



Why

We want to model uncertain outcomes in dynamical systems.¹

Definition

Let $\mathcal{X}_0, \mathcal{X}_1, \dots, \mathcal{X}_T$ and $\mathcal{U}_0, \mathcal{U}_1, \dots, \mathcal{U}_{T-1}$ be sets. Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space. Let $\mathcal{W}_0, \dots, \mathcal{W}_T$. Let $w_t : \Omega \rightarrow \mathcal{W}_t$ for $t = 0, \dots, T$ be random variables. For $t = 0, \dots, T - 1$, let $f_t : \mathcal{X}_t \times \mathcal{U}_t \times \mathcal{W}_t \rightarrow \mathcal{X}_{t+1}$.

We call the sequence

$$\mathcal{D} = ((\mathcal{X}_t)_{t=0}^T), (\mathcal{U}_t)_{t=0}^{T-1}, (w_t)_{t=0}^{T-1}, (f_t)_{t=1}^{T-1})$$

a *stochastic discrete-time dynamical system*. We call w_t the *noise* variables.

Problem

Let $x_0 : \Omega \rightarrow \mathcal{X}_0$ be a random variable. Define $x_1 : \Omega \rightarrow \mathcal{X}_1, \dots, x_T : \Omega \rightarrow \mathcal{X}_T$ by

$$x_{t+1} = f_t(x_t, u_t, w_t)$$

for $t = 0, \dots, T - 1$. Roughly speaking, the state transition functions are nondeterministic. In other words, it is uncertain which state we will arrive in given our current state and action. The choice u_t only determines the distribution of x_{t+1} . Here

¹Future editions will expand.

x_0 is (still) called the *initial state* and is a random variable, usually assumed independent of the w_t .

Let $g_t : \mathcal{X}_t \times \mathcal{U}_t \times \mathcal{W}_t \rightarrow \mathbf{R} \cup \{\infty\}$ for $t = 0, \dots, T-1$ and $g_T : \mathcal{X}_T \times \mathcal{W}_T \rightarrow \mathbf{R} \cup \{\infty\}$. We call $(x_0, \mathcal{D}, (g_t)_{t=0}^T)$ a *stochastic dynamic optimization problem*. As with dynamic optimization problems, we call g_t the *stage cost function* and g_T the *terminal cost function*. It is common for these to not depend on w_T (in other words, to be deterministic). It is also common for these to take infinite values to encode constraints.

As before, a stochastic dynamic optimization problem is just an optimization problem. Define $U = \mathcal{U}_0 \times \mathcal{U}_1 \times \dots \times \mathcal{U}_{T-1}$ and let $u \in U$. Define $C : \Omega \rightarrow \mathbf{R}$ by

$$C = \sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T, w_T).$$

We call C the *total cost* for actions u . It is a random variable.

Define $J : U \rightarrow \mathbf{R} \cup \{\infty\}$ by

$$J = \mathbf{E} \left(\sum_{t=0}^{T-1} g_t(x_t, u_t, w_t) + g_T(x_T, w_T) \right).$$

$J(u)$ is the *expected total cost* for inputs u .

The optimization problem is (U, J) . In other words, the objective is the mean total stage cost plus the terminal cost.

Other terminology

Stochastic dynamic optimization problems are frequently called *stochastic control problems*.

