Control using partial observations

Sometimes we may need to control a system of which we don't know the state. This sounds like a completely impossible problem. But, by using the output of the system as well as possible, it is still possible to do this. How it works is explained below.

1 Information states and the separation principle

Consider a stochastic control system of the form

$$x(t+1) = f(t, x(t), u(t), v(t)), \tag{1.1}$$

$$y(t) = h(t, x(t), u(t), v(t)). \tag{1.2}$$

Our task now is to find a control law g which minimizes a cost function like

$$J(g) = E_g\left[\sum_{s=0}^{t_1-1} b(s, x^g(s), u^g(s)) + b_1(x^g(t_1))\right]. \tag{1.3}$$

To determine the control law g, we need information. We can summarize this information in a so-called **information state**. An example of an information state is

$$z(t) = [y(t-1), \ldots, y(0), u(t-1), \ldots, u(0)]^T. \tag{1.4}$$

There is just one problem: this information state grows with time. An information state which does not grow with time is generally more convenient to use.

The information state is descibed by its recursive relation

$$z(t+1) = f_1(t, z(t), y(t), u(t)). \tag{1.5}$$

We say that an information state is a **sufficient information state for the state given past outputs and inputs** if it provides sufficient information about the state x of the system. To be more precise, if

$$E[\exp(iw^T x(t))|F_{t-1}^y \vee F_{t-1}^u] = E[\exp(iw^T x(t))|F^{z(t)} \vee F^{u(t-1)}]. \tag{1.6}$$

We now say that a control law h_1 is a **separated control law** if it uses a sufficient information state z. That is, if

$$u(t) = h_1(t, z(t), y(t)). \tag{1.7}$$

The set of all separated control laws g is denoted by $G_s \subset G$. An optimal stochastic control problem is said to have the **separation property** if the optimal control law is a separated control law.

It is nice if a problem has the separation property, because separated control laws are relatively easy to find. In fact, sometimes we only search for separated control laws, even when we are not dealing with an optimal stochastic control problem. This means that we do not find the most optimal control law, but we do find a control law close to the optimal one in a relatively easy way. This principle is calle the **separation principle**.

2 Control of partially observed systems

2.1 The problem statement

Consider the system

$$x(t+1) = A(t)x(t) + B(t)u(t) + M(t)v(t), \tag{2.1}$$

$$y(t) = C(t)x(t) + D(t)u(t) + N(t)v(t). \tag{2.2}$$

Now also suppose that we don't know the state x of the system. But our goal is to come up with a control law g which minimizes a cost function $J(g)$. How do we do this?

It turns out that this problem has a very special property. We have

$$F_t^{y,g} \vee F_t^{u,g} = F_t^{y,0} \vee F_t^{u,0}. \tag{2.3}$$

In this equation $F_t^{y,0}$ is the σ-algebra obtained if we use the control law $u = 0$. In other words, it does not matter which control law we use, we always get the same information from the system.

So what does this mean? It means that we can simply use a Kalman filter to estimate the state. And this works for every control law g which we might use. We then get a state estimate $\hat{x}^g(t)$. The Kalman filter used is

$$\begin{aligned}
\hat{x}^g(t+1) &= A(t)\hat{x}^g(t) + B(t)u^g(t) + K(t)\left(y(t) - C(t)\hat{x}^g(t) - D(t)u^g(t)\right), && (2.4)\\
I_{11}(t) &= A(t)Q_f(t)A(t)^T + M(t)Q_v(t)M(t)^T, && (2.5)\\
I_{12}(t) &= A(t)Q_f(t)C(t)^T + M(t)Q_v(t)N(t)^T, && (2.6)\\
I_{22}(t) &= C(t)Q_f(t)C(t)^T + N(t)Q_v(t)N(t)^T, && (2.7)\\
Q_f(t+1) &= I_{11}(t) - I_{12}(t)I_{22}^{-1}(t)I_{12}(t)^T, && (2.8)\\
K(t) &= I_{12}(t)I_{22}^{-1}(t). && (2.9)
\end{aligned}$$

Now all that is left to do is find a good control law g. And for that, we can simply use the state estimate $\hat{x}$.

2.2 Finding the value function

In this problem, we are using Gaussian stochastic parameters. So, given the mean m and the variance Q, these Gaussian parameters have a PDF of

$$G(w; m, Q) = \frac{1}{\sqrt{2\pi \det(Q)}} \exp\left(-\frac{1}{2}(m-w)^T Q^{-1}(m-w)\right), \tag{2.10}$$

where w is the variable of the PDF. We need to use this function to find the value function $V(t, x)$ for our problem. This function is defined recursively. At the final time t_1, it can be found using

$$V(t_1, x) = \int_X b_1(w) G(w; x, Q(t_1))\, dw. \tag{2.11}$$

The subscript X means we integrate over the entire state space X. To find the value function at earlier times, we use the recursive relation

$$V(t, x) = \inf_{u \in U} \left(\int_X b(t, w, u) G(w; x, Q(t))\, dw + \int_X V(t+1, w) G(w; A(t)x + B(t)u, Q_u(t+1)) \right). \tag{2.12}$$

The parameter $Q_u(t)$ is the variance matrix obtained at time $t+1$ if we choose an input u. Now, for every control law $g \in G$, we have $V(t, \hat{x}^g(t)) \leq J(g, t)$.

The question remains, how do we find the optimal control law? Given that we are in a time t and an estimated state $\hat{x}$, we simply need to select the input u which minimizes the above relation. If, for all $x \in X$, there is an input $u^* \in U$ which minimizes the above relation, then the map $h_1(t, x) = u^*$ forms the optimal control law.

3 Special cost functions

3.1 The quadratic cost function

Let's examine some special cases. In fact, let's examine the quadratic cost function

$$J(g) = E_g \left[\sum_{s=0}^{t_1-1} \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L(s) \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} + x^g(t_1)^T L_1 x^g(t_1) \right]. \tag{3.1}$$

Here, the matrix $L(t)$ can be split up according to

$$L(t) = \begin{bmatrix} L_{11}(s) & L_{12}(s) \\ L_{12}(s)^T & L_{22}(s) \end{bmatrix}, \tag{3.2}$$

where it is assumed that $L(t) = L(t)^T \geq 0$ and $L_{22}(t) = L_{22}(t)^T > 0$. The optimal control law is now given by

$$u(t) = F(t)\hat{x}(t), \tag{3.3}$$

where $\hat{x}(t)$ is given by the Kalman filter as defined in equation (2.4). The matrix $F(t)$ is recursively defined as

$$\begin{aligned}
H_{11}(t) &= A^T(t)Q_c(t)A(t) + L_{11}(t), & (3.4)\\
H_{12}(t) &= A^T(t)Q_c(t)B(t) + L_{12}(t), & (3.5)\\
H_{22}(t) &= B^T(t)Q_c(t)B(t) + L_{22}(t), & (3.6)\\
Q_c(t+1) &= H_{11}(t) - H_{12}(t)H_{22}^{-1}(t)H_{12}(t)^T, & (3.7)\\
F(t) &= -H_{22}^{-1}(t)H_{12}^T(t). & (3.8)
\end{aligned}$$

This solution can be proven using the value function. This function is given by

$$V(t,x) = x^T P(t) x + r(x). \tag{3.9}$$

Here, the function $r(x)$ is recursively defined such that $r(t_1) = \mathrm{tr}(L_1 Q(t_1))$ and

$$r(t) = r(t+1) + \mathrm{tr}(L_{11}Q(t) + Q(t+1)P(t+1)) + \mathrm{tr}(K(t)I_{22}(t)K(t)^T P(t+1)). \tag{3.10}$$

By the way, tr(.) is the trace function, being the sum of the diagonal elements of the matrix.

3.2 Infinite horizon cost functions

Let's examine a time-invariant system. Previously, there always was some final time t_1. What happens if $t_1 = \infty$? In this case, our cost function will become infinite, which is not convenient. So we need a different cost function. Options are the **discounted quadratic cost function** and the **infinite-horizon average quadratic cost function**, respectively defined as

$$\begin{aligned}
J_{dc}(g) &= E_g \left[\sum_{s=0}^{\infty} r^s \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} \right], & (3.11)\\
J_{av}(g) &= \lim_{t\to\infty} \frac{1}{t} E_g \left[\begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix}^T L \begin{bmatrix} x^g(s) \\ u^g(s) \end{bmatrix} \right]. & (3.12)
\end{aligned}$$

Another cost function is the **minimum variance cost function**. It is used when one wants to minimize the variance of the state. This cost function is defined as

$$J_{mv}(g) = \lim_{t\to\infty} \frac{1}{t} E_g \left[\sum_{s=0}^{t-1} x^g(s)^T L x^g(s) \right]. \tag{3.13}$$

3.3 Solving the infinite horizon problem

In literature, a certain procedure is often used to solve this problem. However, there is no proof yet that this procedure actually minimizes the cost functions. But we will discuss it anyway.

The idea is that we simply take the limit case of $t \to \infty$ of the previous problem. In this case, Q satisfies the algebraic Riccati equation of Kalman filtering

$$Q = AQA^T + MVM^T - (AQC^T + MVN^T)(CQC^T + NVN^T)^{-1}(AQC^T + MVN^T)^T. \tag{3.14}$$

The matrix P satisfies the algebraic Riccati equation of control

$$P = A^T PA + L_{11} - (A^T PB + L_{12})(B^T PB + L_{22})^{-1}(A^T PB + L_{12})^T. \tag{3.15}$$

Using Q and P, we can find the matrices K and F, according to

$$K = (AQC^T + MVN^T)(CQC^T + NVN^T)^{-1}, \tag{3.16}$$

$$F = (B^T PB + L_{22})(A^T PB + L_{12})^T. \tag{3.17}$$

It can now be shown that we have

$$\mathrm{spec}\,(A - KC) \subset \mathbb{D}_o, \qquad \mathrm{spec}\,(A + BF) \subset \mathbb{D}_o \quad \text{and} \quad \mathrm{spec}\begin{pmatrix} A & BF \\ KC & A + BF - KC \end{pmatrix} \subset \mathbb{D}_o. \tag{3.18}$$

The control law which supposedly minimizes the infinite horizon cost functions is

$$u(t) = Fz(t), \qquad \text{with} \qquad z(t+1) = Az(t) + Bu(t) + K(y(t) - Cz(t) - Du(t)). \tag{3.19}$$