An introduction to dynamic programming in discrete time

Martin Branda

Charles University
Faculty of Mathematics and Physics
Department of Probability and Mathematical Statistics

Computational Aspects of Optimization

Dynamic programming

Finite T or infinite ∞ time horizon

$$\max_{A_t, c_t} \sum_{t=1}^{T \vee \infty} \left(\frac{1}{1+i} \right)^{t-1} u(c_t)$$
$$\text{s.t.} \tag{1}$$
$$A_t = (1+r)A_{t-1} + Y_t - c_t.$$

- u – **utility function**
- A_t – **state variables** representing total amount of resources available to the consumer.
- c_t – **control variables** maximizing the consumer's utility. It affects the resources available in the next period.
- Y_t – **exogenous income**
- $1/(1+i)$ – discount factor, r – exogenous interest rate

Dynamic programming

If we assume that there is a finite terminal period T:

$$
\begin{aligned}
V_1(A_0) &= \max_{A_t, c_t} \sum_{t=1}^{T} \left(\frac{1}{1+i}\right)^{t-1} u(c_t) \\
&= \max_{A_t, c_t} u(c_1) + \frac{1}{1+i} u(c_2) + \cdots + \left(\frac{1}{1+i}\right)^{T-1} u(c_T) \\
&= \max_{A_t, c_t} u(c_1) + \frac{1}{1+i} \left[\sum_{t=2}^{T} \left(\frac{1}{1+i}\right)^{t-2} u(c_t)\right] \\
& \quad \text{s.t.} \\
& \quad A_t = (1+r)A_{t-1} + Y_t - c_t.
\end{aligned}
$$

Dynamic programming

We rewrite the maximization problem recursively and obtain the **Bellman equation**

$$V_t(A_{t-1}) = \max_{A_t, c_t} u(c_t) + \frac{1}{1+i} V_{t+1}(A_t),$$

where $A_t = (1+r)A_{t-1} + Y_t - c_t$.

Moreover, since u does not depend on the time period, we can write

$$\begin{aligned} V(A_{t-1}) &= \max_{A_t, c_t} u(c_t) + \frac{1}{1+i} V(A_t), \\ &= \max_{c_t} u(c_t) + \frac{1}{1+i} V\Big((1+r)A_{t-1} + Y_t - c_t\Big), \end{aligned}$$

with $V_{T+1}(A_T) = V(A_T) \equiv 0$.

Bellman principle of optimality

An optimal policy has the property that whatever the initial state and initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision.

This principle is applied recursively (forward/backward recursion)...

Dynamic programming

First order optimality conditions

$$\begin{aligned}\frac{\partial V(A_{t-1})}{\partial c_t} &= 0,\\ \frac{\partial V(A_{t-1})}{\partial A_{t-1}} &= 0.\end{aligned}$$

In particular,

$$\begin{aligned}\frac{\partial V(A_{t-1})}{\partial c_t} &= u'(c_t) + \frac{1}{1+i} V'(A_t) \frac{\partial A_t}{\partial c_t},\\ \frac{\partial V(A_{t-1})}{\partial A_{t-1}} &= \frac{1}{1+i} V'(A_t) \frac{\partial A_t}{\partial A_{t-1}},\end{aligned}$$

where using $A_t = (1+r)A_{t-1} + Y_t - c_t$ we have

$$\frac{\partial A_t}{\partial c_t} = -1, \quad \frac{\partial A_t}{\partial A_{t-1}} = 1 + r.$$

Example: Cake eating problem

Cake eating problem:

- $u(c) = 2c^{1/2}$,
- $\Pi_0 = 1$, $\Pi_T = 0$,
- $\Pi_t = \Pi_{t-1} - c_t$ ...

Example: Cake eating problem

Bellman equation

$$V(\Pi_{t-1}) = \max_{c_t} \; u(c_t) + \frac{1}{1+i} V(\Pi_t),$$
$$\text{s.t. } \Pi_t = \Pi_{t-1} - c_t.$$

Optimality conditions

$$\begin{aligned}
\frac{\partial V(\Pi_{t-1})}{\partial c_t} &= u'(c_t) + \frac{1}{1+i} V'(\Pi_t) \frac{\partial \Pi_t}{\partial c_t} = 0,\\
\frac{\partial V(\Pi_{t-1})}{\partial \Pi_{t-1}} &= \frac{1}{1+i} V'(\Pi_t) \frac{\partial \Pi_t}{\partial \Pi_{t-1}} = 0,
\end{aligned}$$

From $\Pi_t = \Pi_{t-1} - c_t$

$$\frac{\partial \Pi_t}{\partial c_t} = -1, \; \frac{\partial \Pi_t}{\partial \Pi_{t-1}} = 1. \tag{2}$$

Example: Cake eating problem

Putting them together, we obtain

$$\frac{\partial V(\Pi_{t-1})}{\partial c_t} = u'(c_t) - \frac{1}{1+i} V'(\Pi_t) = 0, \tag{3}$$

$$\frac{\partial V(\Pi_{t-1})}{\partial \Pi_{t-1}} = \frac{1}{1+i} V'(\Pi_t) = 0, \tag{4}$$

Taking (3) for $t-1$

$$u'(c_{t-1}) - \frac{1}{1+i} V'(\Pi_{t-1}) = 0, \tag{5}$$

and plugging it into (4), we have

$$u'(c_{t-1}) = \frac{1}{1+i} u'(c_t),$$

which represents the **optimal path of the cake consumption**.

Example: Cake eating problem

For $u(c) = 2c^{1/2}$, we have

$$\begin{aligned}
u'(c_{t-1}) &= \frac{1}{1+i} u'(c_t), \\
(c_{t-1})^{-1/2} &= \frac{1}{1+i} (c_t)^{-1/2}, \\
c_t &= \left(\frac{1}{1+i} \right)^2 c_{t-1},
\end{aligned}$$

with initial and terminal conditions $\Pi_0 = 1$, $\Pi_T = 0$. If we denote $\beta = 1/(1+i)^2$, we obtain

$$c_t = \beta c_{t-1} = \beta^{t-1} c_1.$$

Using

$$c_t = \beta c_{t-1} = \beta^{t-1} c_1.$$

and

$$\Pi_0 - c_1 - c_2 - \cdots - c_T = \Pi_T = 0,$$

we have

$$(1 - \beta - \cdots - \beta^{T-1}) c_1 = \Pi_0,$$

and finally **optimal consumption**

$$\begin{aligned} \hat{c}_1 &= \frac{1-\beta}{1-\beta^T} \Pi_0, \\ \hat{c}_t &= \beta \hat{c}_{t-1} = \beta^{t-1} \hat{c}_1. \end{aligned}$$

Literature

- P. Kall, S.W. Wallace: **Stochastic Programming**, John Wiley & Sons, Chichester.
- M. C. Sunny Wong: **Dynamic Optimization: An Introduction**, Lecture Notes – University of Houston, 2013.