Optimization and Control

J.R. Norris

November 22, 2007

1 Controllable dynamical systems

Controllable dynamical systems may be considered both in discrete time, with parameter $n \in \mathbb{Z}^+ = \{0, 1, 2, \dots\}$, and in continuous time, with parameter $t \in \mathbb{R}^+ = [0, \infty)$. They may be deterministic or stochastic, that is to say random. They are the basic objects of interest in this course. We now present the four main types.

1.1 Discrete-time, deterministic

Let S and A be sets. A *discrete-time controllable dynamical system* with *state-space* S and *action-space*[1] A is a map $f : \mathbb{Z}^+ \times S \times A \to S$. The interpretation is that, if at time n, in state x, we choose action a, then we move to state $f(n, x, a)$ at time $n + 1$. When f has no dependence on its first argument, we call the system *time-homogeneous*. A *control* is a map $u : \mathbb{Z}^+ \to A$. Given a starting time and state (k, x) and a control u, the *controlled sequence* $(x_n)_{n \geqslant k}$ is defined by $x_k = x$ and the equation[2]

$$x_{n+1} = f(n, x_n, u_n), \quad n \geqslant k.$$

In the time-homogeneous case[3], we shall usually specify only a starting state x and take as understood the starting time $k = 0$.

1.2 Discrete-time, stochastic

Assume for now that S is countable. Write $\mathrm{Prob}(S)$ for the set of probability measures on S. We identify each $p \in \mathrm{Prob}(S)$ with the vector $(p_y : y \in S)$ given by $p_y = p(\{y\})$. A *discrete-time stochastic controllable dynamical system*[4] with *state-space* S and *action-space* A is a map $P : \mathbb{Z}^+ \times S \times A \to \mathrm{Prob}(S)$. The interpretation is that, if at time n, in state

[1] In fact, since the actions available in each state are often different, it is convenient sometimes to specify for each state x an action-space A_x, which may depend on x. Then the product $S \times A$ is replaced everywhere by $\cup_{x \in S} \{x\} \times A_x$. This makes no difference to the theory, which we shall therefore explain in the simpler case, only reviving the notation A_x in certain examples.

[2] This is sometimes called the *plant equation*.

[3] We can always reduce to the time-homogeneous case as follows: define $\tilde{S} = \mathbb{Z}^+ \times S$ and, for $(n, x) \in \tilde{S}$, set $\tilde{f}((n, x), a) = (n + 1, f(n, x, a))$. If $(x_n)_{n \geqslant k}$ is the controlled sequence of f for starting time and state $\tilde{x} = (k, x)$ and control u, and if we set, for $n \geqslant 0$, $\tilde{u}_n = u_{k+n}$ and $\tilde{x}_n = (k + n, x_{k+n})$, then $(\tilde{x}_n)_{n \geqslant 0}$ is the controlled sequence of $\tilde{f}$ for starting state $\tilde{x}$ and control $\tilde{u}$.

[4] The term *Markov decision process* is also used, although P is not a process. However, we shall see that a choice of Markov control associates to P a Markov process.

x, we choose action a, then we move to y at time $n+1$ with probability $P(n,x,a)_y$. We write, for a function F on $\mathbb{Z}^+ \times S$,

$$PF(n,x,a) = \int_S F(n+1,y)P(n,x,a)(dy) = \sum_{y\in S} P(n,x,a)_y F(n+1,y).$$

Thus, also, $PF(n,x,a) = \mathbb{E}(F(n+1,Y))$, where Y is a random variable with distribution $P(n,x,a)$. Often, P will be *time-homogeneous*[5] and will be considered as a function $S\times A \to \mathrm{Prob}(S)$. Then we shall write, for a function F on S,

$$PF(x,a) = \sum_{y\in S} P(x,a)_y F(y).$$

A *control* is a map $u : S^* \to A$, where

$$S^* = \{(n, x_k, x_{k+1}, \dots, x_n) : k, n \in \mathbb{Z}^+, k \leqslant n, x_k, x_{k+1}, \dots, x_n \in S\}.$$

Given a control u and a starting time and state (k,x), we specify the distribution of a random process $(X_n)_{n\geqslant k}$ by the requirement that for all $n \geqslant k$ and all $x_k, \dots, x_n \in S$,

$$\begin{aligned}
&\mathbb{P}(X_k = x_k, X_{k+1} = x_{k+1}, \dots, X_n = x_n) \\
&\quad = \delta_{xx_k} P(k, x_k, u_k(x_k))_{x_{k+1}} \\
&\qquad \times P(k+1, x_{k+1}, u_{k+1}(x_k, x_{k+1}))_{x_{k+2}} \dots P(n-1, x_{n-1}, u_{n-1}(x_k, \dots, x_{n-1}))_{x_n}.
\end{aligned}$$

Thus, we determine the action that we take at each time n as a function u of n and of the history of the process up to that time. When we want to indicate the choice of control u and starting time and state (k,x), we shall write $\mathbb{P}^u_{(k,x)}$ in place of $\mathbb{P}$, and similarly $\mathbb{E}^u_{(k,x)}$ in place of $\mathbb{E}$. We take $k=0$ unless otherwise indicated and then write simply $\mathbb{P}^u_x$. We call $(X_n)_{n\geqslant k}$ the *controlled process*. A function $u : \mathbb{Z}^+ \times S \to A$ is called a *Markov control* and is identified with the control $(n, x_k, \dots, x_n) \mapsto u_n(x_n)$. A function $u : S \to A$ is called a *stationary Markov control*. In the *time-homogeneous* case, the controlled process determined by a stationary Markov control u is a (time-homogeneous) Markov chain on S, with transition matrix $P^u = (p^u_{xy} : x, y \in S)$ given by $p^u_{xy} = P(x, u(x))_y$. More generally, for any Markov control u, the controlled process $(X_n)_{n\geqslant 0}$ is a time-inhomogeneous Markov chain with time-dependent transition matrix $P^u(n) = (p^u_{xy}(n) : x, y \in S)$ given by $p^u_{xy}(n) = P(n, x, u_n(x))_y$.

Here is common way for a stochastic controllable dynamical system to arise: there is given a sequence of independent, identically distributed, random variables $(\varepsilon_n)_{n\geqslant 1}$, with values in a set E, say, and a function $G : \mathbb{Z}^+ \times S \times A \times E \to S$. We can then take $P(n,x,a)$ to be the distribution on S of the random variable $G(n,x,a,\varepsilon)$. Thus, for a function F on $\mathbb{Z}^+ \times S$, we have

$$PF(n,x,a) = \mathbb{E}(F(n+1, G(n,x,a,\varepsilon))). \tag{1}$$

Given a control u, this gives a ready-made way to realise the controlled process $(X_n)_{n\geqslant k}$, using the recursion[6]

$$X_{n+1} = G(n, X_n, U_n, \varepsilon_{n+1}), \quad U_n = u_n(X_k, \dots, X_n).$$

[5] A reduction to the time-homogeneous case can be made by a procedure analogous to that described in footnote 3. The details are left as an exercise.

[6] This is like the deterministic plant equation.

We shall call the pair $(G, (\varepsilon_n)_{n\geqslant 1})$ a *realised stochastic controllable dynamical system.* Every stochastic controllable dynamical system can be realised in this way; sometimes this is natural, at other times not. The notion of a realised stochastic controllable dynamical system provides a convenient way to generalize our discussion to the case where S is no longer countable. We shall consider in detail the case where $S = \mathbb{R}^n$, where the random variables ε_n are Gaussian, and where G is an affine function of x and ε.

1.3 Continuous-time, deterministic

Take now as state-space $S = \mathbb{R}^d$, for some $d \geqslant 1$. A *time-dependent vector field* on $\mathbb{R}^d$ is a map $b : \mathbb{R}^+ \times \mathbb{R}^d \to \mathbb{R}^d$. Given a starting point $x_0 \in \mathbb{R}^d$, we can attempt to define a continuous path $(x_t)_{t\geqslant 0}$ in $\mathbb{R}^d$, called the *flow* of b, by solving the differential equation $\dot{x}_t = b(t, x_t)$ for $t \geqslant 0$, with initial value x_0. In examples, we shall often calculate solutions explicitly. In Section 18 we shall show that continuity of b, or just piecewise continuity in time, together with the Lipschitz condition (4), guarantees the existence of a unique solution, even if we cannot calculate it explicitly. The Lipschitz condition is in turn implied by the existence and boundedness of the gradient $\nabla b = \partial b/\partial x$, which is usually easy to check.

A *continuous-time controllable dynamical system* with *action-space* A is given by a map $b : \mathbb{R}^+ \times \mathbb{R}^d \times A \to \mathbb{R}^d$. We interpret this as meaning that, if at time t, in state x, we choose action a, then we move at that moment with velocity $b(t, x, a)$. A *control* is a map $u : \mathbb{R}^+ \to A$. Given a control u, we obtain a vector field b^u by setting $b^u(t, x) = b(t, x, u_t)$. Then, given a starting time and place (s, x), the *controlled path* $(x_t)_{t\geqslant s}$ is defined by the differential equation $\dot{x}_t = b^u(t, x_t)$ for $t \geqslant s$, with initial value $x_s = x$. More generally, it is sometimes convenient to consider as a control a map $u : \mathbb{R}^+ \times \mathbb{R}^d \to A$. Then we set $b^u(t, x) = b(t, x, u(t, x))$ and solve the differential equation as before.

1.4 Continuous-time, stochastic

The most common continuous-time Markov processes fall into two types, jump processes and diffusions, each of which has a controllable counterpart. For simplicity, we give details only for the time-homogeneous case.

We shall consider jump processes only in the case where the state-space S is countable. In this context, Markov processes are called Markov chains[7]. A Markov chain is specified by a Q-matrix Q. Given a starting point $x_0 \in S$, there is an associated continuous-time Markov chain $(X_t)_{t\geqslant 0}$, starting from x_0, with generator matrix Q.

A *continuous-time jump-type stochastic controllable dynamical system* with *state-space* S and *action-space* A is given by a pair of maps $q : S \times A \to \mathbb{R}^+$ and $\pi : S \times A \to \mathrm{Prob}(S)$. We insist that $\pi(x, a)$ have no mass at x. If action a is chosen, then we jump from x at rate $q(x, a)$ to a new state, chosen with distribution $\pi(x, a)$. A *stationary Markov control* is a map $u : S \to A$, and serves to specify a Q-matrix Q^u, and hence a Markov chain, by

$$q^u_{xx} = -q(x, u(x)), \quad q^u_{xy} = q(x, u(x))\pi(x, u(x))_y, \quad y \neq x.$$

[7]These are one of the subjects of the Part II course Applied Probability.

The problem is to optimize the Markov chain over the controls. Note that this type of system has no deterministic analogue, as the only deterministic continuous-time time-homogeneous Markov process of jump type is a constant.

A diffusion process is a generalization of the differential equation $\dot{x}_t = b(x_t)$. Fix $m \geqslant 1$ and specify, in addition to the vector field b, called in this context the *drift*, m further vector fields $\sigma_1, \ldots, \sigma_m$ on S. Take m independent Brownian motions $B^1, \ldots, B^m$ and attempt to solve the stochastic differential equation[8]

$$dX_t = \sum_i \sigma_i(X_t) dB_t^i + b(X_t) dt.$$

The intuition behind this equation is that we move from x in an infinitesimal time δt by a normal random variable with mean $b(x)\delta t$ and with covariance matrix $\sum_i \sigma_i(x)\sigma_i(x)^T \delta t$. The solution $(X_t)_{t \geqslant 0}$, is a Markov process in S having continuous paths, which is known as a diffusion process.

A *continuous-time diffusive stochastic controllable dynamical system* with *state-space* S and *action-space* A is given by a family of maps $\sigma_i : S \times A \to \mathbb{R}^d, i = 1, \ldots, m$, and $b : S \times A \to \mathbb{R}^d$. We assume that these maps are all continuously differentiable on $\mathbb{R}^d$. If action a is chosen, then, intuitively, we move from x in an infinitesimal time δt by a normal random variable with mean $b(x, a)\delta t$ and with covariance matrix $\sum_i \sigma_i(x, a)\sigma_i(x, a)^T \delta t$. On specifying a *stationary Markov control* $u : S \to A$, we obtain the coefficients for a stochastic differential equation by $\sigma_i^u(x) = \sigma_i(x, u(x))$ and $b^u(x) = b(x, u(x))$, and hence, subject to some regularity conditions, we can define a diffusion process. Stochastic differential equations, diffusions, and the controllable systems and controls just introduced all have straightforward time-dependent generalizations.

[8] This discussion is intended only to sketch the outline of the theory, which is treated in the Part III course Stochastic Calculus and Applications. Provided that $\sigma_1, \ldots, \sigma_m$ and b are all differentiable, with bounded derivative, the equation has a unique maximal local solution, just as in the deterministic case.