Dynamic programming: definitions and theorems

This handout sums up some important theorems and results from Acemoglu (2009), chp.6, p.185-205. It provides foundations to the first handout about dynamic programming (cf Macro1) that shown how to solve maximization problems with Bellman equations.

1 Sequential vs recursive formulations

The canonical dynamic optimization program in discrete time and in a stationary form can be written as

**Problem A**

\[
V^*(x(0)) = \sup_{x(t+1)} \sum_{t=0}^{\infty} \beta^t U(x(t), x(t+1))
\]

s.t

\[x(t+1) \in G(x(t)), \forall t \geq 0\]

where \(x(0)\) given.

where \(\beta \in (0,1)\), \(x(t)\) is a vector of variables and \(G(x)\) is a set-valued mapping or a correspondence\(^1\). The first constraint, also called the transition equation, basically specifies what value of \(x(t+1)\) are allowed given the value of \(x(t)\). For this reason, we can think of \(x(t)\) as the state vector (or variables) while \(x(t+1)\) is the control vector (or variables).

**Problem A** is a sequence problem, that is a choice of an infinite sequence \(\{x(t)\}_{t=0}^{\infty}\) from some vector space of infinite sequence. The basic idea of dynamic programming is to turn the sequence problem into a functional equation (the Bellman equation), that is to transform the problem into one of finding a function rather than a sequence. The relevant functional equation is

**Problem B**

\[
V(x) = \sup_{y \in G(x)} \{U(x, y) + \beta V(y)\}, \forall x \in X
\]

where \(V\) is a real-valued function. Instead of explicitly choosing a sequence, we choose a policy which determines what the control vector \(x(t+1)\) should be for a given value of the state vector \(x(t)\). This choice is true for every period, and so the problem is time-independent. Since

---

\(^1\)In economics, the distinction between a function and a correspondence is as follows: a function \(f\) is a real-valued mapping i.e \(f : X \rightarrow \mathbb{R}\) for some arbitrary set \(X\), while a correspondence is a set-valued mapping, i.e \(F : X \rightarrow P(Z)\) for some set \(Z\), that is the mapping assigns a subset of \(Z\) to each element of \(X\).
the function \( V \) appears both on the left and the right-hand sides of the equation, we call this formulation the **recursive formulation**. The recursive/functional formulation is often preferred to the sequential problem for two reasons: first the function \( V \) is often computationally convenient whereas the solutions of sequence problems are often difficult to characterize, second, the recursive formulation often gives better economic insights because it highlights the comparison between today’s and tomorrow’s values (think about the models of the labor market and the value functions of being unemployed or unemployed!).

In fact, the form of Problem \( B \) suggests itself naturally from the formulation of \( A \). If \( A \) has a maximum starting at \( x(0) \) attained by the optimal sequence denoted \( \{x(t)\}^\infty_{t=0} \) with \( x(0)^* = x(0) \) we have

\[
V^*(x(0)) = \sum_{t=0}^{\infty} \beta^t U(x^*(t), x^*(t+1)) \\
V^*(x(0)) = U(x(0), x^*(1)) + \beta V^*(x^*(1))
\]

This equation encapsulates the basic idea of dynamic programming, the Principle of Optimality, that will become clear in the theorems below.

Essentially, an optimal plan can be broken into two parts, what is optimal to do today, and the optimal continuation path. Dynamic programming exploits this principle and provides us with a set of powerful tools to analyze optimization in discrete-time infinite-horizon problems. As noted above, the particular advantage of this formulation is that the solution can be represented by a time invariant policy function (or policy mapping), \( \pi : X \rightarrow X \) determining which value of \( x(t+1) \) to choose for a given value of the state variable \( x(t) \). In general, however, there will be two complications: first, a control reaching the optimal value may not exist\(^2\); second, we may not have a policy function, but a policy correspondence, because there may be more than one maximizer for a given state variable. These issues will be dealt with the theorems below. Once the value function \( V \) is determined, the policy function is given straightforwardly. In particular, by definition it must be the case that if optimal policy is given by a policy function \( \pi(x) \), then

\[
V(x) = U(x, \pi(x)) + \beta V(\pi(x)), \forall x \in X
\]

which is one way of determining the policy function. This equation simply follows from the fact that \( \pi(x) \) is the optimal policy, so when \( y = \pi(x) \), the right-hand side of the equation of Problem \( B \) reaches the maximal value \( V(x) \).

The next section states a number of theorems about the relationship between the solution to the sequence problem, Problem \( A \), and the recursive formulation, Problem \( B \). The proofs of these theorems can be found in Lucas, Stockey and Prescott (1989), chp.4 and Acemoglu (2009), chp.6.

\(^2\) which was the reason why we originally used the notation \( \sup \) rather than \( \max \)
2 Assumptions and theorems

Consider first a sequence \(\{x^*(t)\}_{t=0}^{\infty}\) which attains the maximum in Problem A. The main purpose is to ensure that this sequence will satisfy the recursive equation of dynamic programming, written here as

\[ V^*(x(t)) = U(x^*(t), x^*(t+1)) + \beta V^*(x^*(t+1)), \forall t = 0, 1, 2... \]

and that any solution to this equation will also be a solution of Problem A, in the sense that it will attain its maximum (or supremum). To answer this question, we need to define first the set of feasible sequences or plans starting with an initial value \(x(t)\) as:

\[ \Phi(x(t)) = \{\{x^*(s)\}_{s=t}^{\infty} : x(s+1) \in G(x(s)), \text{ for } s = t, t+1,... \} \]

A typical element of the set \(\Phi(x(0))\) is denoted by \(x = (x(0), x(1), ...) \in \Phi(x(0))\).

2.1 Assumptions

In order to prove the relevant theorems, we need 5 important assumptions that will become usual in dynamic programming problems.

**Assumption 1.** \(G(x)\) is nonempty for all \(x \in X\), and for all \(x(0) \in X\) and \(x \in \Phi(x(0))\),

\[ \lim_{n \to \infty} \sum_{t=0}^{n} \beta^t U(x(t), x(t+1)) \text{ exists and is finite.} \]

The assumption of a finite limit is stronger than necessary but, in economic applications, we are not interested in optimization problems where households or firms achieve infinite values.

**Assumption 2.** \(X\) is a compact subset of \(\mathbb{R}^K\), \(G\) is nonempty, compact-valued and continuous. \(U : X_G \to \mathbb{R}\) is continuous, where \(X_G = \{(x, y) \in X \times X : y \in G(x)\}\).

According to the Heine-Borel Theorem, a compact set has a simple property: Let \(X \subset \mathbb{R}^K\) be an Euclidean space (with a Euclidean metric or topology), then \(X' \subset X\) is compact if and only if \(X'\) is closed and bounded.\(^3\) The assumption of the state variable lying in a compact set is strong. Indeed, in economic growth many variables -including capital- can be considered to grow steadily (values are not bounded). It is especially true for endogenous growth.\(^4\) Nevertheless, the

---

\(^3\)See appendix A of Acemoglu (2009) for more analysis of compactness, as well as any real analysis textbook.

\(^4\)This is one reason explaining why models of endogenous growth mainly used continuous time methods rather than Bellman functions.
most important results of dynamic programming can be generalized to the case in which X is not compact.

These 2 assumptions will ensure that the supremum (maximal value) is attained for some plans x in both problems A and B.

**Assumption 3.** U is concave and G is convex.

This assumption imposes conditions similar to those used in many economic applications: the constraint set is assumed to be convex and the objective function is concave or strictly concave.

**Assumption 4.** For each $y \in X, U(., y)$ is strictly increasing in each of its first K arguments, and G is monotone in the sense that $x \leq x'$ implies $G(x) \subset G(x')$.

That is the objective function is increasing in the state variables (its first K arguments) and that the greater levels of the state variables are also attractive from the viewpoint of relaxing the constraints (i.e. a greater $x$ means more choice).

**Assumption 5.** U is continuously differentiable on the interior of its domain $X_G$.

### 2.2 Theorems

Given these 5 assumptions, the following sequence of results can be established.

**Theorem 1.** (*Equivalence of values*). Suppose assumptions 1 holds. Then for any $x \in X$, any $V^*(x)$ defined in Problem A is also a solution to Problem B. Moreover, any $V(x)$ defined in B that satisfies $\lim_{t \to \infty} \beta V(x(t)) = 0$ for all $(x, x(1), x(2), ...) \in \Phi(x)$ is also a solution to A, so that $V^*(x) = V(x)$ for all $x \in X$.

**Theorem 2.** (*Principle of optimality*). Suppose assumption 1 holds. Let $x^* \in \Phi(x(0))$ be a feasible plan that attains $V^*(x(0))$ in Problem A. Then, with $x^*(0) = x(0)$ we have that

$$V^*(x^*(t)) = U(x^*(t), x^*(t+1)) + \beta V^*(x^*(t+1)), \forall t = 0, 1, 2...$$

Moreover, if any $x^* \in \Phi(x(0))$ satisfies this equation, then it attains the optimal value in A.
This theorem is the major conceptual result in the theory of dynamic programming. It states that the returns from an optimal plan (sequence) \( x^* \in \Phi(x(0)) \) can be split into two parts: the current return \( (U(x^*(t), x^*(t + 1))) \) and the continuation return \( (\beta V^*(x^*(t + 1))) \), where the continuation return is identically given by the discounted value of a problem starting from the state vector from tomorrow onwards, \( x^*(t + 1) \).

Therefore these theorems states that we can go from the solution of the recursive problem to the solution of the original problem and vice versa. Consequently, under assumptions 1 & 2, there is no risk of excluding solutions in writing the problem recursively.

The next results summarize certain important features of the value function \( V \). They will be useful in characterizing qualitative features of optimal plans in dynamic optimization problems without explicitly finding the solutions.

**Theorem 3.** *(Existence of solutions).* Suppose that assumptions 1 & 2 hold. Then there exists a unique continuous and bounded function \( V : X \to \mathbb{R} \) that satisfies Problem B

\[
V(x) = \sup_{y \in G(x)} \{ U(x, y) + \beta V(y) \}, \forall x \in X.
\]

Moreover, an optimal plan \( x^* \in \Phi(x(0)) \) exists for any \( x(0) \in X \).

This theorem proves the uniqueness of the value function (and hence of the Bellman equation), the existence of an optimal plan and that \( V \) and \( V^* \) are continuous and bounded. But the optimal solution may not be unique, despite the uniqueness of \( V \). As in static optimization problems, non-uniqueness of solutions is a consequence of lack of strict concavity of the objective function. Thus, the following theorems are needed:

**Theorem 4.(Concavity of the Value function).** Suppose that assumptions 1, 2 & 3 hold. Then the unique \( V : X \to \mathbb{R} \) of Problem B is strictly concave.

**Corollary 4-1.** Suppose that assumptions 1, 2 & 3 hold. Then there exists a unique optimal plan \( x^* \in \Phi(x(0)) \) for any \( x(0) \in X \). Moreover, the optimal plan can be expressed as \( x^*(t + 1) = \pi(x^*(t)) \), where \( \pi : X \to X \) is a continuous policy function.

The important result in this corollary is that the "policy function" \( \pi \) is indeed a function and not a correspondence. This is a consequence of the uniqueness of \( x^* \). This result also implies that the policy mapping \( \pi \) is continuous in the state vector. Moreover, if there exists a vector of parameters \( Z \) continuously affecting either the constraint correspondence \( \Phi \) or the instantaneous payoff function \( U \), then the same argument establishes that \( \pi \) is also continuous in this vector.
of parameters. This feature will enable qualitative analysis of dynamic macroeconomic models under a variety of circumstances.

**Theorem 5.** (*Monotonicity of the Value function.*) Suppose that assumptions 1, 2 & 4 hold and let \( V : X \to \mathbb{R} \) be the unique solution to Problem B. Then \( V \) is strictly increasing in all of its arguments.

**Theorem 6.** (*Differentiability of the Value function.*) Suppose that assumptions 1, 2, 3 & 5 hold. Let \( \pi \) be the policy function defined above and assume that \( x' \in iX \) and \( \pi(x') \in iG(x') \), then \( V(.) \) is continuously differentiable at \( x' \), with derivative given by:

\[
DV(x') = D_x U(x', \pi(x'))
\]

This last theorem will enable us to use the dynamic programming techniques in a wide variety of dynamic optimization problems. Indeed, as with static optimization, differentiability will be useful to characterize the solution.

But the important theorems above are not so useful to characterize in practice the solutions to dynamic optimization problems. The following theorem (n° 7) shows that the Euler equations together with the transversality condition are sufficient to characterize the solutions.\(^6\)

Consider the functional equation corresponding to Problem B:

\[
V(x) = \sup_{y \in G(x)} \{U(x, y) + \beta V(y)\}, \forall x \in X
\]

and let assume throughout that assumptions 1 to 5 hold. The first order condition gives:

\[
D_y U(x, y^*) + \beta D_x V(y^*) = 0
\]

and the Envelope theorem gives:

\[
D_x U(x, y^*) = DV(x)
\]

Now, using the notation \( y^* = \pi(x) \) to denote the optimal function, we have the usual **Euler equation**:

\[
D_y U(x, \pi(x)) + \beta D_x U(\pi(x), \pi(\pi(x))) = 0
\]

\(^5\)iX denotes the interior of the set \( X \). \( D \) denotes the gradient of the function with respect to all of its arguments. \( D_x \) denotes the gradient of the function with respect to its argument \( x \).

\(^6\)Until now, when you have worked with Bellman equations, you have always used this theorem implicitly!
When both $x$ and $y$ are scalar, this condition can be rewritten:

$$\frac{\partial U(x, \pi(x))}{\partial y} + \beta \frac{\partial U(\pi(x), \pi(\pi(x)))}{\partial x} = 0$$

In the general case, the transversality condition takes the form:

$$\lim_{t \to \infty} \beta^t D_x U(x^*(t), x^*(t+1)).x^*(t) = 0$$

**Theorem 7.** Euler equations and the transversality condition. Let $X \subset \mathbb{R}^K_+$, and suppose that assumptions 1 to 5 hold. Then a sequence $\{x^*(t+1)\}_{t=0}^{\infty}$ with $x^*(t + 1) \in iG(x^*, t = 0, 1, 2...$ is optimal for Problem A, given $x(0)$, if it satisfies the Euler equation and the transversality condition defined above.\(^7\)

**References**


---

\(^7\)See the proof in Acemoglu(2009), p.204-205.