:

$$\dot{P} = AP + PA^{\mathsf{T}} + FR_V(t)F^{\mathsf{T}}.$$

If A is stable then the first two terms tend to decrease the error covariance, but the third term will increase the covariance (because of the effect of disturbances). The remaining term in the covariance update is

$$-PC^{\mathsf{T}}R_W^{-1}(t)CP,$$

which we can regard as a correction term due to the feedback term $-L(C\hat{x} - y)$. This term decreases the covariance (because we have new data), but the amount to which it does so is limited by the noisiness of the measurement (hence the scaling by R_W^{-1}).

Example 6.1 First-order system
Consider a first-order linear system of the form

$$\dot{X} = -aX + V, \qquad Y = cX + W,$$

where V is white noise with variance σ_V^2 and W is white noise with variance σ_W^2. The optimal estimator has the form

$$\dot{\hat{x}} = -a\hat{x} - L(\hat{x} - y) \quad \text{where} \quad L = p(t)c/\sigma_W^2,$$

and the error covariance $p(t)$ satisfies the differential equation

$$\dot{p} = -2ap - \frac{c^2p^2}{\sigma_W^2} + \sigma_V^2, \qquad p(0) = \mathbb{E}(x(0)^2).$$

Figure 6.1 shows a sample plot of $p(t)$ and the estimate $\hat{x}$ versus x for an instance of the noise and disturbance signals. We see that while there is a large initial error in the state estimate, it quickly reduces the error and then (roughly) tracks the state of the underlying (noisy) process. (Since the disturbances are large and unknown, it is not possible to exactly track the actual system state.) ∇

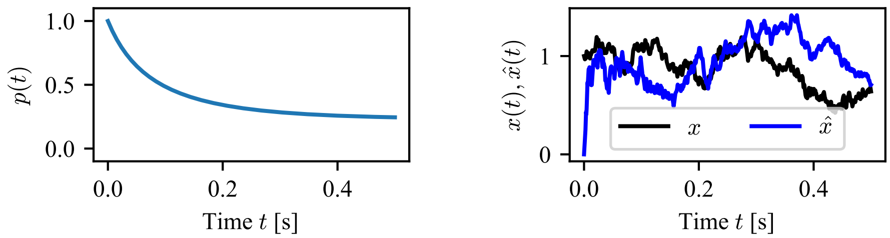
If the noise is stationary (R_V, R_W constant) and if the dynamics for $P(t)$ are stable, then the observer gain converges to a constant and satisfies the *algebraic Riccati equation*:

$$L = PC^{\mathsf{T}}R_W^{-1} \qquad AP + PA^{\mathsf{T}} - PC^{\mathsf{T}}R_W^{-1}CP + FR_VF^{\mathsf{T}}.$$

This is the most commonly used form of the controller since it gives an explicit formula for the estimator gains that minimize the error covariance. The gain matrix for this case can solved use the `control.lqe` command in Python or MATLAB.

Another property of the Kalman filter is that it extracts the maximum possible information about output data. To see this, consider the *residual random process*

$$R = Y - C\hat{X}$$



(a) Covariance $p(t)$ (b) State x and estimate $\hat{x}$

Figure 6.1: Optimal estimator for a first-order linear system with parameter values $a = 1$, $c = 1$, $\sigma_V = 1$, $\sigma_W = 0.1$, starting from initial condition $x(0) = 1$.

(this process is also called the *innovations process*). It can be shown for the Kalman filter that the correlation matrix of R is given by

$$R_R(t_1, t_2) = V(t_1)\delta(t_2 - t_1).$$

This implies that the residuals are a white noise process and so the output error has no remaining dynamic information content.

6.2 Extensions of the Kalman Filter

The Kalman filter has a number of extensions that are used to extend its utility to cases where the noise and disturbances are not white noise and when the process is not linear. We summarize some of these extensions briefly here, with additional extensions provided in the next chapter.

Correlated disturbances and noise

The derivation of the Kalman filter assumes that the disturbances and noise are independent and white. Removing the assumption of independence is straightforward and simply results in a cross term ($\mathbb{E}(V(t)W(s)) = R_{VW}\delta(s-t)$) being carried through all calculations.

To remove the assumption of white noise for the process disturbances or sensor noise, we can construct a filter that takes white noise as an input and produces a random process with the appropriate correlation function (or equivalently, spectral power density function). The intuition behind this approach is that we must have an internal model of the noise and/or disturbances in order to capture the correlation between different times.

Extended Kalman filters

Consider a *nonlinear* system

$$\begin{aligned} \dot{X} &= f(X, u, V), & X &\in \mathbb{R}^n, u \in \mathbb{R}^m, \\ Y &= CX + W, & Y &\in \mathbb{R}^p, \end{aligned}$$

where V and W are Gaussian white noise processes with covariance matrices R_V and R_W. A nonlinear observer for the system can be constructed by using the process

$$\dot{\hat{X}} = f(\hat{X}, u, 0) + L(Y - C\hat{X}).$$

If we define the error as $E = X - \hat{X}$, the error dynamics are given by

$$\begin{aligned}\dot{E} &= f(X, u, V) - f(\hat{X}, u, 0) - LC(X - \hat{X}) \\ &= F(E, \hat{X}, u, V) - LCe,\end{aligned}$$

where

$$F(E, \hat{X}, u, V) = f(E + \hat{X}, u, V) - f(\hat{X}, u, 0).$$

We can now linearize around *current* estimate $\hat{X}$:

$$\begin{aligned}\hat{E} &= \frac{\partial F}{\partial E}E + \underbrace{F(0, \hat{X}, u, 0)}_{=0} + \underbrace{\frac{\partial F}{\partial V}V}_{\text{noise}} - \underbrace{LCe}_{\text{observer gain}} + \text{h.o.t} \\ &\approx \tilde{A}E + \tilde{F}V - LCE,\end{aligned}$$

where the matrices

$$\begin{aligned}\tilde{A} &= \left.\frac{\partial F}{\partial e}\right|_{(0,\hat{X},u,0)} = \left.\frac{\partial f}{\partial X}\right|_{(\hat{X},u,0)}, \\ \tilde{F} &= \left.\frac{\partial F}{\partial V}\right|_{(0,\hat{X},u,0)} = \left.\frac{\partial f}{\partial V}\right|_{(\hat{X},u,0)}\end{aligned}$$

depend on current estimate $\hat{X}$. We can now design an observer for the linearized system around the *current* estimate:

$$\begin{aligned}\dot{\hat{X}} &= f(\hat{X}, u, 0) + L(Y - C\hat{X}), \qquad L = PC^{\mathsf{T}}R_V^{-1}, \\ \dot{P} &= (\tilde{A} - LC)P + P(\tilde{A} - LC)^{\mathsf{T}} + \tilde{F}R_V\tilde{F}^{\mathsf{T}} + LR_WL^{\mathsf{T}}, \\ P(t_0) &= \mathbb{E}(X(t_0)X^{\mathsf{T}}(t_0)).\end{aligned}$$

This is called the (Schmidt) *extended Kalman filter* (EKF).

The intuition in the Kalman filter is that we replace the prediction portion of the filter with the nonlinear modeling while using the instantaneous linearization to compute the observer gain. Although we lose optimality, in applications the extended Kalman filter often works well and it is very versatile, as illustrated in the following example.

Example 6.2 Online parameter estimation
Consider a linear system with unknown parameters ξ

$$\begin{aligned}\dot{X} &= A(\xi)X + B(\xi)u + FV, \qquad \xi \in \mathbb{R}^p, \\ Y &= C(\xi)X + W.\end{aligned}$$

We wish to solve the parameter identification problem: given $u(t)$ and $Y(t)$, estimate the value of the parameters ξ.

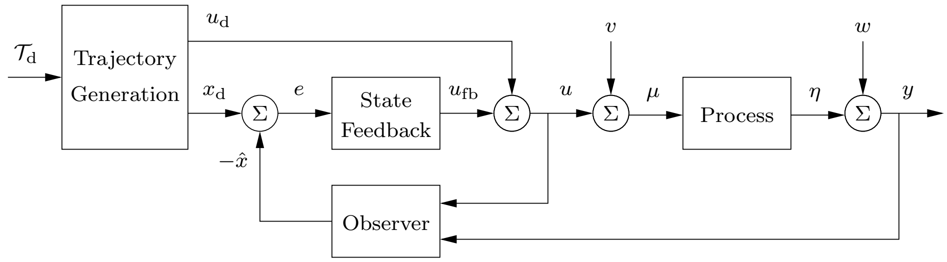


Figure 6.2: Block diagram of a basic feedback loop.

One approach to this online parameter estimation problem is to treat ξ as an unknown *state* that has zero derivative:

$$\dot{X} = A(\xi)X + B(\xi)u + FV, \qquad \dot{\xi} = 0.$$

We can now write the dynamics in terms of the extended state $Z = (X, \xi)$:

$$\frac{d}{dt}\begin{bmatrix} X \\ \xi \end{bmatrix} = \overbrace{\begin{bmatrix} A(\xi) & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} X \\ \xi \end{bmatrix} + \begin{bmatrix} B(\xi) \\ 0 \end{bmatrix} u}^{f(\left[\begin{smallmatrix} X \\ \xi \end{smallmatrix}\right], u, V)} + \begin{bmatrix} F \\ 0 \end{bmatrix} V,$$

$$Y = \underbrace{C(\xi)X + W}_{h(\left[\begin{smallmatrix} X \\ \xi \end{smallmatrix}\right], V)}.$$

This system is nonlinear in the extended state Z, but we can use the extended Kalman filter to *estimate* Z. If this filter converges, then we obtain both an estimate of the original state X and an estimate of the unknown parameter $\xi \in \mathbb{R}^p$.

Remark: need various observability conditions on augmented system in order for this to work. ∇

6.3 LQG Control

We now return to the full control problem, in which we wish to design a controller that uses the estimated state and tracks a trajectory. Figure 1.4 shows the high level view of the system, which we replicate in Figure 6.3, leaving out the unmodeled dynamics for simplicity. We assume that all processes are linear and hence it will suffice to consider the problem of stabilizing the origin.

The model for our process dynamics now must include the control input u and so we write

$$\dot{X} = AX + Bu + FV,$$
$$Y = CX + W,$$

where V and W are white noise processes with appropriate covariances.

The *linear quadratic Gaussian control problem* is to find a controller that will minimize

$$J = \mathbb{E}\left(\int_0^\infty \left[(Y - y_d)^\mathsf{T} Q_y (Y - y_d) + (U - u_d)^\mathsf{T} Q_u (U - u_d)\right] dt\right),$$

where U is now considered as a random variable. While in general we might imagine that the optimal controller could require some complex combination of state estimation and state feedback, it turns out that it can be shown that the structure of the optimal control separates into an optimal controller assuming perfect state knowledge and an optimal estimator that is independent of the control system gains.

Theorem 6.2 (Separation principle)**.** *The optimal controller for a linear system with white noise process disturbances and sensor noise has the form*

$$\dot{\hat{x}} = A\hat{x} + Bu - L(C\hat{x} - y)$$
$$u = u_d - K(\hat{x} - x_d)$$

where L is the optimal observer gain ignoring the controller and K is the optimal controller gain ignoring the noise.

This is called the *separation principle* (for H_2 control). A proof of this theorem can be found in Friedland [Fri04] (and many other textbooks).

6.4 Implementation in Python

Stationary Kalman gains can be computed in python-control using the `lqe` function, which constructs an optimal estimator gain and covariance for a linear system. The Python command

```
L, P, E = ct.lqe(sys, Qv, Qw[, Qvw])
```

computes the optimal estimator gain `L`, steady state error covariance matrix `P`, and closed loop poles for the estimator `E` given the system dynamics and covariance of the process disturbances (`Qv`) and sensor noise (`Qw`), as well as any cross-covariance between the two sets of signals (`Qvw`).

The `create_estimator_iosystem` function can be used to create an I/O system implementing a Kalman filter, including integration of the Riccati ODE. The command has the form

```
estim = ct.create_estimator_iosystem(sys, Qv, Qw)
```

The input to the estimator is the measured outputs y and the system input u. To run the estimator on a noisy signal, use the command

```
resp = ct.input_output_response(est, timepts, [Y, U], [X0, P0])
```

If desired, the `correct` parameter can be set to `False` to allow prediction with no additional sensor information:

```
resp = ct.input_output_response(
    estim, timepts, 0, [X0, P0], param={'correct': False})
```

The `create_statefbk_iosystem` function can be used to combine an estimator with a state feedback controller:

```
K, _, _ = ct.lqr(sys, Qx, Qu)
estim = ct.create_estimator_iosystem(sys, Qv, Qw, P0)
ctrl, clsys = ct.create_statefbk_iosystem(sys, K, estimator=estim)
```

The controller will have the same form as a full state feedback controller, but with the system state x input replaced by the estimated state $\hat{x}$ (output of `estim`):

$$u = u_\mathrm{d} - K(\hat{x} - x_\mathrm{d}).$$

The closed loop controller `clsys` includes both the state feedback and the estimator dynamics and takes as its input the desired state x_d and input u_d:

```
resp = ct.input_output_response(
    clsys, timepts, [Xd, Ud], [X0, np.zeros_like(X0), P0])
```

6.5 Application to a Thrust Vectored Aircraft

To illustrate the use of the Kalman filter, we consider the problem of estimating the state for the Caltech ducted fan, described already in Section 4.6. We use the simplified model described in Example 3.5, with added disturbances and noise.

We begin by defining an extended Kalman filter that uses the nonlinear dynamics to estimate the current state. The dynamics of the system with disturbances on the x and y variables is given by

$$\begin{aligned} m\ddot{x} &= F_1 \cos\theta - F_2 \sin\theta - c\dot{x} + d_x, \\ m\ddot{y} &= F_1 \sin\theta + F_2 \cos\theta - c\dot{y} - mg + d_y, \\ J\ddot{\theta} &= rF_1. \end{aligned} \tag{6.1}$$

The measured values of the system are the position and orientation, with added noise n_x, n_y, and n_θ:

$$\vec{y} = \begin{bmatrix} x \\ y \\ \theta \end{bmatrix} + \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix}. \tag{6.2}$$

We assume that the disturbances are represented by white noise with intensity $\sigma^2 = 0.01$ and that the sensor noise has noise intensity matrix

$$Q_N = \begin{bmatrix} 2\times10^{-4} & 0 & 1\times10^{-5} \\ 0 & 2\times10^{-4} & 1\times10^{-5} \\ 1\times10^{-5} & 1\times10^{-5} & 1\times10^{-4} \end{bmatrix}.$$

To compute the update for the Kalman filter, we require the linearization of the system at a state $\vec{x} = (x, y, \theta, \dot{x}\dot{y}, \dot{z})$, which can be computed from equation (6.1) to be

$$\dot{X} = AX + Bu + FV,$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & -\frac{F_1}{m}s_\theta - \frac{F_2}{m}c_\theta & -\frac{c}{m} & 0 & 0 \\ 0 & 0 & \frac{F_1}{m}c_\theta - \frac{F_2}{m}s_\theta & 0 & -\frac{c}{m} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ \frac{1}{m}c_\theta & -\frac{1}{m}s_\theta \\ \frac{1}{m}s_\theta & \frac{1}{m}c_\theta \\ r/J & 0 \end{bmatrix}, \quad F = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{bmatrix},$$

with $c_\theta = \cos\theta$ and $s_\theta = \sin\theta$.

The state estimate is given by using the nonlinear dynamics for the prediction of the state error with a linear correction term, and the linearized dynamics for the update of the covariance matrix. If we let $\xi = (x, y, \theta, \dot{x}, \dot{y}, \dot{\theta}) \in \mathbb{R}^6$ represent the states and $\eta = (x, y, \theta) \in \mathbb{R}^3$ represent write the the output, the dynamics of the state estimate written as

$$\dot{\hat{\xi}} = f(\xi, u) - L(C\hat{\xi} - \eta),$$

where $f(\xi, u)$ represents the full nonlinear dynamics in equation (6.1) and $C = \begin{bmatrix} I & 0 \end{bmatrix} \in \mathbb{R}^{3\times 6}$ represents the output matrix. The gain matrix $L = P(t)C^\mathsf{T} R_W^{-1}$ is chosen based on the time-varying error covariance matrix $P(t)$, which evolves using the linearized dynamics:

$$\begin{aligned} \dot{P} &= A(\xi)P + PA(\xi)^\mathsf{T} - PC^\mathsf{T} R_W^{-1}(t)CP + FR_V(t)F^\mathsf{T}, \\ P(0) &= \mathbb{E}(X(0)X^\mathsf{T}(0)). \end{aligned}$$

To show how this estimator can be used, consider the problem of stabilizing the system to the origin with an LQR controller that uses the estimated state. We compute the LQR controller as if the entire state ξ were available directly, so the computations are identical to those in Example 3.5. We choose the physically motivated set of weights given by

$$Q_\xi = \begin{bmatrix} 100 & 0 & 0 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 & 0 & 0 \\ 0 & 0 & 36/\pi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \qquad Q_u = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}.$$

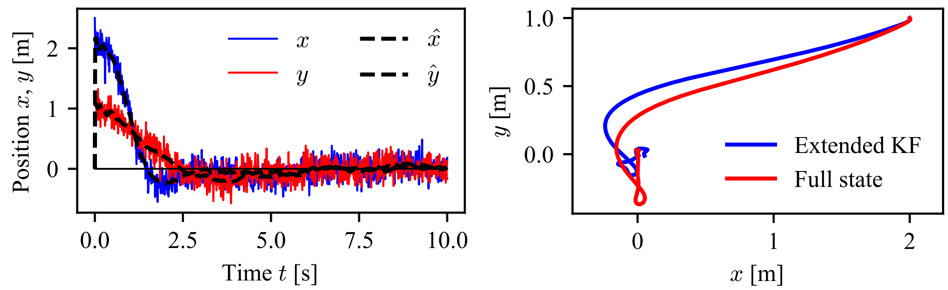
For the (extended) Kalman filter, we model the process disturbances and sensor noise as white noise processes with noise intensities

$$R_V = \begin{bmatrix} 0.01 & 0 \\ 0 & 0.01 \end{bmatrix}, \qquad R_W = \begin{bmatrix} 2\times 10^{-4} & 0 & 1\times 10^{-5} \\ 0 & 2\times 10^{-4} & 1\times 10^{-5} \\ 1\times 10^{-5} & 1\times 10^{-5} & 1\times 10^{-4} \end{bmatrix}$$

Figure 6.4 shows the response of the system starting from an initial position $(x_0, y_0) = (2, 1)$ and with disturbances and noise with intensity 10-100X smaller than the worst case for which we designed the system.

6.6 Further Reading

There is a vast literature on Kalman filtering and linear quadratic Gaussian (LQG) control theory. The treatment in this chapter follows fairly closely to that of Friedland [Fri04]. A compact treatment of LQG theory is given in the books by Anderson and Moore [**?**], Åström [**?**], and Lindquist and Picci [**?**].



(a) Estimated states

(b) Full state versus EKF

Figure 6.3: LQR control of the VTOL system with an extended Kalman filter to estimate the state. (a) The x and y positions as a function of time, with dashed lines showing the estimated values from the extended Kalman filter. (b) The xy path of the system with full state feedback (and no noise) versus the controller using the extended Kalman filter.

Exercises

6.1. Show that if we define the estimated state of a random process X as the conditional mean

$$\hat{x}(t) = \mathbb{E}(X(t) \mid y(\tau), \, \tau \leq t)$$

that $\hat{x}$ minimizes

$$\mathbb{E}(\hat{x}(t) - X(t) \mid y(\tau), \, \tau \leq t).$$

6.2. Consider a scalar control system

$$\dot{X} = \lambda X + u + \sigma_v V$$
$$Y = X + \sigma_w W,$$

where V and W are zero-mean, Gaussian white noise processes with covariance 1 and $\sigma_v, \sigma_w > 0$. Assume that the initial value of X is modeled as a Gaussian with mean X_0 and variance $\sigma^2_{X_0}$.

(a) Assume that we initialize a Kalman filter such that the initial covariance starts near a steady state value p^*. Given conditions on λ such that error covariance is locally stable about this solution.

(b) Suppose that V is no longer taken to be white noise and instead has a correlation function given by

$$\rho_V(\tau) = e^{-\alpha|\tau|}, \qquad \alpha > 0.$$

Write down an estimator that minimizes the mean square error of the output under these conditions. You do not need to explicitly solve the resulting equations, just write them down in a form that is similar to an appropriate Kalman filter equation.

6.3. Consider a discrete-time, scalar linear system with dynamics

$$x[k+1] = ax[k] + v[k], \qquad y[k] = x[k] + w[k],$$

where v and w are discrete-time, Gaussian random processes with mean zero and variances 1 and σ^2, respectively. Assume that the initial value of the state has zero mean and variance σ_0^2.

(a) Compute the optimal estimator for the system using y as a (noisy) measurement. Your solution should be in the form of an explicit, time-varying, discrete-time system.

(b) Assume that $a < 1$. Write down an explicit formula for the mean and covariance of the steady-state error of the optimal estimator.

(c) Suppose the mean value of the initial condition is $\mathbb{E}(x[0]) = 1$ and $a = 1$. Determine the optimal steady-state estimator for the system.

6.4. Consider the dynamics of an inverted pendulum whose dynamics are given by

$$\frac{dx}{dt} = \begin{bmatrix} x_2 \\ \sin x_1 - cx_2 + u\cos x_1 \end{bmatrix},$$

where $x = (\theta, \dot\theta)$ and $c > 0$ is the damping coefficient. We assume that we have a sensor that can measure the offset of a point along the inverted pendulum

$$y = r\sin\theta,$$

where $r \in [0, \ell]$ is a point along the pendulum and ℓ is the length of the pendulum. Assume that the pendulum is subject to input disturbances v that are modeled as white noise with intensity σ_v^2 and that the sensor is subject to additive Guassian white noise with noise intensity σ_w^2.

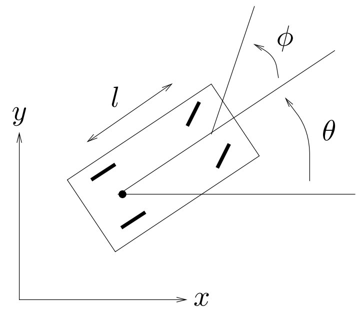
(a) Determine the optimal location of the sensor (r^*) that minimizes the steady state covariance of the error P and justify your answer.

(b) Show that the Kalman filter gain $L = PC^\mathsf{T} R_w^{-1}$ does not depend on the covariance of the error in $\dot\theta$.

(c) Take $c = 0$ and compute the steady state gain for the Kalman filter as a function of the sensor location r.

Note: for the first two parts it is not necessary to solve equations for the steady state covariance of the error. For the last part, your answer should not require a substantial amount of algebra if you organize your calculations a bit (and set $c = 0$).

6.5. Consider the problem of estimating the position of an autonomous mobile vehicle using a GPS receiver and an IMU (inertial measurement unit). The dynamics of the vehicle are given by



$$\dot{x} = \cos\theta\, v$$
$$\dot{y} = \sin\theta\, v$$
$$\dot{\theta} = \frac{1}{\ell}\tan\delta\, v,$$

We assume that the vehicle has disturbances in the inputs v and δ with standard deviation of up to 10% and noisy measurements from the GPS receiver and IMU.

We consider a trajectory in which the car is driving on a constant radius curve at $v = 10$ m/s forward speed with $\delta = 5^\circ$ for a duration of 10 seconds.

(a) Suppose first that we only have the GPS measurements for the xy position of the vehicle. These measurements give the position of the vehicle with approximately 10 cm accuracy. Model the GPS error as Gaussian white noise with $\sigma = 0.1$ meter in each direction. Design an optimization-based estimator for the system and plot the estimated states versus the actual states. What is the covariance of the estimate at the end of the trajectory?

(b) An IMU can be used to measure angular rates and linear acceleration. For simplicity, we assume that the IMU is able to directly measure the angle of the car with a standard deviation of 1 degree. Design an optimal estimator for the system using the GPS and IMU measurements, and plot the estimated states versus the actual states. What is the covariance of the estimate at the end of the trajectory?