INTRO TO DYNAMIC PROGRAMMING

Quantitative Economics 2023

Piotr Żoch

November 6, 2023

A TYPICAL PROBLEM

Many problems seen in economics have a common structure:

- We observe the current state X_t at time t.
- We choose an action A_t at time t.
- We get a reward R_t at time t.
- The state progresses to X_{t+1} at time $t + 1$.

If the largest possible t is $T < \infty$, then we have a finite horizon problem.

Otherwise we have an infinite horizon problem.

EXAMPLE

Consider a problem of a firm that produces a good. The firm wants to maximize the expected present discounted value of profits:

$$\mathbb{E} \sum_{t=0}^{\infty} \left(\frac{1}{1+r} \right) \pi_t$$

- X_t is the current state of the firm. It can be the current level of capital, the current level of inventory, prices set by competitors...
- A_t is the action taken by the firm. It can be the level of production, the level of future inventory, the price of the good...
- R_t is the reward. Here it is the profit of the firm, π_t.
- X_{t+1} is the state of the firm in the next period. It can depend on the current state X_t and the action A_t taken.

EXAMPLE

- This is potentially an extremely complicated problem.
- For example: the state can include the demand for the good – and it could be random.
- Find actions for all possible future states...
- We will learn tools that can help us solve such problems.

PLAN

- Today we will study an example: McCall's job search model (1970).
- Exposition based on Stachurski and Sargent (2023).
- Next time: more general theory of dynamic programming.

MCCALL'S JOB SEARCH MODEL

TWO-PERIOD PROBLEM

- An unemployed agent receives a job offer at wage W_t.
- She can either accept the offer or reject it.
- If she accepts, she gets this wage permenantly.
- If she rejects, she gets unemployment benefit c.
- Wage offers are independent and identically distributed (i.i.d.) and nonnegative, with distribution ϕ:
  - $W \subset \mathbb{R}_+$ is a finite set of possible wages.
  - $\phi : W \to [0, 1]$ is a probability mass function, $\phi(w)$ is the probability of getting a wage w.
- The agent is risk-neutral and impatient. The utlity of getting y tomorrow is $\beta\, y$, with $\beta \in (0, 1)$.

TWO-PERIOD PROBLEM

- The agent lives for two periods and starts unemployed
- The question is: is it better to accept or wait for a better offer?
- What is the lowest wage that the agent should accept?
- We will start analyzing the problem by looking at the second period, $t = 2$: backward induction.

PERIOD T=1

- Suppose the agent is unemployed at $t = 2$.
- She gets a wage offer W_2.
- She can either accept or reject the offer.
- Since this is the last period of her life, she will accept if and only if

$$W_2 \geq c.$$

PERIOD T=2

- The agent gets a wage offer W_1.
- She can either (a) accept and get W_1 forever, or (b) reject and get c in period $t = 1$ and then get the maximum of W_2 and c in period $t = 2$.
- The utility of (a) is $W_1 + \beta W_1$. We call it the stopping value.
- The utility of (b) is $h_1 := c + \beta\, \mathbb{E} \max\{W_2, c\}$. We call it the continuation value.

$$h_1 = c + \beta \sum_{w' \in W} \max v_2(w') \phi(w'), \quad v_2(w') := \max\{w', c\}$$

- The agent will accept if and only if the stopping value is greater than the continuation value:

$$W_1 + \beta W_1 \geq h_1.$$

VALUE FUNCTION

- The key object in dynamic programming is the value function.
- It is a function that maps the state to the maximum expected present discounted value of future rewards.
- In our example, the state is the time t and the wage offer w.
- $v_2(w)$ is the value function at time $t = 2$ and wage w: the largest possible reward that the agent can get if she starts unemployed at $t = 2$ and gets a wage offer w.
- The time 1 value function is

$$v_1(w) := \max\left\{w + \beta w, c + \beta \sum_{w' \in W} v_2(w')\phi(w')\right\}.$$

TWO-PERIOD EXAMPLE

- This particular problem is easy to solve.
- Accept if

$$w \geq \frac{h_1}{1+\beta}$$

so the value function is

$$v_1(w) = \begin{cases} (1+\beta)\,w & \text{if } w \geq \frac{h_1}{1+\beta} \\ h_1 & \text{otherwise.} \end{cases}$$

- In context of this example, we call $w^* := \frac{h_1}{1+\beta}$ the reservation wage.
- We see that since h_1 is increasing in c, the reservation wage is higher when the unemployment benefit is higher.

THREE-PERIOD EXAMPLE

- Extend the model by one period, $t = 0$.
- The value function at $t = 0$ is

$$v_0(w) := \max\left\{w + \beta w + \beta^2 w, c + \beta \sum_{w' \in W} v_1(w')\phi(w')\right\}.$$

  where the formula for v_1 is from the previous slide.
- Key insight: at $t = 0$ it is like a two-period problem.
- All information about the future is summarized in the value function at $t = 1$.
- This is the standard approach: convert a complicated dynamic optimization problem into a sequence of two-period problems.

BELLMAN EQUATION

- Recall we had

$$
\begin{aligned}
v_2(w) &= \max\{w, c\} \\
v_1(w) &= \max\left\{w + \beta w, c + \beta \sum_{w' \in W} v_2\left(w'\right) \phi\left(w'\right)\right\} \\
v_0(w) &= \max\left\{w + \beta w + \beta^2 w, c + \beta \sum_{w' \in W} v_1\left(w'\right) \phi\left(w'\right)\right\}.
\end{aligned}
$$

- The recursive relationships between the value functions are called the Bellman equations.
- Warning: these equations are functional equations. We need to find functions, not numbers.

INFINITE HORIZON

- The three-period problem is also easy to solve.
- In fact, we can use the same approach (backward induction) as before for any finite horizon problem.
- What if the horizon is infinite? We no longer have the terminal period.
- Dynamic programming makes this problem tractable.

INFINITE HORIZON

- What is stopping value?
- If the worker accepts wage w she gets

$$w + \beta w + \beta^2 w + \beta^3 w + \cdots = \frac{w}{1-\beta}.$$

- What is the continuation value?
- If the worker rejects wage w she gets

$$c + \beta \sum_{w' \in W} v(w') \phi(w').$$

note that the value function is the same in all periods – there is always infinite remaining future.

INFINITE HORIZON

- Bellman equation is:

$$v(w) = \max\left\{\frac{w}{1-\beta}, c + \beta \sum_{w' \in W} v\left(w'\right)\phi\left(w'\right)\right\}.$$

- Principle of optimality, Bellman (1960): *An optimal policy has the property that whatever the initial state and the initial decisions it must constitute an optimal policy with regards to the state resulting from the first decision.*
- This is not that trivial, we will return to it (and prove it!) later.
- Intepretation: value function satisfies the Bellman equation.

CHALLENGE

- Bellman equation is:

$$v(w) = \max \left\{ \frac{w}{1-\beta}, c + \beta \sum_{w' \in W} v\left(w'\right) \phi\left(w'\right) \right\}.$$

- Once we have $v(w)$ we can characterize the optimal choice of the agent.
- Q: how to find $v(w)$? It is a function!
- Q: is there a solution?
- Q: is the solution unique?
- Q: what are the properties of the solution?
- A: we will learn how to answer these questions.