Math Camp: Macroeconomics
Department of Economics, Harvard University

Argyris Tsiaras[1]

August 2017

[1]atsiaras@fas.harvard.edu

Contents

Overview

In this four-day mini-course, I will cover some important mathematical topics that are relevant for the material taught in the first-year macro sequence. On the first two days I will go over basic results in the theory of ordinary differential equations. On the second day I will also cover results on difference equations, which are very similar to differential equations with most results about the latter extending to the former. On the third day I will present solution methods for (linearized) systems of expectational difference equations, that is, difference systems involving expectations of future realizations of variables. These types of systems are particularly relevant for the important class of dynamic stochastic general equilibrium (DSGE) models, an important feature of which is the forward-looking nature of agents' policies; you will learn about DSGE models in the third quarter of the macro sequence. On the fourth day I will cover the theory of optimal control, which is an important method for solving (deterministic) optimization problems in continuous time. Optimal control theory is relevant for the second part of the macro sequence on economic growth as theories of economic growth have typically been formulated in continuous rather than discrete time, in contrast to most other areas of macroeconomics.

I will not cover important mathematical topics that are explicitly covered in class during the first-year sequence. In particular, I do not cover at all dynamic programming, by far the most important and widely used optimization method in economics (both in discrete and in continuous time); it deserves a course of its own, the first quarter of the macro sequence.

It is simply not possible for me to cover all of the aforementioned topics in full detail and for you to absorb all of this material, especially if you have not seen it before, within just four days of class. So, a large portion of the material in these notes will not be covered in detail or at all in class. I wrote the following chapters, which are quite comprehensive and include examples and economic applications chosen for their relevance for the first-year macro sequence, with the intention that they serve as a good reference for you during your first year and beyond. With this goal in mind, I also include an index of key terms so that you can use these notes as a reference more effectively. The starred sections and the last chapter on loglinearization methods (relevant for the third quarter of the macro sequence) will not be covered in class; they are included for completeness and as a point of reference for further study.

I have also prepared short, informal problem sets for you to work on after each of the first three days of class. The problem sets are strictly optional; the course is not graded in any explicit or implicit manner. The exercises are meant to incentivize and help guide your review of the material after each day of class. The problem sets are not to be turned in; we will go over the problems and their solutions in the beginning of class the following day.

I have consulted a number of textbooks and papers in preparing these notes. These sources are all included in the bibliography and I provide a list with my main sources for the material in each chapter at the end of this booklet, but I provide minimal citation of my sources within the text (with the exception of the sources of included figures) both for ease of exposition and due to the standard nature of the material covered (so that my source of the material, usually a textbook, does not necessarily reflect origin of the concept or application).

Chapter 1

Continuous Dynamical Systems: Solution Methods

This chapter discusses solution methods for deterministic ordinary differential equation (ODE) systems.

1.1 Introduction

Most models in macroeconomics are formulated in discrete time. That is, there are time periods $t = 0, 1, 2, \ldots$, where the unit of time is in general arbitrary and can refer to a day, a month, or a decade. This arbitrariness suggests that it may be helpful, especially when looking at model dynamics, to make the time unit as small as possible. Thus, a number of models in macroeconomics are formulated in continuous time. When we compare continuous-time dynamical systems with discrete-time dynamical systems in Chapter 3, we will see that continuous systems have a number of advantages: they allow for a more flexible analysis of dynamics and allow for explicit solutions in a wider set of circumstances. This is particularly so for heterogeneous-agent models. In addition, a number of "pathological" results of models formulated in discrete time disappear once we move to the corresponding continuous-time version of the model.

Consider a function $x : \mathcal{T} \to \mathbb{R}$, where $\mathcal{T}$ is an interval in $\mathbb{R}$. Given a real number Δt, function x satisfies

$$x(t + \Delta t) - x(t) = G(x(t), t, \Delta t)$$

where $G(x(t), t, \Delta t)$ is a real-valued function. Divide both sides of this equation by Δt and consider the limit as $\Delta t \to 0$. We obtain the *differential equation*

$$\dot{x}(t) \equiv \frac{dx(t)}{dt} = g(x(t), t) \tag{1.1}$$

where

$$g(x(t), t) \equiv \lim_{\Delta t \to 0} \frac{G(x(t), t, \Delta t)}{\Delta t} \tag{1.2}$$

is assumed to exist.

More generally, a differential equation is an equation for an (unknown) function of one or more independent variables (in the example above, the independent variable is time t) that relates the values of the function, the values of the (possibly higher-order) derivatives of the function, and the values of the independent variables. If the function has a single independent variable it is called an *ordinary differential equation* (ODE). If, instead, the function is multivariate we have a *partial differential equation*. We will only cover ordinary differential equations in math camp.

A differential equation is *explicit* if it is of the form

$$x^{(n)}(t) = g(x^{(n-1)}(t), \dots, x(t), t) \tag{1.3}$$

that is, the highest-order derivative of the differential equation is separated from the other terms. In contrast, an *implicit* differential equation has the form

$$g(x^{(n)}(t), x^{(n-1)}(t), \dots, x(t), t) = 0 \tag{1.4}$$

We will only deal with explicit ODEs in math camp.

A differential equation is *of order n* if the highest derivative appearing in the equation is of order n. It is *autonomous* if it does not explicitly depend on time (the independent variable) as a separate argument. Otherwise, the differential equation is called *nonautonomous*. For example,

$$\dot{x}(t) = g(x(t)) \tag{1.5}$$

is an autonomous first-order ODE.

A differential equation is *linear* if it takes the form

$$a_n(t)x^{(n)}(t) + a_{n-1}(t)x^{(n-1)}(t) + \cdots + a_1(t)\dot{x}(t) + a(t)x(t) + b(t) = 0 \tag{1.6}$$

where $a(t)$, $a_1(t), \dots, a_n(t)$ and $b(t)$ are arbitrary functions of time. It is *nonlinear* otherwise. Clearly, a linear differential equation is autonomous if and only if it has constant coefficients. Finally, a linear differential equation as in (1.6) with $b(t) = 0 \, \forall \, t$ is called *homogeneous*.

Boundary conditions are needed to pin down a specific solution to an ODE of the form (1.3) or (1.4). In general, we need n boundary conditions to pin down a solution to an ODE of order n. The most common form of an ODE problem is the *initial value problem*, whereby an ODE, for example, a first-order ODE of the form (1.1) is specified together with an initial condition $x(0) = x_0$. A *solution* to this initial value problem is a function $x : \mathcal{T} \to \mathbb{R}$ that satisfies (1.1) for all $t \in \mathcal{T}$ with $x(0) = x_0$. A family of functions $\{x : \mathcal{T} \to \mathbb{R} \text{such that} x \text{satisfies} (1.1), \forall \, t \in \mathcal{T}\}$ is often referred to as a *general solution*, while an element of this family that satisfies the boundary condition is called a *particular solution*.

Important ODE problems in economics are associated with boundary conditions other than initial values. For example, a terminal value condition specifies what the value of $x(t)$ should be at some finite horizon $T < \infty$ and a *transversality* condition specifies what $x(t)$ should be at the limit as $t \to \infty$.

An explicit ODE of the form (1.3) can be generalized by taking $x(t)$ and $g(\cdot)$ to be vector-valued functions, that is, $x(t) : \mathbb{R} \to \mathbb{R}^m$. We then have an m-dimensional *system of differential equations* of the form

$$\begin{bmatrix} x_1^{(n)}(t) \\ x_2^{(n)}(t) \\ \vdots \\ x_m^{(n)}(t) \end{bmatrix} = \begin{bmatrix} g_1(x^{(n-1)}(t), \dots, x(t), t) \\ g_2(x^{(n-1)}(t), \dots, x(t), t) \\ \vdots \\ g_m(x^{(n-1)}(t), \dots, x(t), t) \end{bmatrix} \tag{1.7}$$

where $x_i(t)$ refers to the ith component of vector $x(t)$.

A *first-order* ODE system of the form

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \\ \vdots \\ \dot{x}_m(t) \end{bmatrix} = \begin{bmatrix} g_1(x(t), t) \\ g_2(x(t), t) \\ \vdots \\ g_m(x(t), t) \end{bmatrix} \tag{1.8}$$

will be the main focus of our analysis in this chapter. It may at first appear that (1.8) is a restrictive special case of (1.7) but this is not true. Any higher-order differential equation or system can be transformed into an equivalent first-order ODE system by introducing additional variables in vector $x(t)$. For a concrete example, consider the second-order differential equation

$$\frac{1}{2}b^2 x''(t) + a x'(t) - \rho x(t) = 0 \tag{1.9}$$

where b, a, and ρ are constants. Let $y(t)$ denote a two-dimensional vector with $y_1(t) = x(t)$ and $y_2(t) = x'(t)$. Then, (1.9) is equivalent to the first-order system

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \frac{2\rho}{b^2} & -\frac{2a}{b^2} \end{bmatrix} \cdot \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \tag{1.10}$$

Thus, there is no loss of generality in restricting our attention to (1.8). Incidentally, in the same vein one can transform any nonautonomous system like (1.8) into an equivalent autonomous system by introducing the independent variable, t, as an additional component of vector $x(t)$. However, the latter transformation is not that useful. As we will see in Section 1.3, only autonomous systems have explicit solutions. In addition, only autonomous systems have steady states (equilibrium points) in general and are thus amenable to stability analysis, which is the focus of Chapter 2.

1.2 Basic Results in Linear Algebra

In this section we briefly review some concepts and results from linear algebra theory that we will use in our analysis of ODE systems in the following sections of this chapter.

1.2.1 Linear Map, Change of Basis

Recall from the micro part of math camp the definition of a *vector space*. Important examples of vector spaces include the Euclidean space $\mathbb{R}^n$ for $n \in \mathbb{N}$, the set of all infinite sequences $\mathbb{R}^\infty$, and the set of all functions from an arbitrary set S to $\mathbb{R}$ (in all of these examples the underlying *scalar field* is $\mathbb{R}$).

A (linear) *subspace* $W \subset V$ of a vector space is a subset of V that contains the zero vector, and for any $x, y \in W$, $x + y \in W$ and $\lambda x \in W$, where λ is an arbitrary scalar. A subspace of a vector space is itself a vector space.

Let V be a vector space and let S be a nonempty subset of V. The *span* of S, denoted by *span*(S), is defined to be the set consisting of all linear combinations of vectors in S. By convention, *span*$(\emptyset) = \{0\}$. A vector space is *finite-dimensional* if it is spanned by a finite set of vectors, and *infinite-dimensional* otherwise.

A set of vectors $v_1, , v_2, \ldots, v_n \in V$ are *linearly independent* if $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n = 0$, where $a_1, , a_2, \ldots, a_n$ are scalars, implies that $a_1 = a_2 = \cdots = a_n = 0$. If a set $B \subset V$ consists of linearly independent vectors and B spans V then each element $v \in V$ can be *uniquely* expressed as a linear combination of the elements in B. Such a subset is called a *basis* of V. If V is finite-dimensional and B is a basis of V with n elements, then V is said to have *dimension n*. For example, the standard or Euclidean basis of $\mathbb{R}^n$ are the n n-dimensional vectors $(1, 0, \ldots, 0)$, $(0, 1, \ldots, 0)$,..., $(0, 0, \ldots, 1)$.

Definition 1.1 (Linear map)**.** A *linear map* from a vector space V to a vector space U is a function $L : V \to U$ such that

- $L(v + w) = L(v) + L(w)$, $\forall\, v, w \in V$ *(additivity)*
- $L(\lambda v) = \lambda L(v)$, $\forall\, v \in V$, and $\lambda \in \mathbb{F}$, where $\mathbb{F}$ is the underlying scalar field of V *(homogeneity of degree 1)*

A linear ODE system, that is, a system of the form (1.8) where each function $g_i(\cdot, \cdot)$ is a linear function of its arguments, is an example of a linear map. It maps a function (an element of a vector space of differentiable functions) to its derivative.

Once the bases of the (finite-dimensional) vector spaces U and V in the definition above are specified, a *matrix* can capture all of the information of, and thus be identified with, map L. Assume $V \subseteq \mathbb{R}^m$ with basis $\{v_j\}_{j=1}^m$ and $U \subseteq \mathbb{R}^n$ with basis $\{u_i\}_{i=1}^n$. Then, the n-by-m matrix $P = [p_{ij}]$ that corresponds to map L under the specified bases satisfies

$$L(v_j) = p_{1j} u_1 + \cdots + p_{nj} u_n \quad \forall j = 1, \ldots, m \tag{1.11}$$

Once we know P we can find the value of $L(v)$ for any $v \in V$ in the following way. Let $c = [c_j]_{j=1}^m = c(v, \{v_j\}_{j=1}^m)$ be the m-dimensional vector that represents vector v under the specified basis: $v = c_1 v_1 + \cdots + c_m v_m$. Then,

$$\begin{aligned} L(v) &= L(c_1 v_1 + \cdots + c_m v_m) \\ &= c_1 L(v_1) + \cdots + c_m L(v_m) \\ &= \left(\sum_{j=1}^m p_{1j} c_j \right) u_1 + \cdots + \left(\sum_{j=1}^m p_{nj} c_j \right) u_n \\ &= Pc \end{aligned} \tag{1.12}$$

where the last line uses the convention that the ith component of a vector is the coefficient of the ith element of a given basis of its vector space when the vector is written as a linear combination of that basis.

An implication of this is that any (autonomous) linear ODE system can be written as

$$\dot{x}(t) = A\,x(t) \tag{1.13}$$

where $x(t) \in \mathbb{R}^n$, $n \in \mathbb{N}$, A is an $n \times n$ matrix, and $x(t)$ is the representation of the underlying vector with respect to the standard Euclidean basis of $\mathbb{R}^n$, or as

$$\dot{z}(t) = D\,z(t) \tag{1.14}$$

for a different $n \times n$ matrix D where z is the representation of the same underlying vector with respect to another basis of $\mathbb{R}^n$. Matrices A and D, which represent the linear map $M(t) : \mathbb{R}^n \to \mathbb{R}^n$ associated with the same ODE system under different bases, are called *similar*.

How are the representations $x(t)$ and z(t) of equations (1.13) and (1.14) related to each other? In our discussion of equations (1.11) and (1.12), take $V = U = \mathbb{R}^n$, so that $m = n$, and $\{u_i\}_{i=1}^n = \{e_i\}_{i=1}^n$ (the standard Euclidean basis) and $\{v_j\}_{j=1}^n$ are two different bases of $\mathbb{R}^n$, associated with representations $x(t)$ and $z(t)$, respectively. That is, map L now represents the *change of basis* from basis $\{v_j\}_{j=1}^n$ to the standard basis of $\mathbb{R}^n$ (note that map L is different from map M associated with the ODE system in (1.13) and (1.14)). Then, z is precisely vector c in equation (1.12), so that

$$x(t) = Pz(t) \tag{1.15}$$

where the jth column of P corresponds to the standard Euclidean representation of basis vector v_j.

1.2.2 Eigenvalues and eigenvectors

An $n \times n$ (square) matrix A is *nonsingular* or *invertible* if its determinant is not zero, $\det A \neq 0$, or equivalently if the only $n \times 1$ column vector v that is a solution to equation

$$Av = 0 \tag{1.16}$$

is the zero vector $v = (0, \ldots, 0)^T$. In other words, the columns of an invertible matrix are linearly independent. If A is invertible, there exists matrix A^{-1} such that $A^{-1}A = I_n$, where I_n is the $n \times n$ identity matrix. Conversely, if there exists a nonzero solution v to (1.16) or if $\det A = 0$, then A is singular and does not have an inverse.

A complex number λ is an *eigenvalue* of A if

$$\det(A - \lambda I_n) = 0 \tag{1.17}$$

$p_A(\lambda) = \det(A - \lambda I_n)$ is a polynomial of order n in λ, called the *characteristic polynomial* of A. Thus, λ is an eigenvalue of A if and only if it is a root of its characteristic polynomial. A is invertible if and only if none of its eigenvalues are equal to zero.

Given the eigenvalue λ of A, the $n \times 1$ nonzero column vector v_λ is an *eigenvector* of A if

$$(A - \lambda I_n)v_\lambda = 0 \tag{1.18}$$

Clearly, v_λ can only be unique up to a normalization, since if v_λ satisfies (1.18) then so does av_λ for any $a \in \mathbb{R}$.

Lemma 1.1. *If $\lambda_1, \ldots, \lambda_k$ are k distinct eigenvalues of square matrix A, so that $\lambda_i \neq \lambda_j$ for all $i \neq j$, then the associated eigenvectors $v_{\lambda_1}, \ldots, v_{\lambda_k}$ are linearly independent.*

From this lemma, it follows that if A has n distinct eigenvalues, then the associated eigenvectors form a basis of $\mathbb{R}^n$, called the *eigenbasis* for A. A key result is the following.

Theorem 1.1 (Spectral Decomposition)**.** *Suppose $n \times n$ matrix A has n distinct eigenvalues. Then, A satisfies*

$$A = PDP^{-1} \tag{1.19}$$

where D is the diagonal matrix with the eigenvalues $\lambda_1, \ldots, \lambda_n$ on the diagonal, and $P = (v_{\lambda_1}, \ldots, v_{\lambda_n})$ is a matrix with the corresponding eigenvectors as its columns.

Going back to our discussion of change of basis in the previous subsection, P, whose jth column is the (standard Euclidean representation of the) jth eigenvector, is associated with the linear map representing the change of basis from the eigenbasis to the standard Euclidean basis of $\mathbb{R}^n$.

The space spanned by the eigenvectors corresponding to a subset of eigenvalues is called the *eigenspace* of matrix A associated with these eigenvalues and is a linear subspace of $\mathbb{R}^n$. In Section 2.2, we will see that the stable subspace of a (homogeneous) linear system is precisely the eigenspace associated with the system's negative eigenvalues (eigenvalues with negative real parts, if complex).

Finally, note that when A has repeated eigenvalues, diagonalization is still possible through the use of *generalized eigenvectors*, which satisfy $(A - \lambda I_n)^k v_\lambda = 0$ for some $k \in \mathbb{N}$. However, in this case matrix D will be block diagonal rather than diagonal (it has the Jordan form). We will not cover this case in detail, although it is straightforward.

A final result we will make use of is summarized in the following Lemma:

Lemma 1.2. *Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_k$ and $m_1, \ldots, m_k$ denote the multiplicity of the corresponding eigenvalue. Then,*

(i) The determinant of A equals the product of its eigenvalues, repeated according to their multiplicity,

$$det(A) = \lambda_1^{m_1} \cdots \lambda_k^{m_k} \tag{1.20}$$

(ii) The trace of A, $tr(A)$ (defined to be the sum of the diagonal entries of A), equals the sum of its eigenvalues, repeated according to their multiplicity.

(iii) Let $p_A(\lambda) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n$ be the characteristic polynomial of A. Then

$$\begin{aligned} c_1 &= -tr(A) \\ c_n &= (-1)^n det(A) \end{aligned}$$

In particular, when A is a 2×2 matrix with eigenvalues λ_1 and λ_2,

$$p_A(\lambda) = \lambda^2 - tr(A)\lambda + det(A) \tag{1.21}$$

$$= \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1 \lambda_2 \tag{1.22}$$

1.3 Solutions of Autonomous Linear Systems

In this section, we will cover solutions to autonomous linear differential equations and systems.

1.3.1 Homogeneous Systems

A linear first-order differential equation has the general form

$$\dot{x}_t = a(t)x(t) + b(t) \tag{1.23}$$

Recall that a linear differential equation (or ODE system) is autonomous if and only if it has constant coefficients. Thus, an autonomous first-order linear differential equation takes the general form

$$\dot{x}(t) = ax(t) + b \tag{1.24}$$

Let us first consider the homogeneous linear equation

$$\dot{x}(t) = ax(t) \tag{1.25}$$

We can divide both sides with $x(t)$, integrate with respect to t, and recall that for $x(t) \neq 0$,

$$\int \frac{\dot{x}(t)}{x(t)}\,\mathrm{d}t = \log|x(t)| + c_0 \qquad \text{and}$$
$$\int a\,\mathrm{d}t = at + c_1$$

where c_0 and c_1 are constants of integration. Now, taking exponents on both sides, we obtain the general solution to (1.25),

$$x(t) = c\exp(at) \tag{1.26}$$

where c is a constant of integration combining c_0 and c_1. Suppose we are given an initial condition $x(0) = x_0$. This condition then pins down the unique value of the constant of integration. In this case, $c = x_0$.

We can generalize this simple derivation to arrive at the solution of a homogeneous first-order system of the form

$$\dot{x}(t) = Ax(t) \tag{1.27}$$

where $x(t) \in \mathbb{R}^n$ and A is an $n \times n$ matrix.

Under the assumption that A has *n distinct real* eigenvalues, we can transform (1.27) to an equivalent diagonal or “decoupled” system using Theorem 1.1. The transformed diagonal system is then simply a set of independent first-order linear homogeneous equations of the form (1.25), which have the solution (1.26) as we have already shown.

As we discussed in the previous section, we need to perform a change of basis from the standard Euclidean basis to the eigenbasis. The relationship between the representation of the vector under the standard basis, $x(t)$, and the representation of the same vector under the eigenbasis, $z(t)$, is once again given by equation (1.15). We then have

$$\begin{aligned}\dot{z}(t) &= P^{-1}\dot{x}(t)\\ &= P^{-1}Ax(t)\\ &= P^{-1}APz(t)\\ &= Dz(t)\end{aligned} \tag{1.28}$$

whose solution is $z_1(t) = c_1\exp(\lambda_1 t), \ldots, z_n(t) = c_n\exp(\lambda_n t)$, where $\lambda_1, \ldots, \lambda_n$ are the n distinct eigenvalues of matrix A.

We have thus derived the following result:

Theorem 1.2 (Solution to Homogeneous Autonomous Linear ODE Systems)**.** *Suppose $n \times n$ matrix A has n distinct eigenvalues $\lambda_1, \dots, \lambda_n$. Then the unique solution to (1.27), $\dot{x}(t) = Ax(t)$, with initial value $x(0) = x_0$ takes the form*

$$x(t) = \sum_{j=1}^{n} c_j \exp(\lambda_j t) v_{\lambda_j} \tag{1.29}$$

where $v_{\lambda_1}, \dots, v_{\lambda_n}$ are the eigenvectors corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$ and $c_1, \dots, c_n$ denote the constants of integration (pinned down by the initial value condition).

Theorem 1.2 applies only when all eigenvalues of A are real. What happens when some of the eigenvalues are complex (with nonzero imaginary parts)? The method and solution of Theorem 1.2 in fact still applies. Since A is a matrix with real entries, complex eigenvalues will always come in conjugate pairs. For example, assume A has two complex eigenvalues, $\lambda_1 = \alpha + i\mu$ and $\lambda_2 = \alpha - i\mu$, where $i \equiv \sqrt{-1}$ is the *imaginary unit*, and $v_{\lambda_1} = d + if$ and $v_{\lambda_2} = d - if$ are the corresponding eigenvectors. The remaining $n-2$ eigenvalues of A are real. Then, standard results in the theory of complex numbers imply that the general solution of (1.27) has the form

$$\begin{aligned} x(t) &= c_1 \exp(\alpha t)(d \cos(\mu t) - f \sin(\mu t)) \\ &\quad + c_2 \exp(\alpha t)(f \cos(\mu t) + d \sin(\mu t)) \\ &\quad + \sum_{j=3}^{n} c_j \exp(\lambda_j t) v_{\lambda_j} \end{aligned}$$

What happens when A has repeated eigenvalues? Recall our brief mention of generalized eigenvectors and the Jordan form in the previous section. It turns out that if A has k distinct eigenvalues $\lambda_1, \dots, \lambda_k$ with multiplicities $m_1, \dots, m_k$, respectively, the general solution to (1.27) has the form

$$\begin{aligned} x(t) &= \sum_{i=1}^{k} P_i(t) \exp(\lambda_i t) \\ &= \sum_{i=1}^{k} \sum_{j=1}^{m_i} p_{ij} t^{j-1} \exp(\lambda_i t) \end{aligned}$$

where $P_i(t)$ is a polynomial in t with vector-valued coefficients.

As will become clear later in the chapter, the case of n distinct real eigenvalues is by far the most relevant for economic applications.

1.3.2 Nonhomogeneous Systems

Next consider the autonomous but nonhomogeneous first-order linear equation

$$\dot{x}(t) = ax(t) + b \tag{1.30}$$

To derive the solution, use the change of variables $y(t) = x(t) + \frac{b}{a}$. Writing (1.30) in terms of $y(t)$,

$$\dot{y}(t) = ay(t) \tag{1.31}$$

which is a homogeneous linear equation, whose solution (1.26) we derived above. Transforming the equation back into $x(t)$, we obtain the general solution to (1.30) as

$$x(t) = -\frac{b}{a} + c\exp(at) \tag{1.32}$$

Application 1.1 (The Cobb-Douglas Version of the Solow Growth Model)**.** Consider the key equation of the Solow growth model

$$\dot{k}(t) = sf(k(t)) - \delta k \tag{1.33}$$

with initial condition $k(0) = k_0 > 0$. (1.33) says that capital (the capital-labor ratio), $k(t)$, which is the state variable of the model, grows by an amount equal to new investment minus depreciation. Investment equals the exogenously given and constant savings rate of the economy, s (where $0 \leq s \leq 1$), times output at time t, $f(k(t))$. Depreciation equals the exogenously given and constant depreciation rate, δ (where $0 \leq \delta \leq 1$) times the level of capital at time t.

We now solve (1.33) under the Cobb-Douglas specification for the output production function, $f(k(t)) = Ak(t)^\alpha$, where $0 \leq \alpha \leq 1$. Thus, (1.33) becomes

$$\dot{k}(t) = sAk(t)^\alpha - \delta k \tag{1.34}$$

This is a nonlinear differential equation, so it appears that our results above are not applicable to this problem. However, if we let $x(t) \equiv k(t)^{1-\alpha}$ and express equation (1.34) in terms of this new auxiliary variable, we get a linear differential equation.

Differentiating the definition of the auxiliary variable and applying the chain rule, $\dot{x} = (1-\alpha)k^{-\alpha}\dot{k}$. Then (1.34) implies

$$\begin{aligned}
\frac{\dot{x}}{(1-\alpha)k^{-\alpha}} &= sAk^\alpha - \delta k \\
\dot{x} &= (1-\alpha)sA - (1-\alpha)\delta k^{1-\alpha} \\
\dot{x}(t) &= -(1-\alpha)\delta x(t) + (1-\alpha)sA
\end{aligned}$$

which is an autonomous, nonhomogeneous linear first-order differential equation in $x(t)$. A direct application of formula (1.32) then gives

$$x(t) = \frac{sA}{\delta} + c\exp\left(-(1-\alpha)\delta t\right) \tag{1.35}$$

where c is the constant of integration, pinned down by the initial condition $x(0) = x_0 = k_0^{1-\alpha}$:

$$x(t) = \frac{sA}{\delta} + \left[x_0 - \frac{sA}{\delta}\right]\exp\left(-(1-\alpha)\delta t\right) \tag{1.36}$$

Expressing this in terms of $k(t)$, we obtain the solution to our initial value problem

$$k(t) = \left\{\frac{sA}{\delta} + \left[k_0^{1-\alpha} - \frac{sA}{\delta}\right]\exp\left(-(1-\alpha)\delta t\right)\right\}^{\frac{1}{1-\alpha}} \tag{1.37}$$

The solution reveals, in particular, that the economy converges to the steady-state level of capital $\bar{k} = \left\{\frac{sA}{\delta}\right\}^{\frac{1}{1-\alpha}}$ and the gap between $k(t)$ and $\bar{k}$ narrows at the exponential rate $(1-\alpha)\delta$. That is, less diminishing returns to capital (higher α) and slower depreciation (lower δ) imply slower adjustment to the steady state. ■

The derivation of (1.32) illustrates that nonhomogeneous linear equations and systems, whether autonomous or nonautonomous, can be easily transformed into homogeneous systems with a simple change of variables; yet, it is convenient to explicitly derive the solution for nonhomogeneous systems of the form

$$\dot{x}(t) = Ax(t) + B \tag{1.38}$$

where B is an $n \times 1$ vector with constant coefficients.

It turns out that the general solution of such a system can be written as

$$x^N(t) = x^H(t) + x^P(t)$$

where $x^H(t)$ is the general solution of the corresponding homogeneous system, $\dot{x}(t) = Ax(t)$, and $x^P(t)$ is an arbitrary particular solution of the nonhomogeneous system. We will see in the next section that this holds for nonautonomous linear systems as well.

Since we already know how to compute the solution to the homogeneous system, we only need to find one particular solution of the nonhomogeneous system. An obvious choice is the stationary solution of the system, denoted by $\bar{x}$, whenever it exists. The stationary solution by definition satisfies

$$\begin{aligned} \dot{x}|_{x=\bar{x}} &= 0 \\ \Rightarrow A\bar{x} &= -B \\ \Rightarrow \bar{x} &= -A^{-1}B \end{aligned}$$

provided A is invertible (that is, it does not have any zero eigenvalues).

We then obtain the following result, a direct analog to (1.32) for systems:

Theorem 1.3 (Solution to Nonhomogeneous Autonomous Linear ODE Systems)**.** *Suppose $n \times n$ matrix A has n distinct nonzero eigenvalues $\lambda_1, \ldots, \lambda_n$. Then the unique solution to* (1.38), $\dot{x}(t) = Ax(t) + B$, *with initial value $x(0) = x_0$ takes the form*

$$x(t) = \bar{x} + \sum_{j=1}^{n} c_j \exp(\lambda_j t) v_{\lambda_j} \tag{1.39}$$

where $\bar{x} = -A^{-1}B$ is the unique stationary state of the system, $v_{\lambda_1}, \ldots, v_{\lambda_n}$ are the eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$, and $c_1, \ldots, c_n$ denote the constants of integration (pinned down by the initial value condition).

Application 1.2 (Liquidity Traps in the New Keynesian Model)**.** Consider a deterministic, continuous-time, linearized version of the New Keynesian model as in Werning (2012):

$$\dot{x}(t) = \gamma^{-1}\left(i(t) - \pi(t) - r(t)\right) \tag{1.40}$$

$$\dot{\pi}(t) = \rho\pi(t) - \kappa x(t) \tag{1.41}$$

$$i(t) \geq 0 \tag{1.42}$$

where $\rho, \gamma, \kappa > 0$. As you will see in the third quarter of the macro sequence, the key assumption behind the model is that nominal rigidities (sticky prices) imply that inflation has an effect on equilibrium output (another way to put this is that aggregate demand affects equilibrium output). The two key endogenous variables are the output gap, $x(t)$, which represents the log difference between actual output and the hypothetical output that would prevail in an identical economy that is not subject to nominal rigidities (that is, in which prices are fully flexible), and inflation $\pi(t) \equiv \dot{p}(t)$, where $p(t)$ is an economy-wide price index. $i(t)$ is the nominal interest rate, so that $i(t) - \pi(t)$ is, by definition, the equilibrium real interest rate. Finally, the *exogenous* variable $r(t)$ stands for the *natural* real interest rate, that is, the real interest rate that would prevail in an identical economy in which prices are fully flexible. Equation (1.40) is the Dynamic IS (DIS) equation and reflects consumers' optimal intertemporal decision between consumption and saving. The DIS equation represents the demand side of the model. Equation (1.41) is the New Keynesian Phillips Curve (NKPC) and represents the supply side of the model. The monetary authority, which dislikes both output gaps (whether positive or negative) and inflation (whether positive or negative), is free to choose the nominal interest rate, subject to the constraint (1.42). Constraint (1.42), $i(t) \geq 0$, is the Zero Lower Bound (ZLB) constraint, which captures the fact that the monetary authority is unable to set a negative nominal interest rate, or people would exchange their savings for cash, ensuring a (net) nominal return equal to zero.

A liquidity trap refers to a situation where the natural real interest rate $r(t)$ is negative. In particular, assume that, for some $T > 0$,

$$r(t) = \underline{r} < 0 \qquad \text{for } 0 \leq t < T \tag{1.43}$$

$$r(t) = \overline{r} > 0 \qquad \text{for } t \geq T \tag{1.44}$$

That is, the economy exits the liquidity trap at some (known) future date T. Also assume that the monetary authority cannot precommit to a future course of action, so that agents take it as given that, when the economy exits the liquidity trap at time T, monetary policy will return to its "normal-times" optimal action, setting $i^*(t) = r(t) = \overline{r}$, regardless of any promises it previously made to the contrary (any such promises are not credible). This policy is optimal at T as it implies that $x(t) = \pi(t) = 0$ for all $t \geq T$, replicating the flexible-price equilibrium. Note that policy $i(t) = r(t)$ is not feasible for $t < T$ due to the ZLB constraint. Unable to do any better, the monetary policy will set $i^*(t) = 0$ for $t < T$.[1]

Since we have that $(x(t), \pi(t)) = (0, 0)$ for all $t \geq T$, we now need to solve system (1.40)-(1.42) for $t \in [0, T)$ in order to fully characterize the dynamics of an economy faced with a liquidity trap. Given our assumptions on the exogenous process for $r(t)$ and the monetary authority's policy $i^*(t)$ during the liquidity trap, we end up with a *terminal value problem* consisting of the 2×2 linear system

$$\dot{x}(t) = -\pi(t) - \underline{r} \tag{1.45}$$

$$\dot{\pi}(t) = \rho\pi(t) - \kappa x(t) \tag{1.46}$$

for $t \in \mathcal{T} = [0, T]$ and the terminal value condition $x(T) = \pi(T) = 0$. Note that we have assumed for simplicity that $\gamma = 1$.[2]

First write the system in matrix form:[3]

$$\begin{bmatrix} \kappa\dot{x} \\ \dot{\pi} \end{bmatrix} = \overbrace{\begin{bmatrix} 0 & -\kappa \\ -1 & \rho \end{bmatrix}}^{=A} \begin{bmatrix} \kappa x \\ \pi \end{bmatrix} + \begin{bmatrix} -\kappa\underline{r} \\ 0 \end{bmatrix} \tag{1.47}$$

which reveals that our system is a 2×2 autonomous nonhomogeneous linear ODE system, whose solution is given in Theorem 1.3 (with the constants of integration now pinned down by the terminal value condition).

The characteristic polynomial of matrix A is

$$\begin{aligned} p_A(\lambda) &= \det(A - \lambda I_2) \\ &= \lambda^2 - \rho\lambda - \kappa \end{aligned} \tag{1.48}$$

[1] Cochrane (2014) has recently criticized the standard line of argument that lack of commitment pins down $(x(T), \pi(T)) = (0, 0)$ as the only possible equilibrium outcome. Also see the discussion of monetary policy determinacy in Chapter 4.

[2] γ is the inverse of the intertemporal elasticity of substitution of the representative consumer.

[3] We have multiplied the first equation by κ as it turns out to yield simpler algebraic manipulations; this step is not necessary.

where the second line follows directly from part (iii) of Lemma 1.2. The roots of this quadratic equation are the two eigenvalues of the system

$$\lambda_+ = \frac{1}{2}\left(\rho + \sqrt{\rho^2 + 4\kappa}\right) \tag{1.49}$$

$$\lambda_- = \frac{1}{2}\left(\rho - \sqrt{\rho^2 + 4\kappa}\right) \tag{1.50}$$

We have $\lambda_+ > 0 > \lambda_-$, since κ is a positive constant (κ reflects the degree of price rigidity, so that $\kappa = 0$ implies no price rigidity).

Manipulating these expressions and simply recalling from Lemma 1.2 that, for a 2×2 matrix,

$$\begin{aligned} p_A(\lambda) &= \lambda^2 - tr(A)\lambda + det(A) && (1.51)\\ &= \lambda^2 - (\lambda_+ + \lambda_-)\lambda + \lambda_+\lambda_- && (1.52) \end{aligned}$$

we have $\lambda_+ + \lambda_- = \rho$, $\lambda_+\lambda_- = -\kappa$, and $\lambda_+ - \lambda_- = \sqrt{\rho^2 + 4\kappa}$.

To find the eigenvectors corresponding to each eigenvalue, note that the eigenvector for the positive root, which can be represented as $v_{\lambda_+} = (v, 1)$ after a normalization (recall that one component of the eigenvector can always be normalized) must satisfy

$$\begin{aligned} (A - \lambda_+ I_n)v_{\lambda_+} &= 0 && (1.53)\\ \Rightarrow -\lambda_+ v - \kappa &= 0 && (1.54) \end{aligned}$$

Using the relation $\lambda_+\lambda_- = -\kappa$ we find that $v_{\lambda_+} = [\lambda_-, 1]^T$. Similarly, for the negative eigenvalue we get $v_{\lambda_-} = [\lambda_+, 1]^T$.

Next, we compute the steady state of the system, which corresponds to vector $\overline{x}$ in the statement of Theorem 1.3. The steady state by definition satisfies

$$\dot{x} = 0 \Rightarrow \pi = -\underline{r} > 0 \tag{1.55}$$

$$\dot{\pi} = 0 \Rightarrow x = \frac{\rho}{\kappa}\pi \tag{1.56}$$

Combining them, we get

$$\begin{bmatrix} \overline{x} \\ \overline{\pi} \end{bmatrix} = \begin{bmatrix} -\frac{\rho}{\kappa}\underline{r} \\ -\underline{r} \end{bmatrix} > 0$$

Note that this is the steady state associated with the system (1.45)-(1.46), not the steady state of the model, which evolves under a different system after time T (so that the steady state of the model is $(\overline{x}, \overline{\pi}) = (0, 0)$).

We are now able to apply the result of Theorem 1.3, which gives a solution for our system

$$\begin{bmatrix} \kappa x(t) \\ \pi(t) \end{bmatrix} = \begin{bmatrix} -\rho\underline{r} \\ -\underline{r} \end{bmatrix} + c_+ \exp(\lambda_+ t) \begin{bmatrix} \lambda_- \\ 1 \end{bmatrix} + c_- \exp(\lambda_- t) \begin{bmatrix} \lambda_+ \\ 1 \end{bmatrix} \tag{1.57}$$

where c_+ and c_- are constant of integrations pinned down by our boundary condition $(x(T), \pi(T)) = (0,0)$. In particular,

$$x(T) = 0 \Rightarrow c_+\lambda_- \exp(\lambda_+ T) + c_-\lambda_+ \exp(\lambda_- T) = \rho \underline{r} \tag{1.58}$$

$$\pi(T) = 0 \Rightarrow c_+ \exp(\lambda_+ T) + c_- \exp(\lambda_- T) = \underline{r} \tag{1.59}$$

This is a 2×2 system of linear equations in c_+ and c_-, which can be easily solved to yield

$$c_+ = \frac{(\rho - \lambda_+)\underline{r}}{(\lambda_- - \lambda_+)\exp(\lambda_+ T)} = \frac{\lambda_- \underline{r}}{-\sqrt{\rho^2 + 4\kappa}} \exp(-\lambda_+ T) < 0 \tag{1.60}$$

$$c_- = \frac{(\rho - \lambda_-)\underline{r}}{(\lambda_+ - \lambda_-)\exp(\lambda_- T)} = \frac{\lambda_+ \underline{r}}{\sqrt{\rho^2 + 4\kappa}} \exp(-\lambda_- T) < 0 \tag{1.61}$$

Plugging the values for the constants back to (1.57), we get the solution to our terminal value problem:

$$\begin{bmatrix} \kappa x(t) \\ \pi(t) \end{bmatrix} = \begin{bmatrix} -\rho \underline{r} \\ -\underline{r} \end{bmatrix} + \overbrace{\frac{\underline{r}}{\sqrt{\rho^2 + 4\kappa}}}^{(-)} \begin{bmatrix} \lambda_+^2 & -\lambda_-^2 \\ \lambda_+ & -\lambda_- \end{bmatrix} \begin{bmatrix} \exp(-\lambda_-(T-t)) \\ \exp(-\lambda_+(T-t) \end{bmatrix} \tag{1.62}$$

We can now explore the dynamics of our system during the liquidity trap. First, for the output gap, we can calculate the rate of change in $x(t)$ from (1.62) as

$$\begin{aligned} \dot{x}(t) &= \frac{\underline{r}}{\kappa\sqrt{\rho^2 + 4\kappa}} \lambda_- \lambda_+^2 \exp(-\lambda_-(T-t)) - \frac{\underline{r}}{\kappa\sqrt{\rho^2 + 4\kappa}} \lambda_+ \lambda_-^2 \exp(-\lambda_+(T-t)) \\ &> 0 \qquad \forall\ t \in [0, T) \end{aligned} \tag{1.63}$$

This tells us that the output gap is negative but decreasing in absolute value during the liquidity trap until the time it reaches $t = T$, at which time it becomes zero (and stays at zero for all $t \geq T$). Also note that if, *counterfactually*, the economy obeyed system (1.45)-(1.46) even after time T, the derivative computed above implies that the output gap would keep growing to positive territory (towards positive infinity).

Now, consider the dynamics of inflation. We immediately see from (1.62) that $\pi(t) < 0$, that is, the economy faces deflation during the liquidity trap. The rate of change of inflation is

$$\begin{aligned} \dot{\pi}(t) &= \frac{\underline{r}}{\sqrt{\rho^2 + 4\kappa}} [\lambda_+ \lambda_- \exp(-\lambda_-(T-t)) - \lambda_- \lambda_+ \exp(-\lambda_+(T-t))] \\ &= -\frac{\kappa \underline{r}}{\sqrt{\rho^2 + 4\kappa}} [\exp(-\lambda_-(T-t)) - \exp(-\lambda_+(T-t))] \\ &> 0 \qquad \forall\ t \in [0, T) \end{aligned} \tag{1.64}$$

from which we see that $\dot{\pi}(T) = 0$ and, computing the second derivative of inflation, $\ddot{\pi}(t) < 0$ for all $t \leq T$. Thus, as time progresses and we get closer to the time when

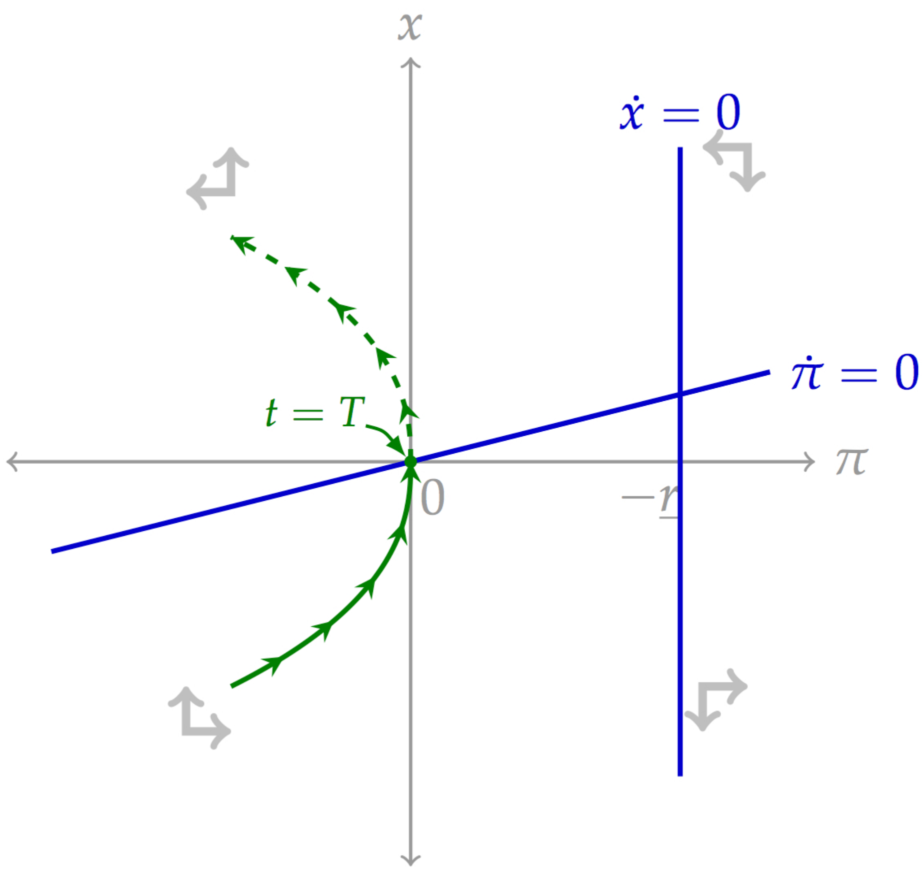


Figure 1.1: Phase diagram for a New Keynesian economy in a liquidity trap when the monetary authority lacks policy commitment. Source: Werning (2012).

the economy exits the liquidity trap, inflation increases (decreases in absolute magnitude), $\dot{\pi}(t) > 0$, but does so at a decelerating pace, $\ddot{\pi}(t) < 0$. If, *counterfactually*, the economy obeyed system (1.45)-(1.46) even after time T, inflation would go back to negative territory after reaching $\pi(T) = 0$, since $\dot{\pi}(T) = 0$ and $\ddot{\pi}(T) = \kappa \underline{r} < 0$.

The dynamics are nicely illustrated graphically in a *phase diagram*, shown in Figure 1.1. This example is, therefore, a good opportunity for an introduction to graphical analysis, which is particularly useful in the stability analysis of planar (2×2) continuous-time systems, covered in Chapter 2.

Let's explain the different elements of this phase diagram. System dynamics are depicted in *state space*, that is, each axis represents the values that each state variable can take (this contrasts with a graphical representation of the system where state variables are plotted over time, that is, time is on the horizontal axis). The two blue lines are called *phase lines*.[4] Each line shows the combination of points that feature no time movement in a given variable. In our case, the phase line for x corresponds

[4]Phase lines are also referred to as nullclines or zero-growth isoclines.

to the points for which $\dot{x} = 0$ and the phase line for π corresponds to the points for which $\dot{\pi} = 0$. As we saw in equation (1.55), $\pi = -\underline{r}$ for all points on the phase line for x, so that it is a vertical line. As we saw in equation (1.56), all points on the phase line for inflation satisfy $x = \frac{\rho}{\kappa}\pi$, so that the phase line is a straight line through the origin with slope $\frac{\rho}{\kappa} > 0$. The steady state or equilibrium point of the system (1.45)-(1.46) (which, once again, is different from the steady state of the economy which is $(0,0)$) is the intersection of the two phase lines. As we saw above, both the output gap and inflation are positive at the steady state, so that the intersection lies in the positive quadrant.

The (π, x) plane is thus divided into four regions by the two phase lines (in non-linear planar models the phase lines, which are curved, may intersect more than once, in which case the plane may be divided into more regions). The little gray arrows are called *arrows of motion* of the system and describe the direction of motion of the system along each of the axes. For example, let's consider the direction of motion in the region to the southwest of the steady state. Fix a given value for inflation in this region. In the DIS equation (1.45), which determines the rate of change of the output gap, we see that the derivative of x depends negatively on the level of π. Since the value of π that we are considering is lower than the value that makes $\dot{x} = 0$, we must have $\dot{x} > 0$. This is portrayed as an arrow pointing upwards, that is, in the direction of increasing x. Now fix a a value for the output gap in this region. In the New Keynesian Phillips Curve (1.46), which determines the rate of change of inflation, we see that the derivative of π depends negatively on the level of x. Since the value of x that we are considering is lower than the value that makes $\dot{\pi} = 0$, we must have $\dot{\pi} > 0$. This is portrayed as an arrow pointing to the right, that is, in the direction of increasing π.

The solid green line depicts the trajectory of the system over time (the arrows on the line represent the direction of time). Note that, since the trajectory is in the region to the southwest of the steady state (the point of intersection of the phase lines), it must obey the direction of the arrows of motion of that region. Hence, we see that both the output gap and inflation are increasing during the liquidity trap, as we derived analytically above.

We now discuss the dynamics of an economy entering a liquidity trap. Suppose that, for all $t < 0$, the economy is at its long-run steady state, $(x(t), \pi(t)) = (0,0)$. At time 0, the agents are informed of an (unanticipated) jump of the real interest rate to the negative value $r = \underline{r}$ for all $t \in [0, T)$ and they also know with certainty that the interest rate will jump back to its positive value $r = \overline{r}$ at time T, so that the economy will be at its long-run steady state $(0,0)$ at time T and stay at point $(0,0)$ forever after time T. What happens at time 0 when the real interest rate shock occurs? The economy will *jump* from $(0,0)$ to a point in the region to the southwest of the steady state. This is the (unique) point that ensures that the economy will go back to point $(0,0)$ at T while obeying the liquidity-trap dynamics (1.45)-(1.46) in the mean time. In other words, it is the terminal value condition that pins down the point at which the economy will "enter" the liquidity-trap dynamics at time 0. As can been by the arrows of motions of the four regions in Figure 1.1, the economy could not reach point $(0,0)$ at time T if it jumped at time 0 to any region other than that to the southwest of the liquidity-trap steady state.[5]

What are the comparative statics of $(x(0),\pi(0))$ with respect to the duration T of the liquidity trap? The figure makes clear that, as T increases, the economy would have to jump to more and more negative values for the output gap and inflation at time 0 in order to still be able to reach point $(0,0)$ at time T. That the magnitude of the recession is increasing in the expected duration of the liquidity trap is one of the key predictions of the liquidity-trap version of the New Keynesian model.

Finally, the dashed green line depicts the trajectory of the economy after time T in the *counterfactual* scenario that the economy continued to obey the system (1.45)-(1.46) even after time t. This would be the case if, as soon as the economy reached its long-term point $(0,0)$ at time T (given the initial expectation that the economy would exit the liquidity trap at time T), agents were unexpectedly informed that the economy would in fact remain in the liquidity trap forever (that is, the real interest rate would be stuck at its negative value, $\underline{r}$). Just as we derived above, in that case inflation would turn back to negative territory after reaching 0 at time T. ■

1.4 Solutions of General (Nonautonomous) Linear Systems

1.4.1 Homogeneous Systems

Unfortunately, nonautonomous linear systems, that is, systems whose coefficients are time-varying, in general do not admit explicit solutions of the form (1.39) with the exception of a few special cases, such as the unidimensional case (that is, an equation rather than a system) and the special class of nonautonomous systems discussed in subsection 1.4.3.

Consider the homogeneous but nonautonomous linear equation

$$\dot{x}(t) = a(t)x(t) \tag{1.65}$$

defined over $t \geq 0$. That is, $a(t)$ is the instantaneous growth rate of variable $x(t)$ at time t. Following the same procedure as for the autonomous analog, (1.25), of this equation, we obtain the general solution

$$x(t) = c\exp\left(R(t)\right) \tag{1.66}$$

[5]It may appear arbitrary that the economy must jump at time 0 and follow a continuous path in state space afterwards. The reason is that in many continuous-time economic models, jumps in certain variables that are *fully expected* are inconsistent with equilibrium because they imply arbitrage opportunities. Therefore, jumps can only occur when new information arrives.

where

$$R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$$

We now consider the homogeneous system

$$\dot{x}(t) = A(t)x(t) \tag{1.67}$$

Recall our discussion of vector spaces in Section 1.2 . The following is an important result about the algebraic structure of the space of solutions of (1.67).

Theorem 1.4. *The set*

$$S^H = \{x(t),\ t \in \mathcal{T},\ \text{such that } \dot{x}(t) = A(t)x(t)\ \forall\, t \in \mathcal{T}\} \tag{1.68}$$

of solutions of (1.67) *is a vector space of dimension n.*

A set $\{x^1(t), \ldots, x^n(t)\}$ of n linearly independent solutions of (1.67), that is, a basis of S^H, is called a *fundamental set of solutions* of (1.67) and

$$X(t) \equiv \left[x^1(t), \ldots, x^n(t)\right] \tag{1.69}$$

is a *fundamental matrix* for the system. Note that the fundamental matrix itself satisfies equation (1.67), that is, $\dot{X}(t) = A(t)X(t)$.

Since a fundamental set of solutions is a basis of S^H any arbitrary solution $\tilde{x}(t)$ of (1.67) can be written uniquely as a linear combination of $x^1(t), \ldots, x^n(t)$, that is, there exist unique scalars $c_1, \ldots, c_n$ such that

$$\tilde{x}(t) = \sum_{i=1}^{n} c_i x^i(t) = X(t)c \tag{1.70}$$

Therefore, (1.70) is the *general solution* of system (1.67). A *particular solution* of the system corresponds to a specific choice of the constants $\{c_i\}_{i=1}^n$, pinned down by a given set of boundary conditions.

Boundary conditions can take the form $\tilde{x}(t_0) = x_0$ for any $t_0 \in \mathcal{T}$. Take $\tilde{x}(0) = x_0$ to be a given set of boundary (initial value) conditions. Then, we have $\tilde{x}(0) = X(0)c = x_0$, so that c is pinned down as $c = [X(0)]^{-1}x_0$ and the solution to a given initial value problem is

$$\tilde{x}(t) = X(t)[X(0)]^{-1}x_0 \tag{1.71}$$

A different way to state the general and particular solutions (1.70) and (1.71) of the homogeneous system (1.67) is by defining the *state transition matrix*, $\Phi(t,s)$, corresponding to $A(t)$. The state transition matrix is defined as the the $n \times n$ matrix function that is differentiable in its first argument and is uniquely defined by

$$\frac{d}{dt}\Phi(t,s) = A(t)\Phi(t,s) \qquad \text{and} \tag{1.72}$$

$$\Phi(t,t) = I_n \tag{1.73}$$

for all $t, s \in \mathcal{T}$. In the one-dimensional case, (1.65), the state transition matrix reduces to the scalar $\Phi(t,s) = \exp(R(t) - R(s))$.

It can be shown that that the state-transition matrix satisfies

$$\hat{x}(t) = \Phi(t,s)\hat{x}(s) \tag{1.74}$$

for any solution $\hat{x}(t)$ of the homogeneous system (1.67), justifying its name.

One can also show that

$$\Phi(t,s) = X(t)[X(s)]^{-1}$$

so that we have arrived the following result.

Theorem 1.5 (Solution to General Homogeneous Linear ODE Systems)**.** *Let $X(t)$ and $\Phi(t,s)$, $\forall t,s \in \mathcal{T}$ be the fundamental matrix and state-transition matrix, respectively, corresponding to the matrix-valued function $A(t)$. Then, a (particular) solution to the homogeneous linear system* (1.67)*, $\dot{x}(t) = A(t)x(t)$, with boundary condition $x(0) = x_0$, is given by*

$$\hat{x}(t) = X(t)[X(0)]^{-1}x_0 \tag{1.75}$$

$$= \Phi(t,0)x_0 \tag{1.76}$$

Note that the theorem above does not offer explicit solutions to a nonautonomous linear system, since it assumes that one already knows a fundamental set of solutions to the system.

1.4.2 Nonhomogeneous Systems

Consider the most general form of a linear first-order differential equation:

$$\dot{x}(t) = a(t)x(t) + b(t) \tag{1.77}$$

The term $b(t)$ is called the *forcing term* of the equation. To derive the explicit solution for this case, we need to use a slightly different argument compared to the previous special cases. Consider the *integrating factor* $\exp(-R(t))$, where again $R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$. Multiply both sides of (1.77) with the integrating factor to obtain

$$\dot{x}(t)\exp(-R(t)) - a(t)x(t)\exp(-R(t)) = b(t)\exp(-R(t))$$

Note from the definition of $R(t)$ that $dR(t)/dt = a(t)$, so that the left-hand side of the equation above is the derivative of $x(t)\exp(-R(t))$,

$$\frac{d}{dt}[x(t)\exp(-R(t))] = b(t)\exp(-R(t))$$

Integrating both sides yields

$$x(t)\exp(-R(t)) = \int_0^t b(s)\exp(-R(s))\,\mathrm{d}s + c$$

where c is the constant of integration. Finally, multiplying both sides with $\exp(R(t))$, we obtain the general solution, given in the following lemma.

Lemma 1.3 (Solution to a General Linear First-Order Differential Equation)**.** *The solution to a general linear first-order differential equation*

$$\dot{x}(t) = a(t)x(t) + b(t)$$

is given by

$$x(t) = \left[c + \int_0^t b(s)\exp(-R(s))\,\mathrm{d}s\right]\exp(R(t)) \tag{1.78}$$

where $R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$ *and* c *is a constant of integration pinned down by a boundary condition.*

Finally, let's consider the most general case of a linear ODE system[6]

$$\dot{x}(t) = A(t)x(t) + B(t) \tag{1.79}$$

Equation (1.79) is a general nonhomogeneous linear system, whose solution space is

$$S^N = \{x(t),\, t \in \mathcal{T},\ \text{ such that } \dot{x}(t) = A(t)x(t) + B(t),\ \forall t \in \mathcal{T}\} \tag{1.80}$$

It turns out that S^N is an affine space of dimension n, over (parallel to) S^H, defined in (1.68). That is, S^N is a translation of S^H and, moreover, the translation factor is an arbitrary particular solution $x^P(t)$ of (1.79). Thus, even in the nonautonomous case, the general solution of the nonhomogeneous system is the sum of the general solution of the corresponding homogeneous system and an arbitrary particular solution of the nonhomogeneous system, that is,

$$x^N(t) = x^H(t) + x^P(t) \tag{1.81}$$

[6]Recall from our discussion in the beginning of the chapter that the fact that an ODE system is written as a first-order system is without loss of generality, as any higher-order system can be transformed into a first-order system.

$x^H(t)$ is sometimes called the *complementary function* of system (1.79).

In the autonomous case, we could find a particularly tractable particular solution of the nonhomogeneous system, the steady state, that we could then add to the complementary function of the system. Nonautonomous systems like (1.79) in general do not have steady states, so it can be quite difficult to find a particular solution. However, our final result states that, once we have a fundamental set of solutions $X(t)$ for the corresponding homogeneous system, we also have the general solution to the nonhomogeneous system.

Theorem 1.6 (Solution to General Linear ODE Systems)**.** *Let $X(t)$ and $\Phi(t,s)$, $\forall t,s \in \mathcal{T}$ be the fundamental matrix and state-transition matrix, respectively, corresponding to the matrix-valued function $A(t)$. Then, a (particular) solution to the linear system* (1.79)*, $\dot{x}(t) = A(t)x(t) + B(t)$, with boundary condition $x(0) = x_0$, is given by*

$$\hat{x}(t) = X(t)[X(0)]^{-1}x_0 + \int_0^t X(t)[X(s)]^{-1}B(s)\,\mathrm{d}s \tag{1.82}$$

$$= \Phi(t,0)x_0 + \int_0^t \Phi(t,s)B(s)\,\mathrm{d}s \tag{1.83}$$

1.4.2.1 Forward Solutions

A solution of the form of Lemma 1.3 and Theorem 1.6 is sometimes called the *backward solution* of the differential equation or system, since we express the solution by integrating over past values of the forcing term. Backward solutions are useful when the system has a natural predetermined initial condition. But they may not be as useful for economic problems that involve boundary conditions at infinity.

Suppose that we are asked to solve the differential equation

$$\dot{x}(t) = a(t)x(t) + b(t) \tag{1.84}$$

where $a(t) > 0$ for all t, $b(t)$ satisfies

$$\left|\int_0^\infty b(s)\exp(-R(s))\,\mathrm{d}s\right| < \infty \tag{1.85}$$

where $R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$, and $x(t)$ satisfies:

1. either a *transversality condition*

$$\lim_{t\to\infty} g(t)x(t) = K \in \mathbb{R}_+ \tag{1.86}$$

   for some function $g(t)$ that satisfies $|g(t)| \geq \exp(-R(t))$

2. or the condition that the solution be *bounded*, which in this (deterministic) case means that there exists $K \in \mathbb{R}_+$ such that

$$x(t) \leq K \tag{1.87}$$

for all t.

Note that in both cases, $\lim_{s\to\infty}[x(s)\exp(-R(s))]$ is finite. In case 1 with $|g(t)| = \exp(-R(t))$, $\lim_{s\to\infty}[x(s)\exp(-R(s))] = K$. In case 2, this limit is zero.

In this type of problem, the *forward solution* of the equation is more useful. Recall from the derivation of Lemma 1.3 that, using the integrating factor $\exp(-R(t))$, where $R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$, the differential equation can be rewritten as

$$\frac{d}{dt}[x(t)\exp(-R(t))] = b(t)\exp(-R(t)) \tag{1.88}$$

Instead of integrating this equation backwards, as we did earlier, we now integrate this forward

$$\int_t^\infty \left\{\frac{d}{ds}[x(s)\exp(-R(s))]\right\}\mathrm{d}s = \int_t^\infty b(s)\exp(-R(s))\,\mathrm{d}s \tag{1.89}$$

which yields

$$\lim_{s\to\infty}[x(s)\exp(-R(s))] - x(t)\exp(-R(t)) = \int_t^\infty b(s)\exp(-R(s))\,\mathrm{d}s \tag{1.90}$$

This can be rewritten as

$$x(t) = \exp(R(t))\lim_{s\to\infty}[x(s)\exp(-R(s))] - \int_t^\infty b(s)\exp(R(t)-R(s))\,\mathrm{d}s \tag{1.91}$$

Define the *fundamental forward solution* of (1.84) to be

$$F(t) \equiv -\int_t^\infty b(s)\exp(R(t)-R(s))\,\mathrm{d}s \tag{1.92}$$

Then the (forward) solution to the problem can be expressed as

$$x(t) = F(t) + K\exp(R(t))$$
$$x(t) = F(t)$$

for cases 1 and 2, respectively. Note that expressing the solution to $x(t)$ as in (1.91) is valid only if the limit and the integral in the expression are well-defined, which is true if and only if $|F(t)| < \infty$ for all $t \in \mathcal{T}$ and $|\lim_{s\to\infty} x(s)\exp(-R(s))| < \infty$, that is, (the absolute value of) variable $x(t)$ cannot growth too fast.

We see that, for case 1, the forward solution is comprised of two terms, the fundamental solution and another term sometimes called the *bubble term* of the forward solution. If we assume that $a(t) > 0$ for all t, so that $R(t) > 0$, then the bubble term increases at an exponential rate of $R(t)$, so that $x(t)$ becomes unbounded even though its fundamental forward solution remains bounded at all times. *If we require our solution to be bounded (case 2), then bubbles are ruled out.*

Also note that, when it exists, the limit $\lim_{s\to\infty} x(s)\exp(-R(s))$ can also be expressed using the backward solution of (1.84) (Lemma 1.3) as

$$\lim_{s\to\infty}\left[x(s)\exp(-R(s))\right] = x(0) + \int_0^\infty b(s)\exp(-R(s))\,\mathrm{d}s \tag{1.93}$$

$$= x(0) - F(0) \tag{1.94}$$

so that the (forward) solution to (1.84) can also be expressed as

$$x(t) = F(t) + [x(0) - F(0)]\exp(R(t)) \tag{1.95}$$

Forward solutions will play an important role when we look at expectational dynamical systems in Chapter 4.

Theorem 1.7 (Forward Solution to Linear ODE Equations)**.** *Consider the linear equation $\dot{x}(t) = a(t)x(t) + b(t)$ subject to the conditions*

$$\lim_{t\to\infty} |x(t)\exp(-R(t))| < \infty$$

and

$$\left|\int_t^\infty b(s)\exp(-R(s))\,\mathrm{d}s\right| < \infty$$

for all $t \geq 0$, where $R(t) \equiv \int_0^t a(s)\,\mathrm{d}s$.

Then, the solution to the equation can be written as

$$x(t) = F(t) + [x(0) - F(0)]\exp(R(t)) \tag{1.96}$$

where

$$F(t) = -\int_t^\infty b(s)\exp(R(t) - R(s))\,\mathrm{d}s \tag{1.97}$$

is the fundamental forward solution of the equation.

1.4.3 A special class of nonautonomous systems

A special case of nonautonomous linear systems that is relevant for economic applications and that admits explicit solutions (like autonomous systems do) is

$$\dot{x} = Ax(t) + B(t) \tag{1.98}$$

That is, matrix A is constant, but the forcing term $B(t)$ is time-varying.

Because matrix A is constant, we can use our diagonalization method for A just like we did in the autonomous case. Assume A has n distinct, non-zero, real eigenvalues and let P denote the matrix whose columns consist of n linearly independent eigenvectors of A. Then, again define $z(t) = P^{-1}x(t)$, as in (1.15). Now, let $E(t) = P^{-1}B(t)$, so that the transformed system

$$\dot{z}(t) = Dz(t) + E(t) \tag{1.99}$$

is simply a set of n independent linear equations that we know how to solve from Lemma 1.3. For example, the equation

$$\dot{z}_i(t) = \lambda_i z_i(t) + E_i(t) \tag{1.100}$$

has solution

$$z_i(t) = \left[c_i + \int_0^t E_i(s)\exp\left(-\lambda_i s\right)\,\mathrm{d}s\right]\exp(\lambda_i t) \tag{1.101}$$

where c_i is a constant of integration.

Once vector $z(t)$ has been computed we can simply apply the inverse of the original transformation, $x(t) = Pz(t)$, to obtain $x(t)$.

1.5 Two Special cases of nonlinear ODEs

Nonlinear ODE systems in general do not admit explicit solutions. However, there are two special cases that appear in economic applications and that allow us to derive explicit solutions: separable differential equations and exact differential equations.

First, a differential equation

$$\dot{x}(t) = g(x(t), t) \tag{1.102}$$

is *separable* if g can be written as

$$g(x,t) \equiv f(x)h(t) \tag{1.103}$$

In other words, the part of g that depends on x is separate from the one that depends on t. Then, we have

$$\frac{dx(t)}{f(x(t))} = h(t)dt \tag{1.104}$$

$$\Rightarrow \int \frac{1}{f(x)}dx = \int h(t)dt \tag{1.105}$$

For example, suppose that $f(x) = 1/x$. Then we can obtain the explicit solution for $x(t)$ as

$$x^2 = \int h(t)dt + c \tag{1.106}$$

$$\Rightarrow x(t) = \pm\sqrt{\int h(t)dt + c} \tag{1.107}$$

Second, a differential equation

$$\dot{x}(t) = g(x(t), t) \tag{1.108}$$

is *exact* if g can be written as

$$g(x(t), t) \equiv \frac{G_1(x(t), t)}{G_2(x(t), t)} \tag{1.109}$$

where

$$G_1(x(t), t) \equiv \frac{\partial F(x(t), t)}{\partial t} \qquad \text{and} \qquad G_2(x(t), t) \equiv -\frac{\partial F(x(t), t)}{\partial x} \tag{1.110}$$

for some function $F(\cdot, \cdot)$.

Then, note that we have

$$\dot{x}(t)\frac{\partial F(x(t), t)}{\partial x} + \frac{\partial F(x(t), t)}{\partial t} = 0 \tag{1.111}$$

$$\Rightarrow \frac{d}{dt}F(x(t), t) = 0 \tag{1.112}$$

where $\frac{d}{dt}$ denotes the *total* derivative of function $F(\cdot, \cdot)$ with respect to t.

Clearly, we have

$$F(x(t), t) = c \tag{1.113}$$

where c is a constant of integration, pinned down by a boundary condition. This equation implicitly defines the solution to the exact differential equation.

We see that the solution method is quite straightforward, although identifying an exact differential equation can sometimes be tricky. Consider the following example.

Example 1.1 (Exact Differential Equation)**.** Let's solve the boundary value problem

$$\dot{x}(t) = -\frac{2x(t)\log x(t)}{t} \tag{1.114}$$

subject to $x(1) = \exp(1)$.

Although not immediately apparent, this is an exact differential equation since it can be written as

$$\dot{x}(t) = -\frac{2t\log x(t)}{\frac{t^2}{x(t)}} = -\frac{\frac{\partial(t^2\log x(t))}{\partial t}}{\frac{\partial(t^2\log x(t))}{\partial x}} \tag{1.115}$$

Therefore, its general solution is given by

$$t^2 \log x(t) = c \tag{1.116}$$

$$\Rightarrow x(t) = \exp(ct^{-2}) \tag{1.117}$$

Finally, the boundary condition implies that $c = 1$, so that the solution to our problem is

$$x(t) = \exp(t^{-2}) \tag{1.118}$$

1.6 General Results on Properties of Solutions*

In the previous sections, we discussed methods for obtaining analytical solutions to certain special cases. When we consider a general nonlinear system of the form

$$\dot{x}(t) = g(x(t), t) \tag{1.119}$$

where $g : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$, we have no hope of obtaining closed-form solutions in general but at least we want to know whether a solution exists and whether it is unique. Moreover, we want to know whether, given a function $g(\cdot,,\cdot)$, we can expect the solution, expressed as a function of the independent variable (time) as well as parameters and initial conditions, to possess certain desirable properties, such as continuity and differentiability with respect to its arguments.

It turns out that the key conditions that function $g(\cdot,\cdot)$ has to satisfy for many of these desirable properties is that it is (locally) Lipschitz continuous in the dependent variable, $x(t)$, and continuous in the independent variable, t.

Definition 1.2 (Lipschitz Continuity)**.** Let $\mathcal{X}$ and $\mathcal{Y}$ be two normed vector spaces. A function $g : \mathcal{X} \times \mathcal{T} \to \mathcal{Y}$, where $\mathcal{T} \subseteq \mathbb{R}$, is *Lipschitz continuous* in x on E or is said to satisfy the *Lipschitz condition* on E, where $E = \mathcal{C} \times \mathcal{D}$, $\mathcal{C} \subseteq \mathcal{X}$ and $\mathcal{D} \subseteq \mathcal{T}$, if there exists a positive real number $L < \infty$ such that

$$\|g(x,t) - g(x',t)\| \leq L\|x - x'\| \tag{1.120}$$

for all $x, x' \in \mathcal{C}$.

A function $g : \mathcal{X} \times \mathcal{T} \to \mathcal{Y}$, $\mathcal{T} \subseteq \mathbb{R}$, is *locally* Lipschitz continuous in x on E, where $E = \mathcal{C} \times \mathcal{D}$, $\mathcal{C} \subseteq \mathcal{X}$ and $\mathcal{D} \subseteq \mathcal{T}$, if for every point $(x_0, t_0) \in E$ there exists some $\varepsilon > 0$ and a positive real number $L(x_0, t_0) = L_0 < \infty$ such that $B_\varepsilon(x_0, t_0) \subseteq E$ and

$$\|g(x,t) - g(x',t)\| \leq L_0\|x - x'\| \tag{1.121}$$

for all $(x,t),(x',t) \in B_\varepsilon(x_0,t_0)$ and for all $t \in \mathcal{D}$. Here, $B_\varepsilon(x_0,t_0)$ denotes a ball of radius ε centered around point (x,t).

The Lipschitz condition ensures that a function is "sufficiently" bounded (formally, it is a locally bounded function). We have the following relationships between different concepts of continuity, some of which you have seen in the micro part of math camp.

Lemma 1.4 (Concepts of Continuity). *If a function $f : \mathcal{X} \to \mathcal{Y}$ is Lipschitz continuous then it is locally Lipschitz continuous and uniformly continuous. If f is locally Lipschitz continuous then it is continuous. If f is locally Lipschitz continuous on a set E that is compact then it is Lipschitz continuous on E. If f is continuously differentiable then it is locally Lipschitz continuous but not necessarily Lipschitz continuous.*

We now state a synthesis of key theoretical results on existence and uniqueness for solutions to ODE systems subject to a boundary condition of the form $x(t_0) = x_0$.

Theorem 1.8 (Existence and Uniqueness of Solutions to ODE Boundary Value Problems). *Let $\mathcal{X}$ and $\mathcal{Y}$ be two normed vector spaces. Consider the ODE system*

$$\dot{x}(t) = g(x(t),t) \tag{1.122}$$

where $g : \mathcal{X} \times \mathcal{T} \to \mathcal{X}$, $\mathcal{X} \subseteq \mathbb{R}^n$ and $\mathcal{T} \subseteq \mathbb{R}$.

(i) Assume that g is locally Lipschitz continuous in x on $E = \mathcal{X} \times \mathcal{D}$, where $\mathcal{D} \subseteq \mathcal{T}$, and continuous in t. Then there exists $\delta > 0$ such that the boundary value problem defined by (1.122) *with $x(t_0) = x_0 \in \mathcal{X}$, $t_0 \in \mathcal{D}$, has a unique solution $x(t)$ over the interval $[t_0 - \delta, t_0 + \delta] \subset \mathcal{D}$.*

(ii) Assume that g is Lipschitz continuous in x on $E = \mathcal{X} \times \mathcal{D}$, where $\mathcal{D} \subseteq \mathcal{T}$, continuous in t, and $\mathcal{D}$ is a (bounded or unbounded) open interval. Then the boundary value problem defined by (1.122) *with $x(t_0) = x_0 \in \mathcal{X}$, $t_0 \in \mathcal{D}$, has a unique solution $x(t)$ over the entire interval $\mathcal{D}$.*

(iii) Assume that $\mathcal{X} = \mathbb{R}^n$, that g satisfies the assumptions of part (i), and that a solution $x(t)$ to the boundary value problem defined by (1.122) *with $x(t_0) = x_0 \in \mathcal{X}$, $t_0 \in \mathcal{D}$, fails to exist for at least some $t \in \mathbb{R}$. Then there exists a time t^* and a sequence $\{t_i\}$ converging to t^* such that for each $t_i \in \mathbb{R}$ the solution exists at t_i and $\|x(t_i)\| \to \infty$. That is, the solution "explodes" to infinity in finite time.*

Part (i) of Theorem 1.8 implies, in particular, that the existence of solutions when g is a continuously differentiable function can be guaranteed locally around the boundary point since continuously differentiable functions are locally Lipschitz continuous, in light of Lemma 1.4. However, the stronger result of global existence and uniqueness requires that g is globally Lipschitz continuous (on the potentially unbounded strip $\mathcal{X} \times \mathcal{D}$). Many continuously differentiable functions that appear to be "well-behaved" may in fact fail to be Lipschitz continuous. As an example consider the boundary value problem

$$\dot{x}(t) = x(t)^2 \tag{1.123}$$

subject to $x(0) = b > 0$. The solution to this problem has the form

$$x(t) = \frac{b}{1 - bt} \tag{1.124}$$

Because the denominator vanishes for $t = \frac{1}{b}$, a solution to the boundary value problem exists only for $t < \frac{1}{b}$. Note that x^2, despite being continuously differentiable, is not Lipschitz continuous on the (unbounded) strip $\mathcal{X} = \mathbb{R}$:

$$|x^2 - y^2| = |x + y| \cdot |x - y| \tag{1.125}$$

and $x + y$ can be arbitrarily large (so there cannot exist a constant $L < \infty$ such that $x + y \leq L$ for all x and y in $\mathbb{R}$ as Definition 1.2 requires).

Part (iii) of Theorem 1.8 implies that the *only* way for a solution to (1.122), with $g(\cdot,\cdot)$ continuously differentiable, to fail to exist is for the solution to "blow up" in finite time, that is, to diverge to infinity as it approaches a finite point in time. Technically, one other (rather uninteresting) way for solutions to fail to exist is possible in the case where the domain $\mathcal{X}$ of g is a strict subset of $\mathbb{R}^n$ so that a solution becomes undefined as $x(t)$ reaches the boundary of the domain of definition of g.

The proof of (parts of) Theorem 1.8 uses the Picard method of successive approximations, an elegant application of the (powerful) contraction mapping theorem, which you will cover in the first part of the macro sequence. In fact, it has been shown that local existence of solutions for boundary value problems is guaranteed when g is merely continuous. However, Lipschitz continuity is indispensable for ensuring uniqueness. Lastly, note that a limitation of Theorem 1.8 is that it only addresses boundary value problems of the form $x(t_0) = x_0$ for some $t_0 \in \mathbb{R}$. An important class of boundary value problems encountered in economic applications involves a transversality condition as part of its boundary conditions, so that the results of Theorem 1.8 do not apply directly to these problems.

Finally, the following theorem states that the solution to a boundary value problem inherits the smoothness properties of function g with respect to the independent variable (t) and also with respect to parameters, including the parameters defining the boundary condition of the problem. We will make use of this result in Section 2.4 on comparative dynamics.

Theorem 1.9 (Smoothness of Solutions to ODE systems). *Consider the ODE system*

$$\dot{x}(t) = g(x(t), t; \alpha, t_0, x_0) \tag{1.126}$$

where $g : \mathcal{X} \times \mathcal{T} \times \Omega \to \mathcal{X}$, $\mathcal{X} \subseteq \mathbb{R}^n$, $\mathcal{T} \subseteq \mathbb{R}$ and Ω is an open subset of $\mathbb{R}^p$ (the set *of possible values of the parameters in* α *as well as the boundary time and boundary value). Assume that the boundary value problem defined by* (1.126) *with* $x(t_0) = x_0 \in \mathcal{X}$, $t_0 \in \mathcal{D}$, *has a unique solution* $x(t)$ *on* $\mathcal{D}$, *which is an open subset of* $\mathcal{T}$.

If function g *is* $\mathcal{C}^k$ *(k-times continuously differentiable),* $k \in \{0 \cup \mathbb{N}\}$, *on* $\mathcal{X} \times \mathcal{D} \times \Omega$ *then the solution* $x(t) = x(t; \alpha, t_0, x_0)$ *of* (1.126) *is a* $\mathcal{C}^k$ *function on* $\mathcal{D} \times \Omega$.

For a proof of Theorems 1.8 and 1.9 and additional results for existence and uniqueness under weaker conditions than those stated here, see chapters 2 and 3 of Walter (1998).

Problem Set 1

1. The Arrow-Pratt measure of relative risk aversion of a twice differentiable utility function $u(\cdot)$ is given by

$$\mathcal{R}_u(c) = -\frac{u''(c)c}{u'(c)}$$

Assume that $c > 0$ and $u'(\cdot) > 0$. Solve for the family of utility functions with a constant coefficient of relative risk aversion (CRRA), $\mathcal{R}_u(c) = \gamma > 0$. Hints: Consider the substitution $v(c) = u'(c)$. Take special care of the case $\gamma = 1$.

2. (Second-Order ODEs) This exercise asks you to solve two kinds of second-order differential equations that will be useful in the first quarter of the macro sequence.

   (a) Derive the general solution to the differential equation

   $$\frac{1}{2}b^2x''(t) + ax'(t) - \rho x(t) = 0$$

   where $\rho > 0$, by transforming it into a 2×2 system with $y_1(t) = x(t)$ and $y_2(t) = x'(t)$. Hint: For a quick calculation of the characteristic polynomial, use Lemma 1.2. Then, use Theorem 1.2. Explain why we do not need to calculate the eigenvectors of the transformed system.

   Note that an alternative way to solve the homogeneous equation of the first part is the following guess-and-verify method: consider the candidate (elementary) solution $x(t) = \exp(\lambda t)$; compute the differential equation at the candidate solution; the resulting equation is a quadratic equation in λ that coincides with the characteristic equation computed in the first part.

   (b) Derive the general solution to the related nonhomogeneous differential equation

   $$\frac{1}{2}b^2x''(t) + ax'(t) - \rho x(t) + K = 0$$

   where $\rho > 0$, and $K > 0$ is a constant. Hint: Use Theorem 1.3.

   (c) Finally, consider the following linear but nonautonomous second-order differential equation:

   $$\frac{1}{2}b^2t^2x''(t) + atx'(t) - \rho x(t) = 0$$

   where $\rho > 0$.

   Recall from the discussion of Theorem 1.5 that we do not have results for explicit solutions of *nonautonomous* linear systems. Instead, Theorem 1.5 requires that we already have a fundamental set of solutions, $Y(t) =$

$\begin{bmatrix} y^1(t) & y^2(t) \end{bmatrix}$ so that the general solution to our system can be writen as $\tilde{y}(t) = Y(t)c$. Thus, the guess-and-verify method is our only option in this case.

Solve the equation by using as a guess the solution $x(t) = t^{\gamma}$ for some constant γ.

Chapter 2

Continuous Dynamical Systems: Stability Analysis

In this chapter we discuss stability analysis for deterministic ODE systems.

2.1 Concepts of Stability

In dynamic macroeconomic models, we are often interested in the behavior of the economy as $t \to \infty$. If the system underlying the model converges to a particular point as $t \to \infty$, we can think of this point as the long-run equilibrium of the model and we can then examine the dynamics of the economy's convergence to this long-run equilibrium.

Definition 2.1 (Steady sate)**.** A *steady state* (or stationary state or fixed point or rest point or equilibrium point) of a system $\dot{x}(t) = g(x(t))$, $t \in \mathcal{T}$, is a constant solution of the system. A point $\overline{x}$ is a steady state if it satisfies $\dot{\overline{x}}(t) = g(\overline{x}) = 0$, for all $t \in \mathcal{T}$, that is, if $\overline{x}$ is a zero of $g(\cdot)$.

Note that we have limited attention to autonomous systems, since nonautonomous systems generically do not have equilibrium points.[1]

Given the existence of a steady state $\overline{x}$, we are often interested in whether or not it is stable. That is, imagine that the economy is initially at rest at its steady state and suffers a shock that causes a deviation from the steady state. Will the economy return to its steady state, remain close to it, or move farther and farther away over time?

[1]Since we restrict our attention to autonomous systems, if a system is stable in the sense of any of the definitions below, then it is *uniformly* stable in the corresponding sense, that is, the constants δ and ϵ in the definitions below are independent of time. Thus, we state the definitions in their uniform version.

Definition 2.2 (Stability)**.** Let $\overline{x}$ be an (isolated) steady state of the system $\dot{x}(t) = g(x(t))$, $x(t) \in X$ and $t \in \mathcal{T}$. $\overline{x}$ is a (Lyapunov) stable steady state of the system, if given any $\varepsilon > 0$, there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\|x(t_0) - \overline{x}\| < \delta \text{ for any } t_0 \in \mathcal{T} \text{ implies that } \|x(t) - \overline{x}\| < \varepsilon \; \forall t \geq t_0$$

Note that Lyapunov stability is a local definition. The constant δ in the definition may be arbitrarily small, so that the system must start arbitrarily close to the steady state in order for it to remain within ϵ of the steady state forever. An equilibrium that is not stable is *unstable*. That is, there exists some $\varepsilon > 0$ and some solution of the system that, while passing arbitrarily close to the steady state, does not remain within the ball of radius ε centered at $\overline{x}$.

Note that this basic concept of stability does *not* imply that that the system always converges to the steady state. It may be simply circling around the steady state forever without getting any closer to it. In most economic applications, we are interested in a stronger concept of stability, asymptotic stability, which does imply convergence to the steady state.

Definition 2.3 (Asymptotic Stability)**.** A steady state $\overline{x}$ of the system $\dot{x}(t) = g(x(t))$, $x(t) \in X$ and $t \in \mathcal{T}$, is *globally asymptotically stable* if it is (Lyapunov) stable and, moreover, if for every $t_0 \in \mathcal{T}$ and $x(t_0) \in X$,

$$\|x(t) - \overline{x}\| \to 0 \quad \text{as } t \to \infty$$

A steady state is *locally asymptotically stable* if it is (Lyapunov) stable and, moreover, there exists some $\delta > 0$ such that

$$\|x(t_0) - \overline{x}\| < \delta \text{ for any } t_0 \in \mathcal{T} \text{ implies that } \|x(t) - \overline{x}\| \to 0 \quad \text{as } t \to \infty$$

Local asymptotic stability implies that if the system starts at a point arbitrarily close to the steady state it will converge to the steady state. Global asymptotic stability implies that, starting from any point in the state space, the system will always converge to the steady state.

Besides asymptotic stability, there is another (weaker) notion of stability, *saddle-path stability*, which is central to many growth models. Saddle-path stability is defined and discussed in Subsection 2.2.2.

Note that for linear systems, which are the focus of the following section, local asymptotic stability is equivalent to global asymptotic stability. This is not true for nonlinear systems.

2.2 Stability Analysis in Autonomous Linear Systems

2.2.1 Asymptotic Stability

Consider once again the differential equation

$$\dot{x}(t) = ax(t) + b \tag{2.1}$$

which has the general solution

$$x(t) = -\frac{b}{a} + c\exp(at) \tag{2.2}$$

It is clear that $x(t)$ will converge to its steady state $\overline{x} = -\frac{b}{a}$ if and only if $a < 0$. Given Theorem 1.3, this result is easily generalized to the case of an autonomous linear system of arbitrary (finite) dimension.

Theorem 2.1 (Asymptotic Stability of Autonomous Linear ODE Systems)**.** *Consider the autonomous linear differential equation system*

$$\dot{x}(t) = Ax(t) + B \tag{2.3}$$

with initial value $x(0) = x_0$, where $x(t) \in \mathbb{R}^n$, A is an $n \times n$ matrix, and B is an $n \times 1$ column vector. Suppose that all eigenvalues of A have negative real parts. Let $\overline{x}$ be the steady state of the system, given by $\overline{x} = -A^{-1}B$. Then, the steady state $\overline{x}$ is (globally) asymptotically stable.

2.2.2 Saddle-path Stability

As with most things in life, too much stability is undesirable. In the case of macroeconomic models, asymptotic stability in systems with dimension $n > 1$ is associated with a multiplicity (nondeterminacy) of equilibria.

Consider the setting of Theorem 2.1 and assume for simplicity that matrix A has n distinct, nonzero eigenvalues. The case with repeated or complex eigenvalues can be handled similarly (see the related discussion in Section 1.3.1). If, as in Theorem 2.1, the eigenvalues of A are all negative, so that the system is (globally) asymptotically stable, steady state $\overline{x}$ is called a *sink*. If all eigenvalues of A are positive, $\overline{x}$ is called a *source*. In that case the system can be thought of as "completely unstable," since all trajectories explode when they start from an initial position other than the steady state itself.

The interesting case occurs when some of the eigenvalues of A are positive and some are negative. In this case, the steady state is a *saddle point* (formal definition below). A saddle point is a (Lyapunov) unstable equilibrium, since there exist solution trajectories that, starting arbitrarily close to the steady state, get arbitrarily far from it as time passes. However, there exists a region of the state space such that any trajectory that starts in the region (or is found in this region at some point in time) converges to the steady state and, moreover, remains inside this region while converging to the steady state.

Consider the homogeneous system associated with matrix A. This system has steady state $\overline{x} = 0$. Partition the eigenvalues of A, $\{\lambda_i;\, i = 1,\dots,n\}$ into two sets S (for "stable") and U (for "unstable"), with $i \in S$ if $\lambda_i < 0$ and $i \in U$ if $\lambda_i > 0$. We can then write the general solution of $\dot{x} = Ax$ as

$$x(t) = \sum_{i \in S} c_i \exp(\lambda_i t) v_{\lambda_i} + \sum_{i \in U} c_i \exp(\lambda_i t) v_{\lambda_i} \tag{2.4}$$

where c_i are the constants of integration.

Now note that if we set $c_i = 0$ for all $i \in U$, then the solution of the system will converge to the steady state for any value of the constants c_i for $i \in S$. Setting some constants to zero is equivalent to choosing a subset of the state space. We now show that this subspace is in fact the linear (vector) subspace of the state space (say, $\mathbb{R}^n$) that is spanned by the eigenvectors associated with the stable eigenvalues.

Because the n linearly independent eigenvectors of A, $\mathbf{v} = \{v_{\lambda_i},\, i = 1,\dots,n\}$, form a basis for $\mathbb{R}^n$ (the eigenbasis of $\mathbb{R}^n$), any point x_0 in the state space can be written as $x_0 = \sum_{i=1}^n b_i v_{\lambda_i}$ where $\{b_i\}$ are constants not all equal to zero. Also take this vector x_0 to be the initial point of the system and note from the general solution (2.4) that

$$x_0 = x(0) = \sum_{i=1}^{n} c_i \exp(\lambda_i 0) v_{\lambda_i} = \sum_{i=1}^{n} c_i v_{\lambda_i} \tag{2.5}$$

But then we have $x_0 - x_0 = \sum_{i=1}^n (c_i - b_i) v_{\lambda_i} = 0$. The linear independence of the eigenvectors in $\mathbf{v}$ imply that $c_i = b_i$ for all $i = 1,\dots,n$. We conclude that the vector $c = \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix}^T$ of the constants of integration constitutes the representation of the system's initial position under the eigenbasis. That is, the c_i's correspond to the coordinates of the system's initial position in the coordinate system defined by the eigenbasis $\mathbf{v}$.

Then, the restriction $c_i = 0$ for $i \in U$ that defines the stable subspace of the system implies that the stable subspace is simply the subspace spanned by the set of the stable eigenvectors, $\mathbf{v}_S = \{v_{\lambda_i},\, i \in S\}$. That is,

$$W^S(0) = \{x \in \mathbb{R}^n \,:\, \exists\{\beta_i\}_{i \in S} \in \mathbb{R} \text{ such that } x = \sum_{i \in S} \beta_i v_{\lambda_i}\} \tag{2.6}$$

is the stable subspace of the homogeneous system associated with matrix A (the argument of W^S refers to the steady state of the system, which is 0 for the homogeneous system).

Similarly, we define the unstable subspace of the homogeneous system as

$$W^U(0) = \{x \in \mathbb{R}^n \,:\, \exists\{\beta_i\}_{i \in U} \in \mathbb{R} \text{ such that } x = \sum_{i \in U} \beta_i v_{\lambda_i}\} \tag{2.7}$$

Our construction also implies that any trajectory that starts within the stable subspace will remain inside the stable subspace: if x_0 is a linear combination of the stable eigenvectors, then $x(t)$ for $t \geq 0$ is also a linear combination of the stable eigenvectors.

Finally, consider once again our nonhomogeneous system $\dot{x} = Ax+B$. It is straightforward to show that the stable and unstable subspaces of this system, $W^S(\overline{x})$ and $W^U(\overline{x})$, are simply translations of the stable and unstable subspaces for the corresponding homogeneous system going through the non-zero steady state $\overline{x}$. That is, $W^S(\overline{x})$ and $W^U(\overline{x})$ are the affine (rather than linear) subspaces of $\mathbb{R}^n$ defined by

$$W^S(\overline{x}) = \overline{x} + W^S(0) \quad \text{and} \quad W^U(\overline{x}) = \overline{x} + W^U(0) \tag{2.8}$$

We now summarize our results.

Definition 2.4 (Saddle-Path Stability)**.** A steady state $\overline{x}$ of an autonomous linear system is called *saddle-path stable* or a *saddle point* if the stable and unstable subspaces $W^S(\overline{x})$ and $W^U(\overline{x})$, defined in equations (2.6)-(2.8), each have dimension greater than 0.

Theorem 2.2 (Saddle-Path Stability in Autonomous Linear ODE Systems)**.** *Consider the autonomous linear ODE system*

$$\dot{x}(t) = Ax(t) + B \tag{2.9}$$

where $x(t) \in \mathbb{R}^n$, A is an $n \times n$ matrix, and B is an $n \times 1$ column vector. Let $\overline{x}$ be the steady state of the system, given by $A\overline{x}+B = 0$. Suppose that $m < n$, $m > 0$ eigenvalues of A have negative real parts. Then, there exists an m-dimensional subspace $W^S(\overline{x})$ of $\mathbb{R}^n$, defined in equations (2.6) *and* (2.8)*, such that, starting from any $x(0) \in W^S(\overline{x})$,* (2.9) *has a unique solution, which satisfies $x(t) \in W^S(\overline{x})$ for all $t \geq 0$ and $x(t) \to \overline{x}$.*

In a two-dimensional system with a saddle point, the one-dimensional stable subspace (which is a line) is also called the *saddle path* of the system.

Note that $W^S(\overline{x})$ and $W^U(\overline{x})$ are lower-dimensional subspaces, so that a "generic" point will not be in either region. Of course, since such a point has a non-zero weight on the unstable eigenvectors, we know that a system starting from such a point will explode, just like it does if it starts in the unstable subspace. Then, what is the significance of the unstable subspace (or "anti-saddle path")? If at some point in time the system lies in the unstable subspace, the solution will "converge" to the steady state as $t \to -\infty$. In initial value problems of the form we usually consider, the unstable subspace has no significance.

2.2.3 The Dynamics of Planar Systems

Now consider the special and important case of a 2×2, or planar, system. Recall from Lemma 1.2 that 2×2 systems satisfy

$$tr(A) = \lambda_1 + \lambda_2 \tag{2.10}$$
$$det(A) = \lambda_1 \lambda_2 \tag{2.11}$$

where λ_1 and λ_2 are the eigenvalues, which solve the characteristic quadratic equation

$$p_A(\lambda) = \lambda^2 - tr(A)\lambda + det(A) = 0 \tag{2.12}$$

$$\Rightarrow \lambda_{1,2} = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2} \tag{2.13}$$

These eigenvalues always exist but they may be complex or repeated (that is, $\lambda_1 = \lambda_2$). Clearly, the sign of the discriminant $\Delta = tr(A)^2 - 4det(A)$ determines whether they are real ($\Delta > 0$), real repeated ($\Delta = 0$) or complex ($\Delta < 0$).

Just by looking at the determinant and trace of matrix A we can determine the nature and stability of the system's steady state. Figure 2.1 is a rich summary of all the possibilities (including knife-edge cases), plotting three different types of two-dimensional plots: stability behavior is classified in the $(tr(A), \det(A))$ plane shown in black; phase diagrams for each distinct case are shown in orange; the eigenvalues for each case are plotted in the complex plane in blue.

We have the following main cases (omitting edge cases):

1. If $\det(A) = \lambda_1\lambda_2 < 0$, the eigenvalues of the system are real numbers of opposite signs; hence, we have a *saddle point*. Note that the slope of the saddle path (stable subspace), which is a straight line for linear systems, is determined by the eigenvector associated with the negative (stable) root.

2. If $\det(A) = \lambda_1\lambda_2 > 0$, the roots are either complex numbers or real numbers of the same sign. In this case, there are two possibilities

   (a) If $tr(A) = \lambda_1 + \lambda_2 < 0$, the two eigenvalues are negative (if real) or have negative real parts; in either case, the system is stable. The steady state is a *sink*.

   (b) If $tr(A) = \lambda_1 + \lambda_2 > 0$, both roots are positive (if real) or have positive real parts; in both cases the system is unstable. The steady state is a *source*.

2.3 Stability Analysis in Autonomous Nonlinear Systems

Many dynamical systems appearing in economic applications are nonlinear, so the stability results of the previous sections do not directly apply. It turns out, however, that, if a certain condition holds, stability in a neighborhood of the steady state can be studied using exactly the same techniques as for linear systems examined in the previous sections, through a linearization of the nonlinear model around its steady state. Unfortunately, global and local stability do not coincide for nonlinear models, so little can be said about the dynamics far from the steady state without analytically (or, sometimes, graphically) solving the nonlinear model.

We first analytically consider and prove the one-dimensional case so as to build intuition for the results.

Lemma 2.1 (Local Stability by Linearization)**.** *Consider the differential equation* $\dot{x}(t) = f(x(t))$*, where* f *is continuously differentiable. Let* $\overline{x}$ *be a steady state of the equation that satisfies* $f'(\overline{x}) \neq 0$*. Then* $\overline{x}$ *is locally asymptotically stable if* $f'(\overline{x}) < 0$ *and is locally asymptotically unstable if* $f'(\overline{x}) > 0$*.*

Proof. We first write the equation in terms of deviation from the steady state (note that $\dot{\tilde{x}} = (\dot{x - \overline{x}}) = \dot{x}$):

$$\dot{\tilde{x}} = f(\overline{x} + \tilde{x}) \tag{2.14}$$

Let $\phi(\tilde{x})$ be the error committed when we take a first-order Taylor approximation of $f(\cdot)$ around the steady state, that is,

$$\phi(\tilde{x}) = f(\overline{x} + \tilde{x}) - f'(\overline{x})\tilde{x} \tag{2.15}$$

and note that $\phi(0) = 0$. Differentiating this equation with respect to the deviation $\tilde{x}$ we get

$$\phi'(\tilde{x}) = f'(\overline{x} + \tilde{x}) - f'(\overline{x}) \tag{2.16}$$

where $\phi'(0) = 0$.

Note that $\phi(\cdot)$ inherits the continuous differentiability of $f(\cdot)$. This implies the following: Fix some $\varepsilon > 0$ such that $\varepsilon < |f'(\overline{x})|$. Then, by the continuous differentiability of $f(\cdot)$ (that is, by the fact that $f'(\cdot)$ is continuous), there exists some $\delta > 0$ such that $|\phi'(\hat{x})| = |f'(\overline{x} + \hat{x}) - f'(\overline{x})| < \varepsilon$ for all $\hat{x}$ such that $|\hat{x}| < \delta$.[2]

From the fundamental theorem of calculus, we have that

$$\begin{aligned}
\phi(\tilde{x}) &= \phi(0) + \int_0^{\tilde{x}} \phi'(\hat{x})\,\mathrm{d}\hat{x} && (2.17)\\
&= \int_0^{\tilde{x}} \phi'(\hat{x})\,\mathrm{d}\hat{x} && (2.18)
\end{aligned}$$

Then, for all $\tilde{x}$ with $|\tilde{x}| < \delta$,

$$\begin{aligned}
|\phi(\tilde{x})| &= |\int_0^{\tilde{x}} \phi'(\hat{x})\,\mathrm{d}\hat{x}| && (2.19)\\
&\leq \int_0^{\tilde{x}} |\phi'(\hat{x})|\,\mathrm{d}\hat{x} && (2.20)\\
&\leq \varepsilon|\tilde{x}| && (2.21)\\
&< |f'(\overline{x})\tilde{x}| && (2.22)
\end{aligned}$$

[2]This statement is a direct application of the definition of continuity.

where the first line follows from (2.18) and the second line follows from Holder's inequality. Finally, from (2.14) and (2.15), we have

$$\dot{x} = \dot{\tilde{x}} = f'(\overline{x})\tilde{x} + \phi(\tilde{x}) \tag{2.23}$$

so that (2.22) implies that for small enough deviations, $|\tilde{x}| < \delta$, the sign of $\dot{x}$ is fully determined by the sign of $f'(\overline{x})\tilde{x}$.

Clearly, in order for steady state $\overline{x}$ to be asymptotically stable, we want the system to approach it from both sides. That is, when $x < \overline{x} \Leftrightarrow \tilde{x} < 0$ we need $\dot{x} > 0$ and when $x > \overline{x} \Leftrightarrow \tilde{x} > 0$ we need $\dot{x} < 0$. In other words, we need $\dot{x}$ and $\tilde{x}$ to have opposite signs. For small enough deviations, we have just shown that this is equivalent to $f'(\overline{x}) < 0$. By the same argument, the steady state is locally asymptotically unstable when $f'(\overline{x}) > 0$. □

The proof of Lemma 2.1 makes clear that in order to be able to infer the local stability properties of the equation from those of its linear approximation we need $f'(\overline{x}) \neq 0$. Otherwise, no such information can be inferred. A steady state $\overline{x}$ with $f'(\overline{x}) \neq 0$ is called *hyperbolic*. Otherwise, it is called *nonhyperbolic*. The phase diagrams for our one-dimensional system in Figures 2.2 and 2.3 illustrate why a steady state needs to be hyperbolic: the *phase line*, that is, the graph of $f(x)$, needs to cross the x-axis transversally in order for the derivative to tell us if the steady state is stable or not. If the phase line crosses the x-axis at $\overline{x}$ transversally from above as we increase x, we know that at least locally steady state $\overline{x}$ is asymptotically stable.

Now consider the general case of the $n \times n$ system

$$\dot{x}(t) = G(x(t)) \tag{2.24}$$

where $G(\cdot)$ is a continuously differentiable mapping.

Definition 2.5 (Hyperbolic Steady State)**.** A steady state $\overline{x}$ of system (2.24) is *hyperbolic* if the matrix $DG(\overline{x})$ does not have eigenvalues with zero real parts.[3]

Then, as long as $\overline{x}$ is a hyperbolic steady state, the behavior of $x(t)$ in the neigborhood of this steady state can be approximately by the linear system

$$\dot{x}(t) = DG(\overline{x})(x(t) - \overline{x}) \tag{2.25}$$

We formalize the notion that systems (2.24) and (2.25) behave very "similarly" through the following definition.

Definition 2.6 (Topological equivalence for Dynamical Systems)**.** Two dynamical systems F and G are *topologically equivalent* if there exists a homeomorphism (a continuous change of coordinates) that maps F trajectories into G trajectories while preserving the sense of direction in time.

The key result of this section is the following.

[3]Note that this condition is stronger than assuming that $DG(\overline{x})$ is invertible (no zero eigenvalues), since it also precludes situations where $DG(\overline{x})$ has purely imaginary eigenvalues. Observe in Figure 2.1 that in this case the trajectories of the linearized system are circles around $\overline{x}$.

Theorem 2.3 (Grobman-Hartman Theorem)**.** *Let $\overline{x}$ be a* hyperbolic *steady state of system* (2.24)*, $\dot{x}(t) = G(x(t))$, where $G : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable mapping. Then, there exists a neighborhood U of $\overline{x}$ such that* (2.24) *is topologically equivalent in U to linear system* (2.25)*,*

$$\dot{x}(t) = DG(\overline{x})(x(t) - \overline{x})$$

A straightforward corollary of this result is the generalization of Lemma 2.1 to autonomous ODE systems.

Theorem 2.4 (Local Stability of Nonlinear Autonomous ODE Systems)**.** *Consider the autonomous nonlinear ODE system* (2.24)*,*

$$\dot{x}(t) = G(x(t))$$

where $G : \mathbb{R}^n \to \mathbb{R}^n$ is a continuously differentiable mapping. Then,

(i) If all eigenvalues of $DG(\overline{x})$ have strictly negative real parts, then $\overline{x}$ is locally asymptotically stable.

(ii) If at least one eigenvalue of $DG(\overline{x})$ has a positive real part, then $\overline{x}$ is (Lyapunov) unstable (and thus asymptotically unstable).

(iii) If at least one eigenvalue of $DG(\overline{x})$ has a zero real part and all other eigenvalues have negative real parts, then $\overline{x}$ may be (Lyapunov) stable, locally asymptotically stable, or (Lyapunov) unstable.

Finally, as we also discussed in the previous section, in economic applications featuring multidimensional systems we are usually not interested in case (i) of Theorem 2.4, since asymptotic stability is associated with a multiplicity of equilibria. Instead, we are interested in unstable systems that exhibit saddle-path stability. The next theorem, another corollary of the Grobman-Hartman Theorem, tells us that if the linear approximation of a nonlinear system around a hyperbolic steady state is saddle-path stable then, at least locally in a neighborhood of the steady state, the nonlinear system possesses a stable subspace that is tangent to the (linear) stable subspace $W^S(\overline{x})$ of the linearized system. The stable subspace of the nonlinear system is called a *manifold*, that is, a topological space that resembles a Euclidean space of the same dimension in the sense that each point of a manifold has a neighborhood that is homeomorphic to the Euclidean space of the same dimension.

Theorem 2.5 (Saddle-Path Stability in Autonomous Nonlinear ODE Systems)**.**

Consider the nonlinear autonomous ODE system (2.24),

$$\dot{x}(t) = G(x(t))$$

where $G : \mathbb{R}^n \to \mathbb{R}^n$ *is a continuously differentiable mapping. Let* $\overline{x}$ *be the steady state of the system, given by* $G(\overline{x}) = 0$. *Suppose that* $m < n$, $m > 0$ *eigenvalues of* $DG(\overline{x})$ *have strictly negative real parts and the rest have strictly positive real parts. Then, there exists an open neighborhood of* x^*, $B(\overline{x}) \subset \mathbb{R}^n$, *and an* m*-dimensional manifold* $M \subset B(\overline{x})$ *such that, starting from any* $x(0) \in M$, (2.24) *has a unique solution, which satisfies* $x(t) \in M$ *for all* $t \geq 0$ *and* $x(t) \to \overline{x}$.

Figure 2.4 shows the phase diagram of a 2×2 system, including the manifolds of the nonlinear systems and the tangent subspaces of the corresponding linearized system.

Application 2.1 (Saddle-Path Stability in the Neoclassical Growth Model)**.** Consider the boundary value problem, corresponding to the equilibrium conditions, of the baseline neoclassical (Ramsey) growth model:

$$\dot{k}(t) = f(k(t)) - \delta k(t) - c(t) \tag{2.26}$$

$$\dot{c}(t) = \frac{1}{\gamma}\left(f'(k(t)) - \delta - \rho\right) c(t) \tag{2.27}$$

subject to $k(t), c(t) \geq 0$, the initial value condition $k(0) = k_0 > 0$ and the *transversality condition*

$$\lim_{t \to \infty} \left[k(t) \exp\left(-\int_0^t \left(f'(k(s)) - \delta\right) \mathrm{d}s \right) \right] = 0 \tag{2.28}$$

where parameters satisfy $\gamma, \delta, \rho > 0$ and $\delta < 1$. We assume that the production function $f(\cdot)$ satisfies $f(0) = 0$, $f'(k) > 0$, $f''(k) < 0$ for all $k > 0$ (so it features diminishing marginal returns to capital), and the two *Inada conditions*, $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$.

Here, equation (2.26) specifies the evolution of capital, the stock variable, while equation (2.27) determines the evolution of consumption, the flow variable. This combination of an initial value condition and a transversality condition is common in economic problems involving the behavior of both state and control variables. At the end of this application, we will show that these two boundary conditions (together with the admissibility condition $k(t), c(t) \geq 0$ for all $t \geq 0$) imply that the system must always be on the saddle path. At time 0 the flow variable, which can be adjusted freely, will jump to the value that places the system on the saddle path. That is, $(k_0, c(k_0))$ is a point of the (one-dimensional) stable manifold, where we have denoted by $c(k)$ the value of consumption as a function of the contemporaneous value of capital.

We now solve for the linear approximation of the system in order to study its local stability properties around the steady state, based on Theorem 2.5.

First, the steady state (k^*, c^*) of the system is (implicitly) given by

$$\dot{k} = 0 \Rightarrow c^* = f(k^*) - \delta k^* \tag{2.29}$$

$$\dot{c} = 0 \Rightarrow f'(k^*) = \delta + \rho \tag{2.30}$$

Now, let $\tilde{c} = c - c^*$ and $\tilde{k} = k - k^*$ denote deviations from the steady state. Then, the first-order approximation of system (2.26)-(2.27) around the steady state is

$$\begin{bmatrix} \dot{\tilde{k}} \\ \dot{\tilde{c}} \end{bmatrix} \approx \begin{bmatrix} f'(k^*) - \delta & -1 \\ \frac{c^* f''(k^*)}{\gamma} & \frac{1}{\gamma}(f'(k^*) - \delta - \rho) \end{bmatrix} \begin{bmatrix} \tilde{k} \\ \tilde{c} \end{bmatrix} \tag{2.31}$$

Making use of the steady-value relations, the linearized model simplifies to

$$\begin{bmatrix} \dot{\tilde{k}} \\ \dot{\tilde{c}} \end{bmatrix} \approx \overbrace{\begin{bmatrix} \rho & -1 \\ \frac{c^* f''(k^*)}{\gamma} & 0 \end{bmatrix}}^{\equiv A} \begin{bmatrix} \tilde{k} \\ \tilde{c} \end{bmatrix} \tag{2.32}$$

Since $f''(k^*) < 0$ and $\rho > 0$ by assumption, we immediately get that $\det(A) \neq 0$ and $\text{tr}(A) \neq 0$ so that our steady state is indeed hyperbolic (none of its eigenvalues have zero real part) and we can, therefore apply Theorem 2.5.

From the fact that $\det(A) = \frac{c^* f''(k^*)}{\gamma} < 0$ and our results for planar systems in Subsection 2.2.3 we know that this system has two real eigenvalues of opposite signs. Therefore, we know from Theorem 2.5 that our system is indeed saddle-path stable, at least in a neighborhood of the steady state.

The characteristic equation is

$$p_A(\lambda) = \lambda^2 - tr(A)\lambda + det(A) = 0 \tag{2.33}$$

$$= \lambda^2 - \rho\lambda + \frac{c^* f''(k^*)}{\gamma} = 0 \tag{2.34}$$

hence we can compute the eigenvalues of the linearized system as

$$\lambda_{\pm} = \frac{1}{2}\left(\rho \pm \sqrt{\rho^2 - 4\frac{c^* f'(k^*)}{\gamma}}\right) \tag{2.35}$$

The general solution of the linearized system is

$$\begin{bmatrix} k(t) \\ c(t) \end{bmatrix} = \begin{bmatrix} k^* \\ c^* \end{bmatrix} + \alpha_1 \exp(\lambda_+ t) v_{\lambda_+} + \alpha_2 \exp(\lambda_- t) v_{\lambda_-} \tag{2.36}$$

$$= \begin{bmatrix} k^* \\ c^* \end{bmatrix} + \alpha_1 \exp(\lambda_+ t) \begin{bmatrix} 1 \\ \lambda_- \end{bmatrix} + \alpha_2 \exp(\lambda_- t) \begin{bmatrix} 1 \\ \lambda_+ \end{bmatrix} \tag{2.37}$$

Finally, let us use the assertion made above (and proved below) that the specified boundary conditions imply that the particular solution of the boundary value problem must lie on the saddle path. This means that the coefficient α_1 of the explosive root λ_+ must be set to $\alpha_1 = 0$. The other condition, $k(0) = k_0$ then pins down α_2:

$$k_0 = k(0) = k^* + \alpha_2 \Rightarrow \alpha_2 = k_0 - k^* \tag{2.38}$$

Thus, the solution to the linearized version of our boundary value problem is

$$\begin{bmatrix} k(t) \\ c(t) \end{bmatrix} = \begin{bmatrix} k^* + (k_0 - k^*)\exp(\lambda_- t) \\ c^* + (k_0 - k^*)\lambda_+ \exp(\lambda_- t) \end{bmatrix} \tag{2.39}$$

The solution tells us that, for $k(0)$ close to k^*, the gap between $k(t)$ and k^* declines approximately at an exponential rate of $\lambda_- = \frac{1}{2}\left(\rho - \sqrt{\rho^2 - 4\frac{c^* f''(k^*)}{\gamma}}\right)$. Consumption at time 0 is set approximately to

$$\begin{aligned} c(0) &\approx c^* + (k_0 - k^*)\lambda_+ && (2.40) \\ &= c^* + (k_0 - k^*)\frac{1}{2}\left(\rho + \sqrt{\rho^2 - 4\frac{c^* f''(k^*)}{\gamma}}\right) && (2.41) \end{aligned}$$

We now establish the claim we made earlier that the solution to our system under the given boundary conditions always lies on the (global) saddle path of the system, depicted in Figure 2.5, the phase diagram of the system.

First, suppose that $(\hat{k}(0), \hat{c}(0))$ is above the saddle path. An example is point $(\hat{k}_0, \hat{c}_0')$ in Figure 2.5. In this case, the arrows of motion tell us that the capital stock would reach $\hat{k}_t = 0$ and in fact would do so finite time,[4] while consumption would remain strictly positive. This implies that in the next instant after the system reaches that point on the y-axis capital would turn negative, which violates admissibility, $k(t) \geq 0$ for all $t \geq 0$, so that it can not be the solution.

Second, suppose that $(\hat{k}(0), \hat{c}(0))$ is below the saddle path, such as point $(\hat{k}_0, \hat{c}_0'')$ in Figure 2.5. Then, the arrows of motion tells us that capital will converge to $\hat{k}^{**}$. We argue that this violates the transversality condition (2.28). First note that this level of capital is to the right of point $\hat{k}_{\text{gold}}$, which maximizes consumption among all points on the concave $\dot{k}(t) = 0$ phase line, so that $f'(\hat{k}_{\text{gold}}) = \delta$. Second, the strict concavity of f (f' strictly decreasing) implies that $f'(\hat{k}^{**}) < f'(\hat{k}_{\text{gold}}) = \delta$. But then we

[4] To see this, note that if the system were to converge to the point with zero capital only asymptotically, consumption and capital should change along the trajectory at a decelerating pace. In other words, the second derivatives of consumption and capital with respect to time should be strictly negative. Differentiating equation (2.27) with respect to time, $\ddot{c}(t) = (f''(k(t))c(t)\gamma)\dot{k}(t) + (f'(k(t)) - \delta - \rho)\dot{c}(t)/\gamma$, which is strictly positive for small k (use the Inada condition $lim_{k\to 0} f'(k) = \infty$). Therefore, the system reaches $\hat{k} = 0$ in finite time.

have

$$\begin{aligned}
&\lim_{t\to\infty}\left\{\hat{k}(t)\exp\left(-\int_0^t\left[f'(\hat{k}(s))-\delta\right]\mathrm{d}s\right)\right\} \\
&=\hat{k}^{**}\lim_{t\to\infty}\left\{\exp\left(-\int_0^t\left[f'(\hat{k}^{**})-\delta\right]\mathrm{d}s\right)\right\} \\
&>0
\end{aligned} \tag{2.42}$$

which violates (2.28).

Thus, under the given boundary conditions the existence of a saddle path is equivalent to the existence of an equilibrium. The assumptions of Theorem 2.5 only ensure the existence of a saddle path (and thus of an equilibrium) *locally*, in a potentially very small neighborhood of the steady state. In order to show that an equilibrium exists for all initial values of capital $k(0)>0$, one needs to show that there exists a saddle path that includes points with all possible capital levels $k(0)>0$, that is, one needs to show that for any $k(0)>0$ there exists a consumption level $c=c(k(0))$ such that $(k(0),c(k(0)))$ is a point on the saddle path. This can be done by analytically solving the model or, informally, by drawing the (exact) phase diagram for the model. In Figure 2.5 we see that the direction of arrows suggests the existence of a saddle-path (labelled $\hat{c}(\hat{k})$ in the graph) that includes points with all possible capital levels $k(0)>0$.

We conclude that an equilibrium exists and is unique (since the saddle-path is a one-dimensional curve) in the neoclassical growth model for all $k(0)>0$. ■

2.4 Comparative Dynamics in Autonomous Systems*

Related to the idea of linearization of a nonlinear model is the application of *comparative statics* to gauge the response of the model's solution to small changes in parameters of the model. Recall from the micro part of math camp that the *implicit function theorem* enables us to perform comparative statics in models of the form

$$g(x;\alpha)=0 \tag{2.43}$$

where $g(\cdot,\cdot)$ is an n-dimensional nonlinear system of model conditions. x is usually a vector of the endogenous variables of the model, and α a vector of model parameters the impact of which on the endogenous variables we intend to gauge.

For a summary, suppose we know that $\overline{x}$ is the solution to (2.43) for a particular value of $\alpha=\alpha_{old}$. That is,

$$g(\overline{x};\alpha_{old})=0 \tag{2.44}$$

The *inverse function theorem* tells us that this solution is locally unique if the matrix of partial derivatives $D_x g(\overline{x},\alpha_{old})$ is nonsingular, where $g(\cdot,\cdot)$ is continuously differentiable in its arguments. Local uniqueness means that there exists a neighborhood of $\overline{x}$ in which there are no solutions to (2.43) other than $x = \overline{x}$. The *implicit function theorem* then tells us that, under the hypothesis that $g(\cdot,\cdot)$ is continuously differentiable, $D_{(x,\alpha)}g(\overline{x},\alpha_{old})$ finite, and $D_x g(\overline{x},\alpha_{old})$ nonsignular, there is also a locally unique, continuously differentiable solution $\overline{x}(\alpha)$ defined for all α close to α_{old}. The solution function $\overline{x}(\alpha)$ is defined implicitly through

$$g(\overline{x}(\alpha);\alpha) = 0 \tag{2.45}$$

for all α and, by totally differentiating (2.45) with respect to α, its derivative at α_{old} is found to be

$$D\overline{x}(a_{old}) = -\left[D_x g(\overline{x},\alpha_{old})\right]^{-1} D_\alpha g(\overline{x},\alpha_{old}) \tag{2.46}$$

In the case of dynamical systems, comparative statics can capture changes in the steady state values of endogenous variables x in response to small changes in parameters α. For example, consider an arbitrary nonlinear autonomous system

$$\dot{x}(t) = g(x(t),\alpha) \tag{2.47}$$

where α is a vector of model parameters and $g(\cdot,\cdot)$ is continuously differentiable. Because the steady state satisfies $g(\overline{x},\alpha) = 0$, we can apply the above procedure to obtain the steady state as a function of the parameters, $\overline{x}(\alpha)$.

Yet, for dynamical systems we are usually interested in more than how the steady state responds to changes in the structural environment. We want to study *comparative dynamics*, which refers to the response of the *entire equilibrium path* of variables to a change in policy or parameters.

To do this we can follow a similar approach as in the comparative statics case under the assumption that $g(\cdot,\cdot)$ is Lipschitz continuous in x on its entire domain and also *twice* continuously differentiable in both of its arguments.

According to Theorem 1.8 in Section 1.6, our assumption that $g(\cdot,\cdot)$ is Lipschitz continuous in x on its entire domain implies that it has a unique solution, $x(t,\alpha)$, everywhere. Then, according to Theorem 1.9, the fact that g is twice continuously differentiable in both x and α implies that the (unique) solution function $x(t,\alpha)$ is also twice continuously differentiable in its two arguments.

The solution function, $x(t,\alpha)$, is defined through

$$\dot{x}(t,\alpha) = g(x(t,\alpha),\alpha) \tag{2.48}$$

for all α, so that we can again differentiate totally with respect to α to get

$$D_\alpha \dot{x}(t,\alpha) = D_x g(x(t,\alpha),\alpha) D_\alpha x(t,\alpha) + D_\alpha g(x(t,\alpha),\alpha) \tag{2.49}$$

Now note that the second partial derivatives of $x(t,\alpha)$ are continuous, since $x(t,a)$ is twice continuously differentiable, so by Schwarz' theorem of the symmetry of second derivatives, we can reverse the order of differentiation so that

$$\begin{aligned} D_a \dot{x}(t,a) &= D_\alpha D_t x(t,\alpha) \\ &= D_t D_\alpha x(t,\alpha) \\ &= \dot{x}_\alpha(t,a) \end{aligned}$$

Thus, we can write (2.49) as

$$\dot{x}_\alpha(t,\alpha) = D_x g(x(t,\alpha),\alpha)x_\alpha(t,\alpha) + D_\alpha g(x(t,\alpha),\alpha) \tag{2.50}$$

Evaluating this equation at $\alpha = \alpha_{old}$, it becomes a *linear* differential equation for x_α, the derivative of interest. It is in general a nonautonomous differential equation since $A(t) \equiv D_x g(x(t,\alpha_{old}),\alpha_{old})$ and $B(t) \equiv D_\alpha g(x(t,\alpha_{old}),\alpha_{old})$ are evaluated along the solution trajectory (under the old parameter values) which is, in general, non-degenerate across time.

However, there is a special case that can be handled particularly easily because it turns (2.50) into an autonomous linear differential equation. This is the case where the system is initially (at $t = 0$) at the (old) steady state, $\overline{x}$. In that case, $x(t,\alpha_{old}) = \overline{x}$ for all $t \geq 0$, so that

$$A(t) = A = D_x g(\overline{x},\alpha_{old}) \tag{2.51}$$

$$B(t) = B = D_\alpha g(\overline{x},\alpha_{old}) \tag{2.52}$$

We can now use Theorem 1.3, which tells us that the general solution of (2.50) is

$$x_\alpha(t) = -[D_x g(\overline{x},\alpha_{old})]^{-1} D_\alpha g(\overline{x},\alpha_{old}) + \sum_{j=1}^{n} c_j \exp(\lambda_j t) v_{\lambda_j} \tag{2.53}$$

where $\{\lambda_i\}$ are the (distinct) eigenvalues of matrix A and and $\{v_{\lambda_j}\}$ are the corresponding eigenvectors.

First, note from equation (2.46) that the steady state of the law of motion of x_α is precisely the comparative static $D\overline{x}(a_{old})$ of the model, that is, the derivative of the steady state value with respect to parameters. Thus, the steady state of x_α can be interpreted as the long-run effect of the parameter change (assuming that the steady state of x is asymptotically stable, so that the system actually converges to the new steady state).

Next, consider the one-dimensional case. The solution for x_α can then be written as

$$x_\alpha(t) = D\overline{x}(\alpha_{old}) + [x_\alpha(0) - D\overline{x}(\alpha_{old})]\exp(g_x(\overline{x},\alpha_{old})t) \tag{2.54}$$

$$= x_\alpha(0)\exp(g_x(\overline{x},\alpha_{old})t) + D\overline{x}(\alpha_{old})(1 - \exp(g_x(\overline{x},\alpha_{old})t)) \tag{2.55}$$

which says that the impact of an (infinitesimal) parameter change on the system is a weighted average of its impact (immediate) effect, $x_\alpha(0)$, and its long-run effect, $D\overline{x}(\alpha_{old})$. Note that $x_\alpha(t) \to D\overline{x}(\alpha_{old})$ if and only if the system is stable, $g_x(\overline{x},\alpha_{old}) < 0$.

The initial conditions for system (2.50) can be pinned down by economic considerations. If a given variable (a component of x) is predetermined (a stock variable), such as capital in the neoclassical growth model, then the impact effect, $x_\alpha(0)$, will be zero. If a variable is a free (flow) variable then, depending on the economics of the problem, it may jump directly to the new steady state, $x_\alpha(0) = D\overline{x}(\alpha_{old})$, or jump to a different value. For example, if there is a shock (a parameter change) in the neoclassical growth model that changes the model's steady state then consumption will immediately jump to the saddle path (at the old steady state value of capital) corresponding to the new steady state.

Application 2.2 (Comparative Dynamics in the Solow Model)**.** As in Application 1.1, consider the key equation of the Solow growth model

$$\dot{k}(t) = sf(k(t)) - (n+\delta)k \tag{2.56}$$

where $k(t)$ is capital (the capital-labor ratio). Note that (2.56) is slightly different from (1.33) because we have also assumed population growth at the (exponential) rate n.[5] We assume that $f(\cdot)$ is twice continuously differentiable, with $f'(k) > 0$ and $f''(k) < 0$, $\forall k \geq 0$.

Our goal is to study the comparative dynamics of the economy with respect to the savings rate (which is exogenous in the Solow model). We can write (2.56) as

$$\dot{k}(t) = g(k,s) \qquad \text{where} \qquad g(k,s) \equiv sf(k) - (n+\delta)k(t) \tag{2.57}$$

and note that $g(\cdot,\cdot)$ is twice continuously differentiable in its arguments because $f(k)$ is twice-continuously differentiable, so that that we can apply the comparative dynamics methods described earlier in the section. Let $k(t,s)$ denote the solution of (2.56) expressed as a function of time and the savings rate. We know from Theorem 1.9 that $k(\cdot,\cdot)$ is twice-continuously differentiable.

The derivative function of interest, $k_s(t,s)$, gives us the effect on the entire trajectory of capital of an infinitesimal increase in the savings rate. From the derivation earlier in the section, $k_s(t,s)$ satisfies the following linear differential equation:

$$\dot{k}_s(t,s) = [s_{old}f'(k(t,s_{old})) - (n+\delta)]k_s(t,s) + f(k(t,s_{old})) \tag{2.58}$$

To get closed-form results, we will assume that the system is initially at the old steady state $\overline{k} = \overline{k}(s_{old})$, so that k_s is now described by the following autonomous linear differential equation:

$$\dot{k}_s(t,s) = \left[s_{old}f'(\overline{k}) - (n+\delta)\right]k_s(t,s) + f(\overline{k}) \tag{2.59}$$

We focus on the relevant case where the steady state $\overline{k}$ is asymptotically stable, which requires that $g_k(\overline{k}, s_{old}) = s_{old}f'(\overline{k}) - (n+\delta) < 0$. Then, we can compute the solution to (2.58) using Theorem 1.3 as

$$\begin{aligned} k_s(t,s_{old}) &= \overline{k}_s(s_{old}) + \left[k_s(0) - \overline{k}_s(s_{old})\right]\exp(\sigma t) && (2.60) \\ &= k_s(0)\exp(\sigma t) + \overline{k}_s(s_{old})(1 - \exp(\sigma t)) && (2.61) \end{aligned}$$

[5] Because $k(t)$ is the capital-labor ratio, n enters in exactly the same way as the depreciation rate δ. Including population growth is not essential for the results of this section.

where

$$\sigma \equiv g_k(\bar{k}, s_{old}) = s_{old} f'(\bar{k}) - (n+\delta) < 0 \tag{2.62}$$

and

$$\bar{k}_s(s_{old}) = -\frac{g_s(\bar{k}, s_{old})}{g_k(\bar{k}, s_{old})} \tag{2.63}$$

$$= -\frac{f(\bar{k})}{s_{old} f'(\bar{k}) - (n+\delta)} > 0 \tag{2.64}$$

Note that because capital is a stock variable it cannot "jump", so that the impact effect of a change in the savings rate on capital is zero, $k_s(0) = 0$. Therefore, $k_s(t)$ grows over time at a (declining) rate of $\frac{\dot{k}_s}{k_s} = \frac{-\sigma}{\exp(-\sigma t)-1}$ and converges asymptotically to the value given by the steady-state impact of the savings rate increase.[6]

The impact of a (discrete) increase in the savings rate from $s_1 = s_{old}$ to s_2, starting from the old steady state and giving rise to a transition towards the new steady state, is illustrated in Figure 2.6. In this figure, $k_1^* = \bar{k}(s_1)$ and, for a small increase in the saving rate, $k(t) - k_1^* = k(t, s_2) - \bar{k}(s_1)$ is approximately equal to $k_s(t, s_1)(s_2 - s_1)$. ■

[6] Alternatively, we can say that the gap between the effect at time t, $k_s(t)$, and the steady-state effect declines at a rate of $|\sigma|$).

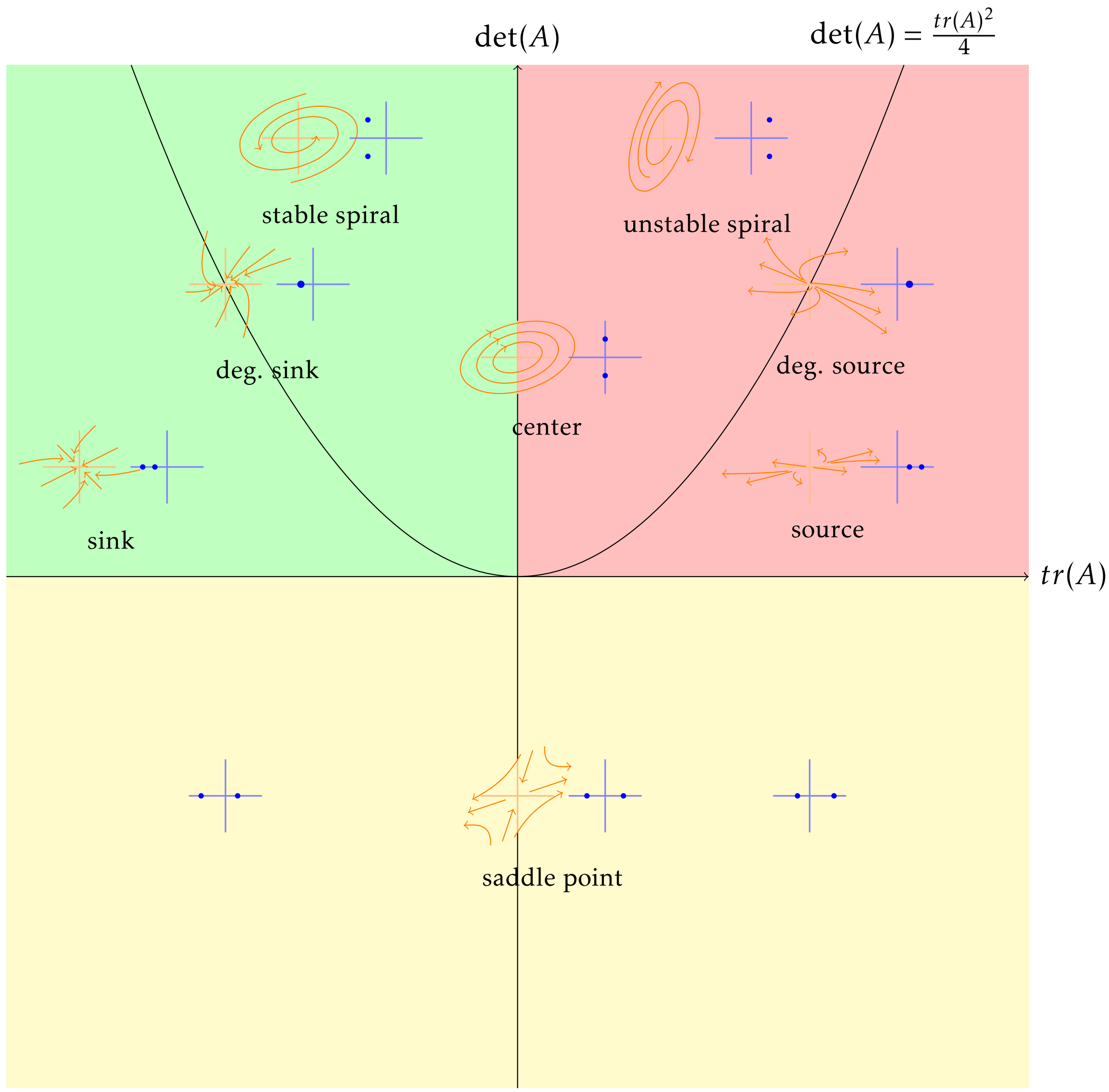


Figure 2.1: Stability map for autonomous linear planar ODE systems. Adapted from source code by Dan Drake (KAIST).

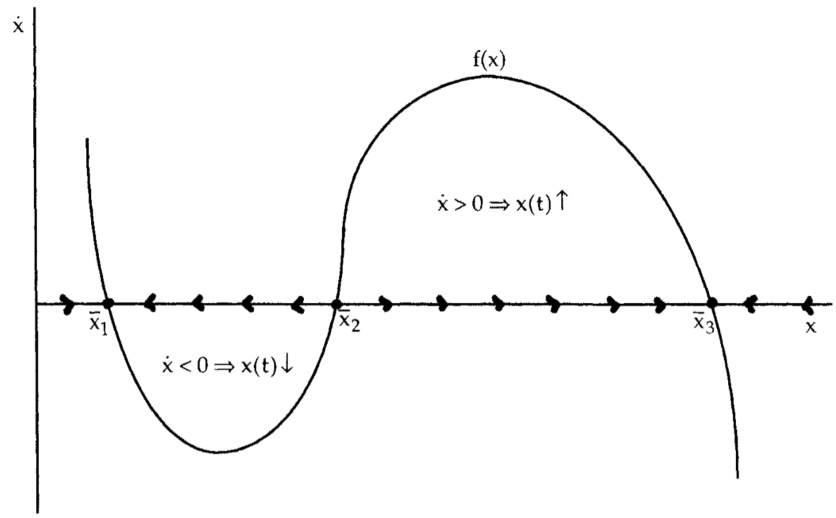


Figure 2.2: Phase diagram for the nonlinear equation $\dot{x}(t) = f(x(t))$. All steady states are hyperbolic. Source: de la Fuente (2000).

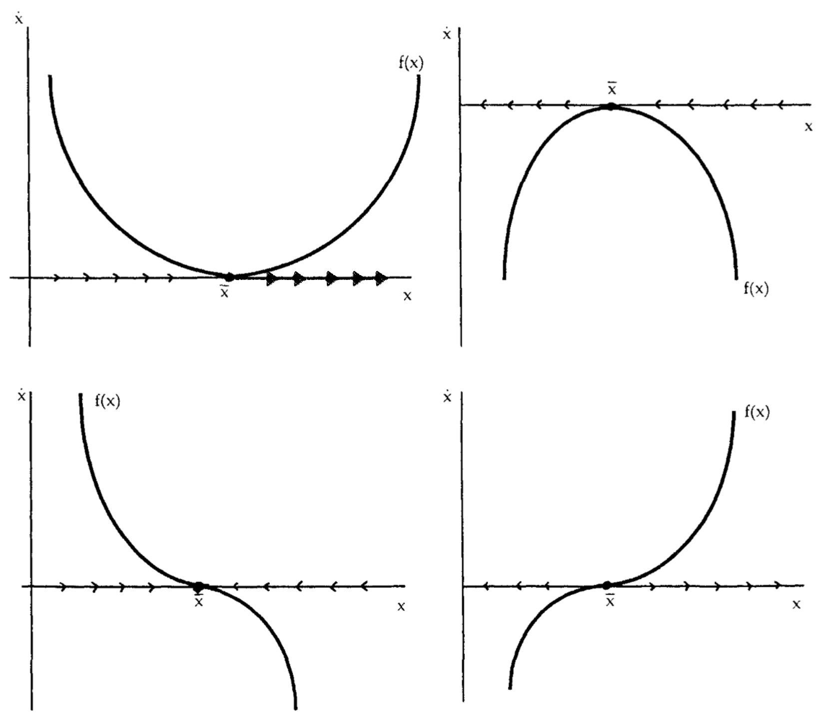


Figure 2.3: Phase diagram for the nonlinear equation $\dot{x}(t) = f(x(t))$. Examples of nonhyperbolic steady states. Source: de la Fuente (2000).

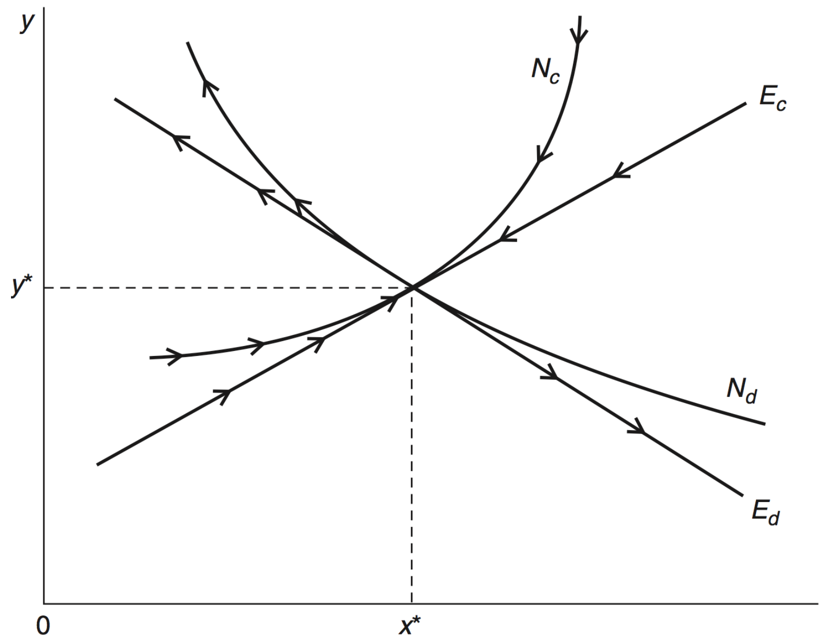


Figure 2.4: Phase diagram of a nonlinear ODE system with its convergent and divergent manifolds, N_c and N_d, and the corresponding linear subspaces, E_c and E_d, of the linearized system. Source: Acemoglu (2009).

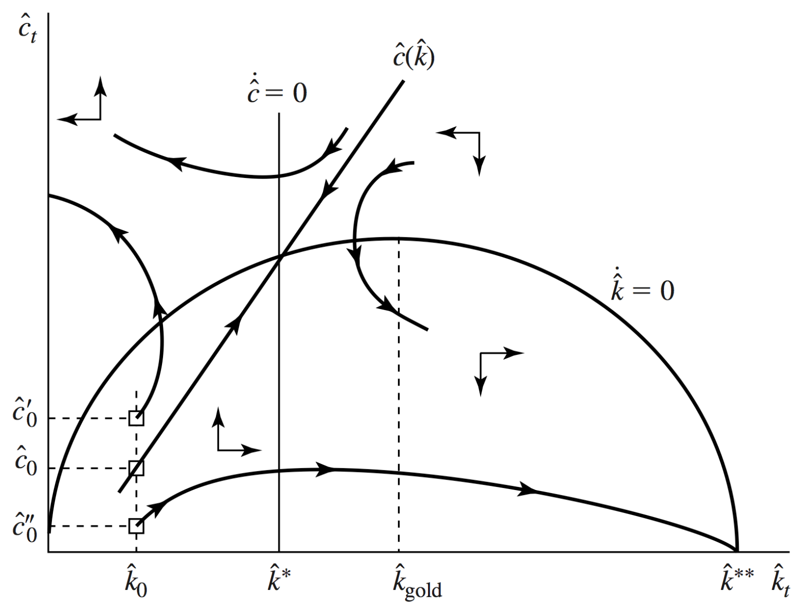


Figure 2.5: Transitional Dynamics in the neoclassical growth model. Source: Barro and Sala-i Martin (2004).

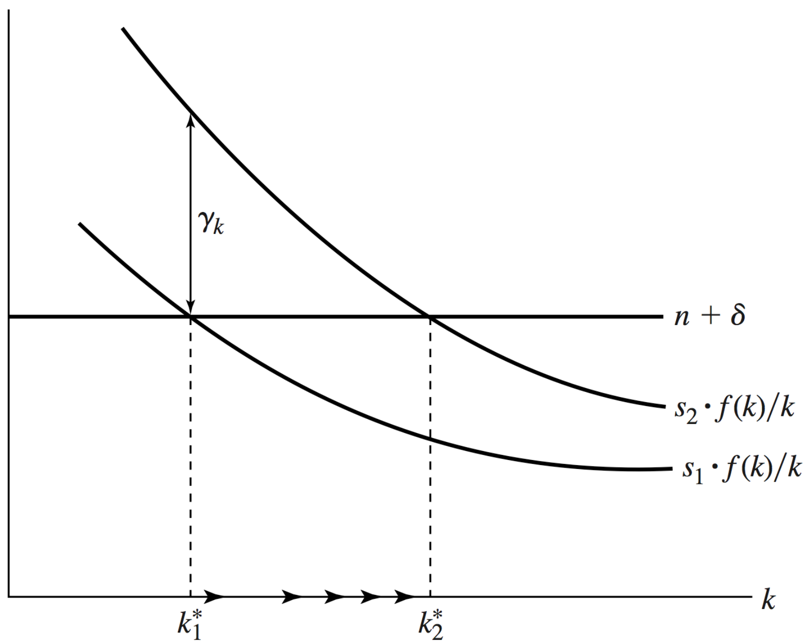


Figure 2.6: Comparative dynamics in the Solow model. The effect of an increase in the savings rate from s_1 to s_2, starting from the old steady state $k_1^* = \overline{k}(s_1)$. $\gamma_k \equiv \dot{k}/k = s_2 f(k)/k - (n+\delta)$ denotes the growth rate of k. Source: Barro and Sala-i Martin (2004).

Chapter 3

Discrete Dynamical Systems: Difference Equations

In this chapter we move to discrete time, again within a deterministic environment, and consider *difference* equations and systems of the form

$$x(t+1) = g(x(t), t) \tag{3.1}$$

One might think that we need a theory distinct from ODE theory to study this type of systems but, in fact, almost all results from Chapters 1 and 2 apply for difference systems of the form (3.1) with a few adjustments. Compared to differential systems, difference systems have certain weaknesses. In particular, graphical analysis is not as useful or flexible for difference systems; in addition, the continuous and discrete-time formulations of a given setting may exhibit different *global* stability properties, a possibility that we explore in Application 3.1 in the context of the Solow growth model. On the other hand, a big advantage of difference systems is that the existence and uniqueness of solutions to boundary value problems is ensured for any arbitrary mapping g in (3.1), while, as we saw in Section 1.6, existence and uniqueness of solutions is far from a trivial matter for differential equations.

To highlight the differences between difference and differential equations, consider the autonomous linear first-order difference equation

$$x(t+1) = ax(t) + b \tag{3.2}$$

subject to an initial value condition $x(0) = x_0$. To solve this system, we can simply proceed by induction on $t \in \{0 \cup \mathbb{N}\}$, so that

$$\begin{aligned}
x(1) &= ax_0 + b \\
x(2) &= a^2 x_0 + ab + b \\
&\vdots \\
x(t) &= \begin{cases} x_0 + bt & \text{if } a = 1 \\ \frac{b}{1-a} + a^t \left(x_0 - \frac{b}{1-a}\right) & \text{if } a \neq 1 \end{cases}
\end{aligned} \tag{3.3}$$

When comparing (3.3) with the solution to the corresponding differential equation,

$$\dot{x}(t) = ax(t) + b \tag{3.4}$$

$$\Rightarrow x(t) = -\frac{b}{a} + (x_0 + \frac{b}{a})\exp(at) \tag{3.5}$$

we can immediately see that the solutions are similar with two adjustments. First, the exponential term $\exp(at)$ is now replaced by a^t. Second, the steady state of the differential equation is $-\frac{b}{a}$ whereas that of the difference equation is $\frac{b}{1-a} = -\frac{b}{a-1}$. To see why this is so, simply write (3.2) as $\Delta x(t) = x(t+1) - x(t) = (a-1)x(t) + b$ and compare it to (3.4).

Related to the second adjustment, we can see that the criterion for asymptotic stability, which is associated with a change in x that is declining over time, is now whether the absolute value of a is less than, equal to, or greater than 1, while it was whether a is less than, equal to, or greater than 0 in the continuous case.

The fact that stability depends on the *absolute value* of a is of particular importance. When we look at systems of difference equations, whose eigenvalues can be complex numbers, what matters for stability is how the *modulus* (a generalization of the concept of absolute value) of the eigenvalue compares to the number 1.

Definition 3.1 (Modulus)**.** The modulus of a complex number $c = a + ib$, where $i = \sqrt{-1}$, is the norm of the vector that represents it in the complex plane. That is,

$$|c| = \sqrt{a^2 + b^2} \tag{3.6}$$

When $|c| < 1 (> 1)$, we say that c lies inside (outside) the unit circle.

In what follows, we sometimes refer to systems of difference equations as matrix difference equations.

Remark 3.1. Autonomous linear matrix difference equations admit the same closed-form solutions as autonomous linear ODE systems after the adjustment $\exp(\lambda t) \rightsquigarrow \lambda^t$.

Stability properties for linear matrix difference equations and *local* stability properties for nonlinear matrix difference equations coincide with their counterparts for ODE systems but under different criteria regarding the stability of a given eigenvalue: in the case of continuous systems, how the *real part* of the eigenvalue compares to number 0; in the case of discrete systems, how the *modulus* of the eigenvalue compares to number 1.

We now state the analogs of Theorem 1.3 and Lemma 1.3.

Theorem 3.1 (Solution to Nonhomogeneous Autonomous Linear Difference Equations)**.** *Suppose $n \times n$ matrix A has n distinct eigenvalues $\lambda_1, \ldots, \lambda_n$, all with moduli not equal to 1. Then the unique solution to*

$$x(t+1) = Ax(t) + B \tag{3.7}$$

with initial value $x(0) = x_0$, takes the form

$$x(t) = \bar{x} + \sum_{j=1}^{n} c_j \lambda_j^t v_{\lambda_j} \tag{3.8}$$

where $\bar{x} = -[A - I_n]^{-1} b$ is the unique stationary state of the system, $v_{\lambda_1}, \ldots, v_{\lambda_n}$ are the eigenvectors corresponding to the eigenvalues $\lambda_1, \ldots, \lambda_n$ and $c_1, \ldots, c_n$ denote the constants of integration pinned down by the initial value condition.

Lemma 3.1 (Solution to a General (Nonautonomous) Linear First-Order Difference Equation)**.** *The solution to a general (nonautonomous) linear first-order difference equation*

$$x(t+1) = a(t)x(t) + b(t)$$

with initial value $x(0) = x_0$ is given by

$$x(t) = \left[x_0 + \sum_{s=0}^{t-1} \left(\tilde{R}(s+1) \right)^{-1} b(s) \right] \tilde{R}(t) \tag{3.9}$$

where $\tilde{R}(t) \equiv \prod_{s=0}^{t-1} a(s)$ and c is a constant of integration pinned down by the initial value condition.

The solution to nonautonomous linear systems of difference equations is again very similar, with the state-transition matrix $\Phi(t,s)$ now defined by

$$\Phi(t+1, s) = A(t)\Phi(t,s) \qquad \text{and} \tag{3.10}$$

$$\Phi(t,t) = I_n \tag{3.11}$$

for all $t, s \in \mathcal{T}$. In the one-dimensional case of Lemma 3.1, the state transition matrix reduces to the scalar $\Phi(t,s) = \tilde{R}(t)/\tilde{R}(s)$.

Theorem 3.2 (Solution to General Linear Difference Equations)**.** *Let $X(t)$ and $\Phi(t,s)$, $\forall t, s \in \mathcal{T}$ be the fundamental matrix and state-transition matrix, respectively, corresponding to the matrix-valued function $A(t)$. Then, a (particular) solution to the linear system*

$$x(t+1) = A(t)x(t) + B(t) \tag{3.12}$$

with boundary condition $x(0) = x_0$, is given by

$$\hat{x}(t) = X(t)[X(0)]^{-1}x_0 + \sum_{s=0}^{t-1} X(t)[X(s+1)]^{-1}B(s) \tag{3.13}$$

$$= \Phi(t,0)x_0 + \sum_{s=0}^{t-1} \Phi(t,s+1)B(s) \tag{3.14}$$

We now discuss stability in discrete systems. Define the hyperbolic state of a nonlinear system as follows

Definition 3.2 (Hyperbolic Steady State)**.** A steady state $\overline{x}$ of system

$$x(t+1) = G(x(t),t) \tag{3.15}$$

is *hyperbolic* if the matrix $DG(\overline{x})$ does not have eigenvalues with moduli equal to 1.

Then, with the adjustment noted in Remark 3.1, Theorems 2.1, 2.2, 2.3, 2.4, and 2.5 directly apply to discrete systems.

Finally, note that solutions to initial value problems for *any* discrete system exist and are unique. This can be shown almost immediately by induction.

Theorem 3.3 (Existence and Uniqueness of Solutions to Difference Equations)**.** *Consider the matrix difference equation*

$$x(t+1) = G(x(t),t) \tag{3.16}$$

where $G: \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$ is an arbitrary mapping. Then, the boundary value problem given by (3.16) *and the boundary condition $x(t_0) = x_0$, $t_0 \in \mathbb{Z}$, has a unique solution for all $t \geq t_0$. If, in addition, mapping G is invertible, then the boundary value problem has a unique solution for all $t \in \mathbb{Z}$.*

Application 3.1 (Global Stability in the Solow Model: Continuous vs. Discrete Time)**.** Consider the continuous-time and discrete-time versions of the key equation of the Solow growth model

$$\dot{k}(t) = sf(k(t)) - \delta k(t) \tag{3.17}$$

$$\equiv g(k(t)) \tag{3.18}$$

and

$$k_{t+1} = sf(k_t) + (1-\delta)k_t \tag{3.19}$$

$$\equiv h(k_t) \tag{3.20}$$

respectively, where $s, \delta \in (0,1)$. The steady state value of capital in both versions is defined by $sf(\overline{k}) = \delta\overline{k}$.

The benchmark Solow model has two main assumptions on the production function, $f(k)$. First, it is increasing and strictly concave in k with $f(0) = 0$. Second, it satisfies the *Inada Conditions*

$$\lim_{k\to 0} f'(k) = \infty \tag{3.21}$$

$$\lim_{k\to\infty} f'(k) = 0 \tag{3.22}$$

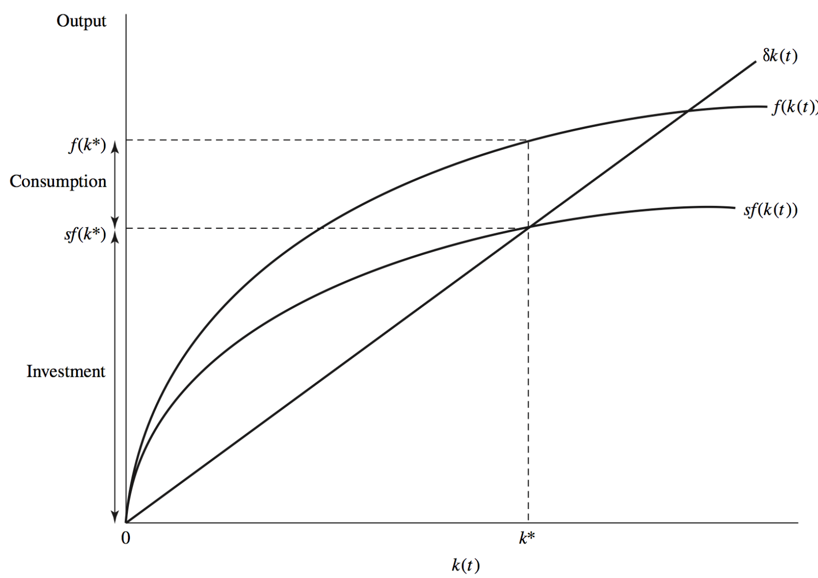


Figure 3.1: Capital-output graph for the Solow growth model (both continuous and discrete versions). Source: Acemoglu (2009).

Figure 3.1 illustrates the implications of the concavity of the production function and the Inada conditions for both continuous-time and discrete-time versions of the Solow model. The derivative of investment $sf(k(t))$ with respect to capital $k(t)$ is large and positive for low values of k and declines monotonically on its entire domain over k. Graphically, this implies that $sf(k)$ will cross the straight line δk only once, so that the steady state in both models is unique, and will do so transversally from above.

Local asymptotic stability for both models follows easily from our previous results and from the two assumptions. We know from stability theory that local asymptotic stability requires $g'(\bar{k}) < 0$ in the continuous case and $|h'(\bar{k})| < 1$ in the discrete case. We just noted that $sf(k)$ crosses δk from above, $sf'(\bar{k}) < \delta$, so that local stability for the continuous case immediately follows. For the discrete case, note that the strict concavity of $f(k)$ implies

$$f(k) > f(0) + kf'(k) \tag{3.23}$$

$$= kf'(k) \tag{3.24}$$

so that

$$h'(\bar{k}) = sf'(\bar{k}) + 1 - \delta \tag{3.25}$$
$$< \frac{sf(\bar{k})}{\bar{k}} + 1 - \delta \tag{3.26}$$
$$= \delta + 1 - \delta$$
$$= 1 \tag{3.27}$$

using the definition of the steady state. Since $h'(k) > 0$ for all k, we have that $|h'(\bar{k})| < 1$.

In contrast, global asymptotic stability is a property with respect to which the discrete-time and continuous-time formulations of (certain versions of) the Solow model may differ. In the continuous case global asymptotic stability requires $g(k) > 0$ for all $k < k^*$ and $g(k) < 0$ for all $k > k^*$. It follows immediately that the steady state is globally asymptotically stable.

In the discrete case global asymptotic stability requires that at the limit as $t \to \infty$, $\left|k_{t+1} - \bar{k}\right| < \left|k_t - \bar{k}\right|$ for all possible values of $k_t > 0$. A *sufficient* condition for global stability in the discrete version would be $|h'(k_t)| < 1$ for all $k_t > 0$ (not just at the steady state) as $t \to \infty$. To see this, assume $|h'(k)| < 1$ for all $k > 0$. Then,

$$\begin{aligned}
|k_{t+1} - \bar{k}| &= |h(k) - h(\bar{k})| \\
&= \left| \int_{\bar{k}}^{k_t} h'(k)\, \mathrm{d}k \right| \\
&\leq \int_{\bar{k}}^{k_t} \left|h'(k)\right| \mathrm{d}k \\
&< |k_t - \bar{k}| \qquad (3.28)
\end{aligned}$$

where the last inequality uses the stated assumption. But this assumption is *not* true for our function $h(k)$: $h'(k) = sf'(k) + 1 - \delta \to \infty$ as $k \searrow 0$.

The discrete-time formulation of the benchmark Solow model, (3.19), does turn out to be globally asymptotically stable, but the argument requires a more direct line of attack. Suppose that $k_t \in (0, \bar{k})$. Then

$$k_{t+1} - \bar{k} = h(k_t) - h(\bar{k}) \tag{3.29}$$
$$= -\int_{k_t}^{\bar{k}} h'(k)\, \mathrm{d}k \tag{3.30}$$
$$< 0$$

since $h'(k) > 1 - \delta > 0$ for all k, which implies that $k_{t+1} \in (k_t, k^*)$ and, by induction over t, that the level of capital at any time after t is bounded above by $\bar{k}$. Moreover,

$$\frac{k_{t+1}-k_t}{k_t} = s\frac{f(k_t)}{k_t} - \delta \tag{3.31}$$

$$> s\frac{f(\bar{k})}{\bar{k}} - \delta \tag{3.32}$$

$$= 0 \tag{3.33}$$

where the second line uses the concavity of f (which implies that $f(k)/k$ is decreasing with k) and the last line uses the definition of the steady state value. Therefore, capital is monotonically approaching its steady-state level.

We have shown that $\{k_t\}$ is a monotonically increasing sequence bounded above by $\bar{k}$. Since $\bar{k}$ is the unique steady state, $\bar{k}$ must be the least upper bound of $\{k_t\}$. By the monotone convergence theorem, it follows that $\{k_t\} \to \bar{k}$ from below. An identical argument proves that $\{k_t\} \to \bar{k}$ from above. Convergence towards the steady state in the discrete model is illustrated in Figure 3.2.

The fact that global asymptotic stability in the benchmark version of the discrete-time Solow model does not immediately follow from the usual assumptions of concavity of $f(k)$ and the Inada conditions illustrates that discrete systems may have different global stability properties from their continuous-time counterparts. Indeed, it can be shown, for example, that in a version of the Solow model with population growth and where the (exogenous) savings rate is a deterministic function of the level of capital, cycles between two capital levels are possible, so that the model's steady state cannot be globally stable.[1] Such cycles are clearly impossible in the continuous-time counterpart of this model.[2] To the extent that such cycles are interpreted as pathological results (and not as desirable features of the model), one can argue that the continuous-time formulation of the model is superior.

■

3.1 The Dynamics of Planar Systems

We end with some tips on stability analysis for linear (linearized) planar systems, just as we did in Section 2.2.3 for continuous planar systems. We will make use of them in Application 4.2.

Consider system (3.7). Recall that

$$tr(A) = \lambda_1 + \lambda_2 \tag{3.34}$$

$$det(A) = \lambda_1 \lambda_2 \tag{3.35}$$

[1] See exersise 2.21 in Acemoglu (2009).

[2] To see this, suppose that such a cycle among different capital levels existed. Because the level of capital must be a continuous function of time, there would exist a value of capital that the system crossed both while going up and while going down. This would imply that f has both a positive and a negative derivative at that level of capital, a contradiction.

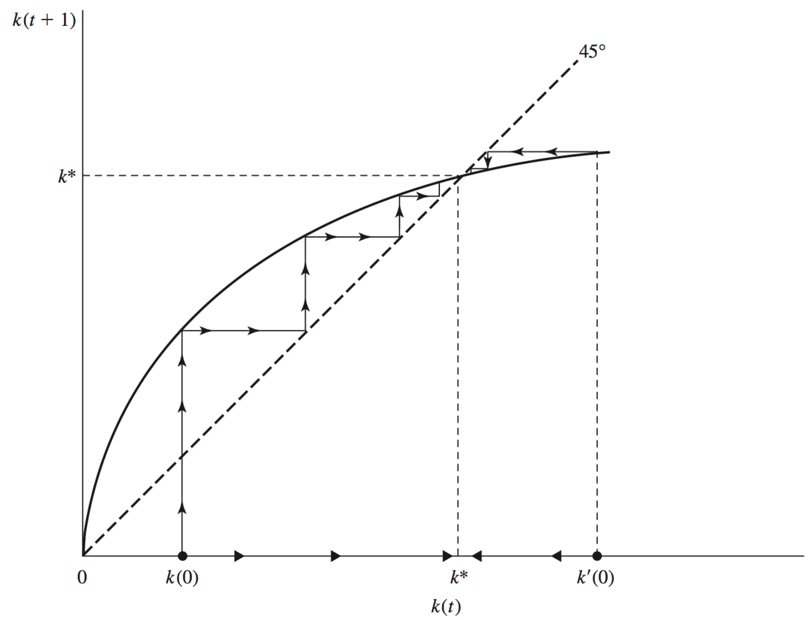


Figure 3.2: Convergence towards the steady state in the discrete-time version of the Solow growth model. Source: Acemoglu (2009).

where λ_1 and λ_2 are the eigenvalues, which solve the characteristic quadratic equation

$$p_A(\lambda) = \lambda^2 - tr(A)\lambda + det(A) = 0 \tag{3.36}$$

$$\Rightarrow \lambda_{1,2} = \frac{tr(A) \pm \sqrt{tr(A)^2 - 4det(A)}}{2} \tag{3.37}$$

We want to derive the conditions on the trace and determinant of matrix A under which the eigenvalues of the matrix lie *inside* the unit circle, that is, they both have modulus less than 1.

First, suppose that λ_1 and λ_2 are real. Then, clearly we must have

$$|\det(A)| = |\lambda_1||\lambda_2| < 1 \tag{3.38}$$

Now, suppose they are both complex (complex eigenvalues always come in pairs since A has real entries), of the form $a + ib$, $a - ib$. Then,

$$\det(A) = \lambda_1\lambda_2 = a^2 + b^2 = |\lambda_1|^2 = |\lambda_2|^2 \tag{3.39}$$

Therefore, we know that, in all cases, $|\det(A)| < 1$. Second, note that the characteristic polynomial can be written as

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) \tag{3.40}$$

Suppose for a moment that the eigenvalues are real and note that it must be the case that

$$p_A(1) > 0 \text{ and} \tag{3.41}$$
$$p_A(-1) > 0 \tag{3.42}$$

since we require that both eigenvalues are on the same side of both 1 and -1 (in the real line). It can be shown that these inequalities must also hold for the case of complex eigenvalues. Also note from (3.36) that $p_A(1) = 1 - tr(A) + det(A)$ and $p_A(-1) = 1 + tr(A) + det(A)$. We have thus derived the following result: [3]

Lemma 3.2. *A 2×2 matrix A has both of its eigenvalues inside the unit circle if and only if all three of the following conditions hold:*

- $det(A) < 1$
- $p_A(1) > 0 \Leftrightarrow det(A) > tr(A) - 1$
- $p_A(-1) > 0 \Leftrightarrow det(A) > -tr(A) - 1$

The result is illustrated in Figure 3.3, which draws the lines $p_A(1) = 0$ and $p_A(-1) = 0$ in trace-determinant space, together with the horizontal line corresponding to $\det(A) = 1$. Matrices with trace-determinant "coordinates" in the shaded area have both of their eigenvalues inside the unit circle.

The important case of saddle-path stable 2×2 systems can also be handled using Lemma 3.2. A discrete planar system is saddle-path stable if matrix A (equal to $DG(\overline{x})$ in the nonlinear case) has one eigenvalue inside the unit circle and one eigenvalue outside the unit circle. To show this is the case for a matrix without explicitly solving for its eigenvalues, one can show that both A and its inverse fail at least one of the three conditions in Lemma 3.2.

[3] Note that $\det(A) > -1$ is reduntant given the conditions $p_A(1) > 0$ and $p_A(-1) > 0$, as we can immediately see in figure 3.3.

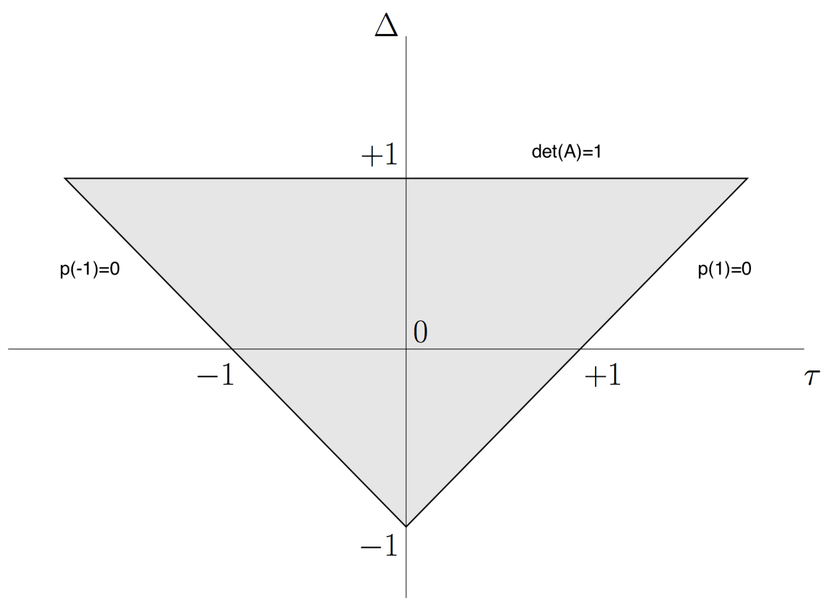


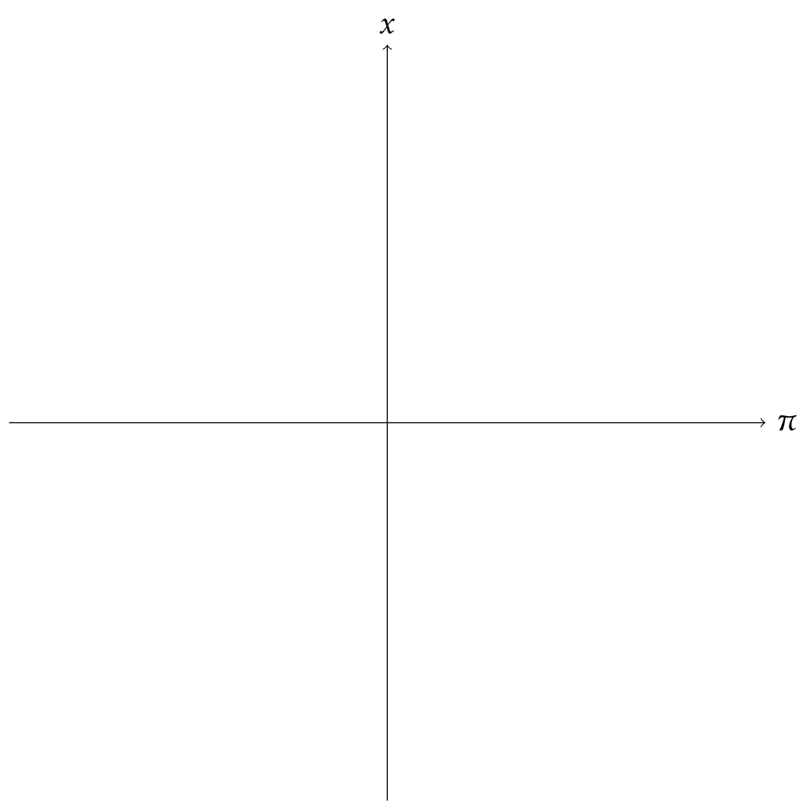
Figure 3.3: Stability map for discrete planar systems, drawn in trace (τ) - determinant (Δ) space. Matrices in the shaded area have both of their eigenvalues inside the unit circle. Adapted from: Diekman and Kuznetsov (2011).

Problem Set 2

1. Consider the linear system (1.45)-(1.46) studied in Application 1.2 and reproduced here:

$$\dot{x}(t) = -\pi(t) - \underline{r}$$
$$\dot{\pi}(t) = \rho\pi(t) - \kappa x(t)$$

In the axes provided below, draw the phase diagram of the model. That is, draw the phase lines in (π, x) space (these are the same as in Figure 1.1) and the stable and unstable subspaces of the system (since the system is linear these are straight lines). Which of the two is the saddle path of the system? You may use the parametrization $\rho = \kappa = -\underline{r} = 0.05$.



2. (Local Stability for Q-theory) Consider the ODE system

$$\dot{K}(t) = C'^{-1}\left(q(t) - 1\right) - \gamma K(t)$$
$$\dot{q}(t) = (r + \gamma)\, q(t) - \pi(K)$$

where $r, \gamma > 0$, $\pi'(\cdot) < 0$, and $C(\cdot)$ is a strictly increasing and strictly convex function, $C'(\cdot) > 0$ and $C''(\cdot) > 0$ (with $C(0) = C'(0) = 0$). Note that $C'^{-1}(\cdot)$ denotes the inverse function of function $C'(\cdot)$.

(i) Show that the system is locally saddle-path stable around its steady state.

(ii) Show that the saddle path of the system is downward sloping in a neighborhood of the steady state.

3. This exercise is intended to offer some insight into the link between autonomous differential and difference equations stated in Remark 3.1, namely, why they have elementary solutions of the form $\exp(\lambda t)$ and λ^t, respectively.

   We will derive the solution to the (simplest possible) differential equation $\dot{x}(t) = \lambda x(t)$, subject to $x(0) = x_0$, where $\lambda \in \mathbb{R}$. Suppose we do not know integration, so we cannot immediately compute the solution as $x(t) = x_0 \exp(\lambda t)$. Instead, we will guess that the solution $x(t)$ is a *power series*, that is, it can be written in the form

$$x(t) = \sum_{n=0}^{\infty} a_n t^n$$

   for a sequence of coefficients $\{a_n\} \in \mathbb{R}^{\infty}$.

   (i) Substitute the guess (3.43) into $\dot{x} = \lambda x$ and transform the differential equation into a difference equation in the coefficients $\{a_n\}$ (that is, with n as the independent variable). Hint: Use the fact that if $\sum_0^\infty C_n t^n = 0$ must hold for all t, for some sequence of coefficients $\{C_n\}$ then it must be that $C_n = 0$, for all $n \in \{0 \cup \mathbb{N}\}$.

   (ii) Solve the difference equation by induction, for given a_0.

(iii) Substitute the solution for $\{a_n\}$ from part (ii) into the guess, (3.43). Use the initial condition $x(0) = x_0$.

(iv) Finally use the fact that the natural exponential function $\exp(t)$ is *defined* as

$$\exp(t) \equiv \sum_{n=0}^{\infty} \frac{t^n}{n!}$$

to conclude that $x(t) = x_0 \exp(\lambda t)$.

4. Consider the Tth-order difference equation

$$x_{t+T} + a_{T-1}x_{t+T-1} + a_{T-2}x_{t+T-2} + \ldots + a_0 x_t = 0$$

for all $t \in \{0 \cup \mathbb{N}\}$, , where $T \in \mathbb{N}$, subject to the initial condition

$$\begin{bmatrix} x_{T-1} & x_{T-2} & \cdots & x_0 \end{bmatrix}^T = y_0$$

for some T-dimensional column vector y_0.

(i) Transform (3.43) into a first-order system of differential equations of the form $y(t+1) = Ay(t)$, where A is a $T \times T$ matrix.

(ii) Assume that A has no repeated eigenvalues. Show that the general solution to (3.43) is $x(t) = \sum_{i=1}^{T} c_i \lambda_i^t$. Show that the eigenvector v_{λ_i} corresponding to the ith eigenvalue of matrix A, λ_i, can be expressed as

$$v_{\lambda_i} = \begin{bmatrix} \lambda_i^{T-1} \\ \lambda_i^{T-2} \\ \vdots \\ \lambda_i \\ 1 \end{bmatrix}$$

(iii) Let $c = \begin{bmatrix} c_1 & c_2 & \ldots & c_T \end{bmatrix}^T$. Express c in terms of y_0.

Chapter 4

Discrete Dynamical Systems: Expectational Difference Equations

Until now, we have only discussed solutions to deterministic models. We now introduce uncertainty to difference systems to study an important class of linear systems[1] associated with a class of economic models called *dynamic stochastic general equilibrium (DSGE)* or *rational expectations equilibrium* models. These stochastic linear discrete-time systems are sometimes called *linear expectational difference equations*.

4.1 The One-Dimensional Case

Our goal is to solve a system of the general form

$$\mathbb{E}_t[x(t+1)] = Ax(t) + Bu(t) \tag{4.1}$$

where $x(t)$ is a n-dimensional vector of *endogenous* variables, $u(t)$ is a k-dimensional vector of, possibly stochastic, *exogenous* variables (or shocks), A is an $n \times n$ matrix, and B is an $n \times k$ matrix. $\mathbb{E}_t[\cdot]$ denotes the expectational operator with respect to all available information up to and including time t. Expectations are called "rational" because they are assumed to be "consistent" with the equilibrium of the model in the sense that agents' perceived probability distribution over economic outcomes coincides with the probability distribution implied by the model.

To build intuition, we begin with the one-dimensional case. Consider the system

$$\mathbb{E}_t[x_{t+1}] = ax_t + u_t \tag{4.2}$$

where $a \in \mathbb{R}$ and where x_{t+1} is *nonpredetermined* at time t, that is it is nondegenerate with respect to information available up to and including time t.

[1]Introducing uncertainty in continuous time requires techniques from stochastic calculus, which will be covered in the first quarter of the macro sequence.

Let us first try to solve this solution "backwards," as we have done for deterministic difference equations in the previous section, given an arbitrary initial condition, $x_{-1} = x_{init}$.

One may be tempted to recursively derive the solution to (4.2) as

$$\begin{aligned} x_0 &= ax_{init} + u_{-1} \\ &\vdots \\ x_t &= ax_{t-1} + u_{t-1} \end{aligned}$$

However, this is only one of many (uncountably infinite) solutions. The set of solutions is described by

$$x_t = ax_{t-1} + u_{t-1} + v_t \tag{4.3}$$

for $t \geq 0$, where $\{v_t\}_{t=1}^{\infty}$, the sequence of *expectational errors*, is an *arbitrary* stochastic process that satisfies $\mathbb{E}_t v_{t+1} = 0$ for all t. Therefore, in contrast to deterministic difference equations, we have no hope of attaining a unique solution to an expectational difference equation subject only to an initial condition. However, for reasons explained in Section 4.2, we are usually interested only in *bounded* solutions of expectational difference systems. That is, we are interested in the following boundary value problem: given the realization of the disturbances u_s for $s = -1, \ldots, t$, solve for the value(s) of x_t such that

$$\begin{aligned} \mathbb{E}_t[x_{t+1}] = ax_t + u_t \quad \text{subject to} \quad & x_{-1} = x_{init} \\ \text{and} \quad & \left|\mathbb{E}_t\left[x_{t+j}\right]\right| < K \in \mathbb{R} \quad \forall j \geq 0 \end{aligned} \tag{4.4}$$

where the exogenous sequence $\{u_t\}_{t=0}^{\infty}$ is bounded, $\left|\mathbb{E}_t\left[u_{t+j}\right]\right| < M \in \mathbb{R}$ for all $j \geq 0$.[2]

We ask under which condition on the parameter a the solution for $\{x_t\}_{t=0}^{\infty}$ in problem (4.4) is unique. As we also saw in Section 1.4.2.1 on forward solutions of differential equations, the forward solution will be more useful for problems like (4.4) involving conditions at infinity. We can solve (4.2) forward noting that

$$\begin{aligned} (4.2) \;\Rightarrow\; & \mathbb{E}_{t+j}x_{t+j+1} = ax_{t+j} + u_{t+j} && (4.5) \\ \Rightarrow\; & \mathbb{E}_t[\mathbb{E}_{t+j}x_{t+j+1}] = a\mathbb{E}_t x_{t+j} + \mathbb{E}_t u_{t+j} && (4.6) \\ \Rightarrow\; & \mathbb{E}_t x_{t+j+1} = a\mathbb{E}_t x_{t+j} + \mathbb{E}_t u_{t+j} && (4.7) \end{aligned}$$

[2] More formally, we define a random variable $y = \{y_t\}_{t=0}^{\infty}$ to be *bounded* if the norm of y_t is finite (almost surely) for all t. y_t can be thought of as a (measurable) function of the history of states of nature, s^t, that is, $y_t = y(s^t)$, and y is an infinite sequence of such measurable functions. Under the L_∞ topology the norm of y, $\|y\|_\infty$, is the least upper upper bound such that $|y_t|$ is below this bound everywhere except at states of measure zero, for all t. Boundedness implies that $\mathbb{E}_t y_{t+j}$ is finite for all $j \geq 0$.

where the last line follows from the Law of Iterated Expectations. Then, (4.2) implies

$$\begin{aligned} x_t &= a^{-1}\mathbb{E}_t[x_{t+1}] - a^{-1}u_t & (4.8)\\ &\overset{(4.7)}{=} a^{-1}\left[a^{-1}\mathbb{E}_t x_{t+2} - a^{-1}\mathbb{E}_t u_{t+1}\right] - a^{-1}u_t & (4.9)\\ &\vdots \\ &= a^{-k}\mathbb{E}_t x_{t+k} - \sum_{j=0}^{k-1} a^{-(j+1)}\mathbb{E}_t[u_{t+j}] & (4.10)\end{aligned}$$

Provided

$$\lim_{k\to\infty} |a^{-k}\mathbb{E}_t x_{t+k}| < \infty \tag{4.11}$$

and

$$\lim_{k\to\infty} \left|\sum_{j=0}^{\infty} a^{-(j+1)}\mathbb{E}_t\left[u_{t+j}\right]\right| < \infty \tag{4.12}$$

we can write the forward solution to (4.2) as

$$x_t = \lim_{k\to\infty} a^{-k}\mathbb{E}_t x_{t+k} - \sum_{j=0}^{\infty} a^{-(j+1)}\mathbb{E}_t[u_{t+j}] \tag{4.13}$$

Just as in the continuous-time case of Section 1.4.2.1, the second term on the right hand side of (4.13) is sometimes referred to as the *fundamental solution* of (4.2) and the first term as the *bubble* term.

Now consider three cases based on the absolute value (modulus) of constant a. First assume $|a| > 1$. Then, for a bounded $\{x_t\}$, we have

$$\lim_{k\to\infty} a^{-k}\mathbb{E}_t x_{t+k} = 0 \tag{4.14}$$

so that the limit in (4.12) exists and thus x_t must satisfy (4.13). This pins down a unique bounded sequence $\{x_t\}$, equal to the fundamental solution of (4.4),

$$x_t = -\sum_{j=0}^{\infty} a^{-(j+1)}\mathbb{E}_t[u_{t+j}] \tag{4.15}$$

Now assume $|a| < 1$. Clearly, the limit $\lim_{k\to\infty} a^{-k}\mathbb{E}_t x_{t+k}$ is not finite, so that (4.13) does not help us pin down a unique bounded solution. Instead, all solutions defined by (4.3) for a bounded sequence $\{v_t\}$ satisfying $\mathbb{E}_t v_{t+1} = 0$, $\forall t$, are bounded solutions. That is, if $|a| < 1$, we have an infinity of bounded solutions.

What happens if $|a| = 1$? From (4.3) ,we can see that not all solutions of that form need be bounded even if both $\{u_t\}$ and $\{v_t\}$ are bounded. However, it can be shown that, if (4.3) has any bounded solution, then it must have an uncountably infinite number of them. For example, suppose that

$$U_t \equiv \sum_{j=0}^{\infty} \mathbb{E}_t u_{t+j} \tag{4.16}$$

is finite for all t .

Then, a bounded solution

$$x_t = U_t + (x_{t-1} - U_{t-1}) + v_t \tag{4.17}$$

exists for arbitrary $\{v_t\}$ satisfying $\mathbb{E}_t v_{t+1} = 0$ and

$$\left| \sum_{t=1}^{\infty} v_t \right| < \infty \tag{4.18}$$

If follows that a unique bounded solution cannot exist when $|a| = 1$.

Theorem 4.1. *The expectational difference equation*

$$\mathbb{E}_t[x_{t+1}] = a x_t + u_t \tag{4.19}$$

where $\{u_t\}$ is bounded, has a unique bounded solution if and only if $|a| > 1$. In this case, the solution is

$$x_t = -\sum_{j=0}^{\infty} a^{-(j+1)} \mathbb{E}_t[u_{t+j}] \tag{4.20}$$

for all $t \geq 0$.

For a concrete example, suppose that $\{u_t\}_{t=0}^{\infty}$ is a martingale difference sequence, that is $\mathbb{E}_t\left[u_{t+j}\right] = 0$ for all $j \geq 1$. Then, with $|a| > 1$, if $\{x_t\}_{t=0}^{\infty}$ is to be bounded, we must have $\mathbb{E}_t[x_{t+1}] = 0$. If not, for example if $\mathbb{E}_t[x_{t+1}] = c \neq 0$ then $\left|\mathbb{E}_t\left[x_{t+j}\right]\right| = \left|a^{j-1}c\right| \to \infty$ as $j \to \infty$. Therefore, with $|a| > 1$, the bounded solution to $\{x_t\}_{t=0}^{\infty}$ is unique and given by $\mathbb{E}_t[x_{t+1}] = 0 \Rightarrow x_t = -u_t/a$ for all $t \geq 0$.

In contrast, if $|a| < 1$, $\mathbb{E}_t[x_{t+1}]$ can be any number $c \in \mathbb{R}$ because this would imply $\left|\mathbb{E}_t\left[x_{t+j}\right]\right| = \left|a^{j-1}c\right| < \infty$. Then, for each $c \in \mathbb{R}$, $x_t = \frac{1}{a}c - \frac{1}{a}u_t$ is a valid bounded solution to problem (4.4).

Recall from the previous chapter that deterministic (autonomous) difference equations of the form (4.19) with $|a| > 1$ were associated with unstable steady states. The idea of inherent *instability* giving rise to equilibrium *determinacy* (uniqueness) is central to the workings of rational-expectations models. We discuss this idea in the context of monetary policy determinacy in Application 4.2.

4.2 Equilibrium Determinacy

In this section we discuss two important reasons why we are interested in unique bounded solutions of models of the form (4.1). The first reason is that in many rational expectations equilibrium models the existence of equilibrium requires that certain *transversality conditions* are satisfied. These conditions preclude explosive paths for variables that affects agents' utility. Therefore, insofar as we are looking at rational-expectations equilibria of such a model, we are necessarily looking at bounded solutions to the equilibrium conditions. Effectively, agents coordinate their expectations of the equilibrium paths of endogenous variables by ruling out any future paths of the variables that would be explosive and thus not consistent with equilibrium.

Then, uniqueness of the bounded solution translates directly into uniqueness of the rational expectations equilibrium. The latter is desirable for obvious reasons, since it implies that models have sharp, and thus testable, predictions. Models with a multiplicity of equilibria emphasize the self-fulfilling nature of expectations, and usually interpret negative economic outcomes as the result of expectational coordination failures. That is, agents at some point in time choose, for some unmodeled reason, to form their expectations according to the "bad" equilibrium, thus bringing about this "bad" equilibrium (and, in this sense, expectations are self-fulfilling). A caveat is that transversality conditions do not necessarily preclude explosive paths for all endogenous variables but only for *real* endogenous variables, since only the latter affect agents' utility. This caveat is particularly relevant for a monetary economy at the "Woodfordian" cashless limit and is discussed further in Application 4.2.

The second reason, which is distinct conceptually from the first, applies to *nonlinear* models. An equilibrium is *locally determinate* if it is the unique equilibrium such that the endogenous variables remain within some neighborhood of their steady-state values for all t. It turns out that an *infinite-dimensional* nonlinear model is locally determinate for sufficiently small disturbances and hence (log)linearization around the steady state and related comparative dynamics exercises are justified only if there exists a unique bounded solution to the *linearized* counterpart of the model for any exogenous processes (disturbances) that are also bounded. Since most microfounded macroeconomic models are nonlinear in their exact form, making sure that we are justified in looking at their tractable (log)linear approximations of the form (4.1) is crucial. Subsection 4.2.1 presents a formal discussion of this point.

4.2.1 Local Determinacy and (Log)linearization*

Consider a nonlinear model of the general form

$$\Phi(x;u) \equiv \{\mathbb{E}_t \phi(x_t, x_{t+1}; u_t)\}_{t=0}^{T} = 0 \in \mathbb{R}^T \tag{4.21}$$

subject to $x_{-1} = x_{\text{init}}$, where we assume that x_t and u_t are both n-dimensional random variables, for some $n \in \mathbb{N}$ (this is without loss of generality). x_t is the vector of endogenous variables and u_t is the vector of exogenous processes (shocks) or model parameters.[3] We assume that ϕ is also n-dimensional and continuously differentiable with respect to its arguments on its entire domain. Denote the equilibrium around which we wish to linearize the model (perform comparative dynamics) by x^*. In principle, x^* can be any equilibrium satisfying (4.21) but usually we are interested in the (deterministic) *steady state equilibrium*, that is, the equilibrium that satisfies, $x_t = \overline{x} \in \mathbb{R}^n$ for all t when $u_t = 0$ for all t (in all states of the world) and $x_{-1} = x_{\text{init}} = \overline{x}$,[4]

$$\Phi(x^*; 0) \equiv \{\phi(\overline{x}, \overline{x}; 0)\}_{t=0}^{T} = 0 \in \mathbb{R}^T \tag{4.22}$$

where x^* is T-dimensional copy of vector $\overline{x}$.

Whether $T < \infty$ or $T = \infty$, we can interpret Φ as a mapping from the sequences $\{x_t\}$ and $\{u_t\}$ to the sequence $\{\phi_t\}$ of values of function ϕ in period t. Similarly, we can interpret the derivative $D_x\Phi(x^*; 0)$ as the *linear* operator that maps perturbations $\{\hat{x}_t\} \equiv \{x_t - \overline{x}\}$ to perturbations $\{\hat{\phi}_t\} \equiv \{\phi_t - 0\}$. That is,

$$\begin{aligned} D_x\Phi(x^*; 0)[\{\hat{x}_t\}] &= \{\hat{\phi}_t\} \\ &= \{\phi_1 \hat{x}_t + \phi_2 \mathbb{E}_t \hat{x}_{t+1}\} \end{aligned} \tag{4.23}$$

where ϕ_1 and ϕ_2 are the derivatives of function ϕ with respect to its first two arguments, evaluated at $x_t = x_{t+1} = \overline{x}$ and $u_t = 0$ for all $t \geq -1$.

Definition 4.1 (Toplinear Isomorphism)**.** A linear map L is a *toplinear isomorphism* if it is continuous and has an inverse that is also a continuous linear map.

We now state the generalized versions of the inverse function and implicit function theorems that apply to the class of models of the form (4.21).

Theorem 4.2 (Inverse Mapping Theorem for REE models)**.** *Consider a Rational Expectations Equilibrium (REE) model of the form* (4.21)*, where ϕ is a $\mathcal{C}^k$ (k-times continuously differentiable) function, $k \in \mathbb{N}$.*

If $D_x\Phi(x^*; 0)$ is a toplinear isomorphism then the steady-state equilibrium x^*, defined by (4.22)*, is* locally determinate*. That is, there exist neighborhoods $\mathcal{X}$ and $\mathcal{X}'$ of x^* in the space of bounded sequences such that Φ is a one-to-one mapping from $\mathcal{X}$ to $\mathcal{X}'$ with a $\mathcal{C}^k$ inverse mapping. Therefore, x^* is the* locally unique *solution to* (4.22) *in the sense that there is no other equilibrium (for $u = 0$) such that x_t remains within some neighborhood of $\overline{x}$ for all t.*

[3] In the latter case, simply replace $u_t = u$ for all t.

[4] That is, a steady state equilibrium associated with steady state $\overline{x} \in \mathbb{R}^n$ is defined to be an equilibrium such that, if the system starts at the steady state, $x_{\text{init}} = \overline{x}$, then, in the absence of disturbances (and of expectations of disturbances in the future), it remains in the steady state forever, $x_t = \overline{x}$ for all $t \geq 1$. Also see Remark 6.1 in Chapter 6.

Theorem 4.3 (Implicit Mapping Theorem for REE models). *Consider a Rational Expectations Equilibrium (REE) model of the form* (4.21), *where ϕ is a $\mathcal{C}^k$ (k-times continuously differentiable) function, $k \in \mathbb{N}$.*

If $D_x\Phi(x^*;0)$ is a toplinear isomorphism then there exist neighborhoods $\mathcal{X}$ of x^* and $\mathcal{U}$ of $u = 0$ in the space of bounded sequences and a unique *$\mathcal{C}^k$ map $f : \mathcal{U} \to \mathcal{X}$ such that $f(0) = x^*$ and $\Phi(f(u), u) = 0$ for all $u \in \mathcal{U}$. Thus, equilibrium at $u \in \mathcal{U}$ is* locally determinate. *In other words, for any sequence $u = \{u_t\}$ of disturbances that are sufficiently close to $u_t = 0$ for all t there exists a* locally unique *rational expectations equilibrium x, that is, it is the unique equilibrium such that x_t remains within some neighborhood of $\overline{x}$ for all t.*

Since f is a $\mathcal{C}^k$ map, $k \geq 1$, the (first-order) approximation to $x = f(u)$ given by

$$f(u) \approx x^* + Df(0)[u] \tag{4.24}$$

where

$$Df(0) = -[D_x\Phi(\overline{x};0)]^{-1} D_u\Phi(\overline{x};0) \tag{4.25}$$

is accurate up to an error term of order $\mathcal{O}(\|u\|^2)$.[a]

[a]The norm is with respect to the L_∞ topology, see footnote 2.

Theorems 4.2 and 4.3 tell us that if the derivative of map Φ possesses a certain property (it is a toplinear isomorphism) then the (exact) rational-expectations equilibrium is locally determinate both when disturbances are completely absent and when disturbances (changes in u) are “small.” Thus, the (log)linear model defined by the (log)linearized versions of the equilibrium conditions in Φ is a good approximation for such small disturbances.

How are the cases $T < \infty$ and $T = \infty$ different? Theorems 4.2 and 4.3 apply to both cases, but the requirements for a linear map (in our case $D_x\Phi(x^*;0)$) to be invertible are more stringent if our problem is infinite-dimensional. If our map Φ has finite dimension, which would be the case if both $T < \infty$ and there was a finite number of states of the world, then a linear map is a toplinear isomorphism if and only if it is nonsingular, that is, the determinant of the matrix $D_x\Phi(x^*;0)$ has no eigenvalues equal to zero. In our framework, this is the case almost by construction, since function ϕ in (4.21) represents a system of n equations in n unknowns (and we assume no equations are redundant). In the latter case, an inverse exists and the map and its inverse are continuous since they are linear (see Definition 4.1).

However, for an infinite-dimensional map, linearity alone does not imply continuity. It can be shown that a linear map is continuous if and only if the map and its inverse map are bounded, that is, they map an arbitrary bounded sequence into a sequence that is also bounded. In our case (see equation (4.23)), we first need that for any sequence of perturbations in x such that $\|\{\hat{x}_t\}\| < \infty$ we must have $\|\{\hat{\phi}_t\}\| < \infty$, where, as before,

$$\hat{\phi}_t \equiv \phi_1 \hat{x}_t + \phi_2 \mathbb{E}_t \hat{x}_{t+1} \quad \forall t \geq -1 \tag{4.26}$$

This condition holds since we have assumed that ϕ is continuously differentiable so that its partial derivatives are finite. More importantly, we also need that map $D_x\Phi(x^*;0)$ has a bounded inverse map,which requires that the sequence of equations

$$\{\hat{\phi}_t\} = \{\phi_1 \hat{x}_t + \phi_2 \mathbb{E}_t \hat{x}_{t+1}\} \tag{4.27}$$

with $\{\hat{x}_t\}$ as its unknowns, has a unique and bounded solution $\{\hat{x}_t\}$ for any bounded sequence $\{\hat{\phi}_t\}$.[5]

This is far from a trivial requirement. Take the one-dimensional case, where $x_t, \phi \in \mathbb{R}$. We can write (4.27) as

$$\mathbb{E}_t \hat{x}_{t+1} = -\frac{\phi_1}{\phi_2}\hat{x}_t + \frac{1}{\phi_2}\hat{\phi}_t \tag{4.28}$$

From Theorem 4.1 we know that (4.28) has a unique bounded solution if and only if

$$\left|\frac{\phi_1}{\phi_2}\right| > 1 \tag{4.29}$$

We conclude that REE models consisting of a single equilibrium equation per period and a single endogenous variable are locally determinate and admit a valid loglinear approximation if and only if (4.29) holds. The condition for the multidimensional case follows directly from the results of the next section, reinterpreting the linear model as the (log)linearization of the nonlinear model (4.21) around its steady state. In particular, assuming that all variables in vector x_t are nonpredetermined at time t (just as we assumed for the one-dimensional case (4.2)), the condition is that matrix $[\phi_2]^{-1}\phi_1$ has all of its eigenvalues outside the unit circle, the natural generalization of (4.29). The more relevant case where some variables are predetermined is discussed in the next section.

Finally, consider a rational-expectations model of the general form

$$\Phi(x;u) \equiv \{\mathbb{E}_t \phi(x_{t-1}, x_t, x_{t+1}; u_t)\}_{t=0}^{\infty} = 0 \tag{4.30}$$

where $x_t, \phi \in \mathbb{R}^n$, the steady state of which is similarly defined by

$$\Phi(x^*;0) \equiv \{\phi(\overline{x}, \overline{x}, \overline{x}; 0)\}_{t=0}^{\infty} = 0 \tag{4.31}$$

[5]Note that the linearized model is defined by $\phi_1 \hat{x}_t + \phi_2 \hat{x}_{t+1} + \phi_3 u_t = 0$, where ϕ_3 is defined analogously to ϕ_1 and ϕ_2 as the derivative of function ϕ with respect to its third argument, u_t, evaluated at $x_t = x_{t+1} = \overline{x}$ and $u_t = 0$ for all $t \geq -1$. Therefore, $\hat{\phi}_t = -\phi_3 u_t$ and the sequence of perturbations $\{\hat{\phi}_t\}$ corresponds to a multiple of the sequence of exogenous disturbances $\{u_t\}$.

where all variables in vector x_t are assumed to be nonpredetermined at time t. The invertibility condition for the linearized model

$$\{\hat{\phi}_t\} = \{\phi_1 \hat{x}_{t-1} + \phi_2 \hat{x}_t + \phi_3 \mathbb{E}_t \hat{x}_{t+1}\} \tag{4.32}$$

turns out to be that the (characteristic) equation

$$\det[\phi_3 \lambda^2 + \phi_2 \lambda + \phi_1] = 0 \tag{4.33}$$

has exactly n roots strictly inside the unit circle and n roots strictly outside the unit circle. An exercise in Problem Set 3 asks you to derive the above condition for $n = 1$ using the Blanchard and Kahn (1980) method discussed in the next section.

An important caveat is that the results above only refer to *local* determinacy of the (nonlinear) equilibrium, not to global determinacy, that is whether the model has only one possible solution for $\{x_t\}$, whether bounded or unbounded. If transversality conditions do not preclude explosive paths for all variables (including nominal variables, in particular), then equilibrium may not be unique even though it is locally so.

4.3 The Blanchard-Kahn solution method

Consider the general system (4.1),

$$\mathbb{E}_t x_{t+1} = A x_t + B u_t \tag{4.34}$$

Given the results of Chapter 3 and Theorem 4.1, one can easily see that, if all variables in the vector x_t are nonpredetermined at time t, system (4.34) has a unique bounded solution if and only if all eigenvalues of matrix A lie outside the unit circle (that is, they have moduli greater than 1).

However, consider a special case of system (4.34) that is more common in economic applications, as in Blanchard and Kahn (1980):

$$\begin{bmatrix} x_{1,t+1} \\ \mathbb{E}_t x_{2,t+1} \end{bmatrix} = A \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} + B u_t \tag{4.35}$$

subject to $x_{1,-1} = x_{\text{init}}$, where $x_{1,t}$ is an $n \times 1$ vector of variables *predetermined* at time t, $x_{2,t}$ is an $m \times 1$ vector of variables that are *nonpredetermined* at time t, and u_t is a $k \times 1$ vector of exogenous variables (which can be deterministic or stochastic). A and B are constant $(n+m)\times(n+m)$ and $(n+m)\times k$ matrices, respectively. We assume that A has $n+m$ distinct eigenvalues, $\overline{n}$ eigenvalues inside the unit circle and $\overline{m}$ eigenvalues outside the unit circle. Note that we must have $n + m = \overline{n} + \overline{m}$. We seek a unique bounded solution, $\{x_t\}$, given that $\{u_t\}$ is a bounded sequence.[6]

[6]Blanchard and Kahn (1980) in fact prove the validity of their method for a more general class of exogenous sequences and corresponding admissible solution sequences, namely, sequences whose expectations grow slower than exponentially. They are able to do so because all we need for the forward solution to be unique is that the multidimensional analog of equation (4.14) holds.

Recall our discussion in Chapters 1 and 3 on the method of eigenvalue decomposition, which underlies the solution methods for autonomous differential and difference systems. In particular, Theorem 1.1 implies that A can be diagonalized as[7]

$$A = PDP^{-1} \tag{4.36}$$

where D is the diagonal matrix with the eigenvalues on the diagonal in increasing order according to their moduli and P is a matrix with the corresponding eigenvectors as its columns. Let P_{11} denote the top-left $n \times \overline{n}$ submatrix of P.

We summarize the key result of Blanchard and Kahn (1980) in the following theorem.

Theorem 4.4 (The Blanchard and Kahn (1980) Method)**.** *Consider the linear expectational difference equation system* (4.35), *subject to $x_{1,-1} = x_{init}$. Assume that matrix P_{11}, the top-left $n \times \overline{n}$ submatrix of P in* (4.36) *is of full rank.*

Let m denote the dimension of the vector of nonpredetermined variables at t, $x_{2,t}$, and let $\overline{m}$ denote the number of eigenvalues of A strictly outside the unit circle.

If $\overline{m} = m$ there exists a unique bounded solution to (4.35).

If $\overline{m} > m$ there exists no bounded solution.

If $\overline{m} < m$ there exists an infinity of bounded solutions.

The proof of the result, which also describes the algorithm for obtaining the unique bounded solution when $\overline{m} = m$, is given in the next subsection. The intuition behind the derivation can be stated as follows. The key to the *existence* of a bounded solution is that the system is not "too explosive" in the sense that it has sufficiently many *stable* eigenvalues (i.e. $\overline{n} \geq n \Rightarrow \overline{m} \leq m$). The key to the *uniqueness* of the bounded solution is that nonpredetermined variables are only forward-looking, in the sense that they depend on the past only indirectly through the effect of the past on the currently predetermined variables; this is ensured by the existence of sufficiently many *unstable* eigenvalues of the system (i.e. $\overline{m} \geq m$).

Remark 4.1. In many models and textbooks, system (4.34) is written as

$$x_t = S\mathbb{E}_t x_{t+1} + T u_t \tag{4.37}$$

where $S = A^{-1}$ and $T = -A^{-1}B$.

In this case, we use the fact from matrix algebra that, if λ_i is an eigenvalue of matrix A, then $\frac{1}{\lambda_i}$ is an eigenvalue of its inverse matrix, $A^{-1} = S$. Therefore, all of our results on the existence and uniqueness of bounded solutions go through but with the *opposite sign*. That is, the stable eigenvalues of S are those that

[7] In the case of a matrix with repeated eigenvalues a similar procedure is possible, where D now has the Jordan form. See subsection 1.3.1.

lie outside the unit circle and its unstable eigenvalues are those inside the unit circle.

Remark 4.2 (System Reduction). The assumption in Theorem that matrix P_{11} is of full rank (i.e. invertible when $\overline{m} = m$) requires that one applies the preliminary step of "system reduction" to the model before applying the Blanchard and Kahn (1980) method, writing the model in terms of a subset of variables that are uniquely determined. However, system reduction may not be possible or straightforward for all models. Moreover, some models take the form

$$M\mathbb{E}_t x_{t+1} = Ax_t + Bu_t \tag{4.38}$$

where the lead matrix M is not invertible, so that the model cannot be cast in the form (4.35) amenable to the Blanchard and Kahn (1980) method. For recently devised solution methods to systems of linear expectational difference equations that do not require invertibility of P_{11} and M, and also possess a number of practical advantages compared to Blanchard and Kahn (1980), see Klein (2000); King and Watson (2002); Sims (2002) and Uhlig (1999).

A related solution method for linear rational-expectations models, the *method of undetermined coefficients*, is discussed in Application 4.1 in the context of the linearized neoclassical growth model.

4.3.1 Proof of Theorem 4.4*

Consider the representation of system (4.35) corresponding to the eigenbasis,

$$\begin{bmatrix} \mathbb{E}_t z_{1,t+1} \\ \mathbb{E}_t z_{2,t+1} \end{bmatrix} = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} + Cu_t \tag{4.39}$$

where

$$\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = P^{-1} \begin{bmatrix} x_{1,t} \\ x_{2,t} \end{bmatrix} \tag{4.40}$$

$$C = P^{-1}B \tag{4.41}$$

$$= P^{-1} \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \tag{4.42}$$

and Λ_1 and Λ_2 are $\overline{n} \times \overline{n}$ and $\overline{m} \times \overline{m}$ diagonal matrices with the eigenvalues of A that are inside and outside the unit circle, respectively, on their diagonal. Also let

$$P \equiv \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \tag{4.43}$$

and

$$P^{-1} \equiv Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix} \tag{4.44}$$

The transformation (4.39) effectively "decouples" the system into two subsystems, one of which is "fully stable" (with dimension $\overline{n}$) and the other "fully explosive" (with dimension $\overline{m}$).

Let us first consider the explosive subsystem of (4.39):

$$\mathbb{E}_t z_{2,t+1} = \Lambda_2 z_{2,t} + (Q_{21} B_1 + Q_{22} B_2) u_t \tag{4.45}$$

Because a bounded solution for x_t requires a bounded solution for (every component of) z_t, the solution of the explosive subsystem must be bounded as well. Thus, proceeding exactly as we did for the unidimensional case (4.2) with $|a| > 1$, we can write the unique bounded solution of (4.45) as

$$z_{2,t} = -\sum_{j=0}^{\infty} \Lambda_2^{-(j+1)} (Q_{21} B_1 + Q_{22} B_2) \mathbb{E}_t u_{t+j} \tag{4.46}$$

for all $t \geq -1$.

Now consider the "fully stable" subsystem

$$\mathbb{E}_t z_{1,t+1} = \Lambda_1 z_{1,t} + (Q_{11} B_1 + Q_{12} B_2) u_t \tag{4.47}$$

Just like its unidimensional counterpart (4.2) with $|a| < 1$, this system has an uncountable infinity of bounded solutions. Our hope is that the model implies enough restrictions on $\{z_{1,t}\}$ so as to pin down a unique solution for this subsystem as well. Moreover, we want the model to imply just enough restrictions, otherwise no bounded solution will exist.

We have two sets of restrictions on $\{z_{1,t}\}$. First, the initial condition $x_{1,-1} = x_{\text{init}}$ on the predetermined variables at time -1 implies restrictions on $\{z_{1,-1}\}$,

$$x_{\text{init}} = P_{11} z_{1,-1} + P_{12} z_{2,-1} \tag{4.48}$$

where $z_{2,-1}$ is already pinned down from (4.46).

Second, we use the fact that $\{x_{1,t+1}\}$ is a sequence of variables predetermined at time t, so that we have the following restrictions on the relationship between $\{z_{1,t}\}$ and $\{z_{2,t}\}$.

$$\begin{aligned} 0 &= x_{1,t+1} - \mathbb{E}_t x_{1,t+1} && (4.49) \\ &= P_{11}(z_{1,t+1} - \mathbb{E}_t z_{1,t+1}) + P_{12}(z_{2,t+1} - \mathbb{E}_t z_{2,t+1}) && (4.50) \end{aligned}$$

With these two sets of restrictions in hand, we are ready to consider the three possible cases: $\overline{m} = m$, $\overline{m} > m$, and $\overline{m} < m$.

Throughout, we will assume that P_{11} is of full rank, $\text{rank}(P_{11}) = \min(n, \overline{n})$. This essentially requires the preliminary step of "system reduction" before applying the Blanchard and Kahn (1980) method, discussed in Remark 4.2.

Consider the case $\overline{m} = m$, that is, we have as many explosive eigenvalues as we have nonpredetermined variables. Then, P_{11} is an $n \times n$ matrix, which, together with the assumption that it is of full rank, implies that it is invertible. Thus, the entire sequence $\{z_{1,t}\}$ is uniquely pinned down as follows: first, $z_{2,-1}$ is pinned down from (4.46); second, $z_{1,-1}$ is pinned down from (4.48) since P_{11} is invertible; third, $\mathbb{E}_t z_{1,0}$ is pinned down from (4.47); fourth, $z_{1,0}$ is pinned down from (4.50) again because P_{11} is invertible. Thus, proceeding recursively with respect to t, we have proved that $\{z_{1,t}\}$ is uniquely determined. It is also bounded by construction, since $\{z_{2,t}\}$ is bounded. We conclude from (4.40) that the original system has a unique bounded solution.

What happens in the case $\overline{m} > m$? As soon as we reach the second step in the recursion described above we note that (4.48) imposes more than $\overline{n}$ restrictions on the $\overline{n}$-dimensional vector $z_{1,-1}$. Similarly, system (4.50), evaluated at $t = -1$, is also overdetermined. Therefore, the system generically has no bounded solution.

What happens when $\overline{m} < m$? Note that both systems of linear restrictions (4.48) and (4.50), evaluated at $t = -1$, are underdetermined. In this case, we cannot rule out any of the infinitely many bounded solutions of (4.47). Therefore, our original system also has an infinity of bounded solutions.

Application 4.1 (Solution to the Loglinearized Stochastic Neoclassical Growth Model)**.** Consider a stochastic version of the neoclassical growth model:

$$
\begin{aligned}
k_{t+1} &= f(k_t, z_t) + (1-\delta)k_t - c_t && (4.51)\\
u'(c_t) &= \mathbb{E}_t \beta \left[f_k(k_{t+1}, z_{t+1}) + (1-\delta) \right] u'(c_{t+1}) && (4.52)\\
\ln z_{t+1} &= \rho \ln z_t + \varepsilon_{t+1} && (4.53)
\end{aligned}
$$

where $\beta, \rho < 1$ and $\delta \leq 1$, subject to the initial conditions $k_0 = k_{\text{init}} > 0$ and $z_0 = z_{\text{init}}$, and the transversality condition

$$
\lim_{t\to\infty} \mathbb{E}\left\{ \beta^t \left[f_k(k_t, z_t) + (1-\delta) \right] u'(c_t) k_t | z_0 = z_{\text{init}} \right\} = 0 \tag{4.54}
$$

We assume the usual parametric forms:

$$
\begin{aligned}
u(c_t) &= \frac{c_t^{1-\gamma}}{1-\gamma} && (4.55)\\
f(k_t, z_t) &= z_t k_t^{\alpha} && (4.56)
\end{aligned}
$$

where $\gamma > 0$ and $\alpha \in (0,1)$, that is, CRRA utility and Cobb-Douglas production technology. For simplicity, we also assume that capital depreciates fully every period, $\delta = 1$.

We are looking for a unique bounded solution to the loglinearized versions of the equilibrium conditions around the model's steady state. The loglinear approximation to (4.51)-(4.52) (see application 6.1 for its derivation) under our parametric assumptions is given by

$$\hat{k}_{t+1} = \frac{\overline{y}}{\overline{k}}\left[\hat{z}_t + \alpha\hat{k}_t\right] - \frac{\overline{c}}{\overline{k}}\hat{c}_t \tag{4.57}$$

$$\mathbb{E}_t\hat{c}_{t+1} = \hat{c}_t + \frac{1}{\gamma}\left[\mathbb{E}_t\hat{z}_{t+1} - (1-\alpha)\hat{k}_{t+1}\right] \tag{4.58}$$

$$\mathbb{E}_t\hat{z}_{t+1} = \rho\hat{z}_t \tag{4.59}$$

where $\overline{x}$ denotes the steady state value of a variable x and $\hat{x}_t \equiv \ln x_t - \ln\overline{x}$ denotes the relative (percentage) deviation of variable x from its steady state value.

In order to use the Blanchard and Kahn (1980) method we need to cast the model into the form given by (4.35). After some algebra we find

$$\begin{bmatrix}\hat{k}_{t+1}\\ \mathbb{E}_t\hat{c}_{t+1}\end{bmatrix} = \overbrace{\begin{bmatrix}\alpha\frac{\overline{y}}{\overline{k}} & -\frac{\overline{c}}{\overline{k}}\\ -\frac{\alpha(1-\alpha)}{\gamma\overline{k}} & 1+\frac{(1-\alpha)\overline{c}}{\gamma\overline{k}}\end{bmatrix}}^{\equiv A}\begin{bmatrix}\hat{k}_t\\ \hat{c}_t\end{bmatrix} + \overbrace{\begin{bmatrix}\frac{\overline{y}}{\overline{k}}\\ \frac{\rho\overline{k}-(1-\alpha)\overline{y}}{\gamma\overline{k}}\end{bmatrix}}^{\equiv B}\hat{z}_t \tag{4.60}$$

where expectations of the exogenous process (the technology shock) evolve according to $\mathbb{E}_t\hat{z}_{t+1} = \rho\hat{z}_t$. [8]

We know from Theorem 4.4 that the system has a unique and bounded solution if and only if matrix A has one eigenvalue strictly inside the unit circle and one eigenvalue strictly outside the unit circle. It can be shown that this is the case for all allowed values of the parameters α, β, and γ. To solve for the system in this case, we can follow the procedure outlined in subsection 4.3.1: transform the system into its representation in terms of its eigenbasis, and solve recursively for the transformed and the original variables. As these steps involve only tedious algebraic manipulations, we do not cover them here.

We now discuss a popular alternative (but related) method to solve system (4.60) called *the method of undetermined coefficients*. Note that existence of a bounded solution by definition requires that capital remains bounded (in the sense of footnote 2). Also recall from the Blanchard and Kahn (1980) approach that uniqueness of the bounded solution requires that variables not predetermined at time t (in our case, consumption) are purely forward-looking in the sense that they depend on the past only indirectly through the current values of the predetermined variables. This motivates us to conjecture that equilibrium consumption is a function only of current

[8] Note that in this model the requirement of system reduction discussed in Remark 4.2 implies that one should not include investment and output in addition to the two main variables, consumption and capital, as the solutions of the former are directly pinned down from the solutions to the sequences of the two main variables.

variables, $c_t = g(k_t, z_t)$. Function g is known as the equilibrium *policy function* for consumption.

Furthermore, given the linearity of system (4.60) it is natural to conjecture that function g is linear so that

$$c_t = a_{ck}k_t + a_{cz}z_t \tag{4.61}$$

for some coefficients a_{ck} and a_{cz}. Also note that once we know $\{c_t\}$ we also know $\{k_{t+1}\}$ from (4.60) and the initial conditions. This implies that the policy function for k_{t+1} depends only on k_t and z_t as well. By the same logic as above, we also conjecture that it is a linear function of these two variables,

$$k_{t+1} = a_{kk}k_t + a_{kz}z_t \tag{4.62}$$

for some coefficients a_{kk} and a_{kz}.

Coefficients a_{ck}, a_{cz}, a_{kk}, and a_{kz} are "undetermined." After plugging in our conjecture to (4.60) for a given t, we will end up with a system of the form

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ \hat{z}_t \end{bmatrix} = 0 \tag{4.63}$$

where b_{ij} are (possibly nonlinear) functions of the coefficients a_{ck}, a_{cz}, a_{kk}, and a_{kz}.

At this point, we note that, if our conjecture is correct, the system above must hold for all possible values of $\hat{k}_t$ and $\hat{z}_t$. Therefore, each b_{ij} must be zero. We thus obtain a (nonlinear) system of 4 algebraic equations in 4 unknowns (the coefficients a_{ck}, a_{cz}, a_{kk}, and a_{kz}).

It will turn out that one of the resulting equations will be quadratic in a_{kk}, which has two roots. We must pick the root with absolute value less than 1, since we are looking for a bounded solution, as discussed above.

For a formal description and application of the method of undetermined coefficients to more complex problems, see Uhlig (1999). ■

Application 4.2 (Equilibrium Determinacy and Monetary Policy*)**.** Consider the loglinearized New Keynesian model in discrete time:

$$\mathbb{E}_t\tilde{y}_{t+1} = \tilde{y}_t + \sigma^{-1}\left(i_t - \mathbb{E}_t\pi_{t+1} - r_t^n\right) \tag{4.64}$$

$$\mathbb{E}_t\pi_{t+1} = \beta^{-1}\left[\pi_t - \kappa\tilde{y}_t\right] \tag{4.65}$$

Here, the two endogenous variables are the output gap, $\tilde{y}_t$, and inflation, π_t. The system is affected by exogenous shocks to the natural interest rate, $\{r_t^n\}$, which we assume to be a bounded sequence, and by an "exogenous" (with respect to private agents) sequence $\{i_t\}$ of the value of the nominal interest rate every period, which the monetary authority is assumed to control. We know (from the economics of the problem) that the output gap and inflation are *not predetermined* at a given time period.

First, assume that the monetary authority simply sets the nominal interest rate per period, so that $i_t = i_t^*$, where i_t^* is some target for the monetary authority. We will not be concerned with what the optimal target is. In order to see whether our system has a unique bounded solution, so that the rational-expectations equilibrium of the New Keynesian model is locally unique, let us write (4.64)-(4.65) in matrix form:

$$\begin{bmatrix} \mathbb{E}_t \tilde{y}_{t+1} \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \overbrace{\frac{1}{\beta}\begin{bmatrix} \beta + \frac{\kappa}{\sigma} & -\frac{1}{\sigma} \\ -\kappa & 1 \end{bmatrix}}^{\equiv A} \begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} + \frac{1}{\sigma}\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} i_t^* \\ r_t^n \end{bmatrix} \tag{4.66}$$

From Theorem 4.4, we know that since we have $n = 0$ and $m = 2$, a unique bounded solution exists if we have $\overline{m} = 2$, that is, if both eigenvalues of A are strictly outside the unit circle. It turns out, however, that one of the eigenvalues of A lies inside the unit circle,[9] $1 = \overline{m} < m$, and we have an infinity of bounded solutions, so that equilibrium is indeterminate even locally. This is the "classical" determinacy problem associated with monetary policy, first observed by Sargent and Wallace (1975).

Instead, consider a monetary policy rule, often called the *Taylor rule*, of the form

$$i_t = i_t^* + \phi_\pi \pi_t \tag{4.67}$$

We want to see if such a rule helps us achieve determinacy. In terms of the mathematics, are there values of ϕ_π such that both eigenvalues of matrix A are outside the unit circle?

Our system now becomes

$$\begin{bmatrix} \mathbb{E}_t \tilde{y}_{t+1} \\ \mathbb{E}_t \pi_{t+1} \end{bmatrix} = \overbrace{\frac{1}{\beta}\begin{bmatrix} \beta + \frac{\kappa}{\sigma} & \frac{\beta\phi_\pi - 1}{\sigma} \\ -\kappa & 1 \end{bmatrix}}^{\equiv B} \begin{bmatrix} \tilde{y}_t \\ \pi_t \end{bmatrix} + \frac{1}{\sigma}\begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} i_t^* \\ r_t^n \end{bmatrix} \tag{4.68}$$

where the only difference of matrix B from matrix A is the inclusion of a term in ϕ_π on the top-right element of B.

We want to know which values, if any, of ϕ_π correspond to A having both of its eigenvalues outside of the unit circle. As is clear from Figure 3.3, it turns out to be much easier to derive the region where both eigenvalues are inside the unit circle. To use this shortcut, we can make use of Remark 4.1 and look at the inverse of matrix B,

$$B^{-1} = \frac{1}{\sigma + \kappa\phi_\pi}\begin{bmatrix} \sigma & 1 - \beta\phi\pi \\ \sigma\kappa & \kappa + \beta\sigma \end{bmatrix} \tag{4.69}$$

with

$$tr(A) = \frac{\sigma + \kappa + \beta\sigma}{\sigma + \kappa\phi_\pi} \tag{4.70}$$

$$\det(A) = \frac{\beta\sigma}{\sigma + \kappa\phi_\pi} \tag{4.71}$$

[9] This can be seen as a special case of the derivation below with $\phi_\pi = 0$.

and use the three necessary and sufficient conditions in Lemma 3.2. We get the three corresponding restrictions

$$\phi_\pi > -\frac{\sigma(1-\beta)}{\kappa} \tag{4.72}$$

$$\phi_\pi > 1 \tag{4.73}$$

$$\beta\sigma > -\kappa(1+\phi_\pi) - \sigma(2+\beta) \tag{4.74}$$

Since κ, σ, and β are positive, the third condition places no restrictions, while the first restriction is implied by the second. We conclude that the equilibrium is locally unique (around the deterministic steady state) if and only if

$$\phi_\pi > 1 \tag{4.75}$$

In the simplest version of the New Keynesian model studied here the steady-state (and optimal) value of inflation is zero. Then, the determinacy rule (4.75) says that monetary policy must promise to "more than" offset any deviation of inflation from its optimal value of zero via a change in the nominal interest rate.

Importantly, this response will *not* be observed on the equilibrium path. It is an off-equilibrium response or "threat" that is present only to rule out other (bounded) equilibria. An implication of this is that, in principle, we cannot identify empirically the Taylor rule coefficient, ϕ_π.

This fact contrasts sharply with the predictions of "old Keynesian" models with backward-looking (and "non-microfoundable") equations of the form

$$\begin{bmatrix} y_t \\ \pi_t \end{bmatrix} = C \begin{bmatrix} y_{t-1} \\ \pi_{t-1} \end{bmatrix} + D \begin{bmatrix} i_t^* \\ u_t \end{bmatrix} \tag{4.76}$$

Note that expectational terms are absent. This is a linear difference system of the kind that we studied in Chapter 3.[10] When monetary policy follows the same Taylor rule, (4.67), it is usually the case that an "active" monetary policy, $\phi_\pi > 1$, is associated with a matrix C with eigenvalues *inside* the unit circle. This implies that the system follows stable dynamics in equilibrium (see Remark 3.1). For example, if i_t^* and u_t are constant over time, the system will converge asymptotically to its steady state.

Therefore, $\phi_\pi > 1$ in (4.76) ensures *stable* dynamics *in equilibrium*. There is no issue of indeterminacy since there is a unique (global) equilibrium path, as we know from Theorem 3.3. In contrast, in New Keynesian models *unstable* dynamics *off the equilibrium path* ensure local determinacy.

Let us finally return to the loglinearized New Keynesian model (4.64)-(4.65) and briefly discuss global determinacy (uniqueness) of equilibrium in the exact (nonlinear) model. Consistent with our discussion in section 4.2, an "active" Taylor rule and

[10] The presence of potentially stochastic *exogenous* variables, i_t^* or u_t, does not affect this conclusion.

the associated unstable eigenvalues of the system only ensure local determinacy, that is, uniqueness of the *bounded* solutions. As discussed in section 4.2, transversality conditions usually rule out explosive equilibrium paths for real variables, but this is not the case for nominal variables at the benchmark New Keynesian economy (an economy at the "cashless" limit).

New Keynesian economists have offered a number of responses. First, Woodford (2003) suggests a behavioral equilibrium selection criterion, motivated by the fact that large expectational adjustments, such as those associated with explosive paths, are implausible. A second, more direct response is to assume that the government pairs a monetary policy based on an active Taylor rule with other types of monetary or fiscal policies that ensure that equilibrium is impossible under explosive inflation or deflation.[11] Again, all of these policies serve to rule out other equilibria, and the government will never need to actually implement them in equilibrium.[12] ■

[11]For example, the government could lower taxes so much during a low-inflation (deflation) state that the real value of public debt explodes, so that the consumers' transversality condition cannot be satisfied.

[12]For more on equilibrium determinacy in the New Keynesian model, see Woodford (2003), Cochrane (2011), and references therein.

Problem Set 3

1. (Second-Order Expectational Difference Equation) Consider the second-order expectational difference equation

$$a_1 x_{t-1} + a_2 x_t + a_3 \mathbb{E}_t x_{t+1} = u_t$$

subject to the initial condition $x_{-1} = x_{\text{init}}$, where $x_t, u_t \in \mathbb{R}$ and $\{u_t\}$ is a bounded exogenous process. You are asked to derive the condition on the coefficients a_1, a_2, and a_3 under which equation (4.77) has a unique bounded solution.

   (i) Transform the equation into a first-order system

$$\mathbb{E}_t y_{t+1} = A y_t + B u_t$$

   where $y_t \equiv \begin{bmatrix} x_t \\ x_{t-1} \end{bmatrix}$.

   (ii) Compute the characteristic polynomial of matrix A, $p_A(\lambda)$.

   (iii) How many elements of vector y_t are predetermined at time t? What conditions does the Blanchard-Kahn method (Theorem 4.4) impose on the roots of $p_\lambda(A)$ in order for the system to have a unique bounded solution? (You do not need to solve explicitly for the restrictions in terms of the coefficients.)

2. (Stochastic Lagrange Multipliers) Consider the monetary policy problem

$$\min_{\{x_t,\pi_t\}_{t=0}^{\infty}} \frac{1}{2}\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t(\pi_t^2 + a_x x_t^2) \tag{4.77}$$

for given π_{-1} and x_{-1}, subject to the sequence of constraints

$$\pi_t = \beta\mathbb{E}_t\pi_{t+1} + \kappa x_t + u_t \tag{4.78}$$

for all $t \geq 0$, where $\{u_t\}$ is an exogenous stochastic process. The exact realization of u_t is not revealed until time t. The optimal policies x_t^* and π_t^* are allowed to depend on the history of the exogenous process up to time t.

(i) State the Lagrangian for this problem.

(ii) State the first-order conditions for x_t and π_t.

(iii) It can be shown that the policies defined by the first-order conditions satisfy certain conditions (including a transversality condition) ensuring that they are indeed the optimal policies x_t^* and π_t^*. Combine the first-order conditions to eliminate the Lagrange multipliers and derive a relationship between the optimal policies of the form

$$x_t^* = c\hat{p}(t) \tag{4.79}$$

where $\hat{p}(t) \equiv \pi_0^* + \pi_1^* + \ldots + \pi_t^*$ and c is a function of the parameters of the problem. What is c?

Chapter 5

Dynamic Optimization: Optimal Control Theory

5.1 The Lagrange Method in Discrete Time

Recall the method of *Lagrange multipliers* from the micro part of math camp. As a first example, consider a version of the Kuhn-Tucker theorem for finite-dimensional problems with equality constraints:

$$\max_{x \in \mathbb{R}^K} f(x) \tag{5.1}$$

subject to

$$g(x) = 0 \tag{5.2}$$

where $f : \mathbb{R}^K \to \mathbb{R}^N$, $g : \mathbb{R}^K \to \mathbb{R}^M$, for some $K, N, M \in \mathbb{N}$, and f and g are differentiable.

The Kuhn-Tucker theorem tells us that if there exists an interior solution to this problem and provided a certain regularity condition is satisfied (the constraint qualification condition) we can convert the problem into an unconstrained optimization problem by constructing the Lagrangian function

$$\mathcal{L}(x, \lambda) = f(x) + \lambda \cdot g(x) \tag{5.3}$$

where $\lambda \in \mathbb{R}^M$ are the Lagrange multipliers. The first-order conditions of this unconstrained problem are

$$\begin{aligned} D_x f(x^*) + \lambda^* \cdot D_x g(x^*) &= 0 && (5.4) \\ g(x^*) &= 0 && (5.5) \end{aligned}$$

where the second line is the first-order condition with respect to the additional set of variables of the unconstrained problem, the vector of Lagrange multipliers.

The saddle-point and Kuhn-Tucker theorems provide necessary and sufficient conditions for a wide range of constrained optimization problems. Importantly, the Lagrange multipliers have an intuitive and useful interpretation as the shadow value of relaxing their respective constraints. The Lagrange multiplier method is extremely powerful; it has its foundations in duality theory and has generalized counterparts that provide existence and uniqueness results for problems where the argument with respect to which the maximization or minimization takes place is an infinite-dimensional vector or even a function.[1]

In this chapter we study dynamic optimization problems. Consider the following general formulation of an infinite-horizon, deterministic optimization problem:

$$\max_{\{x_t,y_t\}_{t=0}^{\infty}} W(\{x_t\},\{y_t\}) = \sum_{t=0}^{\infty} f(t,x_t,y_t) \tag{5.6}$$

such that

$$x_{t+1} = g(t,x_t,y_t) \quad \forall t \geq 0 \tag{5.7}$$

where $x_t \in \mathcal{X} \subseteq \mathbb{R}$, $y_t \in \mathcal{Y} \subseteq \mathbb{R}$, and given the boundary conditions $x_0 = x_{init}$ and $\lim_{\to\infty} b_t x_t \geq \overline{x} \in \mathbb{R}$.

Assuming that appropriate regularity conditions are satisfied, we can construct the Lagrangian function of this problem:

$$\begin{aligned}
&\max_{\{x_t,y_t\}_{t=0}^{\infty}} \min_{\{\lambda_t\}_{t=0}^{\infty}} \mathcal{L}(\{x_t\},\{y_t\},\{\lambda_t\}) \\
&\qquad = \sum_{t=0}^{\infty} f(t,x_t,y_t) + \sum_{t=0}^{\infty} \lambda_t (g(t,x_t,y_t) - x_{t+1})
\end{aligned} \tag{5.8}$$

where $x_t \in \mathcal{X} \subseteq \mathbb{R}$, $y_t \in \mathcal{Y} \subseteq \mathbb{R}$, $\lambda_t \in \mathbb{R}$, and given the boundary conditions $x_0 = x_{init}$ and $\lim_{\to\infty} b_t x_t \geq \overline{x} \in \mathbb{R}$.

Note that we have turned our constrained optimization problem into a formally unconstrained one. The Lagrange multipliers λ_t are not arbitrary but take the values that ensure that the constraint $x_{t+1} = g(t,x_t,y_t)$ is satisfied for all $t \geq 0$. In the original formulation the constraint on the law of motion of x_t captures the intertemporal

[1] As one would expect, there are also limitations when it comes to infinite-dimensional problems. A well-known issue is related to the fact that in optimization problems involving an infinite number of commodities (time periods), the corresponding sequence of Lagrange multipliers may not be bounded, that is, it may "place all of its weight at infinity." The most important example of this issue in economics is the literature on the existence of equilibrium in infinite-dimensional economies (see chapter 15 of Stokey and Lucas (1989) for an excellent exposition). In that case, the question is essentially whether the Lagrange multipliers of the social planner's problem can function as market-clearing equilibrium prices. Care should be taken to prove the existence of a bounded sequence of multipliers that has the usual inner-product representation. In the case of infinite-horizon problems this is usually accomplished by assuming that the future is sufficiently discounted, which ensures that an appropriate transversality condition is satisfied.

tradeoffs that the agent faces; in the Lagrangian formulation it is the law of motion of the Lagrange multipliers that captures these dynamic tradeoffs.

The first-order conditions to problem 5.8 are[2]

$$
\begin{aligned}
&\text{wrt } y_t: \quad && f_y(t,x_t,y_t) + \lambda_t g_y(t,x_t,y_t) = 0 && \forall\, t \geq 0 && (5.9)\\
&\text{wrt } x_t: \quad && f_x(t,x_t,y_t) + \lambda_t g_x(t,x_t,y_t) - \lambda_{t-1} = 0 && \forall\, t \geq 1 && (5.10)\\
& && f_x(0,x_0,y_0) + \lambda_0 g_x(0,x_0,y_0) = 0 && t = 0 && (5.11)\\
&\text{wrt } \lambda_t: \quad && x_{t+1} = g(t,x_t,y_t) && \forall\, t \geq 0 && (5.12)
\end{aligned}
$$

For reasons that will become clear in the next section, it is useful to reformulate the Lagrangian in term of a function H as follows:

$$
\begin{aligned}
\max_{\{x_t,y_t\}_{t=0}^{\infty}} \min_{\{\lambda_t\}_{t=0}^{\infty}} & \mathcal{L}(\{x_t\},\{y_t\},\{\lambda_t\})\\
&= \sum_{t=0}^{\infty} f(t,x_t,y_t) + \sum_{t=0}^{\infty} \lambda_t (g(t,x_t,y_t) - x_{t+1})\\
&= \sum_{t=0}^{\infty} [f(t,x_t,y_t) + \lambda_t (g(t,x_t,y_t) - x_{t+1})]\\
&= \sum_{t=0}^{\infty} [f(t,x_t,y_t) + \lambda_t (g(t,x_t,y_t) - x_t) - \lambda_t (x_{t+1} - x_t)]\\
&= \sum_{t=0}^{\infty} [H(t,x_t,y_t,\lambda_t) - \lambda_t (x_{t+1} - x_t)] \qquad (5.13)
\end{aligned}
$$

where

$$H(t,x_t,y_t,\lambda_t) \equiv f(t,x_t,y_t) + \lambda_t (g(t,x_t,y_t) - x_t) \qquad (5.14)$$

The first order conditions can be restated in terms of function H as

$$
\begin{aligned}
&\text{wrt } y_t: \quad && H_y(t,x_t,y_t,\lambda_t) = 0 && \forall\, t \geq 0 && (5.15)\\
&\text{wrt } x_t: \quad && \lambda_t - \lambda_{t-1} = -H_x(t,x_t,y_t,\lambda_t) && \forall\, t \geq 1 && (5.16)\\
& && \lambda_0 = -H_x(0,x_0,y_0,\lambda_0) && t = 0 && (5.17)\\
&\text{wrt } \lambda_t: \quad && x_{t+1} = g(t,x_t,y_t) && \forall\, t \geq 0 && (5.18)
\end{aligned}
$$

5.2 The Optimal Control Problem

Our main focus in this chapter is a continuous-time optimization problem called the *optimal control problem*. The dynamic constraint in this problem takes the form of a differential equation. The problem is:

[2] In fact, as in every infinite-horizon optimization problem, there exists a fourth optimality condition, the transversality condition, restricting the behavior of the optimal policy as $t \to \infty$. See section 5.3 for a discussion of this condition.

$$\max_{[x(t),y(t)]_{t=0}^{\infty}} W(x(t),y(t)) \equiv \int_0^{\infty} f(t,x(t),y(t))\,\mathrm{d}t \tag{5.19}$$

subject to the constraints

$$\dot{x}(t) = g(t,x(t),y(t)) \tag{5.20}$$

and

$$x(t) \in \mathcal{X} \tag{5.21}$$

$$y(t) \in \mathcal{Y} \tag{5.22}$$

for all $t \geq 0$,[3] where $\mathcal{X} \subset \mathbb{R}$ and $\mathcal{Y} \subset \mathbb{R}$ are nonempty and convex, and given the boundary conditions

$$x(0) = x_0 \in \mathbb{R} \tag{5.23}$$

$$\lim_{t\to\infty} b(t)x(t) \geq x_1 \in \mathbb{R} \tag{5.24}$$

where $b: \mathbb{R}_+ \to \mathbb{R}_+$, $\lim_{t\to\infty} b(t)$ exists and is finite. Moreover, f and g are continuously differentiable functions of x, y and t.

Function $x(t)$ is called the *state variable* and $y(t)$ the *control variable* of problem (5.19)-(5.24). Also note that (5.24) is a feasibility (in certain settings, no-Ponzi) condition, not a transversality condition.

An *admissible* policy pair $(x(t),y(t))$ is defined to be a pair satisfying all conditions (5.20)-(5.24).[4] A *solution* $(\hat{x}(t),\hat{y}(t))$ is defined to be a pair of functions of t such that

$$W(\hat{x}(\cdot),\hat{y}(\cdot)) \geq W(x(\cdot),y(\cdot)) \tag{5.25}$$

for any other admissible pair $(x(\cdot),y(cdot))$.

The *Maximum Principle* states that if problem (5.19)-(5.24) has an *interior* solution, that is, if $\hat{x}(t) \in \mathrm{Int}\mathcal{X}$ and $\hat{y}(t) \in \mathrm{Int}\mathcal{Y}$, then one can reformulate problem (5.19)-(5.24) as the "unconstrained" problem [5]

$$\max_{[y(t)]_{t=0}^{\infty}} \int_0^{\infty} H(t,\hat{x}(t),y(t),\lambda(t))\,\mathrm{d}t \tag{5.26}$$

where

$$H(t,x(t),y(t),\lambda(t)) \equiv f(t,x(t),y(t)) + \lambda(t)g(t,x(t),y(t)) \tag{5.27}$$

---

[3]Formally, $y(t)$ must be a Lebesgue measurable function and $x(t)$ an absolutely continuous function.

[4]This includes the requirements on $y(t)$ and $x(t)$ outlined in footnote 3.

[5]Note that in (5.26) we maximize only with respect to $[y(t)]_{t=0}^{\infty}$; the Hamiltonian is evaluated at the optimal policy for the state variable, $[\hat{x}(t)]_{t=0}^{\infty}$, and at the process $[\lambda(t)]_{t=0}^{\infty}$ associated with the optimal policy.

for a continuously differentiable function $\lambda(t)$ satisfying certain conditions.

We have effectively transformed our original dynamic problem into a sequence of static problems, since (5.26) implies that, at each t, $\hat{y}(t)$ maximizes the Hamiltonian at time t:

$$H(t, \hat{x}(t), \hat{y}(t), \lambda(t)) \geq H(t, \hat{x}(t), y(t), \lambda(t)) \tag{5.28}$$

It also follows that $\hat{y}(t)$ satisfies the usual first-order condition since we have assumed an interior solution:

$$H_y(t, \hat{x}(t), \hat{y}(t), \lambda(t)) = 0 \tag{5.29}$$

$$\Leftrightarrow f_y(t, \hat{x}(t), \hat{y}(t)) + \lambda(t) g_y(t, \hat{x}(t), \hat{y}(t)) = 0 \tag{5.30}$$

Function H is called the *Hamiltonian* of problem (5.19)-(5.24), and function λ is called the *costate variable* associated with the solution pair $(\hat{x}(t), \hat{y}(t))$. We will see that the dynamic tradeoffs inherent in the original problem are now captured by the dynamics of the costate variable.

Observe the similarities between problem (5.19)-(5.24) and the discrete-time problem (5.6)-(5.7) of the previous section. The first-order condition with respect to the control variable, (5.29), coincides with condition (5.15) with the discrete-time problem. It is therefore clear that the formulation of discrete-time problems in terms of the Lagrangian and the formulation of optimal control problems in terms of the Hamiltonian are profoundly related to each other.[6]

5.3 Necessary and Sufficient Conditions for Optimality

Let us define the *value function* of problem (5.19)-(5.24) as

$$V(t_0, x_0) = \sup_{[x(t), y(t)]_{t=0}^{\infty}} \int_{t_0}^{\infty} f(t, x(t), y(t))\, \mathrm{d}t \tag{5.31}$$

subject to $x(t_0) = x_0$, (5.20)-(5.22), and (5.24).

That is, the value function $V(t_0, x(t_0))$ gives the optimal value of the dynamic optimization problem starting at time t_0 with state variable $x(t_0)$.[7] We focus on cases where the value function is finite (otherwise, the problem is economically uninteresting).

[6] In fact, the type of problem considered in this chapter, and in economics more generally, is referred to as the "Lagrange problem" in optimal control theory.

[7] This value function is the link between optimal control theory and continuous-time dynaming programming theory, which you will study in the first quarter of the macro sequence. The so-called Hamilton-Jacobi-Bellman (HJB) equation can be proved as a direct corollary of the Maximum Principle, Theorem 5.1.

We are now ready to formally state the Maximum Principle, providing *necessary* conditions for an interior solution to problem (5.19)-(5.24).

Theorem 5.1 (Pontryagin's Maximum Principle)**.** *Suppose that problem* (5.19)-(5.24) *has a piecewise continuous interior solution* $(\hat{x}(t), \hat{y}(t)) \in Int\mathcal{X}\times\mathcal{Y}$. *Let* $H(t,x,y,\lambda)$ *be as defined in* (5.27)*. Then, given* $(\hat{x}(t), \hat{y}(t))$*, the Hamiltonian* $H(t,x,y,\lambda)$ *satisfies the Maximum Principle*

$$H(t,\hat{x}(t),\hat{y}(t),\lambda(t)) \geq H(t,\hat{x}(t),y(t),\lambda(t)) \tag{5.32}$$

for all $y(t) \in \mathcal{Y}$ *and for all* $t \in \mathbb{R}_+$.

Moreover, for all $t \in \mathbb{R}_+$ *for which* $\hat{y}(t)$ *is continuous, the following necessary conditions are satisfied:*

$$H_y(t,\hat{x}(t),\hat{y}(t),\lambda(t)) = 0 \tag{5.33}$$

$$\dot{\lambda}(t) = -H_x(t,\hat{x}(t),\hat{y}(t),\lambda(t)) \tag{5.34}$$

$$\dot{x}(t) = H_\lambda(t,\hat{x}(t),\hat{y}(t),\lambda(t)) \tag{5.35}$$

with $\hat{x}(0) = x_0$ *and* $\lim_{t\to\infty} b(t)\hat{x}(t) \geq x_1$.

Additionally, suppose that the value function $V(t,\hat{x}(t))$ *is differentiable in* x *and* t *for* t *sufficiently large, and that* $\lim_{t\to\infty} \partial V(t,\hat{x}(t))/\partial t = 0$. *Then, the pair* $(\hat{x}(t), \hat{y}(t))$ *also satisfies the transversality condition*

$$\lim_{t\to\infty} H(t,\hat{x}(t),\hat{y}(t),\lambda(t)) = 0 \tag{5.36}$$

First-order conditions (5.33)- (5.35) are the analogous conditions to (5.15)- (5.18) for the discrete-time problem. Once again, note that the costate variable, $\lambda(t)$, appearing in the Maximum Principle equation (5.32) and in the necessary conditions (5.33)-(5.36) is associated with the particular solution pair $(\hat{x}(t), \hat{y}(t))$, as can be seen from the specification of its law of motion in necessary condition (5.34).

The intuition for necessary conditions (5.33) and (5.35) should be straightfoward, as it is exactly analogous to the first-order conditions (5.4) and (5.5). Condition (5.33) is the first-order condition of the "unconstrained" problem (5.26). Condition (5.35) is simply a restatement of the constraint $\dot{x} = g(\cdot) = H_\lambda$.

For further insight on equation (5.33) note that, just as for the Lagrange multipliers of a finite-dimensional problem, it can be shown that

$$\lambda(t) = \frac{\partial V(t,\hat{x}(t))}{\partial x} \tag{5.37}$$

so the costate variable measures the effect of a marginal increase in x on the optimal value of the problem. In other words, $\lambda(t)$ can be interpreted as the shadow value of "relaxing" the constraint (5.20) by increasing the value of $x(t)$ at time t. We can, therefore, think of the state variable $x(t)$ as a "stock" variable and of the control variable $y(t)$ as a "flow" variable.[8]

Note that the second term in the Hamiltonian $\lambda(t)g(t,x(t),y(t))$ equals $\lambda(t)\dot{x}(t)$; this is the shadow value of a marginal increase in x times a marginal increase in x. So, the Maximum Principle (equations (5.32) and (5.33)) can be interpreted as saying that the original maximization problem is equivalent to maximizing the instantaneous (flow) return $f(t,x(t),y(t))$ plus the value of the change in the stock of the state variable at each instant.

Let us now turn to equation (5.34). This necessary condition specifying the law of motion of the costate variable is unique to the optimal control problem (it has no counterpart in the finite-dimensional Lagrange problem). It states

$$\dot{\lambda}(t) = -f_x(t,\hat{x}(t),\hat{y}(t)) - \lambda(t)g_x(t,\hat{x}(t),\hat{y}(t)) \tag{5.38}$$

Since $\lambda(t)$ is the value of a (marginal) unit of the stock of the state variable, $\dot{\lambda}(t)$ can be interpreted as the appreciation in this value. A marginal increase in x affects the current flow return plus the value of the change in the stock by the total amount H_x, but it also affects the valuation of a unit of the stock by the amount $\dot{\lambda}(t)$. Then, optimality requires that, at the optimal policy, the instantaneous gain of increasing the stock at time t, H_x, should equal the loss in the value of a marginal unit of the stock over the next instant, $-\dot{\lambda}(t)$; otherwise, it would be possible to pick a different $x(t)$ (through an appropriate change in the control policy) and increase the value of the problem.[9]

The *transversality condition* is a necessary condition for an optimal policy to exist.[10] We want our problems to have a finite maximum attainable value (so that they are economically meaningful) and the transversality condition is then a consequence of the finiteness of the value function. As we saw in Application 2.1, in the neoclassical growth model the transversality condition implies that the (unique) optimal policy must lie on the saddle path at all times.

In the context of competitive equilibria, the transversality condition should be thought of as a necessary condition for the existence of equilibrium, since the latter requires the existence of an optimal policy to every agent's optimization problem. Although it can be thought of as an equilibrium condition, it is important to understand that it is not related to feasibility. Another terminal value (limiting value)

---

[8] In fact, this interpretation will be quite literal in Application 5.1.

[9] Imagine that at time t you have a stock $\hat{x}(t)$ and you consider whether to purchase an additional quantity dx of the stock. You will have to pay $\lambda(t)dx$ to buy the stock at time t; you will gain $H_x dxdt$ from using the stock over the next instant dt; and you will also be left with this additional stock at $t+dt$, at which time it will be worth $\lambda(t+dt)dx$. You will decide not to purchase the additional amount of stock (i.e. you are at an optimum at $\hat{x}(t)$) if your net gain, $(H_x dt + \lambda(t+dt) - \lambda(t))dx = \left(H_x + \dot{\lambda}(t)\right)dxdt$, is equal to zero.

[10] Note that condition $\lim_{t\to\infty} \partial V(t,\hat{x}(t))/\partial t = 0$ in the statement of Theorem 5.1 is a fairly weak assumption, as it is satisfied for all economically interesting problems. It is only slightly stronger than assuming that $\lim_{t\to\infty} V(t,\hat{x}(t))$ is finite.

condition, in our case condition (5.24), should be imposed to restrict the set of admissible policies to those that are economically feasible. In the context of competitive equilibria, the feasibility constraint usually takes the form of a no-Ponzi (or natural debt limit) condition.

We now discuss the special version of problem (5.19)-(5.24) that is most common in economic applications: *exponentially-discounted infinite-horizon problems*, of the form

$$\max_{[x(t),y(t)]_{t=0}^{\infty}} W(x(t),y(t)) \equiv \int_0^{\infty} \exp(-\rho t) f(x(t),y(t))\,\mathrm{d}t \tag{5.39}$$

subject to the same constraints and boundary conditions (5.20)-(5.24).[11]

The Hamiltonian in this case is:

$$\begin{aligned} H(t,x(t),y(t),\lambda(t)) &= \exp(-\rho t) f(x(t),y(t)) + \lambda(t) g(t,x(t),y(t)) && (5.40)\\ &= \exp(-\rho t)\left[f(x(t),y(t)) + \mu(t) g(t,x(t),y(t))\right] && (5.41)\end{aligned}$$

where

$$\mu(t) \equiv \exp(\rho t)\lambda(t) \tag{5.42}$$

It is, therefore, convenient to work with the *current value Hamiltonian*, defined as

$$\hat{H}(t,x(t),y(t),\mu(t)) = \left[f(x(t),y(t)) + \mu(t) g(t,x(t),y(t))\right] \tag{5.43}$$

Theorem 5.2 (Maximum Principle for Exponentially Discounted Problems). *Suppose that the problem defined by* (5.39) *subject to* (5.20)-(5.24) *has a piecewise continuous interior solution* $(\hat{x}(t),\hat{y}(t)) \in Int\mathcal{X}(t)\times\mathcal{Y}(t)$.

Suppose that the value function $V(t,\hat{x}(t))$ *is differentiable in* x *and* t *for* t *sufficiently large, that* $V(t,\hat{x}(t))$ *exists and is finite for all* t, *and that* $\lim_{t\to\infty} \partial V(t,\hat{x}(t))/\partial t = 0$.

Let $\hat{H}(t,x,y,\mu)$ *be as defined in* (5.43). *Then, given* $(\hat{x}(t),\hat{y}(t))$, *the current-value Hamiltonian* $\hat{H}(t,x,y,\mu)$ *satisfies the Maximum Principle*

$$\hat{H}(t,\hat{x}(t),\hat{y}(t),\mu(t)) \geq \hat{H}(t,\hat{x}(t),y(t),\mu(t)) \tag{5.44}$$

for all $y(t)\in\mathcal{Y}$ *and for all* $t\in\mathbb{R}$.

Moreover, for all $t\in\mathbb{R}_+$ *for which* $\hat{y}(t)$ *is continuous, the following necessary conditions are satisfied:*

$$\begin{aligned} &\hat{H}_y(t,\hat{x}(t),\hat{y}(t),\mu(t)) = 0 && (5.45)\\ &\dot{\mu}(t) - \rho\mu(t) = -\hat{H}_x(t,\hat{x}(t),\hat{y}(t),\mu(t)) && (5.46)\\ &\dot{x}(t) = \hat{H}_\mu(t,\hat{x}(t),\hat{y}(t),\mu(t)) && (5.47)\end{aligned}$$

[11] For this type of problems, we can allow the codomain of x and y to vary over time, that is, we can replace (5.21) and (5.22) with $x(t)\in\mathcal{X}(t)$ and $x(t)\in\mathcal{Y}(t)$.

with $\hat{x}(0) = x_0$ *and* $\lim_{t\to\infty} b(t)\hat{x}(t) \geq x_1$.

Also, the following transversality *condition holds*

$$\lim_{t\to\infty}\left\{\exp(-\rho t)\hat{H}(t,\hat{x}(t),\hat{y}(t),\mu(t))\right\} = 0 \tag{5.48}$$

Note that condition (5.46) is simply a reformulation of condition (5.34), since

$$\begin{aligned}\dot{\lambda}(t) &= \frac{d[\exp(-\rho t)\mu(t)]}{dt} &(5.49)\\ &= \exp(-\rho t)[\dot{\mu}(t) - \rho\mu(t)] &(5.50)\\ &= -\exp(-\rho t)\hat{H}_x(t,\hat{x}(t),\hat{y}(t),\mu(t)) &(5.51)\\ &= -H_x(t,\hat{x}(t),\hat{y}(t),\lambda(t)) &(5.52)\end{aligned}$$

where the third line is condition (5.34) and the fourth line follows from the definition of the current-value Hamiltonian.[12]

Remark 5.1 (Transversality Conditions). The necessary transversality condition (5.48),

$$\lim_{t\to\infty}\left\{\exp(-\rho t)\hat{H}(t,\hat{x}(t),\hat{y}(t),\mu(t))\right\} = 0$$

is usually hard to check.

A stronger transversality condition that implies (5.48) and is usually satisfied in economic applications is

$$\lim_{t\to\infty}[\exp(-\rho t)\mu(t)\hat{x}(t)] = 0 \tag{5.53}$$

To see that (5.53) implies (5.48), note that

$$V(t,\hat{x}(t)) = \int_t^\infty \exp(-\rho s) f(\hat{x}(s),\hat{y}(s))\,\mathrm{d}s \tag{5.54}$$

Since Theorem 5.2 assumes that $\lim_{t\to\infty} V(t,\hat{x}(t))$ exists and finite, it must be that

$$\lim_{t\to\infty}\exp(-\rho t) f(\hat{x}(t),\hat{y}(t)) = 0 \tag{5.55}$$

Thus, (5.48) is equivalent to

$$\begin{aligned}&\lim_{t\to\infty}\left\{\exp(-\rho t)\hat{H}(t,\hat{x}(t),\hat{y}(t),\mu(t))\right\} = 0 &(5.56)\\ \Leftrightarrow &\lim_{t\to\infty}\{\exp(-\rho t)\mu(t) g(t,\hat{x}(t),\hat{y}(t))\} = 0 &(5.57)\\ \Leftrightarrow &\lim_{t\to\infty}\left\{\exp(-\rho t)\mu(t)\dot{\hat{x}}(t)\right\} = 0 &(5.58)\end{aligned}$$

[12]The intuition of footnote 9 still goes through with the adjustment that the current value of the stock at time $t+dt$, $\mu(t+dt)dx$, must be discounted through multiplication with $1-\rho dt$ to be expressed in time-t present value terms.

where the second line follows from the definition of the (current-value) Hamiltonian, (5.43), and the third line follows from (5.20).

Now note that (5.53) can be written as

$$\lim_{t\to\infty}\left[\exp(-\rho t)\mu(t)\int_0^t \dot{\hat{x}}(s)\,\mathrm{d}s\right]=0 \tag{5.59}$$

since $\hat{x}(t)=\hat{x}(0)+\int_0^t \dot{\hat{x}}(s)\,\mathrm{d}s$ (from the fundamental theorem of calculus), which implies (5.58).

A limitation of the Maximum Principle, as stated in Theorems 5.1 and 5.2, is that the conditions it provides are *neither necessary nor sufficient for a solution in general*. On the one hand, the solution to an optimal control problem may not be interior or piecewise continuous, so that (5.45)-(5.48) need not apply in this case. On the other hand, a pair satisfying (5.45)-(5.48) may be only a *local*, rather than global, optimum (or no optimum of any kind in some cases) and thus not a solution to the optimal control problem. The following theorem is, therefore, a key theoretical result, as it provides sufficient conditions for a candidate policy to be a (unique) global optimum.

Theorem 5.3 (Sufficient Optimality Conditions for Discounted Problems). *Consider the problem defined by* (5.39) *subject to* (5.20)-(5.24). *Define the current-value Hamiltonian $\hat{H}(t,x,y,\mu)$ as in* (5.43). *Suppose that some admissible $\hat{y}(t)$ and the corresponding path of state variable $\hat{x}(t)$ satisfy the necessary conditions* (5.45)-(5.48). *Define the* maximized Hamiltonian

$$M(t,x,\mu)\equiv\max_{y(t)\in\mathcal{Y}(t)}\hat{H}(t,x,y,\mu) \tag{5.60}$$

Suppose that

(i) the value function $V(t,\hat{x}(t))$ exists and is finite for all t.

(ii) for any admissible pair $(x(t),y(t))$,

$$\lim_{t\to\infty}\{\exp(-\rho t)\mu(t)x(t)\}\geq 0 \tag{5.61}$$

where $\mu(t)$ is the costate variable corresponding to the candidate optimal path $(\hat{x}(t),\hat{y}(t))$.

(iii) $\mathcal{X}(t)$ is convex for all t

(iv) and $M(t,x,\mu)$ is concave in $x\in\mathcal{X}(t)$ for all t when evaluated at the costate variable $\mu(t)$ corresponding to the candidate optimal path $(\hat{x}(t),\hat{y}(t))$.

Then, the pair $(\hat{x}(t),\hat{y}(t))$ achieves the global maximum of (5.39).

Moreover, if $M(t,x,\mu)$ is strictly concave in x, $(\hat{x}(t),\hat{y}(t))$ is the unique *solution to* (5.39).[a]

[a] It can also be shown that, if the assumptions of this theorem hold with $M(t,x,\mu)$ strictly concave in x and $\mathcal{Y}(t)=\mathcal{Y}$ is compact, then $\hat{y}(t)$ must be a continuous function of t on $\mathbb{R}_+$.

As in finite-dimensional constrained optimization, the key assumption for conditions (5.45)-(5.48) to also be sufficient for $(\hat{x}(t),\hat{y}(t))$ to be the (unique) solution is that the problem is (strictly) concave, assumption (iv). Also note that proving condition (i), the finiteness of the value function, is typically quite complicated; we will simply assume that it is the case in the examples that we will discuss.

Remark 5.2 (Solution Strategy). In light of Theorems 5.2, 5.3 and the preceding remark, the following strategy can be applied to most optimal control problems in economics to prove existence (and uniqueness) of an optimal policy:

1. Use the necessary conditions (5.45)-(5.47) and the easy-to-check transversality condition (5.53) to locate a candidate interior solution $(\hat{x}(t),\hat{y}(t))$.

2. Verify the (strict) concavity condition of Theorem 5.3 and check that

$$\lim_{t\to\infty}\{\exp(-\rho t)\mu(t)x(t)\}\geq 0 \tag{5.62}$$

   holds for all admissible pairs $(x(t),y(t))$, where $\mu(t)$ is the costate variable associated with our candidate policy. This is usually a direct implication of the boundary (feasibility, no-Ponzi) condition (5.24).

Application 5.1 (Hotelling Rule for nonrenewable resources)**.** We now study an example of an optimal control problem that is quite instructive, as it clearly (and quite literally) illustrates the "stock-flow" intuition for optimal control problems that we discussed above.

We want to solve for the optimal time path of consuming a nonrenewable resource. Suppose that the social planner has access to a nonrenewable resource of mass 1. Denote the stock of the resource at time t by $x(t)$. The instantaneous utility of consuming a flow of resources $y(t)$ is $u(y)$, and the planner discounts the future exponentially at rate $\rho>0$, so that he solves

$$\max_{[x(t),y(t)]_{t=0}^{\infty}}\int_0^{\infty}\exp(-\rho t)u(y(t))\,\mathrm{d}t \tag{5.63}$$

subject to the constraints

$$\dot{x}(t)=-y(t) \tag{5.64}$$

$$x(t)\in[0,1]\quad\forall\, t\geq 0 \tag{5.65}$$

and given the boundary conditions

$$x(0) = 1 \tag{5.66}$$
$$\lim_{t\to\infty} x(t) \geq 0 \tag{5.67}$$

where $u : \mathbb{R}_+ \to \mathbb{R}$ is strictly increasing, continuously differentiable, and strictly concave. Note that the feasibility (no-Ponzi) condition (5.67) is obviously implied by constraint (5.65), but we explicitly include it to illustrate how our example maps to the formulation of the general problem, (5.39) subject to (5.20)-(5.24). The constraint (5.64) captures the fact that the nonrenewable resource becomes depleted as more of it is consumed.

Let us now follow the steps outlined in Remark 5.2 to solve this problem.

First, set up the Hamiltonian:

$$\hat{H}(x(t), y(t), \mu(t)) = u(y(t)) - \mu(t)y(t) \tag{5.68}$$

We look for a candidate optimal solution satisfying (5.45)-(5.47):

$$u'(\hat{y}(t)) = \mu(t) \tag{5.69}$$
$$\rho\mu(t) - \dot{\mu}(t) = 0 \tag{5.70}$$
$$\dot{\hat{x}}(t) = -\hat{y}(t) \tag{5.71}$$

Note that the second FOC, (5.70), follows since neither the payoff function (5.63) nor the constraint (5.64) depend on x, so that $\hat{H}_x$ is zero. In fact, (5.70) yields the *Hotelling rule* for the exploitation of exhaustible resources:

$$\frac{\dot{\mu}(t)}{\mu(t)} = \rho \tag{5.72}$$

The rule states that the shadow current value of a nonrenewable resource (the value of having an extra marginal unit of the resource "today") must grow at the same rate as the discount rate. From section 1.3.1 we know that the solution to this simple differential equation is

$$\mu(t) = \mu(0)\exp(\rho t) \tag{5.73}$$

FOC (5.69) then implies that our candidate control function is

$$\hat{y}(t) = u'^{-1}[\mu(0)\exp(\rho t)] \tag{5.74}$$

Since function $u(\cdot)$ is strictly concave, its derivative $u'(\cdot)$ is strictly decreasing, which in turn implies that the derivative of the inverse $u'^{-1}[\cdot]$ is also strictly decreasing. Therefore, (5.74) implies that the amount of the resource consumed is monotonically decreasing over time. This is intuitive: due to discounting, there is preference for early consumption; on the other hand, the entire resource is not consumed immediately, as there is also a preference for a smooth consumption path over time (since $u(\cdot)$ is strictly concave).

Now solving for $\hat{x}(t)$:

$$\dot{x}(t) = -u'^{-1}[\mu(0)\exp(\rho t)] \tag{5.75}$$

Using the initial value condition $x(0) = 1$ to pin down the constant of integration, out candidate state function is

$$\hat{x}(t) = 1 - \int_0^t u'^{-1}[\mu(0)\exp(\rho t)]\,\mathrm{d}s \tag{5.76}$$

To complete the description of the candidate policy $(\hat{x}(t), \hat{y}(t))$ and the corresponding costate variable $\mu(t)$, all we need is to pin down $\mu(0)$. We use the "easy-to-check" transversality condition (5.53):

$$\lim_{t\to\infty}[\exp(-\rho t)\mu(t)\hat{x}(t)] = 0 \tag{5.77}$$
$$\Rightarrow \lim_{t\to\infty}[\exp(-\rho t)\mu(0)\exp(\rho t)\hat{x}(t)] = 0 \tag{5.78}$$
$$\Rightarrow \lim_{t\to\infty}[\mu(0)\hat{x}(t)] = 0 \tag{5.79}$$
$$\Rightarrow \lim_{t\to\infty}\hat{x}(t) = 0 \tag{5.80}$$

Therefore, in light of (5.76), $\mu(0)$ must be chosen so that (5.80) is satisfied:

$$\int_0^\infty u'^{-1}[\mu(0)\exp(\rho t)]\,\mathrm{d}s = 1 \tag{5.81}$$

We now move to the second step in Remark 5.2. We need to verify conditions (ii)-(iv) of Theorem 5.3 to show that the candidate policy pair $(\hat{x}_t, \hat{y}_t)$ identified above is in fact the solution to our problem.

After plugging in for the candidate costate variable, condition (ii) becomes

$$\lim_{t\to\infty} x(t) \geq 0 \tag{5.82}$$

which is precisely the boundary constraint (5.67).

$\mathcal{X}(t) = [0,1]$ is convex, so assumption (iii) holds.

Finally, the key condition (iv) is trivially satisfied in our case since

$$M(x,\mu) = u\left(u'^{-1}(\mu)\right) - \mu u'^{-1}(\mu) \tag{5.83}$$

is independent of x and thus weakly concave in x.

Therefore, by Theorem 5.3, our candidate solution is indeed a solution to our problem. Note that, since $M(x,\mu)$ is not *strictly* concave in x, we cannot invoke Theorem 5.3 to prove that our identified policy is the unique solution (although it turns out that this is indeed the case). ■

Application 5.2 (Solution to the Neoclassical Growth Model)**.** We have already discussed aspects of the neoclassical (Ramsey) growth model in Applications 2.1 and 4.1, taking the key equations of the model as given. We have finally developed the tools to derive these key equations for the deterministic, continuous-time version of the model.

In the baseline continuous-time version of the model without population growth and without technological progress, the social planner solves the following problem

$$\max_{[k(t),c(t)]_{t=0}^{\infty}} \int_0^{\infty} \exp(-\rho t) u(c(t))\,\mathrm{d}t \tag{5.84}$$

subject to

$$\dot{k}(t) = f(k(t)) - \delta k(t) - c(t) \tag{5.85}$$
$$k(t) \geq 0 \tag{5.86}$$
$$c(t) \geq 0 \tag{5.87}$$

and given the initial condition $k(0) = k_0 > 0$.

We assume that the utility function $u(c)$, $u : \mathbb{R}_+ \to \mathbb{R}$,[13] is strictly increasing, continuously differentiable, strictly concave and satisfies $\lim_{c\to 0} u'(c) = \infty$. The production function $f(k)$, $f : \mathbb{R}_+ \to \mathbb{R}_+$, is strictly increasing, continuously differentiable, strictly concave and satisfies the Inada conditions (3.21) and (3.22). Capital $k(t)$ is the state (stock) variable and $c(t)$ is the control (flow) variable.

We again follow the steps outlined in Remark 5.2 to solve this problem. The (current-value) Hamiltonian is:

$$\hat{H}(k,c,\mu) = u(c(t)) + \mu(t)\left[f(k(t)) - \delta k(t) - c(t)\right] \tag{5.88}$$

Note that since the constraint function in (5.85) does not depend directly on time, the current-value Hamiltonian also does not depend directly on time.

We once again look for a candidate optimal solution satisfying (5.45)-(5.47):

$$\hat{H}_c(\hat{k}(t), \hat{c}(t), \mu(t)) = u'(\hat{c}(t)) - \mu(t) = 0 \tag{5.89}$$
$$\hat{H}_k(\hat{k}(t), \hat{c}(t), \mu(t)) = \mu(t)\left[f'(\hat{k}(t)) - \delta\right] = \rho\mu(t) - \dot{\mu}(t) \tag{5.90}$$
$$\hat{H}_\mu(\hat{x}(t), \hat{y}(t), \mu(t)) = f(\hat{k}(t)) - \delta\hat{k}(t) - \hat{c}(t) = \dot{\hat{k}}(t) \tag{5.91}$$

We wish to substitute out the costate variable in the law of motion of the candidate consumption policy. Since the first-order conditions hold for all $t \geq 0$, we can differentiate the first FOC, (5.89), with respect to time:

$$\dot{\mu}(t) = \dot{u}'(\hat{c}(t)) \tag{5.92}$$
$$= u''(\hat{c}(t))\hat{c}'(t) \tag{5.93}$$

[13] When $u(c) = \log c$ we exclude zero from the domain of u, that is, $u : \mathbb{R}_+ \backslash \{0\} \to \mathbb{R}$.

Combining the first two first-order conditions,

$$-u''(\hat{c}(t))\hat{c}'(t) = \mu(t)\left[f'(\hat{k}(t)) - \delta - \rho\right] \tag{5.94}$$

$$\Rightarrow -\frac{u''(\hat{c}(t))\hat{c}'(t)}{u'(\hat{c}(t))} = f'(\hat{k}(t)) - \delta - \rho \tag{5.95}$$

$$\Rightarrow -\frac{u''(\hat{c}(t))\hat{c}(t)}{u'(\hat{c}(t))}\frac{\hat{c}'(t)}{\hat{c}(t)} = f'(\hat{k}(t)) - \delta - \rho \tag{5.96}$$

Denoting by $\mathcal{R}_u(c) \equiv -\frac{u''(c)c}{u'(c)}$ the coefficient of relative risk aversion of function u evaluated at point c,[14] we obtain the equilibrium law of motion of consumption, known as the *Euler equation*,

$$\frac{\hat{c}'(t)}{\hat{c}(t)} = \frac{1}{\mathcal{R}_u(\hat{c}(t))}\left[f'(\hat{k}(t)) - \delta - \rho\right] \tag{5.97}$$

This condition fully pins down the consumption path given some $\hat{c}(0)$ that we need to specify,

$$\hat{c}(t) = \hat{c}(0)\exp\left(\int_0^t \frac{f'(\hat{k}(s)) - \delta - \rho}{\mathcal{R}_u(\hat{c}(s))}\right)\mathrm{d}s \tag{5.98}$$

We can also integrate the second FOC, (5.90), to get an expression for our costate variable at time t,

$$\mu(t) = \mu(0)\exp\left(-\int_0^t \left[f'(\hat{k}(s)) - \delta - \rho\right]\mathrm{d}s\right) \tag{5.99}$$

$$= u'(\hat{c}(0))\exp\left(-\int_0^t \left[f'(\hat{k}(s)) - \delta - \rho\right]\mathrm{d}s\right) \tag{5.100}$$

where the second line follows by evaluating the first FOC, (5.89), at $t = 0$.

Since we already have the law of motion for capital, (5.85) and $k(0) = k_0 > 0$ is given, all that remains in order to complete our construction of the candidate optimal policy and the associated costate variable is to pin down the initial value of the control variable, $\hat{c}(0)$. To do this, we will use the "easy-to-check" transversality condition (5.53),

$$\lim_{t\to\infty}\left\{\exp(-\rho t)\mu(t)\hat{k}(t)\right\} = 0 \tag{5.101}$$

$$\Rightarrow \lim_{t\to\infty}\left\{\exp(-\rho t)u'(\hat{c}(0))\exp\left(-\int_0^t \left[f'(\hat{k}(s)) - \delta - \rho\right]\mathrm{d}s\right)\hat{k}(t)\right\} = 0 \tag{5.102}$$

Since $u'(\hat{c}(0)) > 0$, this yields the second boundary condition for our dynamical system

$$\lim_{t\to\infty}\left\{\hat{k}(t)\exp\left(-\int_0^t \left[f'(\hat{k}(s)) - \delta\right]\mathrm{d}s\right)\right\} = 0 \tag{5.103}$$

[14] Recall question 1 in Problem Set 1. In the benchmark case where our utility has the CRRA form, $\mathcal{R}_u(c) = \gamma > 0$ for all c.

We showed in Application 2.1 that this transversality condition pins down $\hat{c}(0)$ since it implies that the solution lies on the saddle path of the system, as depicted in Figure 2.5. That is, at time 0 the system jumps to the (unique) value $\hat{c}(0)$ such that $(k_0, \hat{c}(0))$ lies on the saddle path.

In summary, our candidate solution is described by the laws of motion (5.85) for capital, (5.97) for consumption, and (5.93) for the associated costate variable, together with the two boundary conditions, $k(0) = k_0 > 0$ and (5.103).

We now move to showing that our necessary conditions are also sufficient for identifying the unique solution to our problem, by using Theorem 5.3. Our value function $V(t, \hat{x}(t))$ can be shown to be finite for all t, and $\mathcal{X} = [0, \infty]$ is convex. Moreover, since $\mu(t) = u'(c(t)) > 0$ and $k(t) \geq 0$, (5.86), it immediately follows that any admissible path must satisfy

$$\lim_{t\to\infty} \{\exp(-\rho t)\mu(t)k(t)\} \geq 0 \tag{5.104}$$

since all terms in this expression are non-negative.

Finally, we turn to the crucial condition, condition (iv). Note that the current value Hamiltonian, given by (5.88), is jointly strictly concave in the pair $(k(t), c(t))$ since $\mu(t) = u'(c(t)) > 0$ and both u and f are strictly concave. A useful fact is that, when the Hamiltonian $\hat{H}(t, x, y, \mu)$ is jointly strictly concave in the state-control pair $(x(t), y(t))$, the maximized Hamiltonian, $M(t, x, \mu)$, is also strictly concave in x. To see this fact in our case, note from the definition of the maximized Hamiltonian (at arbitrary k but under the $\mu(t)$ of the candidate policy) that

$$\begin{aligned} M(k, \mu) &\equiv \max_{c\in[0,\infty]} \hat{H}(k, c, \mu) && (5.105) \\ &= \hat{H}(k, \tilde{c}, \mu) && (5.106) \end{aligned}$$

where $\tilde{c}$ must satisfy

$$\hat{H}_c(k, \tilde{c}, \mu) = 0 \tag{5.107}$$

since the Hamiltonian is strictly concave in c, so that a control $\tilde{c}$ is its (global) maximizer if and only if the usual first-order condition holds.[15]

Now, (5.107) implies that

$$u'(\tilde{c}(t)) = \mu(t) \tag{5.108}$$

so that

$$M(k(t), \mu(t)) = u\left(u'^{-1}(\mu(t))\right) + \mu(t)\left[f(k(t)) - \delta k(t) - u'^{-1}(\mu(t))\right] \tag{5.109}$$

is strictly concave in $k(t)$, holding μ fixed (at the strictly positive value corresponding to our candidate optimal policy).

We have thus proved that our candidate policy function is the *unique* solution to our problem. ∎

[15] Note that (5.107) is not exactly the same equation as FOC (5.89) for our candidate optimal policy, since (5.107) is evaluated at an arbitrary $k(t)$.

Chapter 6

Perturbation Methods*

6.1 Loglinearization Methods

In section 4.2 we discussed the theory behind linear and loglinear approximations of discrete-time rational expectations models, deriving the conditions under which such approximations are valid. In section 4.3 we also discussed how to solve linear expectations models and, therefore, how to solve the linear or nonloglinear approximations of rational expectations equilibrium models. In this section, we study the missing step: how to derive the loglinear (or linear) approximation to a non-linear model.

Remark 6.1 (Concepts of Steady State). There are three distinct definitions of steady state equilibrium.[a]

First, a *deterministic* steady state, defined in Section 4.2, coincides with the steady state of the deterministic version of the economy. The corresponding steady state equilibrium is sometimes referred to as the *perfect foresight equilibrium*. That is, not only are shocks (exogenous disturbances) absent but this is also common knowledge among agents a priori. This is the steady state around which most loglinearizations as well as higher-order approximations (see next section) of models take place.

Second, a *stochastic* steady state is the *distribution* to which economic variables converge in the long-run (that is, the ergodic distribution) in the stochastic version of the economy. This is not a single vector (such as $\overline{x}$ in (4.22)) but a distribution over vectors.

Finally, a *risky* steady state, as defined, for example, in Coeurdacier, Rey, and Winant (2011), is the point where agents choose to stay at a given date if they expect future shocks and if the realization of shocks is 0 at this date. Note that if shocks have a continuous distribution agents assign zero probability to the risky steady state being realized in the next period.

We will only consider approximations around the deterministic steady state.

[a] Note that there is some ambiguity in the literature about the meaning of these terms.

Consider the nonlinear model

$$\Phi(x;u) \equiv \{\mathbb{E}_t \phi(x_t, x_{t+1}; u_t)\}_{t=0}^T = 0 \in \mathbb{R}^T \tag{6.1}$$

whose (deterministic) steady state $\overline{x}$ and corresponding steady state equilibrium x^* are given by[1]

$$\Phi(x^*;1) \equiv \{\phi(\overline{x},\overline{x};1)\}_{t=0}^T = 0 \in \mathbb{R}^T \tag{6.2}$$

Linearization of ϕ around the steady state involve a simple first-order approximation around $x = \overline{x}$:

$$\begin{aligned}\mathbb{E}_t \phi(x_t, x_{t+1}, u_t) &\approx \mathbb{E}_t \left[\phi(\overline{x},\overline{x},1) + \phi_1(x_t - \overline{x}) + \phi_2(x_{t+1} - \overline{x}) + \phi_3(u_t - 1)\right] &(6.3)\\ &= \phi_1(x_t - \overline{x}) + \phi_2 \mathbb{E}_t(x_{t+1} - \overline{x}) + \phi_3(u_t - 1) &(6.4)\end{aligned}$$

where ϕ_i is the derivative of function ϕ with respect to its ith set of arguments, all evaluated at their steady state values $(\overline{x},\overline{x},1)$.

Alternatively, we can *loglinearize* the model around the steady state. That is, we rewrite $x_t = \exp(\ln x_t)$, reinterpret the function as having $\ln x$, rather than x, as its argument and take a first-order Taylor approximation around $\ln \overline{x}$:

$$\begin{aligned}\mathbb{E}_t \phi(x_t, x_{t+1}, u_t) &\approx \mathbb{E}_t \left[\phi(\overline{x},\overline{x},1) + \phi_1 e^{\ln \overline{x}}(\ln x_t - \ln \overline{x}) + \phi_2 e^{\ln \overline{x}}(\ln x_{t+1} - \ln \overline{x}) + \phi_3 \ln u_t\right] &(6.5)\\ &= \phi_1 \overline{x} \hat{x}_t + \phi_2 \overline{x} \mathbb{E}_t \hat{x}_{t+1} + \phi_3 \hat{u}_t &(6.6)\end{aligned}$$

where

$$\hat{x}_t \equiv \ln x_t - \ln \overline{x} \tag{6.7}$$

is called the *log-deviation* of variable x_t (from its steady state value).

In general we have

$$f(x_t) \approx f(\overline{x}) + f'(\overline{x})\overline{x}\hat{x}_t \tag{6.8}$$

and

$$\ln f(x_t) \approx \ln f(\overline{x}) + \frac{f'(\overline{x})\overline{x}}{f(\overline{x})}\hat{x}_t \tag{6.9}$$

Loglinearization is used more often than linearization because loglinearized equations are easier to interpret. First, note that $\hat{x}_t$ represents the relative (percentage) deviation of x_t from the steady state, a measure of deviation that is independent of the unit of measurement. Second, in loglinear equations, such as (6.9), the coefficient of $\hat{x}_t$, $\frac{f'(\overline{x})\overline{x}}{f(\overline{x})}$, can be interpreted as the elasticity of $f(x_t)$ with respect to x_t.

[1] Here we have normalized the state-state value of u_t to 1 (rather than 0 as we did in section 4.2.1).

Remark 6.2 (Loglinearization Strategy). To loglinearize a given equilibrium equation, follow these steps:

1. Check if the equation fits exactly into one of two common special cases (or if it is a combination of both). If yes, you are done.
   - The equation is of the form $x_t = a y_t^b z_t^c$, where a, b, and c are constants. In this case, the equation can be loglinearized *exactly* as $\hat{x}_t = b\hat{y}_t + c\hat{z}_t$
   - The equation is of the form $x_t = y_t + z_t$. In this case, the loglinear approximation is $\overline{x}\hat{x}_t \approx \overline{y}\hat{y}_t + \overline{z}\hat{z}$.
2. Before taking any approximations, write all variables as $x_t = \overline{x}e^{\hat{x}_t}$ (this is an identity) and try to group as many variables together as possible; log both sides of the equation.
3. If applicable, pass the log through the expectation (certainty equivalence for loglinearized models). The relation becomes an approximation due to Jensen's inequality.
4. If needed, take a first-order approximation of each (logged) side of the equation with respect to the log-deviations (around 0), using

$$\ln f(x_t) = \ln f(\overline{x}e^{\hat{x}_t}) \approx \ln f(\overline{x}) + \frac{f'(\overline{x})\overline{x}}{f(\overline{x})}\hat{x}_t \tag{6.10}$$

Let's go through some examples.

First, consider

$$y_t = k_t + c_t \tag{6.11}$$

This falls under the second case of step 1, so we have

$$\hat{y}_t \approx \frac{\overline{k}}{\overline{y}}\hat{k}_t + \frac{\overline{c}}{\overline{y}}\hat{c}_t \tag{6.12}$$

A second example:

$$\frac{1}{1-l_t} = \frac{1-b}{b}\frac{w_t}{c_t} \tag{6.13}$$

This does not fall under either case in step 1, so we proceed with step 2,

$$\frac{1}{1-\overline{l}e^{\hat{l}_t}} = \frac{(1-b)\overline{w}e^{\hat{w}_t}}{b\overline{c}e^{\hat{c}_t}} \tag{6.14}$$

$$-\ln(1-\overline{l}e^{\hat{l}_t}) = \ln\left(\frac{(1-b)\overline{w}}{b\overline{c}}\right) + \hat{w}_t - \hat{c}_t \tag{6.15}$$

Making use of the steady state relation

$$\frac{1}{1-\bar{l}} = \frac{1-b}{b}\frac{\overline{w}}{\overline{c}} \tag{6.16}$$

we simplify this to

$$\ln\left(\frac{1-\bar{l}}{1-\bar{l}e^{\hat{l}_t}}\right) = \hat{w}_t - \hat{c}_t \tag{6.17}$$

Note that, up to this point, the relation remains exact. We now take a first-order approximation of the left-hand side (step 4) to get our loglinearized equation:

$$\frac{\bar{l}}{1-\bar{l}}\hat{l}_t \approx \hat{w}_t - \hat{c}_t \tag{6.18}$$

Application 6.1 (Loglinearization of the Stochastic Neoclassical Growth Model)**.** Consider the neoclassical growth model introduced in Application 4.1, described by the equations

$$k_{t+1} = z_t k_t^{\alpha} + (1-\delta)k_t - c_t \tag{6.19}$$
$$c_t^{-\gamma} = \mathbb{E}_t\beta\left[az_{t+1}k_{t+1}^{\alpha-1} + (1-\delta)\right]c_{t+1}^{-\gamma} \tag{6.20}$$
$$\ln z_{t+1} = \rho\ln z_t + \varepsilon_{t+1} \tag{6.21}$$

Note from (6.21) that $\bar{z} = 1$.

We first loglinearize (6.19), noting that it is a combination of the two special cases of step 1 in Remark 6.2.

$$\bar{k}\hat{k}_{t+1} \approx \bar{y}\left[\hat{z}_t + \alpha\hat{k}_t\right] + (1-\delta)\bar{k}\hat{k}_t - \bar{c}\hat{c}_t \tag{6.22}$$

where $\bar{y} = \bar{z}\bar{k}^{\alpha} = \bar{k}^{\alpha}$.

Now consider the Euler Equation, (6.20). Let

$$R_{t+1} \equiv az_{t+1}k_{t+1}^{\alpha-1} + (1-\delta) \tag{6.23}$$

Applying step 2 we can write (6.20) as

$$1 = \beta\overline{R}\mathbb{E}_t\left[e^{\hat{R}_{t+1}-\gamma(c_{t+1}-c_t)}\right] \tag{6.24}$$

We use the steady state relation $1 = \beta\overline{R}$ and we log both sides to get

$$0 = \ln\mathbb{E}_t\left[e^{\hat{R}_{t+1}-\gamma(c_{t+1}-c_t)}\right] \tag{6.25}$$

We now pass the log through the expectation (certainty equivalence holds for loglinearized models) to get the approximation:

$$0 \approx \mathbb{E}_t\left[\hat{R}_{t+1} - \gamma(c_{t+1}-c_t)\right] \tag{6.26}$$
$$\Rightarrow \mathbb{E}_t c_{t+1} \approx c_t + \frac{1}{\gamma}\mathbb{E}_t\hat{R}_{t+1} \tag{6.27}$$

Finally, we loglinearize (6.23) using step 1 as

$$\overline{R}\hat{R}_{t+1} \approx a\overline{k}^{\alpha-1}\left(\hat{z}_{t+1} + (\alpha-1)\hat{k}_{t+1}\right) + (1-\delta)\cdot 0 \tag{6.28}$$

to get the loglinearized counterpart of (6.20) as

$$\mathbb{E}_t\hat{c}_{t+1} \approx \hat{c}_t + \frac{1}{\gamma}\frac{\alpha\overline{k}^{\alpha-1}}{a\overline{k}^{\alpha-1} + (1-\delta)}\left[\mathbb{E}_t\hat{z}_{t+1} + (\alpha-1)\hat{k}_{t+1}\right] \tag{6.29}$$

Note that for $\delta = 1$ (6.20) simplifies to (4.58).

Finally, the (exact) loglinearization to (6.21) is

$$\mathbb{E}_t\hat{z}_{t+1} = \rho\hat{z}_t \tag{6.30}$$

■

6.2 Nonlinear Perturbation Methods

First-order approximations, such as the loglinearization method that we discussed in the previous section, abstract from aspects of the exact model that may be crucial. For example, we saw that we could ignore Jensen's inequality when taking the loglinear approximation of a given equation. For models that focus on the time-variability of risk or risk aversion, a first-order approximation would not be useful. Fernández-Villaverde and Rubio-Ramírez (2005) also emphasize the importance of preserving model nonlinearities when conducting likelihood-based analysis. In optimal policy problems involving second-order approximations to the objective (welfare) function, the decision rules must also be approximated to second order in order for the approach to be valid. These are just some examples of reasons why first-order approximations of nonlinear models may not be sufficient. In this section we briefly sketch the idea behind an important class of nonlinear approximation techniques, *higher-order perturbation methods*, that is in fact quite related to the issues we discussed in Chapter 4 and the preceding section.

To give some structure to the discussion, consider a nonlinear model of the form

$$\Phi(x;u) \equiv \{\mathbb{E}_t\phi(x_t, x_{t+1}; u_t)\}_{t=0}^{\infty} = 0 \tag{6.31}$$

Denote by s_t the *state variables* of the model. These include all of the exogenous variables in u_t as well as the subset of endogenous variables in x_t that are predetermined at time t. Denote by c_t the *control variables* of the model.

Assume that model (6.31) has *exact* solutions of the form

$$c_t = c(s_t) \tag{6.32}$$
$$s_t = f(s_{t-1}, u_t) \tag{6.33}$$

$c(s_t)$ is called the policy function and describes the agents' actions given the state the economy is in, and $f(s_{t-1}, u_t)$ describes the evolution of the state variables of the economy.

In Chapter 4 as well as the previous section we saw how to approximate (6.31) by a (log) linear model of the form

$$A\mathbb{E}_t x_{t+1} = Bx_t + Cu_t \tag{6.34}$$

This approximation implies, in particular, that the control variables contained in x are approximated by linear functions of the state variables (recall Application 4.1).

Instead, assume that we wish to derive a more accurate numerical approximation of the policy function, $c(s_t)$. Perturbation methods build on the very same foundations as first-order approximation approaches, which we discussed in Section 4.2. To illustrate the basic idea, consider the simple case of a one-dimensional control variable c and a one-dimensional exogenous shock u_t, the only state variable, following

$$u_{t+1} = \rho u_t + \sigma \varepsilon_{t+1} \tag{6.35}$$

where ε_{t+1} is iid with zero mean.

Note that the implicit mapping theorem, Theorem 4.3, under the very same conditions discussed in Section 4.2 allows us to reformulate (6.31) as

$$\{\mathbb{E}_t \phi(c_t(u_t, \sigma), u_t(u_{t-1}, \sigma), \sigma)\} = 0 \tag{6.36}$$

and also tells us that if ϕ is k-times continuously differentiable in its arguments, then the implicitly defined function $c_t(u_t, \sigma)$ is k-times continuously differentiable as well. Then, assuming $k \geq 2$ we can obtain a second-order approximation to $c(u_t, \sigma)$ around the deterministic steady state, characterized by $\sigma = 0$ and $u_t = 0$, for all t .

In practical terms, we proceed as follows. Let

$$G(u_t, \sigma) \equiv \mathbb{E}_t \phi(c_t(u_t, \sigma), u_t(u_{t-1}, \sigma), \sigma) = 0 \tag{6.37}$$

Using the implicit function theorem, we have

$$G_u(0,0) = 0 \tag{6.38}$$
$$G_\sigma(0,0) = 0 \tag{6.39}$$

and also, totally differentiating the two expressions above (and again evaluating at the deterministic steady state),

$$G_{uu}(0,0) = 0 \tag{6.40}$$
$$G_{u\sigma}(0,0) = 0 \tag{6.41}$$
$$G_{\sigma\sigma}(0,0) = 0 \tag{6.42}$$

We can then employ a method of undetermined coefficients (recall Application 4.1) to identify the first and second derivatives of the policy function $c(u_t, \sigma)$ evaluated at the steady state. Then, the second-order approximation to the policy function would be given by

$$c(u_t, \sigma) = c(0,0) + c_u(0,0)u_t + c_\sigma(0,0)\sigma \quad (6.43)$$
$$+\frac{1}{2}c_{uu}(0,0)u_t^2 + c_{u\sigma}(0,0)u_t\sigma + \frac{1}{2}c_{\sigma\sigma}(0,0)\sigma^2 \quad (6.44)$$

In fact, Schmitt-Grohé and Uribe (2004) show that in a class of models of which our simple example is a special case, we have $c_\sigma(0,0) = c_{u\sigma}(0,0) = 0$, so that the procedure described can be further simplified taking this finding into account.

In principle, there is no reason why we should stop at the second order. We could once again totally differentiate equations (6.40)-(6.42) in order to arrive at a third-order approximation around the steady state. However, the curse of dimensionality quickly kicks in, so that approximations of higher than third order are too computationally demanding to be worth pursuing in practice.

There are two other main classes of nonlinear approximation techniques: projection methods, and iteration techniques. The latter are used quite frequently in macroeconomic applications and build directly on the theory of dynamic programming that you will study during the first quarter of the macro sequence.

Main Sources

The following is a list of the major sources for the material in each chapter:

Chapter 1 The exposition of solution methods to ODE systems mainly draws on chapters 9 and 10 in de la Fuente (2000) and chapter 2 and Appendix B in Acemoglu (2009). Section 1.2 is based on Lang (1987). Application 1.2 is based on Werning (2012). Section 1.6 draws on Walter (1998), Luenberger (1979), and Edwards and Penney (2008).

Chapter 2 The exposition of results on stability of ODE systems mainly draws on chapters 9 and 10 in de la Fuente (2000) and chapter 7 and Appendix B in Acemoglu (2009). Applications 2.1 and 2.2 are based on chapter 8 in Acemoglu (2009) and chapter 9 in de la Fuente (2000), respectively.

Chapter 3 The exposition of results on difference systems mainly draws on chapters 9 and 10 in de la Fuente (2000) and Appendix B in Acemoglu (2009). Application 3.1 is based on chapter 2 in Acemoglu (2009).

Chapter 4 Sections 4.1 and 4.2 are based on Appendix A in Woodford (2003). Section 4.3 is based on Blanchard and Kahn (1980) and chapter 4 in DeJong and Dave (2011). Application 4.2 draws on chapter 2 in Woodford (2003).

Chapter 5 The exposition of optimal control theory draws extensively on chapter 7 in Acemoglu (2009). Applications 5.1 and 5.2 are based on chapters 7 and 8 in Acemoglu (2009).

Chapter 6 Section 6.2 mainly draws on chapter 5 in DeJong and Dave (2011).

Index

Bibliography

Acemoglu, Daron. 2009. *Introduction to Modern Economic Growth*. Princeton: Princeton University Press.

Aliprantis, Charalambos D. and Kim C. Border. 2006. *Infinite Dimensional Analysis: A Hitchhker's Guide*. Springer, Third ed.

Barro, Robert J. and Xavier Sala-i Martin. 2004. *Economic Growth*. Cambridge: The MIT Press, Second ed.

Blanchard, Olivier Jean and Charles M. Kahn. 1980. "The Solution of Linear Difference Models under Rational Expectations." *Econometrica* 48 (5):1305–1311.

Cochrane, John H. 2011. "Determinacy and Identification with Taylor Rules." *Journal of Political Economy* 119 (3):565–615.

———. 2014. "The New-Keynesian Liquidity Trap."

Coeurdacier, Nicolas, Hélène Rey, and Pablo Winant. 2011. "The Risky Steady State." *American Economic Review* 101 (3):398–401.

de la Fuente, Angel. 2000. *Mathematical Methods and Models for Economists*. Cambridge University Press.

DeJong, Dave and Chetan Dave. 2011. *Structural Macroeconometrics*. Princeton University Press, Second ed.

Diekman, Odo and Yuriy Kuznetsov. 2011. "Lecture: Linear Maps and ODEs."

Edwards, Henry C. and David E. Penney. 2008. *Elementary Differential Equations*. Pearson Prentice Hall, Sixth ed.

Fernández-Villaverde, Jesús and Juan F. Rubio-Ramírez. 2005. "Estimating Dynamic Equilibrium Economies: Linear versus Nonlinear Likelihood." *Journal of Applied Econometrics* 20 (7):891–910.

Gali, Jordi. 2009. *Monetary Policy, inflation, and the Business Cycle: An introduction to the New Keynesian Framework*. Princeton University Press.

Judd, Kenneth L. 1998. *Numerical Methods in Economics*. Cambridge: The MIT Press.

King, Robert G. and Mark W. Watson. 2002. "System Reduction and Solution Algorithms for Singular Linear Difference Systems under Rational Expectations." *Computational Economics* 20 (1-2):57–86.

Klein, Paul. 2000. "Using the generalized Schur form to solve a multivariate linear rational expectations model." *Journal of Economic Dynamics and Control* 24 (10):1405–1423.

Lang, Serge. 1987. *Linear Algebra*. Undergraduate Texts in Mathematics. Springer, Third ed.

———. 1993. *Real and Functional Analysis*. Springer, Third ed.

Luenberger, David G. 1969. *Optimization by Vector Space Methods*. Wiley Professional Paperback Series. New York: John Wiley & Sons, Ltd.

———. 1979. *Introduction to Dynamic Systems: Theory, Models, and Applications*. John Wiley & Sons, Ltd.

Sargent, Thomas J. and Neil Wallace. 1975. ""Rational" Expectations, the Optimal Monetary Instrument, and the Optimal Money Supply Rule." *The Journal of Political Economy* 83 (2).

Schmitt-Grohé, Stephanie and Martıń Uribe. 2004. "Solving dynamic general equilibrium models using a second-order approximation to the policy function." *Journal of Economic Dynamics and Control* 28 (4):755–775.

Sims, Christopher A. 2002. "Solving Linear Rational Expectations Models." *Computational Economics* 20 (1-2):1–20.

Stokey, Nancy L. and Robert E. Lucas. 1989. *Recursive Methods in Economic Dynamics*. Harvard University Press.

Uhlig, Harald. 1999. "A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily." In *Computational Methods for the Study for Dynamic Economics*, edited by Ramon Marimon and Andrew Scott. New York: Oxford University Press.

Walter, Wolfgang. 1998. *Ordinary Differential Equations*. Graduate Texts in Mathematics. Springer.

Werning, Ivan. 2012. "Managing a Liquidity Trap: Monetary and Fiscal Policy."

Woodford, Michael. 2003. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.