1 Introduction

In this section I will present some basic results in optimal control theory, which we will use to do dynamic optimization in continuous time. This theory is older than the more modern dynamic programming you saw with Prof. Laibson, in fact it is a predecessor. However, we will study some results of optimal control theory for two reasons. First, in the growth literature it is common to see models being setup and solved using these tools. Second, some problems become simpler in this setting. In the end, learning a new tool can’t hurt you.

In continuous time, optimization is with respect to an infinite-dimensional object: a function $y: [t_0, t_1] \to \mathbb{R}$. The canonical problem can be written as

$$\max_{x(t), u(t)} W(x(t), u(t)) \equiv \int_{t_0}^{t_1} f(t, x(t), u(t)) \, dt$$

subject to

$$\dot{x}(t) = g(t, x(t), u(t))$$

and

$$x(t) \in \mathcal{X}, u(t) \in \mathcal{U} \quad \text{for all } t, \text{ and } x(0) = x_0$$

The objective is to maximize the real valued functional $W$, which takes as arguments the vector of state variables $x$ and control variables $u$. The behavior of the state variables $x$ is governed by a system of differential equations given by the function $g$. We will refer to $f$ as the payoff function. In most economic applications, $f$ depends on time only through a discounting factor. For simplicity, we will assume $W$, $f$, $g$, $x$ and $u$ are real

---

1 Contact: yu8@fas.harvard.edu.
2 This notes are based on Chapter 7 in Acemoglu (2008). Also see Aghion and Howitt (1998) and Barro and Sala-i-Martin (1995).
valued. This will be the case in most applications, although in models with multiple state variables, \( x(t) \) will be a vector. Also, assume that the sets \( X \subset \mathbb{R} \) and \( U \subset \mathbb{R} \) are nonempty and convex and that \( f \) and \( g \) are continuously differentiable in all its arguments.

A pair of functions \((x(t), u(t))\) that satisfy (2) and (3) are referred to as an admissible pair. I will ignore the argument \( t \) when possible.

2 Variational approach

The variational approach assumes there exists a continuous solution \( \hat{u} \) that lies everywhere in the interior of the set \( U \) and then characterizes the properties of this solution to find a set of necessary conditions.

So let us assume that \((\hat{x}, \hat{u})\) is an admissible pair such that \( \hat{u} \) is continuous on \([0, t_1]\), \((\hat{x}, \hat{u}) \in \text{Int } X \times U \) and that

\[
W(\hat{x}, \hat{u}) \geq W(x, u)
\]

for any other admissible pair \((x, u)\).

Take an arbitrary continuous function \( \eta(t) \) and \( \varepsilon \in \mathbb{R} \). Define a variation of the function \( \hat{u}(t) \) by

\[
u(t, \varepsilon) \equiv \hat{u}(t) + \varepsilon \eta(t)\]

We will consider only feasible variations, that is, those with sufficiently small \( \varepsilon \) so that they lie in the interior of \( U \). This is always possible since \( \hat{u} \) and \( \eta \) are continuous functions on compact sets, therefore bounded. Let us also define \( x(t, \varepsilon) \) as the path of the state variable corresponding to the path of the control variable:

\[
\dot{x}(t, \varepsilon) = g(t, x(t, \varepsilon), u(t, \varepsilon)) \quad \text{for all } t \in [0, t_1], \text{ with } x(0, \varepsilon) = x_0
\] (4)

We can choose \( \varepsilon \) sufficiently small so that \( x(t, \varepsilon) \in X \), since solutions to differential equations are continuous. Now define:

\[
W(\varepsilon) \equiv W(x(t, \varepsilon), u(t, \varepsilon)) = \int_0^{t_1} f(t, x(t, \varepsilon), u(t, \varepsilon)) \, dt
\] (5)

Since \( \hat{u} \) is optimal and the variations \( u(t, \varepsilon) \) and \( x(t, \varepsilon) \) are feasible, we must have

\[
W(\varepsilon) \leq W(0)
\]

for sufficiently small \( \varepsilon \).

The objective is to write \( W(\varepsilon) \) in a convenient way and study the derivative of this function at \( \varepsilon = 0 \).

Equation (4) implies that for any function \( \lambda : [0, t_1] \rightarrow \mathbb{R} \) we have

\[
\int_0^{t_1} \lambda(t)[g(t, x(t, \varepsilon), u(t, \varepsilon)) - \dot{x}(t, \varepsilon)] \, dt = 0
\]
This function, when chosen suitably, is the costate variable, with an interpretation similar to the Lagrange multipliers in standard constrained optimization. In what follows, suppose the function $\lambda(\cdot)$ is continuously differentiable. Adding this integral to (5) yields:

$$\mathcal{W}(\varepsilon) = \int_0^{t_1} \left\{ f(t, x(t, \varepsilon), u(t, \varepsilon)) + \lambda(t) \left[ g(t, x(t, \varepsilon), u(t, \varepsilon)) - \dot{x}(t, \varepsilon) \right] \right\} \, dt$$

First integrate the last term by parts to get:

$$\int_0^{t_1} \lambda(t) \dot{x}(t, \varepsilon) \, dt = \lambda(t_1)x(t_1, \varepsilon) - \lambda(0)x_0 - \int_0^{t_1} \dot{\lambda}(t)x(t, \varepsilon) \, dt$$

Plugging back we get:

$$\mathcal{W}(\varepsilon) = \int_0^{t_1} \left\{ f(t, x(t, \varepsilon), u(t, \varepsilon)) + \lambda(t)g(t, x(t, \varepsilon), u(t, \varepsilon)) + \dot{\lambda}(t)x(t, \varepsilon) \right\} \, dt$$

$$- \lambda(t_1)x(t_1, \varepsilon) + \lambda(0)x_0$$

Recall that $f$ and $g$ are continuously differentiable, and $u(t, \varepsilon)$ is continuously differentiable in $\varepsilon$ by construction, therefore $x(t, \varepsilon)$ is continuously differentiable in $\varepsilon$. Differentiating $\mathcal{W}(\varepsilon)$ and evaluating the derivative at $\varepsilon = 0$ gives:

$$\mathcal{W}'(0) = \int_0^{t_1} \left[ f_x(t, \dot{x}(t), \dot{u}(t)) + \lambda(t)g_x(t, \dot{x}(t), \dot{u}(t)) + \ddot{\lambda}(t) \right] x_\varepsilon(t, 0) \, dt$$

$$+ \int_0^{t_1} \left[ f_u(t, \dot{x}(t), \dot{u}(t)) + \lambda(t)g_u(t, \dot{x}(t), \dot{u}(t)) \right] \eta(t) \, dt$$

$$- \lambda(t_1)x_\varepsilon(t_1, 0)$$

(6)

Now note that if there exists a function $\eta(t)$ such that $\mathcal{W}'(0) \neq 0$ then $\mathcal{W}(x, u)$ can be increased and thus the pair $(\dot{x}, \dot{u})$ cannot be a solution. Therefore optimality requires that

$$\mathcal{W}'(0) = 0 \text{ for all } \eta(t)$$

This condition applies for any continuously differentiable $\lambda(t)$ function, but not all such functions will play the role of a costate variable. Let us choose the function $\lambda(t)$ as a solution to the differential equation

$$\ddot{\lambda}(t) = - \left[ f_x(t, \dot{x}(t), \dot{u}(t)) + \lambda(t)g_x(t, \dot{x}(t), \dot{u}(t)) \right]$$

(7)

with boundary condition $\lambda(t_1) = 0$. This sets the first and last terms of (6) to zero. The condition

$$\lambda(t_1) = 0$$

will be referred to as the transversality condition.

Since $\eta(t)$ is arbitrary, the second term is set to zero when

$$f_u(t, \dot{x}(t), \dot{u}(t)) + \lambda(t)g_u(t, \dot{x}(t), \dot{u}(t)) = 0 \text{ for all } t \in [0, t_1]$$

(8)

We have established the following:
**Theorem 1** (Necessary conditions). Consider the problem of maximizing (1) subject to (2) and (3), with \( f \) and \( g \) continuously differentiable. Suppose that this problem has an interior continuous solution \((\hat{x}, \hat{u}) \in \text{Int } X \times U\). Then there exists a continuously differentiable costate function \( \lambda(t) \) defined on \( t \in [0, t_1] \) such that (2), (7) and (8) hold, and also \( \lambda(t_1) = 0 \).

So far we have treated the terminal value of the state variable, \( x_1 \), as a choice variable. To see the role of the transversality condition, suppose the value of \( x_1 \) is given. The only difference will now be that \( x(t_1, \varepsilon) \) must equal \( x_1 \) to be feasible. Therefore \( x_\varepsilon(t_1, 0) = 0 \) in equation (6) and therefore \( \lambda(t_1) \) is unrestricted.

**Theorem 2** (Necessary conditions with fixed endpoint). Consider the problem stated in Theorem 1. If the terminal value of the state variable, \( x_1 \), is fixed instead of being a control variable, then the necessary conditions remain the same except that the terminal value of the costate variable, \( \lambda(t_1) \), is unrestricted.

Let us illustrate the previous results with an example.

**Example** (Life-cycle consumption). Consider an agent that lives between dates 0 and 1 and is trying to solve

\[
\max_{[c(t), a(t)]} \int_0^1 e^{-\rho t} u(c(t)) \, dt
\]

subject to

\[
\dot{a}(t) = ra(t) - c(t)
\]

Here \( c \) denotes consumption, \( a \) is the stock of assets, \( \rho \) is the discount rate and \( r \) is the interest rate paid on asset holdings. The agent cannot take debt, \( a(t) \geq 0 \), and it starts with initial assets \( a(0) > 0 \).

Economic intuition suggests that the agent will have zero assets at \( t = 1 \). This gives us a terminal condition \( a(1) = 0 \). Therefore we can apply Theorem 1 with an unrestricted value for \( \lambda(1) \). Condition (8) takes the form:

\[
e^{-\rho t} u'(c(t)) - \lambda(t) = 0
\]  

(9)

Condition (7) implies that

\[
\dot{\lambda}(t) = -r \lambda(t)
\]  

(10)

Differentiating (9) and using (10) yields a differential equation, the Euler equation for consumption:

\[
\frac{\dot{c}}{c} = \frac{1}{\varepsilon_u} (r - \rho)
\]

where

\[
\varepsilon_u \equiv -\frac{u''(c(t))c(t)}{u'(c(t))}
\]
is the elasticity of marginal utility. Integrating (10) gives
\[ \lambda(t) = \lambda(0)e^{-\rho t} \]
Combining this with (9) implies:
\[ \dot{c}(t) = u^{-1}[\lambda(0)e^{(\rho-r)t}] \]
To fully determine the path for consumption we need the value \( \lambda(0) \). This comes from the terminal condition and the equation of motion of assets:
\[ \dot{a}(t) = ra(t) - u^{-1}[\lambda(0)e^{(\rho-r)t}] \]
The initial value of \( \lambda(0) \) must be chosen such that \( a(1) = 0 \).

**Example** (Life-cycle consumption continued). What happens if we do not impose the terminal condition \( a(1) = 0 \) and apply Theorem 1? Then the first order conditions still imply
\[ \lambda(t) = \lambda(0)e^{-\rho t} \]
However, since \( \lambda(1) = 0 \) this implies \( \lambda(t) = 0 \) for all \( t \). But the necessary conditions also imply
\[ e^{-\rho t}u'(c(t)) = \lambda(t) \]
which cannot be true if \( u' > 0 \). Thus when the terminal value of assets is a choice variable, there is no interior solution. The problem is the additional constraint \( a(t) \geq 0 \). The problem is simplified if we observe that the terminal value of assets must be equal to zero and apply Theorem 2.

A version of Theorem 1 with inequality constraints is given as follows:

**Theorem 3** (Necessary conditions with inequality constraint). Consider the problem stated in Theorem 1 with the additional constraint that the terminal value of the state variable satisfies \( x(t_1) \geq x_1 \). Then the necessary conditions remain the same except that the transversality condition is replaced with \( \lambda(t_1)(x(t_1) - x_1) = 0 \).

### 3 The Maximum Principle

The calculus of variations, the classical method for tackling problems of dynamic optimization, like the ordinary calculus, requires for its applicability the differentiability of the functions that enter in the problem. More importantly, only interior solutions can be handled. A more modern development that can deal with other features such as corner solutions, is found in optimal control theory. In this theory, a control path does not have to be continuous in order to become admissible; it only needs to be piecewise continuous. This means that it is allowed to contain jump discontinuities. The control variable has to be continuous, but it is permissible to have a finite number of
sharp points, or corners. That is to say, to be admissible, a state path only needs to be piecewise differentiable. In most economic applications, however, solution pairs are indeed continuous and interior. In the treatment that follows, we will assume this is the case. We will present the main results using the optimal control language since it is the standard method for solving many growth problems.

By analogy with the Lagrangian, a more economical way to expressing our first order conditions is to construct the Hamiltonian:

\[ H(t, x(t), u(t), \lambda(t)) \equiv f(t, x(t), u(t)) + \lambda(t)g(t, x(t), u(t)) \]  

We will still assume that \( f \) and \( g \) are continuously differentiable, and so is \( H \). However, this is stronger than what we really need.

We will present these tools in the more general case of infinite-horizon optimization. In infinite-horizon, there is usually a terminal value condition on the state variable of the form

\[ \lim_{t \to \infty} b(t)x(t) \geq x_1 \]

The infinite-horizon optimal control problem is:

\[ \max_{x(t), u(t)} W(x(t), u(t)) \equiv \int_0^\infty f(t, x(t), u(t)) \, dt \]  

subject to

\[ \dot{x}(t) = g(t, x(t), u(t)) \]  

and

\[ x(t) \in X, u(t) \in U \quad \text{for all } t, x(0) = x_0 \quad \text{and} \quad \lim_{t \to \infty} b(t)x(t) \geq x_1 \]  

In this setting an admissible pair can now be a piecewise continuous function, which makes \( x(t) \) continuous and differentiable almost everywhere.

**Theorem 4** (Infinite-Horizon Maximum Principle). Suppose that the problem of maximizing (12) subject to (13) and (14), with \( f \) and \( g \) continuously differentiable, has a piecewise continuous interior solution \((\hat{x}, \hat{u}) \in \text{Int } X \times U\). Let \( H \) be as defined in (11). Then there exists a continuously differentiable function \( \lambda(t) \) such that, given \((\hat{x}, \hat{u})\), the Hamiltonian \( H \) satisfies the Maximum Principle

\[ H(t, \hat{x}(t), \hat{u}(t), \lambda(t)) \geq H(t, x(t), u(t), \lambda(t)) \]

for all \( u(t) \in U \) and all \( t \). Moreover, for all \( t \) where \( \hat{u}(t) \) is continuous, the following necessary conditions are satisfied:

\[ H_u(t, \hat{x}(t), \hat{u}(t), \lambda(t)) = 0 \]  

\[ \dot{\lambda}(t) = -H_x(t, \hat{x}(t), \hat{u}(t), \lambda(t)) \]  

and

\[ \dot{x}(t) = H_x(t, \hat{x}(t), \hat{u}(t), \lambda(t)) \quad \text{with } x(0) = x_0 \quad \text{and} \quad \lim_{t \to \infty} b(t)x(t) \geq x_1 \]
The proof of this simplified version of the Maximum Principle is rather long. We will instead use a sufficiency theorem that applies in almost all cases. For now, just note that the necessary conditions stated in the second part of the theorem are equivalent to the conditions found with the variational approach.

4 Economic intuition

To interpret the maximum principle, consider the problem of maximizing
\[
\int_{t_1}^t H(t, \dot{x}(t), u(t), \lambda(t)) \, dt = \int_{t_0}^{t_1} [f(t, \dot{x}(t), u(t)) + \lambda(t)g(t, \dot{x}(t), u(t))] \, dt
\] (18)
with respect to the entire function \(u(t)\) for given \(\lambda(t)\) and \(\dot{x}(t)\), where \(t_1\) can be finite or equal to \(\infty\). Assuming an interior solution, this implies condition (15):
\[
H_u(t, \dot{x}(t), \dot{u}(t), \lambda(t)) = 0
\]
Therefore the optimal principle involves the maximization of the two integrals on the right hand side of (18). An envelope theorem argument implies
\[
\lambda(t) = \frac{\partial V(t, \dot{x}(t))}{\partial x}
\]
which gives the interpretation of the costate variable as the impact of a marginal increase in \(x\) on the optimal value of the program (or the shadow value of \(x\)). Moreover, using the equation of motion for the state variable, note that the second term on the right hand side of (18) is equal to
\[
\int_{t_0}^{t_1} \lambda(t) \dot{x}(t) \, dt
\]
Therefore, maximizing (18) is equivalent to maximizing instantaneous returns \(f\), plus the value of stock of \(x(t)\), given by \(\lambda(t)\), times the increase in the stock, \(\dot{x}(t)\). The objective of the Maximum Principle is to maximize the flow return plus the value of the current stock of the state variable.

5 Transversality conditions

It is tempting to guess that the transversality condition in infinite-horizon should be analogous to that in Theorem 1, with \(t_1\) replaced with the limit of \(t \to \infty\). However, this conjecture is not true, as the following example extracted from Ramsey (1928) shows.

Example (Growth without discounting). Consider the following problem without discounting:
\[
\max \int_0^\infty [\log(c(t)) - \log(c^*)] \, dt
\]
subject to 
\[ \dot{k}(t) = k(t)^\alpha - c(t) - \delta k(t) \]
with \( k(0) = 1 \) and \( \lim_{t \to \infty} k(t) \geq 0 \). Here \( c^* \equiv (k^*)^\alpha - \delta k \) and \( k^* \equiv (\alpha/\delta)^{1/(1-\alpha)} \) are the maximum level of steady state consumption and the corresponding steady state level of capital. The Hamiltonian is:

\[ H(k, c, \lambda) = \log c(t) - \log c^* + \lambda(t)[k(t)^\alpha - c(t) - \delta k(t)] \]

And a necessary condition is:

\[ H_c(k, c, \lambda) = \frac{1}{c(t)} \frac{\partial f(x(t), u(t))}{\partial x} + \lambda(t) \frac{\partial f(x(t), u(t))}{\partial u} = 0 \]

It can be shown that an optimal path features \( c(t) \to c^* \) and \( k \to k^* \), which implies

\[ \lim_{t \to \infty} \lambda(t) = \frac{1}{c^*} > 0 \]

The finite-horizon transversality condition in this case would have been \( \lambda(t_1)k(t_1) = 0 \) but here \( \lim_{t \to \infty} \lambda(t)k(t) = k^*/c^* > 0 \). It can be verified, however, that the following condition holds instead:

\[ \lim_{t \to \infty} H(k(t), c(t), \lambda(t)) = 0 \]

Under some assumptions, it can be shown that this condition is necessary in a more general setting.

### 6 Discounted Infinite-Horizon Optimal Control

Our main interest is in growth models in which utility is discounted exponentially. This models often take the following form:

\[ \max_{x(t), u(t)} \int_0^\infty e^{-\rho t} f(x(t), u(t)) \, dt \quad \text{with} \quad \rho > 0 \]

subject to

\[ \dot{x}(t) = g(t, x(t), u(t)) \]

and

\[ x(t) \in \mathcal{X}(t), u(t) \in \mathcal{U}(t) \quad \text{for all} \quad t, x(0) = x_0 \quad \text{and} \quad \lim_{t \to \infty} b(t)x(t) \geq x_1 \]

The special feature of this problem is that the payoff function \( f \) depends on time only through exponential discounting. The Hamiltonian in this case is

\[ H(t, x(t), u(t), \lambda(t)) = e^{-\rho t} f(x(t), u(t)) + \lambda(t)g(t, x(t), u(t)) \]

\[ = e^{-\rho t}[f(x(t), u(t)) + \mu(t)g(t, x(t), u(t))] \]
where the second line uses the definition
\[ \mu(t) \equiv e^{\rho t} \lambda(t) \]

In fact, in this case, instead of working with the standard Hamiltonian, we can work with the \textit{current-value Hamiltonian}
\[ \hat{H}(t, x(t), u(t), \mu(t)) \equiv f(x(t), u(t)) + \mu(t)g(t, x(t), u(t)) \]

Also define the \textit{value function} \( V(t_0, x(t_0)) \) as the value attained in (19) subject to the constraints.

The next result establishes the necessity of a stronger transversality condition under some additional assumptions, which are typically met in economic applications.

**Assumption 1.** In the maximization of (19) subject to (20) and (21):

1. \( f \) is weakly monotone in \( x \) and \( u \), and \( g \) is weakly monotone in \((t, x, u)\).
2. there exists \( m > 0 \) wich that \( |g_u(t, x(t), u(t))| \geq m \) for all \( t \) and for all admissible pairs \((x(t), u(t))\).
3. there exists \( M < \infty \) such that \( |f_u(x, u)| \leq M \) for all \( x \) and \( u \).

**Theorem 5** (Necessary Conditions for Discounted Infinite-Horizon Problems). Suppose that the problem of maximizing (19) subject to (20) and (21), with \( f \) and \( g \) continuously differentiable, has an interior piecewise continuous optimal control \( \hat{u}(t) \in \text{Int} \ U(t) \) with state variable \( \hat{x}(t) \in \text{Int} \ X(t) \). Suppose that the value function \( V(t, \hat{x}(t)) \) exists and is finite, is differentiable in both arguments for sufficiently large \( t \) and \( \lim_{t \to \infty} \delta V(t, \hat{x}(t))/\delta t = 0 \). Then except at points of discontinuity of the control, the optimal pair \( (\hat{x}, \hat{u}) \) satisfies the following necessary conditions:

\[ \hat{H}_u(t, \hat{x}, \hat{u}, \mu) = 0 \]  \hspace{1cm} (22)
\[ \rho \mu(t) - \dot{\mu}(t) = \hat{H}_x(t, \hat{x}, \hat{u}, \mu) \]  \hspace{1cm} (23)

and

\[ \dot{x}(t) = \hat{H}_\mu(t, \hat{x}, \hat{u}, \mu), \quad x(0) = x_0 \quad \text{and} \quad \lim_{t \to \infty} b(t)x(t) \geq x_1 \]  \hspace{1cm} (24)

and the transversality condition

\[ \lim_{t \to \infty} e^{-\rho t} \hat{H}(t, \hat{x}, \hat{u}, \mu) = 0 \]

Moreover, suppose that Assumption 1 holds and that either \( \lim_{t \to \infty} \dot{x}(t) = x^* \in \mathbb{R} \) or \( \lim_{t \to \infty} \dot{x}(t)/\dot{x}(t) = \chi \in \mathbb{R} \). Then the transversality condition can be strengthened to

\[ \lim_{t \to \infty} e^{-\rho t} \mu(t) \dot{x}(t) = 0 \]  \hspace{1cm} (25)
In many applications, it is rather cumbersome to check Assumption 1. Instead, under concavity assumptions, we can make use of the following sufficient conditions, due to Arrow (1968).

**Theorem 6** (Sufficiency Conditions for Discounted Infinite-Horizon Problems). Consider the problem of maximizing (19) subject to (20) and (21), with \( f \) and \( g \) continuously differentiable. Suppose that some \( \hat{u}(t) \) and the corresponding path for the state variable \( \hat{x}(t) \) satisfy (22)-(25). Define
\[
M(t, x, \mu) \equiv \max_{u(t) \in \text{Int} \, u(t)} \bar{H}(t, x, u, \mu)
\]
Suppose that
1. \( V(t, \hat{x}(t)) \) exists and is finite for all \( t \).
2. for any admissible pair \( (x(t), u(t)) \), \( \lim_{t \to \infty} e^{-\rho t} \mu(t)x(t) \geq 0 \).
3. \( X(t) \) is convex and \( M(t, x, \mu) \) is concave in \( x(t) \in \text{Int} \, X(t) \) for all \( t \).

Then the pair \( (\hat{x}(t), \hat{u}(t)) \) achieves the global maximum of (19). Moreover, if \( M(t, x, \mu) \) is strictly concave in \( x \), the pair is the unique solution to (19).

**Proof.** For the proof we will work with the present-value Hamiltonian. Let \( M \) be the maximized present-value Hamiltonian. Since \( f \) and \( g \) are differentiable, \( \bar{H} \) and \( M \) are also differentiable in \( x \) at time \( t \). Since \( M \) is concave, we have:
\[
M(t, x(t), \lambda(t)) - M(t, \hat{x}(t), \lambda(t)) \leq M_x(t, \hat{x}(t), \lambda(t))(x(t) - \hat{x}(t))
\]
We also have:
\[
M_x(t, \hat{x}(t), \lambda(t)) = H_x(t, \hat{x}(t), \hat{u}(t), \lambda(t)) = -\hat{\lambda}(t)
\]
In the first equality we use an envelope theorem argument together with \( H_u = 0 \) from (22) and the last equality uses (23). Now using the definition of the maximized Hamiltonian:
\[
\int_0^\infty M(t, x(t), \lambda(t)) \, dt \geq W(x(t), u(t)) + \int_0^\infty \lambda(t)g(t, x(t), u(t)) \, dt
\]
and
\[
\int_0^\infty M(t, \hat{x}(t), \lambda(t)) \, dt = W(\hat{x}(t), \hat{u}(t)) + \int_0^\infty \lambda(t)g(t, \hat{x}(t), \hat{u}(t)) \, dt
\]
In this last equality we are using a stronger condition than (22) and assuming that the Maximum Principle holds, that is, \( u \) is chosen every period to maximize the Hamiltonian. Subtracting the two equations we have:
\[
W(x(t), u(t)) - W(\hat{x}(t), \hat{u}(t)) \leq \int_0^\infty \lambda(t)[g(t, \hat{x}(t), \hat{u}(t)) - g(t, x(t), u(t))] \, dt
\]
\[
+ \int_0^\infty [M(t, x(t), \lambda(t)) - M(t, \hat{x}(t), \lambda(t))] \, dt
\]
Using (26) and (27), we can bound the last term:
\[
\int_0^\infty [M(t, x(t), \lambda(t)) - M(t, \hat{x}(t), \lambda(t))] \, dt \leq - \int_0^\infty \dot{\lambda}(t)(x(t) - \hat{x}(t)) \, dt \quad (29)
\]
Integrating by parts and using \(x(0) = \hat{x}(0) = x_0\), we get:
\[
\int_0^\infty \dot{\lambda}(t)(x(t) - \hat{x}(t)) \, dt = \lim_{t \to \infty} [\lambda(t)(x(t) - \hat{x}(t))] - \int_0^\infty \lambda(t)(\dot{x}(t) - \dot{\hat{x}}(t)) \, dt
\]
By the transversality condition we have:
\[
\lim_{t \to \infty} e^{-\rho t} \lambda(t)\hat{x}(t) = \lim_{t \to \infty} \lambda(t)\hat{x}(t) = 0
\]
Also, by hypothesis, \(\lim_{t \to \infty} \lambda(t)x(t) \geq 0\) and therefore
\[
\int_0^\infty \lambda(t)(x(t) - \hat{x}(t)) \, dt \geq - \int_0^\infty \lambda(t)(\dot{x}(t) - \dot{\hat{x}}(t)) \, dt \quad (30)
\]
Combining (29) and (30) and plugging into (28) we get
\[
W(x(t), u(t)) - W(\hat{x}(t), \hat{u}(t)) \leq \int_0^\infty \lambda(t)[g(t, \hat{x}(t), \hat{u}(t)) - g(t, x(t), u(t))] \, dt \\
+ \int_0^\infty \lambda(t)(\dot{x}(t) - \dot{\hat{x}}(t)) \, dt
\]
The right hand side cancels out since by the definition of the admissible pairs we have:
\[
\dot{x}(t) = g(t, \hat{x}(t), \hat{u}(t))
\]
and
\[
\dot{\hat{x}}(t) = g(t, x(t), u(t))
\]
which establishes the first part of the theorem. If \(M\) is strictly concave in \(x\), then the inequality is strict, which establishes the second part. \(\square\)

7 Summing up: a cookbook recipe

Given Theorem 6, in most problems we will use the following strategy:

1. Look for a solution pair \((\hat{x}(t), \hat{u}(t))\) that satisfy (22)-(25).

2. Verify that \(M(t, x, \mu)\) is concave.

3. Check that for any admissible pair \((x(t), u(t))\), \(\lim_{t \to \infty} e^{-\rho t} \mu(t)x(t) \geq 0\), where \(\mu(t)\) is the costate variable associated with the candidate solution.

The main condition to check here is the concavity of \(M(t, x, \mu)\). Assuming concavity of \(f\) and \(g\) is not sufficient since we still need to know the sign of \(\lambda(t)\). But in most economic applications, we can show that \(\lambda(t)\) is always positive.

If these conditions are satisfied, we will have characterized a global maximum. Note that for this sufficiency result, we do not need to check Assumption 1.

Fernando Yu
yu8@fas.harvard.edu
8 Application: Optimal Growth in continuous time

Let’s apply these methods to the problem of optimal growth.

\[
\max_{\{k(t),c(t)\}_{t=0}^{\infty}} \int_{0}^{\infty} e^{-\rho t} u(c(t)) \, dt
\]

subject to

\[
\dot{k}(t) = f(k(t)) - \delta k(t) - c(t)
\]

The current-value Hamiltonian can be written as:

\[
\hat{H}(k, c, \mu) = u(c(t)) + \mu(t)[f(k(t)) - \delta k(t) - c(t)]
\]

Condition (22) takes the form

\[
\hat{H}_c(k, c, \mu) = u'(c(t)) - \mu(t) = 0 \tag{31}
\]

and condition (23) is

\[
\hat{H}_k(k, c, \mu) = \mu(t)[f'(k(t)) - \delta] = \rho \mu(t) - \dot{\mu}(t) \tag{32}
\]

In addition, the stronger transversality condition (25) is:

\[
\lim_{t \to \infty} e^{-\rho t} \mu(t) k(t) = 0 \tag{33}
\]

One can show that this transversality condition is necessary by verifying Assumption 1. Instead, a simpler strategy is to use Theorem 6 and show that these conditions are sufficient for a global maximum.

Condition (31) implies that \(\mu(t) > 0\) along this candidate path, since \(u' > 0\). Therefore, the current-value Hamiltonian is strictly concave in \(k\) since we assume \(f\) is strictly concave. This implies that \(M(k, \mu) = \max_c \hat{H}(k, c, \mu)\) is strictly concave in \(k\).

Second, since \(\mu(t) > 0\) and \(k(t) \geq 0\), any alternative path satisfies

\[
\lim_{t \to \infty} e^{-\rho t} \mu(t) k(t) \geq 0
\]

As a result, the candidate solution is a global maximum.

References


