Stochastic control

In this chapter, we're going to examine how we can control a system using stochastic control. First, we examine some basic principles of stochastic control. After that, we'll look at an example of a field where stochastic control theory can often be applied: statistical decision problems.

1 Basic principles of stochastic control

1.1 Information structures

Let's examine the system

$$x(t+1) = Ax(t) + Bu(t) + Mv(t), \tag{1.1}$$
$$y(t) = Cx(t) + Du(t) + Nv(t). \tag{1.2}$$

We can now use the input $u(t)$ to control the system. To control the system, we need information. But luckily, information is available. We say that, at every time $t \in T$, the σ-algebra G_t specifies the available information. The family of all such σ-algebras G_t, being $\{G_t, t \in T\}$, is called the **information structure**.

There are different types of information structures. The type depends on which data is available.

- The **past-output information structure** is $\left\{F_{t-1}^{y-}, t \in T\right\}$. So, we have all previous outputs available as data. (But the states x are not available.)
- The **current output information structure** is $\left\{F^{y(t)}, t \in T\right\}$. So, only the current output is available.
- The **past-state information structure** is $\left\{F_t^{x-}, t \in T\right\}$. So, all previous states are available.
- The **Markov information structure** (also called the **current state information structure**) is $\left\{F^{x(t)}, t \in T\right\}$. So, only the current state is available.

1.2 Control laws

Based on the information structure, we can make a control law. A **control law** (also called **control policy**) is a measurable mapping from the available data to the input space U. How the control law is called depends on the information structure that is used. We will examine the most important control laws now.

- The **past-output control law** uses the past-output information structure. So, for every $t \in T$, g_t is a measurable map $g_t : Y^t \to U$. This implies that we can also see g itself as a mapping $g : T \times Y^t \to U$.
- The **output control law** uses the current output information structure. So we now have $g : T \times Y \to U$.
- The **past-state control law** uses the past-state information structure. So now $g : T \times X^{t+1} \to U$.
- The **Markov control law** uses the Markov information structure. So now $g : T \times X \to U$. We denote the set of all possible Markov control laws by G_M.
- The **stationary Markov control law** also uses the Markov information structure. But now the control law g does not depend on time. So, $g : X \to U$.

Let's suppose that we use a control law g. For example, we use the Markov control law. The resulting control system parameters are then written with the superscript g. So, we have

$$x^g(t+1) = Ax^g(t) + Bu^g(t) + Mv(t), \tag{1.3}$$
$$y^g(t) = Cx^g(t) + Du^g(t) + Nv(t). \tag{1.4}$$

In this equation, we have $u^g(t) = g(t, x^g(t))$. The above system representation is called a **closed-loop stochastic control system**.

1.3 Control objectives

The question remains, which control law do we use? We usually choose a control law such that control objectives are met. A **control objective** is a property of the closed-loop control system which we should strive to attain. Examples of control objectives are

- Making the system **stable**.
- **Suppressing noise**.
- Optimizing a **performance measure**. For example, we might want to choose g such that a cost function J is minimized. The cost function j can then have a form like

$$J = E\left[\sum_{t=0}^{t_1} (c_1 x(t) + c_2 u(t))\right]. \tag{1.5}$$

- Making the system **robust**. Robustness means that, even when deviations are applied in the model, the system still has a satisfactory performance.

The **stochastic control problem** is now defined as the problem of finding a control law g such that the control objectives are satisfied as well as possible. Solving this problem consists of two steps. First, in **control synthesis**, possible control laws g need to be generated. Then, in **control design**, the best of these control laws g needs to be chosen. In practice, this often means that the numerical parameters of the control law need to be chosen.

2 Statistical decision problems

2.1 Statistical decision problems

Statistical decision problems are often good examples of stochastic control problems. Let's suppose that we have x money. We can invest this in 2 investment opportunities. (It works the same when there are more investment opportunities.) The two investment opportunities return $y_1 = r_1 x_1$ and $y_2 = r_2 x_2$, respectively, where r_1 and r_2 are random variables and x_1 and x_2 are the amount of money invested in r_1 and r_2, respectively. The total return which you get is thus $y = y_1 + y_2$. The question now is how to invest x. Which fraction u should we invest in r_1 and which fraction $(1 - u)$ should we invest in r_2?

To solve this problem, we need a **utility function** $U(y)$. This utility function is a measure of how 'happy' you are with a return y. For most normal people, this is a concave function. (That is, $d^2U/dy^2 < 0$. Initially, people are very happy when they get more money. But, as people get richer, the extra happiness decreases if they get more money.) Of course, y is a random variable as well. So, we need to select u such that the expected utility $E[U(y)]$ is maximized.

2.2 An example

Let's demonstrate the above procedure with an example. Let's say that opportunity 1 is a 'sure' investment opportunity: r_1 is always 1.5. On the other hand, opportunity 2 is a 'risky' investment opportunity: there is a chance of 50% that $r_2 = 3$, but also a chance of 50% that $r_2 = 1$. We also define the utility function as $U(y) = 6y - y^2$. We now have

$$y = r_1 x_1 + r_2 x_2 = r_1 u x + r_2 (1 - u) x. \tag{2.1}$$

We should thus maximize

$$E[U(y)] = E\left[6(r_1ux + r_2(1-u)x) - (r_1ux + r_2(1-u)x)^2\right]. \tag{2.2}$$

Using the data given for r_1 and r_2 gives

$$E[U(y)] = \frac{1}{2}\left(6(1.5ux + (1-u)x) - (1.5ux + (1-u)x)^2\right) + \frac{1}{2}\left(6(1.5ux + 3(1-u)x) - (1.5ux + 3(1-u)x)^2\right). \tag{2.3}$$

We can find the maximum of this equation by differentiating for u. This then shows that the maximum occurs at $u = 2/5$.

2.3 Risk

Risk plays an important role in statistical decision problems. We say that a decision makes with a utility function $U(y)$ is **risk averse** if $E[U(y)] < U(E[y])$. That is, he prefers the certain pay-off $U(E[y])$ above the uncertain pay-off $E[U(y)]$. Similarly, the decision maker is **risk preferring** if $E[U(y)] > U(E[y])$ and **risk neutral** if $E[U(y)] = U(E[y])$.

We can also define the **index of absolute risk aversion** $r(y)$. Assuming that U is twice differentiable, it is defined as

$$r(y) = -\frac{U''(y)}{U'(y)}. \tag{2.4}$$

The index $r(E[y])$ is roughly proportional to the amount of money one would pay to avoid risks. Thus, if $r(E[y]) > 0$, then the person would pay to avoid risks and is thus risk averse. Similarly, if $r(E[y]) < 0$, the person would pay to have risks and is thus risk preferring.