Linear Systems, Estimation, and Control

Linear quadratic regulator and linear quadratic Gaussian control

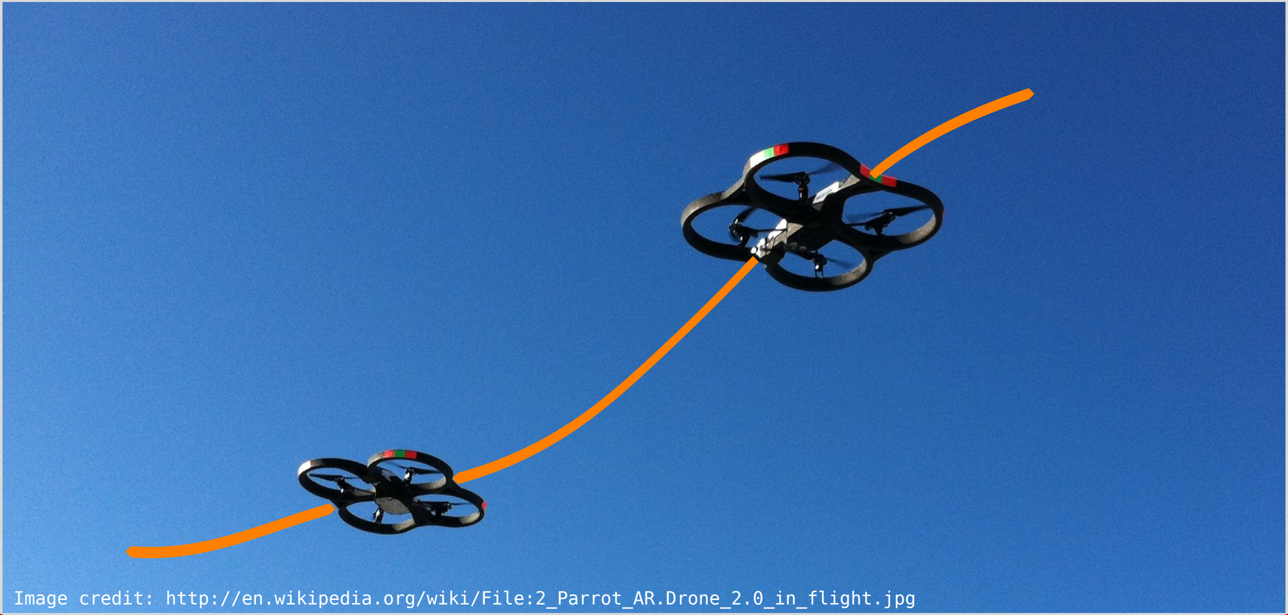
Aditya Mahajan

McGill University

Lecture Notes for ECSE 506: Stochastic Control and Decision Theory

December 24, 2013

The MDP setup — State feedback



Linear Quadratic Regulation (LQR)

Notation
State : $X_t \in \mathbb{R}^n$
Action: $U_t \in \mathbb{R}^m$

Dynamics
$X_{t+1} = \mathbf{A}_t X_t + \mathbf{B}_t U_t,$ where $\mathbf{A}_t \in \mathbb{R}^{n \times n}$, $\mathbf{B}_t \in \mathbb{R}^{n \times m}$.

Cost
Per step cost : $c_t(x_t, u_t) = \|x_t\|^2_{\mathbf{Q}_t} + \|u_t\|^2_{\mathbf{R}_t}$
Terminal reward: $c_T(x_T) = \|x_T\|^2_{\mathbf{Q}_T}$
where $\|X\|_{\mathbf{Q}} = X^\top \mathbf{Q} X$ and for all t, $\mathbf{Q}_t = \mathbf{Q}_t^\top \succeq 0$ and $\mathbf{R}_t = \mathbf{R}_t^\top \succ 0$.

Control objective
Choose $U_t = g_t(X_{1:t}, U_{1:t-1})$ so as to minimize

$$J(\mathbf{g}) = \sum_{t=1}^{T-1} c_t(X_t, U_t) + c_T(x_T)$$

- Regulation problem: keep the state of the system close to the origin.
- Tracking problem: To keep the state of the system close to a reference trajectory $\{x_t^\circ\}$, use the cost

$$c_t(x_t, u_t) = \|x_t - x_t^\circ\|_{\mathbf{Q}_t} + \|u_t\|_{\mathbf{R}_t}, \qquad c_T(x_T) = \|x_T - x_T^\circ\|_{\mathbf{Q}_T}.$$

Deterministic LQR is a MDP

	MDP Dynamic Model	Deterministic LQR
System Dynamics	$X_{t+1} = f_t(X_t, U_t, W_t)$	$X_{t+1} = \mathbf{A}_t X_t + \mathbf{B}_t U_t$
Information Structure	$U_t = g_t(X_{1:t}, U_{1:t-1})$	$U_t = g_t(X_{1:t}, U_{1:t-1})$
Objective Function	$\mathbb{E}\left[\sum_{t=1}^{T-1} c_t(X_t, U_t) + c_T(X_T)\right]$	$\sum_{t=1}^{T-1} c_t(X_t, U_t) + c_T(X_T)$

Structure of Controller

Using Markov strategies does not entail any loss of optimality

$$U_t = g_t(X_t)$$

Dynamic program

$$V_T(x_T) = c_T(x_T);$$

$$V_t(x_t) = \max_{u_t \in \mathcal{U}_t(x_t)} \Big\{ c_t(x_t, u_t) + \mathbb{E}[V_{t+1}(f_t(x_t, u_t, W_t))] \Big\},$$

$$t = T-1, \dots, 1.$$

Structure of optimal deterministic LQR

Theorem The value function at time t is

$$V_t(X_t) = \|X_t\|^2_{\mathbf{S}_t}$$

and the optimal control action is

$$U_t = -\mathbf{H}_t X_t$$

where the gain matrices $\mathbf{H}_t$ are determined recursively as follows:

$$\mathbf{H}_T = 0$$

$$\mathbf{H}_t = [\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t]^{-1} \mathbf{\Lambda}_t$$

where

$$\mathbf{\Lambda}_t = \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{A}_t$$

and $\mathbf{S}_t$ are determined by the backward Riccati difference equations:

$$\mathbf{S}_T = \mathbf{Q}_T$$

$$\mathbf{S}_t = \mathbf{A}_t^\top \mathbf{S}_{t+1} \mathbf{A}_t + \mathbf{Q}_t - \mathbf{\Lambda}_t^\top [\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t]^{-1} \mathbf{\Lambda}_t$$

Riccati equations are named after Count Jacopo Francesco Riccati (1670–1754) who studied the differential equations of the form

$$\dot{x} = a x^2 + bt + ct^2$$

and its variations. In modern control, such equations arise in the calculus of variations and optimal filtering. The discrete-time version of these equations are also named after Riccati.

Completing the squares lemma

Lemma Let

- $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$
- $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{B} \in \mathbb{R}^{n \times m}$
- $\mathbf{R} \in \mathbb{R}^{m \times m}$ and $\mathbf{Q} \in \mathbb{R}^{n \times n}$ such that $\mathbf{R} = \mathbf{R}^\top \succ 0$ and $\mathbf{Q} = \mathbf{Q}^\top \succeq 0$.

Then

$$\|u\|^2_{\mathbf{R}} + \|\mathbf{A}x + \mathbf{B}u\|^2_{\mathbf{Q}} = \|u + \mathbf{H}x\|^2_{\mathbf{K}} + \|x\|^2_{\mathbf{L}}.$$

where

$$\mathbf{K} = \mathbf{R} + \mathbf{B}^\top \mathbf{Q}\mathbf{B}, \qquad \mathbf{K} = \mathbf{K}^\top \succ 0$$

$$\mathbf{H} = \mathbf{K}^{-1}\mathbf{\Lambda}, \qquad \text{where } \mathbf{\Lambda} = \mathbf{B}^\top \mathbf{Q}\mathbf{A}$$

$$\mathbf{L} = \mathbf{A}^\top \mathbf{Q}\mathbf{A} - \mathbf{\Lambda}^\top \mathbf{K}\mathbf{\Lambda}$$

Proof The result follows by completing the squares in two different ways.

$$\text{LHS} = u^\top \mathbf{R}u + u^\top \mathbf{B}^\top \mathbf{Q}\mathbf{B}u + x^\top \mathbf{A}^\top \mathbf{Q}\mathbf{A}x + x^\top \mathbf{A}^\top \mathbf{Q}\mathbf{B}u + u^\top \mathbf{B}^\top \mathbf{Q}\mathbf{A}x$$

$$\text{RHS} = u^\top \mathbf{K}u + x^\top \mathbf{H}^\top \mathbf{K}u + u^\top \mathbf{K}\mathbf{H}x + x^\top \mathbf{H}^\top \mathbf{K}\mathbf{H}x + x^\top \mathbf{L}x$$

Compare the coefficients of $u^\top \cdots u$, $u^\top \cdots x$, and $x^\top \cdots x$,

Proof of the structure of optimal determistic LQR

Proof Proceed by backward induction.

- **Basis**: $V_T(x) = c_T(x) = \|x\|^2_{Q_T}$.
- **Induction hypothesis**: $V_{t+1}(x) = \|x\|^2_{S_{t+1}}$
- **Induction step**:

$$V_t(x) = \min_u \Big[\|x\|^2_{Q_t} + \|u\|^2_{R_t} + V_{t+1}(A_t x + B_t u) \Big]$$

$$= \min_u \Big[\|x\|^2_{Q_t} + \underbrace{\|u\|^2_{R_t} + \|A_t x + B_t u\|^2_{S_{t+1}}}_{\text{completion of squares}} \Big]$$

$$= \min_u \Big[\|x\|^2_{Q_t} + \overbrace{\|u + H_t x\|^2_{K_t} + \|x\|^2_{L_t}} \Big]$$

where

$$H_t = [R_t + B_t^\top S_{t+1} B_t]^{-1} \Lambda_t, \qquad \text{where } \Lambda_t = B_t^\top S_{t+1} A_t.$$

$$L_t = A^\top S_{t+1} A_t - \Lambda_t^\top [R_t + B_t^\top S_{t+1} B_t]^{-1} \Lambda_t.$$

Thus, the optimal control action is $u = -H_t x$ and the optimal cost is

$$V_t(x) = \|x\|^2_{Q_t} + \|x\|^2_{L_t} = \|x\|^2_{S_t}, \qquad \text{where } S_t = Q_t + L_t.$$

Note that the update equation of S_t is same as that in the Theorem.

Linear Quadratic regulator example

Generalized LQR: Cross-term in cost

Minimizing output error

Suppose that instead of minimizing the norm of the state X_t, we are interested in minimizing the norm of the output $Y_t = \mathbf{C}_t X_t + \mathbf{D}_t U_t$. In such a case, the per-step cost function will be of the form

$$c_t(X_t, U_t) = \|X_t\|^2_{\mathbf{Q}_t} + \|U_t\|^2_{\mathbf{R}_t} + 2X_t^\top \mathbf{N}_t U_t$$

Assume that the terminal cost function does not change, and

$$\begin{bmatrix} \mathbf{Q}_t & \mathbf{N}_t \\ \mathbf{N}_t^\top & \mathbf{R}_t \end{bmatrix} \succeq 0, \qquad \text{or equivalently} \qquad \mathbf{Q} - \mathbf{N}\mathbf{R}^{-1}\mathbf{N}^\top \succeq 0.$$

Key Lemma

$$\|x\|^2_{\mathbf{Q}} + \|u\|^2_{\mathbf{R}} + 2x^\top \mathbf{N} u = \|x\|^2_{\tilde{\mathbf{Q}}} + \|u + \mathbf{R}^{-1}\mathbf{N}^\top x\|^2_{\mathbf{R}}.$$

where $\tilde{\mathbf{Q}} = \mathbf{Q} - \mathbf{N}\mathbf{R}^{-1}\mathbf{N}^\top$.

Change of variables

Let $\tilde{U}_t = U_t + \mathbf{R}_t^{-1}\mathbf{N}_t^\top X_t$. Then

$$X_{t+1} = \tilde{\mathbf{A}}_t X_t + \mathbf{B}_t \tilde{U}_t, \qquad \text{where } \tilde{\mathbf{A}}_t = \mathbf{A}_t - \mathbf{B}_t \mathbf{R}_t^{-1}\mathbf{N}_t^\top$$

- Thus, the system is in the same form as the standard LQR.

Generalized LQR: Cross-term in cost

Theorem The value function at time t is

$$V_t(X_t) = \|X_t\|_{S_t}$$

and the optimal control action is

$$U_t = -\mathbf{H}_t X_t$$

where the gain matrices $\mathbf{H}_t$ are computed recursively as follows:

$$\mathbf{H}_T = 0$$

$$\mathbf{H}_t = \left[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t\right]^{-1} \mathbf{\Lambda}_t$$

where

$$\mathbf{\Lambda}_t = \mathbf{N}_t^\top + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{A}_t$$

and $\mathbf{S}_t$ are determined by the modified backward Riccati equations:

$$\mathbf{S}_T = \mathbf{Q}_T$$

$$\mathbf{S}_t = \mathbf{A}_t^\top \mathbf{S}_{t+1} \mathbf{A}_t + \mathbf{Q}_t - \mathbf{\Lambda}_t^\top \left[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t\right]^{-1} \mathbf{\Lambda}_t$$

- Note that the only change from the standard LQR equations is in the definition of $\mathbf{\Lambda}_t$.

Generalized LQR: Proof for cross-term in cost

Proof Consider the system with the change of variables. The structure of the optimal controller and the form of the value function are given as before. Recall that $U_t = \tilde{U}_t - \mathbf{R}_t^{-1}\mathbf{N}_t^\top X_t$. Hence,

$$\mathbf{H}_t = \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{B}_t\big]^{-1}\mathbf{B}_t^\top \mathbf{S}_{t+1}\tilde{\mathbf{A}}_t + \mathbf{R}_t^{-1}\mathbf{N}_t = \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{B}_t\big]^{-1}\mathbf{\Lambda}_t$$

where

$$\begin{aligned}
\mathbf{\Lambda}_t &= \mathbf{B}_t^\top \mathbf{S}_{t+1}\tilde{\mathbf{A}}_t + \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{B}_t\big]\mathbf{R}_t^{-1}\mathbf{N}_t \\
&= \mathbf{B}_t^\top \mathbf{S}_{t+1}\big[\mathbf{A}_t - \mathbf{B}_t\mathbf{R}_t^{-1}\mathbf{N}_t^\top\big] + \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{B}_t\big]\mathbf{R}_t^{-1}\mathbf{N}_t \\
&= \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{A}_t + \mathbf{N}_t^\top
\end{aligned}$$

Furthermore, since the terminal cost is the same as before, the initial condition of the backward Riccati equation does not change. The Riccati update is given by

$$\mathbf{S}_t = \tilde{\mathbf{A}}_t^\top \mathbf{S}_{t+1}\tilde{\mathbf{A}}_t + \tilde{\mathbf{Q}}_t - \big[\mathbf{B}_t^\top \mathbf{S}_{t+1}\tilde{\mathbf{A}}_t\big]^\top\big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1}\mathbf{B}_t\big]^{-1}\big[\mathbf{B}_t^\top \mathbf{S}_{t+1}\tilde{\mathbf{A}}_t\big]$$

Substituting the value of $\tilde{\mathbf{A}}_t$ and $\tilde{\mathbf{Q}}_t$ and some (messy) algebraic manipulation gives the result (see next page).

Generalized LQR: Proof for cross-term in cost (cont.)

Proof (cont.) Ignore the subscripts for ease of notation.

1. Let $\mathbf{K} = \mathbf{R} + \mathbf{B}^\top \mathbf{S} \mathbf{B}$. Thus, $\mathbf{B}^\top \mathbf{S} \mathbf{B} = \mathbf{K} - \mathbf{R}$.
2. $\tilde{\mathbf{A}} \mathbf{S} \tilde{\mathbf{A}} = \mathbf{A} \mathbf{S} \mathbf{A} + \mathbf{N} \mathbf{R}^{-1} (\mathbf{B}^\top \mathbf{S} \mathbf{B}) \mathbf{R}^{-1} \mathbf{N}^\top - 2 \mathbf{A}^\top \mathbf{S} \mathbf{B} \mathbf{R}^{-1} \mathbf{N}^\top$.
3. $\tilde{\mathbf{Q}} = \mathbf{Q} - \mathbf{N} \mathbf{R}^{-1} \mathbf{N}^\top$.
4. $\mathbf{B}^\top \mathbf{S} \tilde{\mathbf{A}} = \mathbf{B}^\top \mathbf{S} \mathbf{A} - (\mathbf{B}^\top \mathbf{S} \mathbf{B}) \mathbf{R}^{-1} \mathbf{N}^\top = \mathbf{\Lambda} - \mathbf{K} \mathbf{R}^{-1} \mathbf{N}^\top$.
5. $(\mathbf{B}^\top \mathbf{S} \tilde{\mathbf{A}})^\top \mathbf{K}^{-1} (\mathbf{B}^\top \mathbf{S} \tilde{\mathbf{A}}) = \mathbf{\Lambda}^\top \mathbf{K}^{-1} \mathbf{\Lambda} + \mathbf{N} \mathbf{R}^{-1} \mathbf{K} \mathbf{R}^{-1} \mathbf{N}^\top - 2 \mathbf{\Lambda}^\top \mathbf{R}^{-1} \mathbf{N}^\top$.
6. (2) + (3) - (5) = Result - $2[\mathbf{N} + \mathbf{A}^\top \mathbf{S} \mathbf{B} - \mathbf{\Lambda}^\top] \mathbf{R}^{-1} \mathbf{N}^\top$, where the last term is zero by definition of $\mathbf{\Lambda}$.

LQR Tracking problem

Tracking setup

Suppose that we want to ensure that the output signal $Y_t = \mathbf{C}_t X_t$ is close to a reference trajectory $\{y^\circ_t\}_{t=1}^T$. Then, the cost functions are

$$c_t(X_t, U_t) = \|\mathbf{C}_t X_t - y^\circ_t\|^2_{\mathbf{Q}_t} + \|U_t\|^2_{\mathbf{R}_t}, \qquad c_T(X_T) = \|\mathbf{C}_T X_T - y^\circ_T\|^2_{\mathbf{Q}_T}.$$

Theorem

The value function at time t is

$$V_t(X_t) = \|X_t\|^2_{S_t} + \alpha_t$$

and the optimal control action is

$$U_t = -H_t X_t + H^\circ_t r_{t+1}$$

Recursive computations

- $\{\mathbf{S}_t\}_{t=1}^T$ and $\{\mathbf{H}_t\}_{t=1}^T$ follow the same recursion as before;
- The gain matrices H°_t are given by $H^\circ_t = \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t\big]^{-1} \mathbf{B}_t^\top$
- The correction terms r_t are given by

$$r_T = \mathbf{C}_T^\top Q_T y^\circ_T, \qquad r_t = \big[\mathbf{A}_t - \mathbf{B}_t \mathbf{H}_t\big]^\top r_{t+1} + \mathbf{C}_t^\top \mathbf{Q}_t y^\circ_t$$

- The tracking error α_t is given by

$$\alpha_T = \|y^\circ_T\|^2_{\mathbf{Q}_T}, \qquad \alpha_t = \|y^\circ_t\|^2_{\mathbf{Q}_t} - 2 r_{t+1}^\top \mathbf{B}_t \big[\mathbf{R}_t + \mathbf{B}_t^\top \mathbf{S}_{t+1} \mathbf{B}_t\big]^{-1} \mathbf{B}_t^\top r_{t+1} + \alpha_{t+1}$$

Linear Quadratic Guassian (LQG) setup

Stochastic dynamics

$$X_{t+1} = \mathbf{A}_t X_t + \mathbf{B}_t U_t + W_t, \qquad \text{where } W_t \sim \mathcal{N}(0, \Sigma_W).$$

The above model is similar to the LQR setup with the exception that the state evolution is stochastic. By using an argument similar to the LQR setup, the structure of the optimal controller is as given below.

Theorem

The value function at time t is

$$V_t(X_t) = \|X_t\|^2_{\mathbf{S}_t} + \alpha_t$$

and the optimal control action is

$$U_t = -\mathbf{H}_t X_t$$

where $\mathbf{S}_t$ and $\mathbf{H}_t$ follow the same recursion as before and

$$\alpha_T = 0$$

$$\alpha_t = \alpha_{t+1} + \mathrm{Tr}[\mathbf{\Sigma}_W \mathbf{S}_{t+1}] = \sum_{\tau=t+1}^{T} \mathrm{Tr}[\mathbf{\Sigma}_W \mathbf{S}_\tau]$$

Thus, the optimal controller is the same as in the deterministic case. The only effect of the noise is to increase the value function. (This phenomenon is unique to LQG systems).

Certainty Equivalence principle (Simon, 1948)

Let $f\colon \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ and X be a $\mathbb{R}^n$ valued random variable. If f is quadratic in its arguments, then

$$u^* = \arg\min_{u \in \mathbb{R}^m} \mathbb{E}[f(X, u)] = \arg\min_{u \in \mathbb{R}^m} f(\mathbb{E}[X], u)$$

Note: f is quadratic means $f(x, u) = \mathbf{A}x + \mathbf{B}u + \|x\|^2_{\mathbf{Q}} + \|u\|^2_{\mathbf{R}} + x^\top \mathbf{G} u + \alpha$ where all matrices are of appropriate dimensions.

The POMDP setup — Output feedback

Problem formulation

Notation

State : $X_t \in \mathbb{R}^n$
Action : $U_t \in \mathbb{R}^m$
Observation: $Y_t \in \mathbb{R}^p$

Dynamics $X_{t+1} = \mathbf{A}_t X_t + \mathbf{B}_t U_t + W_t$, where $\mathbf{A}_t \in \mathbb{R}^{n \times n}$, $\mathbf{B}_t \in \mathbb{R}^{n \times m}$, $W_t \in \mathbb{R}^n$.

Observations $Y_t = \mathbf{C}_t X_t + N_t$, where $\mathbf{C}_t \in \mathbb{R}^{p \times n}$, $N_t \in \mathbb{R}^p$.

Random variables Primitive R.V.s $\{X_1, N_{1:T}, W_{1:T}\}$ are independent and Gaussian with

$$X_1 \sim \mathcal{N}(0, \Sigma_X), \quad N_t \sim \mathcal{N}(0, \Sigma_N), \quad W_t \sim \mathcal{N}(0, \Sigma_W).$$

Cost

Per step cost : $c_t(x_t, u_t) = \|x_t\|^2_{\mathbf{Q}_t} + \|u_t\|^2_{\mathbf{R}_t}$
Terminal reward: $c_T(x_T) = \|x_T\|^2_{\mathbf{Q}_T}$
where $\|X\|_{\mathbf{Q}} = X^\top \mathbf{Q} X$ and for all t, $\mathbf{Q}_t = \mathbf{Q}_t^\top \succeq 0$ and $\mathbf{R}_t = \mathbf{R}_t^\top \succ 0$.

Control objective Choose $U_t = g_t(Y_{1:t}, U_{1:t-1})$ so as to minimize

$$J(\mathbf{g}) = \mathbb{E}^{\mathbf{g}}\Big[\sum_{t=1}^{T-1} c_t(X_t, U_t) + c_T(x_T)\Big]$$

Sufficient statistic for control

Theorem (Sufficient statistic) The state-estimate $\hat{X}_t = \mathbb{E}[X_t \mid Y_{1:t}, U_{1:t-1}]$ is a sufficient statistic for control, i.e., there is no loss of optimality in restricting attention to control laws of the form: $U_t = g_t(\hat{X}_t)$.

Kalman filtering This sufficient statistic is updated using the Kalman filtering equations

$$\hat{X}_{t+1} = \mathbf{A}_t \hat{X}_t + \mathbf{B}_t \hat{u}_t + \mathbf{K}_{t+1} \big[Y_{t+1} - \mathbf{C}_{t+1} \hat{X}_t \big]$$

where $\mathbf{K}_t$ is the Kalman gain given by

$$\mathbf{K}_{t+1} = \mathbf{L}_t \big[\mathbf{\Sigma}_N + \mathbf{C}_t \mathbf{P}_t \mathbf{C}_t^\top \big]^{-1} \qquad \text{with} \quad \mathbf{L}_t = \mathbf{A}_t \mathbf{P}_t \mathbf{C}^\top.$$

The initial mean is given by

$$\hat{X}_1 = \dots$$

and the covariance matrices $\mathbf{P}_t$ are precomputable and are given by forward Riccati difference equation

$$\mathbf{P}_{t+1} = \mathbf{A}_t \mathbf{P}_t \mathbf{A}_t^\top + \mathbf{\Sigma}_W - \mathbf{L}_t \big[\mathbf{\Sigma}_N + \mathbf{C}_t \mathbf{P}_t \mathbf{C}_t^\top \big]^{-1} \mathbf{L}_t^\top$$

with

$$\mathbf{P}_1 = \dots$$

Structure of optimal controller

Theorem The value function at time t is

$$V_t(\hat{X}_t) = \|\hat{X}_t\|^2_{\mathbf{S}_t} + \alpha_t$$

and the optimal control action is

$$U_t = -\mathbf{H}_t \hat{X}_t$$

where $\mathbf{S}_t$ and $\mathbf{H}_t$ follow the same recursion as before and

$$\alpha_T = \mathrm{Tr}[\mathbf{P}_T \mathbf{Q}_T]$$

$$\alpha_t = \alpha_{t+1} + \mathrm{Tr}[\mathbf{P}_t \mathbf{Q}_t + (\mathbf{\Sigma}_W + \mathbf{A}_t \mathbf{P}_t \mathbf{A}_t^\top - \mathbf{P}_{t+1}) \mathbf{S}_{t+1}]$$

$$= \mathrm{Tr}[\mathbf{P}_T \mathbf{Q}_T] + \sum_{\tau=t}^{T-1} \mathrm{Tr}[\mathbf{P}_t \mathbf{Q}_t + (\mathbf{\Sigma}_W + \mathbf{A}_t \mathbf{P}_t \mathbf{A}_t^\top - \mathbf{P}_{t+1}) \mathbf{S}_{t+1}]$$

Further Reading

1. The Kalman filter (and its generalization to non-linear systems called extended Kalman filter) was used by NASA in the Ranger, Mariner, and Apollo missions, including the lunar module of Apollo 11. For a history of Kalman filtering in NASA see:

    Leonard A. McGee and Stanley F. Schmidt, "Discovery of the Kalman Filter as a practical tool for aerospace and industry," NASA Technical Memorandum TM-86847, Nov 1985.

    http://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/19860003843_1986003843.pdf

2. A slightly more cumbersome form of "Kalman filter" was derived in:

    Peter Swerling, "A proposed stagewise differential correlation procedure for satellite tracking and prediction," RAND Corporation report P-1292, Jan 1958; also published in Journal of Astronomical Science, vol 6, 1959

Appendices on background material

Appendix: Positive definite matrices

Positive definite

A $n \times n$ symmetric matrix $\mathbf{M}$ is

- positive definite (written as $\mathbf{M} \succ 0$) if

$$\text{for any } x \neq 0, x \in \mathbb{R}^n, \quad x^\top \mathbf{M} x > 0.$$

- positive semi-definite (written as $\mathbf{M} \succeq 0$) if

$$\text{for any } x \neq 0, x \in \mathbb{R}^n, \quad x^\top \mathbf{M} x \geqslant 0.$$

Eigenvalues characterization

A symmetric matrix is positive definite (resp., positive semi-definite) if and only if all of its eigenvalues are positive (resp., non-negative).

Examples

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3x_1^2 + 2x_2^2 \implies \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \succ 0.$$

$$\begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2x_2^2 \implies \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix} \succeq 0.$$

Appendix: Linear estimation of Gaussian signals

Conditional expectation of Gaussian vectors

Let (X, Y) be jointly Gaussian with mean $\mu = (\mu_X, \mu_Y)$ and covariance $\Sigma = \begin{bmatrix} \mathbf{\Sigma}_{XX} & \mathbf{\Sigma}_{XY} \\ \mathbf{\Sigma}_{XY}^{\top} & \mathbf{\Sigma}_{YY} \end{bmatrix}$. Then,

$$\mathbb{E}[X|Y] = \mu_X + \mathbf{\Sigma}_{XY}\mathbf{\Sigma}_{YY}^{-1}(Y - \mu_Y)$$

is a Gaussian random variable with mean μ_X and covariance $\mathbf{\Sigma}_{XY}\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{XY}^{\top}$.

The mean square error $(X - \mathbb{E}[X|Y])^2$ is $\mathrm{Tr}[\mathbf{\Sigma}_{XX} - \mathbf{\Sigma}_{XY}\mathbf{\Sigma}_{YY}^{-1}\mathbf{\Sigma}_{XY}^{\top}]$.