3. Finite-horizon dynamic optimization

Recall that the infimal cost function satisfies

$$V(k,x) = \inf_a \{ c(k,x,a) + V(k+1, f(k,x,a)) \}$$

or

$$V(k,x) = \inf_a \{ c(k,x,a) + \sum_y P(k,x,a)_y V(k+1,y) \}$$

So, if we know $V(k+1,\cdot)$, we can in principle calculate $V(k,\cdot)$.

In the <u>finite-horizon case</u>, there is a time horizon $n \in \mathbb{Z}^+$ and no costs are incurred after time n. Thus

$$c(m, x, a) = 0, \quad m \geq n+1.$$

Then we have

$$V(n, x) = C(x) = \inf_a c(n, x, a)$$

and we can compute the infimal cost function by a backwards recursion.

Proposition 3.1

Suppose we can find a control u, with controlled sequence $(x_0, \dots, x_n)$ such that

$$V(k, x_k) = c(k, x_k, u_k) + V(k+1, f(k, x_k, u_k)), \quad 0 \le k \le n-1.$$

Then u is optimal for $(0, x_0)$.

Proof Consider

$$M_k = \sum_{j=0}^{k-1} c(j, x_j, u_j) + V(k, x_k), \quad 0 \le k \le n.$$

For $0 \le k \le n-1$, we have $x_{k+1} = f(k, x_k, u_k)$, so

$$M_{k+1} - M_k = c(k, x_k, u_k) + V(k+1, x_{k+1}) - V(k, x_k) = 0.$$

Hence

$$V(0, x_0) = M_0 = M_n = \sum_{j=0}^{n-1} c(j, x_j, u_j) + C(x_n) = V^u(0, x_0). \quad \square$$

Example – managing spending and saving

You hold a capital sum in a building society.

At times $k = 0, 1, \ldots, n-1$ the building society pays you interest, which is $100 \times \theta$ % of the current capital sum.

You can choose to reinvest any fraction of the interest, adding to the capital sum.

No amounts can be withdrawn.

How do you maximize consumption by time $n-1$?

Proposition 3.2

Suppose we can find a Markov control u such that

$$V(k,x) = (c + PV)(k, x, u_k(x)), \qquad 0 \le k \le n-1, \quad x \in S.$$

Then u is optimal for $(0,x)$, for all x.

Proof Consider

$$M_k = \sum_{j=0}^{k-1} c(j, X_j, U_j) + V(k, X_k), \qquad 0 \le k \le n,$$

where $(X_0, \dots, X_n)$ is the (P,u)-Markov chain, starting from $(0,x)$.

For $0 \le k \le n-1$,

$$M_{k+1} - M_k = c(k, X_k, U_k) + V(k+1, X_{k+1}) - V(k, X_k),$$

so, for all $y \in S$,

$$\mathbb{E}^u_{(0,x)}(M_{k+1} - M_k \mid X_k = y) = (c + PV)(k, y, u_k(y)) - V(k,y) = 0.$$

Hence

$$V(0,x) = \mathbb{E}^u_{(0,x)}(M_0) = \mathbb{E}^u_{(0,x)}(M_n) = \mathbb{E}^u_{(0,x)}\left(\sum_{j=0}^{n-1} c(j, X_j, U_j) + C(X_n)\right). \qquad \square$$

A stochastic controllable dynamical system P can always be <u>realized</u> using a function

$$G : \mathbb{Z}^+ \times S \times A \times E \longrightarrow S$$

and a sequence of independent, identically distributed random variables $(\varepsilon_n)_{n \geq 1}$, such that

$$P(n, x, a)_y = \mathbb{P}(\, G(n, x, a, \varepsilon_{n+1}) = y \,).$$

Then, for any control u, if we set $X_k = x$ and

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

then $(X_n)_{n \geq k}$ is a (P, u)-controlled process starting from (k, x).

In many examples we are simply given the equation

$$X_{n+1} = G(n, X_n, U_n, \varepsilon_{n+1}), \qquad n \geq 0.$$

To write the optimality equation, we can then use

$$PV(n,x,a) = \mathbb{E}\left(V(n+1, X_{n+1}) \,\middle|\, X_n = x, U_n = a \right)$$
$$= \mathbb{E}\left(V(n+1, G(n,x,a,\varepsilon_{n+1})) \right).$$

Example – exercising a stock option

You hold an option to buy a share at a fixed price p, which can be exercised at any time $k = 0, 1, \dots, n-1$.

The share price satisfies

$$Y_{k+1} = Y_k + \varepsilon_{k+1}$$

where $(\varepsilon_k)_{k \geq 1}$ is a sequence of independent, identically distributed random variables, with $\mathbb{E}|\varepsilon_k| < \infty$.

How should you act to maximize your expected return?

Time to go

The *time to go* is $m = n - k$.

Assume that P is time-homogeneous and write

$$V_m(x) = V(k, x), \qquad C_m(x, a) = C(k, x, a).$$

Then $V_0 = C$ and the optimality equation can be written

$$V_{m+1}(x) = \inf_a \left(C_m + PV_m\right)(x, a), \qquad 0 \le m \le n-1, \quad x \in S.$$

In particular, if both P and c are time-homogeneous, and if we define

$$V_n^u(x) = \mathbb{E}_{(0,x)}^u \sum_{j=0}^{n-1} c(X_j, U_j), \qquad V_n(x) = \inf_u V_n^u(x),$$

where $(X_n)_{n \geq 0}$ is the (P, u)-controlled process starting from $(0, x)$, then $V_0 = 0$, and, for $n \geq 0$

$$V_{n+1}(x) = \inf_a (c + PV_n)(x, a), \quad x \in S.$$

Interchange arguments

Suppose we can control the order in which we perform a sequence of n tasks, labelled $\{1, \ldots, n\}$.

Write $C(\sigma)$ for the expected cost of performing the task in the order $\sigma = (\sigma_1, \ldots, \sigma_n)$.

The problem of optimizing over σ can be formulated as a finite-horizon dynamic optimization problem, but the calculation of V may be laborious.

The following alternative approach is called an 'interchange argument....

Look for a function f on $\{1,\dots,n\}$ such that, for all σ and all $0 \le i \le n-1$,

$$c(\sigma') < c(\sigma) \quad \text{whenever} \quad f(\sigma_i) < f(\sigma_{i+1})$$

where σ' is obtained from σ by interchanging the order of tasks σ_i and σ_{i+1}.

Then <u>the condition</u> $f(\sigma_1) \le \dots \le f(\sigma_n)$ <u>is necessary for the optimality of</u> σ.

This may already reduce the number of possible optimal orders to 1.

In any case, if we have also, for all σ and all $0 \le i \le n-1$,

$$c(\sigma') = c(\sigma) \quad \text{whenever} \quad f(\sigma_i) = f(\sigma_{i+1}),$$

then our optimality condition is also sufficient.

Example - quiz show tactics

Contestants in a quiz show are presented with n questions which they may attempt in any order. You believe that you can answer question i correctly with probability p_i.

A correct answer to question i is rewarded with a prize worth £a_i.

A wrong answer to any question ends the game.