

## Introduction to Bellman equations

We will introduce the general idea of Bellman equations by considering a standard example from consumption theory.

Consider the following intertemporal optimization problem of an economic agent who lives two periods: In  $t = 0$  he is born and inherits wealth  $a_0$ . During the first period of his life he earns  $y_0$  and has to choose the level of consumption  $c_0$  from which he gets utility  $u(c_0)$ . He can take to period  $t = 1$  what he does not consume in the first period. His wealth in the second period consists of the previous period's "savings" and his income  $y_1$ . He has to consume  $c_1$  and then he dies.

Formalizing this problem, we get

$$\max_{c_0, c_1} \{u(c_0) + \beta u(c_1)\} \quad (0.1)$$

subject to the budget constraints

$$\begin{aligned} a_1 &= R_0(a_0 + y_0 - c_0) \\ a_2 &= R_1(a_1 + y_1 - c_1) \end{aligned} \quad (0.2)$$

where  $u(\cdot)$  is a standard utility function with all the desirable properties,  $\beta$  is the utility discount factor and  $R_0$  is the discount factor for his savings. We have the additional condition that  $a_2 \geq 0$  in order to rule out Ponzi schemes.

The standard way of solving this problem is by combining the budget constraints, setting up the Lagrangean and solving it.

$$L = u(c_0) + \beta u(c_1) + \lambda(a_0 + y_0 + R_0^{-1}y_1 - c_0 - R_0^{-1}c_1 - R_0^{-1}R_1^{-1}a_2) \quad (0.3)$$

The FOCs are

$$\begin{aligned} \frac{\partial L}{\partial c_0} &= u'(c_0) = \lambda \\ \frac{\partial L}{\partial c_1} &= \beta u'(c_1) = \lambda R_0^{-1} \end{aligned} \quad (0.4)$$

Combining the FOCs we get the standard Euler equation

$$u'(c_0) = \beta R_0 u'(c_1) \quad (0.5)$$

So far so good, but what if the consumer lived infinitely many periods? The problem we are now facing is

$$\max_{\{c_t\}_0^\infty} \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t) \right\} \quad (0.6)$$

subject to

$$a_{t+1} = R_t(a_t + y_t - c_t), \text{ for } t = 0, \dots, \infty \quad (0.7)$$

We are virtually facing infinitely many choices and budget constraints. In particular, this means that we have to find a path of optimal consumption  $\{c_0^*, c_1^*, \dots\}$ . One way of solving that problem is the method of Lagrange multipliers but there is an alternative and somewhat shorter way.

We will now introduce this approach due to Richard Bellman. It was his idea to subdivide complicated intertemporal problems into many two-period problems, in which the trade-off is between the "now" and "later".

Define the *value function*  $V(a_0)$  to be an indirect utility function

$$V(a_0) = \max_{\{c_t\}_0^\infty} \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t) \right\} \quad (0.8)$$

It determines the present value of having chosen the optimal path of consumption  $\{c_0^*, c_1^*, \dots\}$  in function of the initial wealth  $a_0$ . Note that we can rewrite

$$\begin{aligned} V(a_0) &= \max_{\{c_t\}_0^\infty} \left\{ \sum_{t=0}^{\infty} \beta^t u(c_t) \right\} \\ &= \max_{\{c_t\}_0^\infty} \left\{ u(c_0) + \beta u(c_1) + \beta^2 u(c_2) + \dots \right\} \\ &= \max_{c_0} \left\{ u(c_0) + \beta \max_{\{c_t\}_1^\infty} \left( u(c_1) + \beta u(c_2) + \beta^2 u(c_3) + \dots \right) \right\} \\ &= \max_{c_0} \left\{ u(c_0) + \beta \max_{\{c_t\}_1^\infty} \sum_{t=0}^{\infty} \beta^t u(c_{t+1}) \right\} \\ &= \max_{c_0} \left\{ u(c_0) + \beta V(a_1) \right\} \end{aligned} \quad (0.9)$$

So that the value function gives the value of being at the optimum at time 0 (first term) and forever thereafter (second term). We have transformed a infinitely-many-period problem into a two-period problem.

An important building block of this approach is the *optimality principle*: "An optimal policy has the property that whatever the state and optimal first decision may be, the remaining decisions constitute an optimal policy with respect to the state originating from the first decisions." (Bellman 1957). In other words, this means that we do not need to know what the agent will do with his wealth after the current period, but it suffices to know that he acts optimally and thus generates utility given by  $V(a_1)$ . Another implication is that time does not matter in the sense that we can just rewrite the problem in terms of "now" and "later", where the apostrophe denotes "next period"

$$\begin{aligned} V(a) &= \max_c \left\{ u(c) + \beta V(a') \right\} \\ \text{s.t. } a' &= R(a + y - c) \end{aligned} \quad (0.10)$$

The above expression is called *Bellman equation* (or functional equation). There are some practical aspects we should keep in mind when setting up the Bellman equation:

- When constructing the Bellman equation we always have to express  $V(\square)$  as a function of the *state variable(s)*. The state completely summarizes all information from the past, which is needed to solve the forward-looking optimization problem. It depends on decisions made in the previous period(s). In this example it is obvious that  $a_t$  is the state variable.
- The *choice or control variable(s)* is the variable that the agent can influence in the present period. In this example  $c_t$  is the choice variable.
- Of course, we mustn't forget the *transition equation(s)* which is essential to describe how tomorrow's state depends on today's state and control. In the example we have

$$a' = R(a + y - c) \quad (0.11)$$

- When it comes to solving the optimization problem, we always have to maximize (0.10) either with respect to tomorrow's state or with respect to the choice variable and substitute one or the other through the transition equation.

Here, we will maximize with respect to tomorrow's state, so that we set up the problem as follows

$$V(a) = \max_{a'} \{u(a + y - R^{-1}a') + \beta V(a')\} \quad (0.12)$$

Note two things: The real unknown in this equation is the value function itself, which appears on both sides of the equal sign. We will simply assume that it exists and find our way around having to find an explicit solution for it. The optimality principle states that the value function is the only function that satisfies this equation.

The FOC of the maximization problem is

$$\begin{aligned} \frac{\partial V(a)}{\partial a'} &= -R^{-1}u'(c) + \beta V'(a') = 0 \\ u'(c) &= R\beta V'(a') \end{aligned} \quad (0.13)$$

We don't know the expression for  $V'(a')$ , but we can use the fact  $V(a)$  satisfies (0.12) for all  $a$  and use the implication of the Envelope theorem to get around that problem. [The envelope theorem states that at the optimum the total and partial derivatives are the same.] Let us compute the derivative of  $V(a)$  with respect to  $a$ , evaluated at optimal  $a'_*$

$$\begin{aligned} \left. \frac{\partial V(a)}{\partial a} \right|_{a'=a'_*} &= \frac{\partial}{\partial a} \left( u(a + y - R^{-1}a'_*) + \beta V(a'_*) \right) \\ V'(a) &= u'(c) \end{aligned} \quad (0.14)$$

Since it holds for all  $a$ , it will hold for  $a'$  as well and we can "shift (0.14) forward by one unit in time" to get the desired expression for  $V'(a') = u'(c')$ . Substitute it in the FOC

(0.13) to get

$$\begin{aligned} u'(c) &= R\beta V'(a') \\ u'(c) &= R\beta u'(c') \end{aligned} \quad (0.15)$$

and we again find the standard Euler equation. Expression (0.15) is a necessary condition of optimality for any  $t$ . If the path of consumption  $\{c_t\}_0^\infty$  is chosen such that the Euler equation holds  $\forall t$ , then it is impossible to increase utility by moving consumption from

one period to an adjacent one. What about moving consumption from one period to another one lying more than one unit in time away? Well, just iterate forward the Euler equation  $t$  times and we still have that optimal consumption has to satisfy  $u'(c_t) = (R, \beta)^k u'(c_{t+1+k})$ . Note that (0.15) is not a sufficient condition for optimality; in addition to it, we need the condition that the consumer's wealth is bigger or equal to zero at the end of his life. In the introductory two-period example we used the terminal condition  $a_2 \geq 0$ . In the infinite horizon case we use a *transversality condition* which usually takes the form

$$\lim_{t \rightarrow \infty} a_t V(a_t) = 0$$

We want that the agent has either zero wealth at the end of his life or that the wealth does not contribute anything to maximized utility. Note that this condition rules out that the agent has debts at the end of his life.

In a final step we could actually find a *policy function*  $\phi(\square)$  that links the choice of the control (today's consumption) to the state (inherited wealth)

$$c = \phi(a)$$

As stated above, the optimality principle implies that this function is constant over time (hence no time indices).

Since we will not use policy functions in our class, we ignore how to derive them.