Dynamic Programming Introduction[1]

Kenneth L. Judd, Hoover Institution

January 26, 2014

[1] Joint work with Yongyang Cai.

Dynamic Programming

Powerful tool for solving dynamic stochastic optimization problems

- Based on principle of recursion due to Bellman and Isaacs
- Replaces multiperiod optimization problems with a sequence of two-period problems

Applications

- Economics
  - Business investment
  - Life-cycle decisions: labor, consumption, education, portfolios
  - Economic policy
- Operations Research
  - Scheduling, queueing
  - Inventory
- Climate change
  - Economic response to climate policies
  - Optimal policy response to global warming problems

Discrete-Time Dynamic Programming

- Objective:

$$E\left\{\sum_{t=1}^{T} \pi(x_t,\ u_t,\ t) + W(x_{T+1})\right\}$$

- X: set of states
  - $\mathcal{D}$: the set of controls
  - $\pi(x,\ u,\ t)$ payoffs in period t, for $x \in X$ at the beginning of period t, and control $u \in \mathcal{D}$ is applied in period t. Time-dependent features such as discounting are included in $\pi(x,\ u,\ t)$.
  - $D(x,t) \subseteq \mathcal{D}$: controls which are feasible in state x at time t.
  - $F(A; x, u, t)$: probability that $x_{t+1} \in A \subset X$ conditional on time t control and state
- Value function

$$V(x,t) \equiv \sup_{\mathcal{U}(x,t)} E\left\{\sum_{s=t}^{T} \pi(x_s,\ u_s,\ s) + W(x_{T+1}) | x_t = x\right\}$$

- Bellman equation

$$V(x,t) = \sup_{u \in D(x,t)} \pi(x,\ u,\ t) + E\left\{V(x_{t+1},\ t+1) | x_t = x, u_t = u\right\}$$

- Existence: boundedness of π is sufficient

Autonomous, Infinite-Horizon Problem

- Objective:

$$\max_{u_t} E \left\{ \sum_{t=1}^{\infty} \beta^t \pi(x_t,\, u_t) \right\}$$

  - X: set of states
  - $\mathcal{D}$: the set of controls
  - $D(x) \subseteq \mathcal{D}$: controls which are feasible in state x.
  - $\pi(x,\, u)$ payoff in period t if $x \in X$ at the beginning of period t, and control $u \in \mathcal{D}$ is applied in period t.
  - $F(A; x, u)$: probability that $x^+ \in A \subset X$ conditional on current control u and current state x.
- Value function definition: if $\mathcal{U}(x)$ is set of all feasible strategies starting at x.

$$V(x) \equiv \sup_{\mathcal{U}(x)} E \left\{ \sum_{t=0}^{\infty} \beta^t \pi(x_t,\, u_t) \middle| x_0 = x \right\}$$

- Bellman equation for $V(x)$

$$V(x) = \sup_{u \in D(x)} \pi(x,\, u) + \beta\, E\, \left\{V(x^+)|x, u\right\} \equiv (TV)(x)$$

- Optimal policy function, $U(x)$, if it exists, is defined by

$$U(x) \;\in\; \arg \max_{u \in D(x)} \pi(x,\, u) + \beta\, E \left\{V(x^+)|x, u\right\}$$

- Standard existence theorem:
  If X is compact, $\beta < 1$, and π is bounded above and below, then the map

$$TV = \sup_{u \in D(x)} \pi(x, u) + \beta E \left\{V(x^+) \mid x, u\right\}$$

  is monotone in V, is a contraction mapping with modulus β in the space of bounded functions, and has a unique fixed point.

Deterministic Growth Example

- Problem:
$$V(k_0) = \max_{c_t} \quad \begin{array}{l} \sum_{t=0}^{\infty} \beta^t u(c_t), \\ k_{t+1} = F(k_t) - c_t \\ k_0 \text{ given} \end{array}$$
  - Bellman equation
$$V(k) = \max_{c} \; u(c) + \beta V(F(k) - c)$$
  - First-order condition
$$0 = u'(c) - \beta V'(F(k) - c)$$
  - Envelope theorem implies
$$V'(k) = \beta V'(F(k) - c) F'(k)$$
  - Solutions is policy function $C(k)$ and value function $V(k)$ satisfying
$$\begin{aligned} V'(k) &= u'(C(k)) F'(k) \\ V(k) &= u(C(k)) + \beta V(F(k) - C(k)) \end{aligned}$$
- The second equation defines the value function for an arbitrary policy function $C(k)$, not just for the optimal $C(k)$.

Stochastic Growth Accumulation

- Problem:

$$V(k,\theta) = \max_{c_t,\ell_t} E\left\{\sum_{t=0}^{\infty} \beta^t\, u(c_t)\right\}$$
$$k_{t+1} = F(k_t,\theta_t) - c_t$$
$$\theta_{t+1} = g(\theta_t,\varepsilon_t)$$
$$\varepsilon_t : \text{ i.i.d. random variable}$$
$$k_0 = k,\ \theta_0 = \theta.$$

- State variables:
    - k: productive capital stock, endogenous
    - θ: productivity state, exogenous
- Applications
    - Economic growth
    - Firm growth, monopoly or competitive
    - Wealth management: k is the vector of assets, $F(k,\theta)$ is the vector of gross income, $F - c$ is the vector of net changes in the states, $u(c)$ is the payoff from the c decisions.

- The dynamic programming formulation is

$$\begin{aligned} V(k,\theta) = \max_{c} \qquad & u(c) + \beta E\{V(F(k,\theta) - c, \theta^{+})|\theta\} \\ & \theta^{+} = g(\theta, \varepsilon) \end{aligned}$$

- The control law $c = C(k,\theta)$ satisfies the first-order conditions

$$0 = u_c\left(C(k,\theta)\right) - \beta\, E\left\{u_c(C(k^{+},\theta^{+}))F_k(k^{+},\theta^{+}) \mid \theta\right\},$$

where

$$k^{+} \quad \equiv F(k, L(k,\theta), \theta) - C(k,\theta),$$

General Stochastic Accumulation

- Problem:

$$V(k,\theta) = \max_{c_t,\ \ell_t} \ E\left\{\sum_{t=0}^{\infty} \beta^t \ u(c_t,\ell_t)\right\}$$
$$k_{t+1} = F(k_t,\ell_t,\theta_t) - c_t$$
$$\theta_{t+1} = g(\theta_t,\varepsilon_t)$$
$$k_0 = k,\ \ \theta_0 = \theta.$$

- State variables:
  - k: productive capital stock, endogenous
  - θ: productivity state, exogenous
- The dynamic programming formulation is

$$V(k,\theta) = \max_{c,\ \ell} \ u(c,\ell) + \beta E\{V(F(k,\ell,\theta) - c,\theta^+)|\theta\},$$

where θ^+ is next period's θ realization.