Tools for Macroeconomic Theory
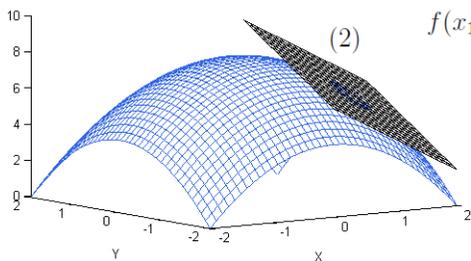
$$f(x_1, \ldots, x_n) \approx f(\bar{x}_1, \ldots, \bar{x}_n)$$
$$+ \nabla f(\bar{x}_1, \ldots, \bar{x}_n) \cdot \langle x_1 - \bar{x}_1, \ldots, x_n - \bar{x}_n \rangle$$

(2)

$$x_0 e^{\hat{x}_t} \approx \underbrace{x_0 e^{\hat{x}_0}}_{f(\bar{x})} + \underbrace{x_0 e^{\hat{x}_0}}_{f'(\bar{x})} \underbrace{(\hat{x}_t - \hat{x}_0)}_{(x - \bar{x})}$$

Local Solution Techniques

Seton Leonard

for Calvin

Contents

Introduction

Like any academic field macroeconomics is constantly evolving. Since the late 1970s and early 1980s this evolution has been particularly rapid. Blanchard and Kahn's (1980) solution approach to rational expectations models (chapter 5) lead to the nearly universal adoption of micro founded representative agent models.

While a number of excellent macroeconomic theory texts (Romer (2011) or Obstfeld and Rogoff (1996) for example) discuss applications of modern modelling tools, there are few texts on how these tools are constructed and used at a level accessible to beginning graduate students or advanced undergraduate students. This book aims to fill that gap.

As an introductory text the coverage of this book is broad but not deep; each chapter therefore contains suggestions on where to find more depth for readers particularly interested in a given subject. At the same time, I've written this book keeping in mind the fact that graduate students are always time constrained. My approach is therefore as concise and to the point as I feel I can be without sacrificing clarity; there is little to no "chat" in any of the chapters. This text assumes the reader is comfortable with basic calculus and has some familiarity with linear algebra. A mathematical economics text should provide more than enough background.

The constant evolution of macroeconomics makes it an exciting field to study. In many ways new developments reinvent older theories — the demand impact of risk shocks (when risk goes up demand falls as agents increase their precautionary wealth) in numerical models have a very classical Keynesian flavor. The tools available to the modern macroeconomist are always improving, and for those familiar with their use there is tremendous scope for innovation.

Note: I often use the pronoun 'we' though this is a single author text. This comes from the fact that many of these chapters come from review session notes in which 'we' is the class, as in 'we find the first order conditions...' at which point I typically ask a student to do so. In the text 'we' then refers to you, the reader, and me, the author on the assumption that you will also work through the derivations herein and not simply take my word for the results. If you do find a mistake please let me know so I can correct it!

Tools for Macroeconomic Theory

$$(2) \quad \begin{aligned} f(x_1, \ldots, x_n) &\approx f(\bar{x}_1, \ldots, \bar{x}_n) \\ &+ \nabla f(\bar{x}_1, \ldots, \bar{x}_n) \cdot \langle x_1 - \bar{x}_1, \ldots, x_n - \bar{x}_n \rangle \end{aligned}$$
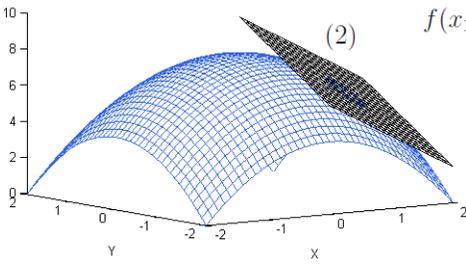
$$x_0 e^{\hat{x}_t} \approx \underbrace{x_0 e^{\hat{x}_0}}_{f(\bar{x})} + \underbrace{x_0 e^{\hat{x}_0}}_{f'(\bar{x})} \underbrace{(\hat{x}_t - \hat{x}_0)}_{(x - \bar{x})}$$

Local Solution Techniques

Preliminaries

Before introducing the various steps involved in solving macroeconomic models we might first ask what is a macro model and why bother solving it? A macroeconomic model is simply a collection or system of equations. In its simplest (micro-founded) form these equations are an objective function or utility function, our measure of what consumers would like to maximize, and a budget constraint, describing the resources with which consumers try to obtain utility. In most models we'll assume that our economy consists of a number of consumers who can be represented by a single agent, and that the decisions of this hypothetical agent can describe the movements of the economy in aggregate.[1] Solving the model consists of finding the conditions that must be met if the consumer is optimizing his (or its) utility function and using these conditions to describe how variables in the model (capital or consumption for example) respond

---

[1]This assumption works well when we assume there are no frictions, that is, no imperfections in the market that prevent agents from obtaining a first best outcome. However, such frictions can undermine the validity of the representative agent model. Consider, for example, that adjustments to employment can only occur in the extensive margin (hiring and firing works) as opposed to the intensive margin (more or fewer hours worked). In a representative agent model employment only adjusts in the intensive margin (you can't fire your workers if there is only one of them), while in the real world employment tends to be in the extensive margin — a much higher risk environment.

to exogenous shocks. This solution, typically in the form of an impulse response function, allows us to compare our model with the real economy and, if the model does well, conduct counter factual experiments.

The traditional (since Blanchard and Kahn (1980)) paper and pencil approach to solving maroeconomics models has three steps: (1) finding the first order conditions for a maximum, (2) locally approximating these conditions around a point, typically the deterministic steady state, and (3) solving the system of difference equations (assuming the model is in discrete time) arising from this local approximation. Chapters 1 through 3 deal with finding the first order conditions for a maximum, the conditions that must be met if our consumer is maximizing his objective function subject to the relevant constraints. This step is typically part of any solution technique, though software such as Dynare may be able to do it for you.

The first order conditions for a maximum are typically a non-linear system of difference equations. In order to solve this system around a given point with pen and paper we need to make it linear. Chapter 4 deals first with finding our point of approximation. This point is typically the deterministic steady state, the rest point of the system in the absence of shocks. Chapter 4 then goes on to introduce two approaches to finding a log linear approximation of our model around the deterministic steady state.

Once we have our linear system of equations we need to solve it, the subject of chapter 5. In this context solving means finding the next period state variables (variables which are predetermined) and current period jump variables (those which can change within the period) as a function of the current period state.[2] These laws of motion for variables allow us to construct

---

[2]In a continuous time context this means finding the derivatives of state variables and the levels of jump variables as a function of the levels of state

impulse response functions or to simulate the model over time and thus to look at how shocks coming from outside the model impact the model economy.

---

variables.

Chapter 1

Lagrange's Method for Constrained Maximization

Lagrange's Method is a simplified means of solving constrained optimization problems. The easiest way to understand how this method works is (in my opinion) with gradients. Imagine a problem in 3-space where a shape is defined by the function $z = f(x, y)$. A **basis vector** is a vector of magnitude one that points in the direction of one of our axes. Thus in this problem our two basis vectors for the x and y axes (call them **i** and **j**) are (1,0) and (0,1). Basis vectors are a simple tool we can use to describe any vector in the x, y space. A **gradient** is just the sum of the partial derivatives of $f(x, y)$ multiplied by each partial derivative's associated basis vector. The gradient operator is typically denoted by the symbol ∇. Thus $\nabla f(x, y) = f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j}$. Alternatively, we could denote this same gradient as $\langle f_x(x, y), f_y(x, y) \rangle$ (they are the same thing). Thus for a given point $P(x, y)$ the gradient gives us the change in slope of the surface z as we move along the x axis times **i** plus

the change in slope of the surface z as we move along the y axis times $\mathbf{j}$. We can use the gradient to find the rate of change of z as we move in any direction from point $P(x, y)$ by simply taking the dot product [1] of a unit vector (magnitude 1) in the direction of interest.
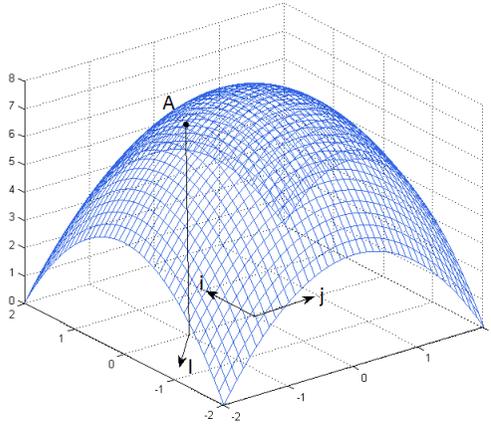

Figure 1.1: $f(x, y) = 8 - x^2 - y^2$

For example, figure 1.1 illustrates the surface $f(x, y) = 8 - x^2 - y^2$ over $x \in [-2, 2]$ and $y \in [-2, 2]$. The gradient of $f(x, y)$ is given by $\nabla f(x, y) = -2x\mathbf{i} - 2y\mathbf{j}$ or $\nabla f(x, y) = \langle -2x, -2y \rangle$. The bases vectors $\mathbf{i}$ and $\mathbf{j}$ are illustrated at point $(0, 0)$ (where the gradient is zero), and a unit vector $\mathbf{l} = \langle -\sqrt{.5}, -\sqrt{.5} \rangle$ is illustrated at point $(0, -1)$. Thus the dot product $\mathbf{l} \bullet \nabla f(0, -1)$ gives the slop of the surface at point $\mathbf{A}$ moving in the direction of vector $\mathbf{l}$, in this case $\langle -\sqrt{.5}, -\sqrt{.5} \rangle \bullet \langle 0, 2 \rangle = -2\sqrt{.5}$.

---

[1] Recall the dot product is the sum of the products of each element in one vector times is corresponding element in the second. That is, $\langle 2, 1 \rangle \bullet \langle 5, 3 \rangle = 2 \times 5 + 1 \times 3 = 13$

If we move along a **level curve**, the space in the (x, y) plane where $z = k$ (i.e. letting z equal some constant) then the rate of change of z will be zero (since that's what defines a level curve). If $\mathbf{u}$ is a unit vector tangent to the level curve (i.e. pointing along the level curve), then since z is not changing in that direction

$$\nabla f(x, y) \bullet \mathbf{u} = 0$$

thus $\nabla f(x, y)$ is orthogonal (perpendicular) to the level curve. [2] Figure 1.2 illustrates the function $f(x, y) = 8 - x^2 - y^2$ with
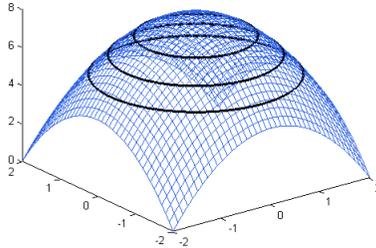
Figure 1.2: Level curves at $f(x, y) = 7$, $f(x, y) = 6$, and $f(x, y) = 5$

level curves drawn at $f(x, y) = 7$, $f(x, y) = 6$, and $f(x, y) = 5$.

For a constrained optimization problem we have two functions, say $z = f(x, y)$ subject to $0 = g(x, y)$. By setting z equal to a series of constants we can produce a series of level curves. The largest (smallest) constant k which we can assign to z will be the one which is tangent (that is, has a shared tangent line in the x, y plane) to the constraint $0 = g(x, y)$, and

---

[2]Recall that if the dot product of two vectors is zero than they are orthogonal.

thus we have identified the maximum (minimum) value z can take subject to $0 = g(x, y)$. This is illustrated in figure 1.3, a
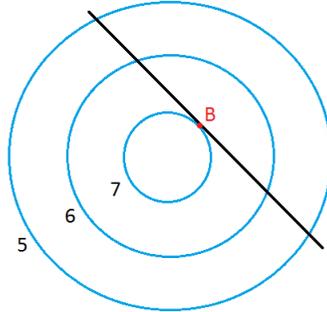
Figure 1.3: Level curves with constraint

top down view of the function (ie looking down onto the x, y plane) $f(x, y) = 8 - x^2 - y^2$ with the level curves given by $f(x, y) = 7$, $f(x, y) = 6$, and $f(x, y) = 5$. In this example the function $f(x, y)$ is subject to the constraint $0 = x + y - 2\sqrt{.5}$ or $2\sqrt{.5} = x + y$. The maximum occurs where the constraint is tangent to the level curve $f(x, y) = 7$.

If the constraint curve, $0 = g(x, y)$, and the level curve of $z = f(x, y)$ share the same tangent line then their gradients (which are perpendicular to the tangent line) must also have the same direction though they may have different magnitudes. That is, at the constrained maximum (minimum)

$$\nabla f(x, y) = \lambda \nabla g(x, y)$$

or

$$f_x(x, y)\mathbf{i} + f_y(x, y)\mathbf{j} = \lambda g_x(x, y)\mathbf{i}$$
$$+ \lambda g_y(x, y)\mathbf{j}$$

which implies that $f_x(x, y) = \lambda g_x(x, y)$ and $f_y(x, y) = \lambda g_y(x, y)$ or, letting $\mathcal{L} = f(x, y) - \lambda g(x, y)$ the constrained maximum (minimum) is identified by

$$\mathcal{L}_x(x, y) = 0$$

$$\mathcal{L}_y(x, y) = 0$$

and

$$g(x, y) = 0$$

For the example $f(x, y) = 8 - x^2 - y^2$ subject to $2\sqrt{.5} = x + y$ the Lagrangian is

$$\mathcal{L} = 8 - x^2 - y^2 - \lambda(x + y - 2\sqrt{.5})$$

thus the constrained maximum is identified by

$$\mathcal{L}_x(x, y) = -2x - \lambda = 0$$

$$\mathcal{L}_y(x, y) = -2y - \lambda = 0$$

thus $2x = 2y$ and $x = y$, and

$$2\sqrt{.5} = x + y$$

Therefore $2\sqrt{.5} = 2x$ and the solution is $x = \sqrt{.5}$, $y = \sqrt{.5}$. The maximized value of $f(x, y)$ subject to $2\sqrt{.5} = x + y$ is 7.

1.1 Economic Application

The Lagrangian, as the function $\mathcal{L}$ is called, is frequently used in economics. One of its primary applications is maximizing

consumer utility subject to a budget constraint. For example, if a consumer's utility (our $f(x, y)$ function from above, called the **objective function**) is defined by

$$U(x_1, x_2) = ln(x_1) + ln(x_2)$$

where x_1 and x_2 are two consumption goods and our constraint, $g(x, y)$ from the previous section, is defined by

$$g(x_1, x_2) = p_1 x_1 + p_2 x_2 - I = 0$$

where I is income and p_n is the price of good n then we can set up our Lagrangian as

$$\mathcal{L} = U(x_1, x_2) - \lambda(p_1 x_1 + p_2 x_2 - I)$$

Applying the above conditions to find the point at which utility is maximized we have

$$\mathcal{L}_{x_1} = U_{x_1}(x_1, x_2) - \lambda g_{x_1}(x_1, x_2) = \frac{1}{x_1} - \lambda p_1 = 0$$

$$\mathcal{L}_{x_2} = U_{x_2}(x_1, x_2) - \lambda g_{x_2}(x_1, x_2) = \frac{1}{x_2} - \lambda p_2 = 0$$

and the constraint

$$g(x_1, x_2) = p_1 x_1 + p_2 x_2 - I = 0$$

From the first two conditions we have $x_2 p_2 = x_1 p_1$. Plugging this into the budget constraint, $g(x_1, x_2)$ allows us to solve for quantities: $x_1 = \frac{I}{2p_1}$ and $x_2 = \frac{I}{2p_2}$.

In an unconstrained multivariate optimization problem we would use a **Hessian** matrix to determine whether our solution meets the criteria for a maximum or minimum. For example, if we wanted to optimize

$$y = x_1 x_2 - x_1^2 - 2x_2^2 + 7x_1$$

the first order conditions are

$$x_2 - 2x_1 + 7 = 0$$

$$x_1 - 4x_2 = 0$$

thus at the optimum $x_1 = 4$ and $x_2 = 1$. The Hessian matrix for this problem is

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix}$$

where y_{nm} refers to $\frac{\partial^2 y}{\partial x_n \partial x_m}$ or

$$\begin{bmatrix} -2 & 1 \\ 1 & -4 \end{bmatrix}$$

An optimum is a minimum if all the determinants of the **principle minors** of the Hessian are positive where the principle minors are $\mathbf{H_1}$, the first upper left element, y_{11} above, $\mathbf{H_2}$, the square matrix of the four upper left most elements, the entire matrix in the above example, $\mathbf{H_3}$, the square matrix of the nine upper left most elements, and so on. An optimum is a maximum if the determinant of the first principle minor is negative, the second positive, the third negative, fourth positive, and so on alternating signs. For the above example with only two variables the determinant of the first principle minor is -2 and the second 7 thus our solution meets the criteria for a maximum. The only difference with a constrained optimization problem is that we use a "bordered Hessian" and that we start from the second principle minor. The "border" is just the first partials of the constraint function with a zero in place of element (1,1). That is, our bordered Hessian for the above constrained optimization problem is

$$\begin{bmatrix} 0 & g_1 & g_2 \\ g_1 & U_{11} & U_{12} \\ g_1 & U_{21} & U_{22} \end{bmatrix}$$

where g_1 is shorthand for g_{x_1}. Plugging in the values for each element we have

$$
\begin{bmatrix}
0 & p_1 & p_2 \\
p_1 & -\frac{1}{x_1^2} & 0 \\
p_2 & 0 & -\frac{1}{x_2^2}
\end{bmatrix}
$$

The determinant of the second principle minor is $0 - p_1^2$, clearly negative, and of the third principle minor $-p_1(p_1(-\frac{1}{x_2^2}) - 0) + p_2(0 - p_2(-\frac{1}{x_1^2}))$, which is positive, thus our solution meets the criteria for a maximum.

1.2   The Envelope Theorem for Constrained Optimization

The envelope theorem is a handy short-cut which allows us to evaluate the effect of a change in an exogenous variable on the objective function (utility above) *at the optimum*. From the previous example, maximizing

(1.1) $U(x_1, x_2,) = ln(x_1) + ln(x_2)$

subject to

(1.2) $g(x_1, x_2; p_1, p_2, I) = p_1 x_1 + p_2 x_2 - I = 0$

where p_1, p_2, and I, are exogenous variables we can write maximized utility as $U(x_1^*, x_2^*; p_1, p_2, I)$ where x^* denotes the optimal value of x. Then if we were interested in, for example, $\frac{dU^*}{dp_1}$ (the change in U at the optimum with respect to p_1) we have (using the chain rule for derivatives)

$$
\frac{dU^*}{dp_1} = U_{x_1}^* x_{1p_1}^* + U_{x_2}^* x_{2p_1}^* + U_{p_1}^*
$$

From the first order conditions of the Lagrangian associated with 1.1 and 1.2[3] we know that at the optimum $U^*_{x_1} = \lambda^* g_{x_1}$ and $U^*_{x_1} = \lambda^* g_{x_1}$ thus

(1.3) $$\frac{dU^*}{dp_1} = \lambda^* [g_{x_1} x^*_{1p_1} + g_{x_2} x^*_{2p_1}] + U^*_{p_1}$$

From the constraint we know that

$$g(x^*_1, x^*_2) = 0$$

thus differentiating both sides with respect to p_1

$$g_{x_1} x^*_{1p_1} + g_{x_2} x^*_{2p_1} + g_{p_1} = 0$$

or

$$-g_{p_1} = g_{x_1} x^*_{1p_1} + g_{x_2} x^*_{2p_1}$$

Plugging this into 1.3 we have our result,

$$\frac{dU^*}{dp_1} = -\lambda^* g_{p_1} + U^*_{p_1} = L_{p_1}$$

The envelope theorem also illustrates the interpretation of the Lagrange multiplier. The derivative of utility with respect to income at the optimum is simply

$$\frac{dU^*}{dI} = \lambda^*$$

implying that the marginal utility of income is λ^*. Because of this in economics λ^* is often called the "shadow price" since it measures the price, in utility, of the constraint (in this case income).

---

[3]As a reminder, the associated Lagrangian is $\mathcal{L} = ln(x_1) + ln(x_2) - \lambda(p_1 x_1 + p_2 x_2 - I)$

1.3   Summary

Lagrange's Method is a technique for solving constrained opti-
mization problems. Given an objective function $z = f(x_1, x_2, \ldots)$
and a constraint $0 = g(x_1, x_2, \ldots)$ we can write the Lagrangian
as

$$\mathcal{L} = f(x_1, x_2, \ldots) - \lambda g(x_1, x_2, \ldots)$$

where λ is the Lagrange multiplier. The first order conditions
for an optimum are

$$\mathcal{L}_{x_1} = 0$$

$$\mathcal{L}_{x_2} = 0$$

$$\vdots$$

$$\mathcal{L}_{x_i} = 0$$

$$\vdots$$

for every control variable and

$$\mathcal{L}_\lambda = 0$$

which is just the constraint $0 = g(x_1, x_2, \ldots)$.
    If the constraint need not hold with equality, that is, if our
problem is given by $z = f(x_1, x_2)$ subject to $g(x, y) \geq 0$, $x_1 \geq 0$,
$x_2 \geq 0$ then our Lagrangian is unchanged, that is

$$\mathcal{L} = f(x_1, x_2) - \lambda g(x_1, x_2)$$

However, the conditions for an optimum are given by

$$\mathcal{L}_{x_1} \leq 0$$

$$\mathcal{L}_{x_2} \leq 0$$

$$x_1 \mathcal{L}_{x_1} = 0$$

$$x_2 \mathcal{L}_{x_2} = 0$$

and the constraints

$$g(x, y) \geq 0$$
$$x_1 \geq 0$$
$$x_2 \geq 0$$

Instead of working with inequality constraints in practice it is a good idea to figure out which constraints (if there are more than one) bind and use those constraints with equality, dropping the constraints that do not bind.

Finally, the envelope theorem provides a useful short cut for finding the impact of a change in an exogenous variable on the objective function at the optimum. The envelope theorem states that for the Lagrangian

$$\mathcal{L} = f(x_1, x_2, \ldots ; p_1, p_2, \ldots) - \lambda g(x_1, x_2, \ldots)$$

where the objective function is $z = f(x_1, x_2, \ldots ; p_1, p_2, \ldots)$ with endogenous variables $x_1, x_2, \ldots$ and exogenous variables $p_1, p_2, \ldots$

$$\frac{dz^*}{dp_j} = \frac{\partial \mathcal{L}}{\partial p_j}$$

where z^* indicates the objective function *at the optimum*.

Further Reading

The Lagrangian is covered in most calculus textbooks. The primary reference for this chapter is Varberg, Purcell, and Rigdon's (2000) *Calculus, Eighth Edition*, a very clear text. Mathematical economics texts should also cover the Lagrangian, for example Dowling (2011) *Schaum's Outline of Mathematical Economics* or Chaing and Wainright's (2004) *Fundamental Methods of Mathematical Economics*. The Wikipedia article on the Lagrange Multiplier is also quite good.

Chapter 2

Dynamic Programming

In this chapter we will consider the problem of a consumer maximizing utility subject to a budget constraint, but the techniques are applicable to a wide range of problems. As a starting point let the consumer's utility (which we assume is well behaved) in any one period be

$$u(c)$$

subject to the budget constraint

$$0 = I - pc$$

The solution to this problem is determined by the exogenous variables I and p thus we can say the consumer attaches a unique utility value to a given income and set of prices (p and c can be vectors), in this case

$$V(I, p) = u(c^*)$$

That is, as long as the consumer is optimizing consumption then $V(I, p)$, the value of a certain income and set of prices, is determined by those exogenous variables. This simple observation is the foundation of dynamic programming.

2.1   The Sequential Approach

We can easily extend the idea of a **value function**, $V(\cdot)$, to a dynamic setting. As a simple example imagine a consumer who lives for infinite periods with a cake (his only consumption good) of size W_0 in period 0. If our consumer's utility is defined by $\sum_{t=0}^{\infty} \beta^t u(c_t)$ then the sequential approach to solving this problem is to note the budget constraint is in this case $W_{t+1} = W_t - c_t$ (since the amount of cake in period $t+1$ is the amount of cake in period t less what he eats) and then to maximize the Lagrangian

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t u(c_t) + \lambda_t(W_t - W_{t+1} - c_t)$$

This gives us the first order conditions

$$\mathcal{L}_{c_t} = \beta^t u'(c_t) - \lambda_t = 0$$

$$\mathcal{L}_{w_{t+1}} = -\lambda_t + \lambda_{t+1} = 0$$

which combine to give us our "Euler equation"

$$u'(c_t) = \beta u'(c_{t+1})$$

Thus W_0 completely identifies the solution to our problem (though we haven't actually solved for it since that would require some functional form for $u(c_t)$) and we can denote the value of a size W_0 cake as

(2.1) $$V_\infty(W_0) = \sum_{t=0}^{\infty} \beta^t u(c_t^*)$$

2.2   The Dynamic Programming Approach

Dynamic programming provides an alternative and more versatile approach to solving problems like the above cake eating

example. The insight of dynamic programming, developed by
Richard Bellman in the 1960s, is that so long as we know our
agent is optimizing his utility we can break the above problem
down into a simple two period problem. Using the above exam-
ple, we can thus write equation 2.1 as

$$V_\infty(W_0) = \max_{c_0}\{u(c_0) + \beta V_\infty(W_0 - c_0)\}$$

which, recalling our budget constraint $W_{t+1} = W_t - c_t$ is simply

$$V_\infty(W_0) = \max_{W_1}\{u(W_0 - W_1) + \beta V_\infty(W_1)\}$$

Thus our problem becomes a matter of, given the size of the cake
today (W_0), choosing the optimal size of tomorrow's cake (W_1).
Because we are only now dealing with two periods (instead of
an infinite number of periods) dynamic programming problems
conventionally do not include time subscripts and instead use $'$
to denote the future period (this can get a bit annoying since $'$
also denotes a derivative). Thus our problem would typically be
written

$$(2.2) \qquad V(W) = \max_{W'}\{u(W - W') + \beta V(W')\}$$

Equation 2.2 is called the **Bellman equation** for this problem.

2.3   An Example

Suppose in our above example that $u(c) = \ln(c)$ and our indi-
vidual only consumes for two periods. To find the value function
we need to solve

$$(2.3) \qquad V_2(W) = \max_{W'}\{\ln(W - W') + \beta \ln(W')\}$$

Taking the first order condition (FOC) with respect to our choice variable, W', we have

$$\frac{1}{W - W'} = \beta \frac{1}{W'}$$

thus

$$W' = \frac{\beta W}{1 + \beta}$$

Plugging this back into (3)

$$V_2(W) = \ln(W(1 - \frac{\beta}{1 + \beta})) + \beta \ln(\frac{\beta W}{1 + \beta})$$

$$= \ln(W) + \ln(\frac{1}{1 + \beta}) + \beta \ln(W) + \beta \ln(\frac{\beta}{1 + \beta})$$

$$= \ln(\frac{1}{1 + \beta}) + \beta \ln(\frac{\beta}{1 + \beta}) + (1 + \beta) \ln(W)$$

The first two terms are just a function of the exogenous parameter β so we'll call that bit A_2. Thus our solution is

$$(2.4) \qquad V_2(W) = A_2 + (1 + \beta) \ln(W)$$

Now that we know $V_2(W)$ we can extend the model one period. Thus our new problem becomes

$$(2.5) \qquad V_3(W) = \max_{W'}\{\ln(W - W') + \beta V_2(W')\}$$

where in 2.4 W is the amount of cake our consumer begins with and W' is the amount he has left after the first period (that is, for the second two periods). The FOC with respect to W' is now

$$\frac{1}{W - W'} = \beta V_2'(W') = \beta(1 + \beta)\frac{1}{W'}.$$

Recall for the above FOC that W' in this three period problem is W from our two period problem. The above condition comes

from plugging in the solution to our two period problem for $V_2(W')$. Thus we have

$$W - W' = \frac{W'}{\beta(1 + \beta)}$$

$$W' = \frac{\beta + \beta^2}{1 + \beta + \beta^2} W$$

Plugging this back into (4) (again using the solution to the two period problem for $V(W')$)

$$V_3(W) = \ln\left(W - \frac{\beta + \beta^2}{1 + \beta + \beta^2} W\right) + \beta\left[(1+\beta)\ln\left(\frac{\beta + \beta^2}{1 + \beta + \beta^2} W\right) + A_2\right]$$

Which we can write as

$$\begin{aligned}V_3(W) &= \ln\left(\frac{1}{1+\beta+\beta^2}\right) + \beta(1 + \beta)\ln\left(\frac{\beta+\beta^2}{1+\beta+\beta^2}\right) \\ &+ \beta A_2 + (1 + \beta + \beta^2)\ln(W)\end{aligned}$$

Everything before the last term is again a function of the parameter β so we can call it A_3 to make things neater. Thus we have

(2.6) $$V_3(W) = A_3 + (1 + \beta + \beta^2)\ln(W)$$

From these two examples we might posit that the solution to the infinite horizon problem (when the number of periods $n \to \infty$) has the same form, that is, $V(W) = A + B\ln(W)$. With this guess our problem becomes

(2.7) $$A + B\ln(W) = \max_{W'}\{\ln(W - W') + \beta(A + B\ln(W'))\}$$

Optimizing with respect to our choice variable W' our first order condition is

$$\frac{1}{W - W'} = \beta B \frac{1}{W'}$$

thus

$$W' = \frac{\beta B}{1 + \beta B} W$$

Plugging this back into our value function or Bellman equation, (7), we have

$$
\begin{aligned}
A + B\ln(W) &= \ln(W) - \ln(1 + \beta B) + \beta A \\
&+ \beta B \ln(\frac{\beta B}{1 + \beta B}) + \beta B \ln(W)
\end{aligned}
$$

To solve this equation we can use the method of undetermined coefficients; equating terms multiplied by $\ln(W)$ we have

$$B = 1 + \beta B$$

thus

$$B = \frac{1}{1 - \beta}$$

and for the constant term (which is less interesting but we can solve for it anyhow)

$$A = \frac{1}{1 - \beta}(\beta B \ln(\frac{\beta B}{1 + \beta B}) - \ln(1 + \beta B)).$$

We can get an interesting result by plugging the solution for B back into the FOC. We then have

$$\frac{1 - \beta}{\beta} W' = W - W'$$

$$\frac{1}{\beta} W' = W$$

or

(2.8) $$W' = \beta W$$

which, recalling the budget constraint, is equivalent to saying

(2.9) $$c = W(1 - \beta)$$

Equations 2.8 and 2.9 (they say the same thing) provide a rule that our consumer (with infinite periods) will follow every period: he will save β of the cake for the future, that is, he will eat $1 - \beta$ of the cake every period. This rule, how to act given the state variable(s) (in this case there is only one state variable, W, the amount of cake in the current period) is called the **policy function** and is often of interest in dynamic programming problems.

2.4 Constrained Maximization

Finding an expression for the value function is often difficult and does not usually yield a closed form solution. Luckily, however, we don't need an expression for the value function to solve many economic problems; instead we typically seek the conditions which maximize utility subject to any applicable constraints. The sequential approach of section 2.1 is one way to derive these conditions. Dynamic programming provides a second method by which to derive first order conditions. While in the section 2.1 example the benefits to the dynamic programming approach my seem scant, in certain contexts (typically when dealing with financial markets and/or contingent claims) using dynamic programming can vastly simplify deriving FOCs.

A Simple Example

To find the FOCs for a maximum given lifetime utility

$$\mathcal{U} = \sum_{t=0}^{\infty} \beta^t u(c_t)$$

subject to a sequential budget constraint

$$(1 + r)b_t + y = c_t + b_{t+1}$$

where r is the interest rate, b_t are bonds owned at time t (and b_{t+1} bonds to carry into the next period), y income in each period (which does not vary with time), and c_t consumption in period t we first have to define our state variables (the arguments of our value function). In this case the **state**, the variables our theoretical agent takes as given in solving his utility maximization problem, is determined by b_t and y. The **control variables**, those which our agent chooses in order to solve his maximization problem, are c_t and b_{t+1}. Thus the Bellman equation for this problem is (using time subscripts for clarity)

$$V(b_t; y) = \max_{c_t, b_{t+1}} \{u(c_t) + \beta V(b_{t+1}; y)\}$$

subject to the above budget constraint. We can then proceed in two ways; either we can plug the budget constraint into the Bellman equation and take the FOCs, or we can solve our constrained maximization problem using a Lagrangian, as in section 4.2.[1] Rewriting the Bellman equation using the budget constraint we have

$$V(b_t; y) = \max_{b_{t+1}} \{u((1+r)b_t + y - b_{t+1}) + \beta V(b_{t+1}; y)\}$$

Note that this eliminates the control variable c_t. The FOC for this problem is then

(2.10) $$u'(c_t) = \beta V_b(b_{t+1}; y)$$

The next step is to find $V_b(b_{t+1}; y)$ by again using the Bellman equation. Taking the derivative with respect to b_t

$$V_b(b_t; y) = u'(c_t)(1+r)$$

---

[1] For more complex problems solving the budget constraint and plugging this into the Bellman equation is often infeasible.

Iterating this forward one period (since we're looking for $V_b(b_{t+1}; y)$)

$$V_b(b_{t+1}; y) = u'(c_{t+1})(1 + r)$$

and thus, plugging this back into equation 2.10, we arrive at our first order condition for a maximum:

$$u'(c_t) = \beta(1 + r)u'(c_{t+1})$$

A Slightly More Involved Example

As a modification to the previous problem, suppose that bonds are not available, that production is a function of capital $y_t = f(k_t)$ and that investment in capital entails a cost $h(i_t)$ where $i_t = k_{t+1} - k_t$. While the form of lifetime utility is unchanged, our budget constraint is now

$$f(k_t) = c_t + i_t + h(i_t)$$

Substituting out i_t from the above expression the Bellman equation is

$$V(k_t) = \max_{c_t, k_{t+1}} \{(c_t) + \beta V(k_{t+1})\}$$

subject to the above constraints. Note that c_t and k_{t+1} are control variables and k_t is our only state variable. The variable λ_t in this problem is called a **costate variable** (its value is implied by the state and cannot be chosen by our agent). We can write the Lagrangian for this constrained maximization problem as

(2.11)
$$\begin{aligned} \mathcal{L} = \quad & u(c_t) + \beta V(k_{t+1}) + \\ & \lambda_t \Big(f(k_t) - (c_t + k_{t+1} - k_t + h(k_{t+1} - k_t)) \Big) \end{aligned}$$

with the FOCs

(2.12)
$$u_c(c_t) = \lambda_t$$

(2.13) $\qquad \beta V_k(k_{t+1}) = \lambda_t(1 + h_i(k_{t+1} - k_t))$

We then need to find $V_k(k_{t+1})$. Noting that $V(k_t)$ is our objective function we can use the envelope theorem for our Lagrangian which states that

$$\frac{dV(k_t)}{dk_t} = \frac{\partial \mathcal{L}}{\partial k_t}$$

Thus the **envelope condition** is

$$V_k(k_t) = \lambda_t\big(f_k(k_t) + 1 + h_i(k_{t+1} - k_t)\big)$$

Iterating this forward one period and plugging into equation 2.13 we thus have the FOC for our constrained maximum

$$\beta u_c(c_{t+1})\bigg(f_k(k_{t+1}) + 1 + h_i(i_{t+1})\bigg) = u_c(c_t)(1 + h_i(i_t))$$

Further Reading

The physical sciences, engineering, and computer sciences tend to be ahead of economics when it comes to advances in dynamic programming and optimal control theory. Though not oriented towards economics, Dimitri Bertsekas' *Dynamic Programming and Optimal Control* is probably the definitive (two volume) text. Professor Bertsekas at MIT has a whole range of texts on optimal control and numerical methods with many potential applications to economics; you can have a look at his work at www.athenasc.com. For a more economics oriented approach Adda and Cooper (2003) is a good place to start. Ljungqvist and Sargent (2012) also cover dynamic programming though their treatment is less approachable to the beginning student.

Chapter 3

Constrained Optimization in Continuous Time

3.1 Derivation

The current value[1] Hamiltonian addresses problems of the form

$$\max \int_0^\infty v(k_t, c_t) e^{-\rho t} dt$$

subject to

$$\dot{k}_t = g(k_t, c_t)$$

---

[1]The current value of an asset refers to its un-discounted value, whereas present value refers to its value today (the terminology is confusing as they sound like they ought to mean the same thing) discounted by the interest rate or time preference. For example, with a period interest rate of 2% the current value of a bond that pays \$100 next period is \$100 while the present value of the same bond this period is $98 \approx \frac{1}{1+.02} 100$. The current valued Hamiltonian is so called because the first order conditions will not include the time discount $e^{-\rho t}$.

Here k_t denotes the state at time t, c_t is our control variable, and ρ is the rate at which we discount the objective function $v(k_t, c_t)$. In macroeconomics $v(k_t, c_t)$ is typically a utility function (k_t doesn't normally enter utility but we can include it to keep our derivation more general) and $g(k_t, c_t)$ is a budget constraint. Following Kamien and Schwartz (2012), it is useful to consider discrete time periods of length τ and then take the limit as $\tau \to 0$. Thus we can write the problem as

$$\max \sum_{t=0}^{\infty} v(k_t, c_t)\tau e^{-\rho t}$$

subject to

$$k_{t+\tau} = g(k_t, c_t)\tau + k_t$$

In discrete time our problem becomes a dynamic programming problem for which we can write the value function as[2]

$$V(k_t) = \max_{c_t} \left\{ v(k_t, c_t)\tau + e^{-\rho \tau} V(g(k_t, c_t)\tau + k_t) \right\}$$

Shifting $V(k_t)$ to the right hand side of the equation

$$0 = \max_{c_t} \left\{ v(k_t, c_t)\tau + e^{-\rho \tau} V(g(k_t, c_t)\tau + k_t) - V(k_t) \right\}$$

In order to make the problem more workable we approximate $e^{-\rho \tau}$ around $\tau = 0$, thus

$$e^{-\rho \tau} = 1 - \rho \tau + O(\tau)$$

where $O(\tau) = \frac{1}{2}\rho^2(\tau - 0)^2 + \frac{1}{6} - \rho^3(\tau - 0)^3 + ...$, that is, $O(\tau)$ contains the higher order terms of the approximation.[3] Importantly, $O(\tau)$ converges to zero faster than the first order term in

---

[2]The one period ahead discount is $e^{-\rho \tau}$ since the time elapsed between t and $t + 1$ is τ.

[3]See section 4.3 for a discussion of linear approximation

the approximation due to the squared, cubed, and higher order terms. Thus we have

$$0 = \max_{c_t} \{v(k_t, c_t)\tau + (1 - \rho\tau + O(\tau))V(g(k_t, c_t)\tau + k_t) - V(k_t)\}$$

The next trick is to get $V(g(k_t, c_t)\tau + k_t) - V(k_t)$ to look like a derivative as $\tau \to 0$. To this end we divide by τ,

$$
\begin{aligned}
0 &= \max_{c_t} \left\{ v(k_t, c_t) + \frac{V(g(k_t,c_t)\tau+k_t)-V(k_t)}{\tau} \right. \\
&+ \left. \left(-\rho + \frac{O(\tau)}{\tau}\right) V(g(k_t, c_t)\tau + k_t) \right\}
\end{aligned}
$$

and multiply and divide the fraction $\frac{V(g(k_t,c_t)\tau+k_t)-V(k_t)}{\tau}$ by $g(k_t, c_t)$

$$
\begin{aligned}
0 &= \max_{c_t} \left\{ v(k_t, c_t) + \frac{(V(g(k_t,c_t)\tau+k_t)-V(k_t))g(k_t,c_t)}{\tau g(k_t,c_t)} \right. \\
&+ \left. \left(-\rho + \frac{O(\tau)}{\tau}\right) V(g(k_t, c_t)\tau + k_t) \right\}
\end{aligned}
$$

Then, taking the limit as $\tau \to 0$ we have

(3.1) $$0 = \max_{c_t} \{v(k_t, c_t) + V_k(k_t)g(k_t, c_t) - \rho V(k_t)\}$$

If we define $\lambda = V_k(k_t)$, the first order condition for c_t in equation 3.1 is

(3.2) $$v_c(k_t, c_t) + \lambda g_c(k_t, c_t) = 0$$

Differentiating 3.1 with respect to time yields

$$
\begin{aligned}
0 &= v_k(k_t, c_t)\dot{k}_t + v_c(k_t, c_t)\dot{c}_t + \dot{\lambda}_t g(k_t, c_t) \\
&+ \lambda(g_k(k_t, c_t)\dot{k}_t + g_c(k_t, c_t)\dot{c}_t) - \rho\lambda\dot{k}_t
\end{aligned}
$$

or

$$
\begin{aligned}
0 &= v_k(k_t, c_t)\dot{k}_t + [v_c(k_t, c_t) + \lambda g_c(k_t, c_t)]\dot{c}_t + \dot{\lambda}_t g(k_t, c_t) \\
&+ \lambda g_k(k_t, c_t)\dot{k}_t + -\rho\lambda\dot{k}_t
\end{aligned}
$$

But from the first order condition for c_t, equation (2), $v_c(k_t, c_t) + \lambda g_c(k_t, c_t) = 0$ thus

$$0 = v_k(k_t, c_t)\dot{k}_t + \dot{\lambda}_t g(k_t, c_t) + \lambda g_k(k_t, c_t)\dot{k}_t + -\rho\lambda\dot{k}_t$$

or, dividing by $\dot{k}_t = g(k_t, c_t)$

(3.3) $$\dot{\lambda}_t = \rho\lambda - [v_k(k_t, c_t) + \lambda g_k(k_t, c_t)]$$

Equations 3.2 and 3.3, along with the law of motion for the state variable $\dot{k}_t = g(k_t, c_t)$, give the conditions for a solution to our problem. If we define the Hamiltonian as

$$\mathcal{H} = v(k_t, c_t) + \lambda_t g(k_t, c_t)$$

then the conditions for a solution are simply

$$\frac{\partial\mathcal{H}}{\partial c_t} = 0$$

$$\dot{\lambda}_t = \rho\lambda_t - \frac{\partial\mathcal{H}}{\partial k_t}$$

$$\dot{k}_t = g(k_t, c_t)$$

3.2   An Example

Suppose our problem (known as the AK model) was to maximize lifetime utility defined as

$$\mathcal{U} = \int_0^\infty e^{-\rho t}\frac{c_t^{1-\sigma}}{1-\sigma}dt$$

$$\dot{k}_t = ak_t - c_t - \delta k_t$$

Our current value Hamiltonian is then

$$\mathcal{H} = \frac{c_t^{1-\sigma}}{1-\sigma} + \lambda_t(ak_t - c_t - \delta k_t)$$

The first order condition with respect to c_t, equation 3.2 above, is

$$c_t^{-\sigma} = \lambda_t$$

and our condition for k_t, equation 3.3 above, is

$$\dot{\lambda}_t = \lambda_t(\rho + \delta - a)$$

Taking the derivative of the first condition with respect to time we have

$$\dot{\lambda}_t = -\sigma c_t^{-\sigma-1}\dot{c}_t$$

thus

$$\frac{\dot{c}_t}{c_t} = \frac{a - \rho - \delta}{\sigma}$$

That is, the growth rate of consumption is constant and determined by the above parameters. Accepting for now that

$$\frac{\dot{c}_t}{c_t} = \frac{\dot{k}_t}{k_t}$$

and using the law of motion for capital we have

$$\frac{a - \rho - \delta}{\sigma} = \frac{\dot{c}_t}{c_t} = \frac{\dot{k}_t}{k_t} = a - \frac{c_t}{k_t} - \delta$$

or

$$c_t = \frac{(a - \delta)(\sigma - 1) + \rho}{\sigma} k_t$$

which gives consumption as a function of capital (that is, our choice variable as a function of our state variable).

Proving that $\frac{\dot{c}_t}{c_t} = \frac{\dot{k}_t}{k_t}$

One approach to proving that $\frac{\dot{c}_t}{c_t} = \frac{\dot{k}_t}{k_t}$ is to calculate c_t and k_t as a function of t alone and to compare their growth rates. Beginning with c_t, we have from the first order conditions

$$\dot{c}_t - \frac{a - \rho - \delta}{\sigma} c_t = 0$$

We can find the general solution to a differential equation of the form

$$\dot{x}_t - p(t)x_t = q(t)$$

by noticing that

$$\frac{d}{dt} e^{-\int p(x)dt} x_t = e^{-\int p(x)dt} \dot{x}_t - p(t)e^{-\int p(x)dt} x_t$$

thus

$$e^{-\int p(x)dt} x_t = \int e^{-\int p(x)dt} q(t)$$

or

$$x_t = e^{\int p(x)dt} \int e^{-\int p(x)dt} q(t)$$

Thus in this specific case we have

$$\int \frac{d}{dt} c_t e^{-\frac{a-\rho-\delta}{\sigma} t} dt = \int 0 dt$$

thus

$$c_t e^{-\frac{a-\rho-\delta}{\sigma} t} = c_0$$

or

$$c_t = e^{\frac{a-\rho-\delta}{\sigma} t} c_0$$

For the law of motion for capital we have

$$\dot{k}_t - (a - \delta)k_t = -c_t$$

thus

$$k_t e^{-(a-\delta)t} = -\int e^{-(a-\delta)t} c_t$$

Plugging in for c_t:

$$k_t e^{-(a-\delta)t} = -\int e^{-(a-\delta)t} e^{\frac{a-\rho-\delta}{\sigma}t} c_0$$

or

$$k_t e^{-(a-\delta)t} = -c_0 \int e^{[\frac{a-\rho-\delta}{\sigma}-(a-\delta)]t}$$

thus

$$k_t e^{-(a-\delta)t} = -c_0 \left[\frac{a-\rho-\delta}{\sigma} - (a-\delta)\right]^{-1} e^{[\frac{a-\rho-\delta}{\sigma}-(a-\delta)]t}$$

or

$$k_t = -c_0 \left[\frac{a-\rho-\delta}{\sigma} - (a-\delta)\right]^{-1} e^{\frac{a-\rho-\delta}{\sigma}t}$$

The rate of growth of capital is therefore

$$\frac{\dot{k}_t}{k_t} = \frac{a-\rho-\delta}{\sigma}$$

which is, of course, the same as the rate of growth of consumption.

3.3 Summary

The Hamiltonian is a technique for solving dynamic constrained optimization problems in continuous time. Supposing an objective function

$$\max \int_0^\infty e^{-\rho t} v(k_t, c_t) dt$$

subject to a flow constraint (i.e. a constraint that determines the rate of change of state variables such as assets or capital)

$$\dot{k}_t = g(k_t, c_t)$$

the current value Hamiltonian is

$$\mathcal{H} = v(k_t, c_t) + \lambda_t g(k_t, c_t)$$

The first order conditions for a maximum are

$$\frac{\partial \mathcal{H}}{\partial c_t} = 0$$

$$\dot{\lambda}_t = \rho \lambda_t - \frac{\partial \mathcal{H}}{\partial k_t}$$

$$\dot{k}_t = g(k_t, c_t)$$

Further Reading

The derivation in this chapter draws primarily on Kamien and Schwartz (2012). However, Dowling (2011) has several examples of solving constrained optimization problems in continuous time. Other mathematical economics texts also cover optimal control theory including, for example, Chaing and Wainwright (2005). For more depth see Weber's (2011) *Optimal Control Theory with Applications in Economics*.

Chapter 4

The Steady State and Log Linearization

After deriving the first order conditions for a Macroeconomic model, the next step towards solving the model is to find the steady state or rest point of the system and to take an approximation around this point. So far, however, there has been no uncertainty in the model, which doesn't sound like an appealing way to model the real economy. Thus I begin by introducing uncertainty before moving on to the steady state and log linearisation.

4.1   Introducing Uncertainty

Uncertainty can enter a macroeconomic model in a number of ways, but for now I'll only introduce one form, an AR(1) shock process that enters linearly, which is fairly standard in dynamic stochastic general equilibrium (DSGE) models.[1] I will assume

---

[1] Dynamic because these models involve more than one period in times, stochastic because there is uncertainty (variables are subject to random

the shock process can be described by

$$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

where $\hat{a}_t = \log(a_t) - \log(\bar{a})$ and ϵ is a zero mean random variable. The form of the shock process means that it is already log linear and thus will fit in easily with the rest of our log linear model. Since the steady state of this equation, the point at which it comes to rest when $\epsilon_{t+1} = 0$ $\forall t$ and $t \to \infty$ is zero, the steady state of a_t itself is $\bar{a} = 1$ (the overbar here denotes steady state). Since there is uncertainty in the model the objective function (formerly lifetime utility) will now be in expectation (thus now expected lifetime utility) with E_t denoting the expectation operator conditional on all available information at time t.[2] Since we will be treating t as the current period we will have to add a new index to measure periods forward from the current period. Thus our lifetime utility function becomes

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s u(c_{t+s})$$

If our budget constraint is

$$a_t f(k_t) + (1 - \delta)k_t = c_t + k_{t+1}$$

then our current value[3] Lagrangian (using the sequential approach) is

$$\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \left[u(c_{t+s}) + \lambda_{t+s} \left(a_{t+s} f(k_{t+s}) + (1-\delta)k_{t+s} - c_{t+s} - k_{t+s+1} \right) \right]$$

---

shocks), general equilibrium because both supply and demand are endogenously determined

[2] For an excellent treatment of the expectations operator see Hsu (2011).

[3] current value because the first order condition for c_t does not include the time discount β.

which we need to optimize with respect to the two choice variables c_t and k_{t+1}. The first order condition with respect to c_t is (recalling that $E_t(x_t) = x_t$)

$$u'(c_t) = \lambda_t$$

Our second first order condition, that for k_{t+1} is

$$\lambda_t = E_t \beta \lambda_{t+1} \big[f'(k_{t+1}) + (1 - \delta) \big]$$

Combining the two for our Euler equation[4] we have

$$u'(c_t) = E_t \beta u'(c_{t+1}) \big[f'(k_{t+1}) + (1 - \delta) \big]$$

A note of caution: suppose we have the Euler equation

$$\frac{1}{c_t} = E_t \frac{1}{c_{t+1}} (1 + r)$$

It is important to remember that $E(\frac{1}{x}) = y$ does *not* imply $E(x) = \frac{1}{y}$ (nor does $E(xy) = 1$ imply $E(x)E(y) = 1$). Thus we can write our Euler as

$$1 = E_t \frac{c_t}{c_{t+1}} (1 + r)$$

but we can *not* write it as $E_t \frac{c_{t+1}}{c_t} = \beta(1 + r)$.

---

[4]In economics the Euler equation typically refers to the first order condition for a maximum that describes the evolution of consumption over time. It's called an Euler equation simply because it is a difference equation and Leonhard Euler did much of the original work on differential and difference equations. However, this can be confusing as in other disciplines the label Euler may have different meanings — there is an entire Wikipedia article titled "Things named after Leonhard Euler" (he was fairly prolific). For example, in time series econometrics the Euler relations are $e^{ibx} = \cos(bx) + i\sin(bx)$ and $e^{-ibx} = \cos(bx) - i\sin(bx)$.

4.2   Finding the Deterministic Steady State

The Deterministic Steady State

The steady state in a macroeconomic model is the point at which the variables of the model do not change from one period to the
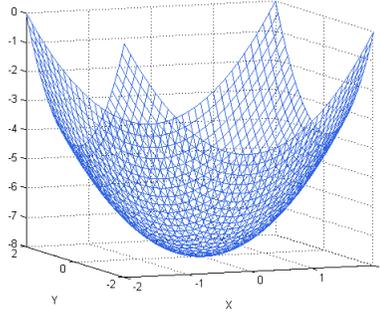

Figure 4.1: A globally stable system, $f(x, y) = x^2 + y^2 - 8$

next. In continuous time, this means that derivatives of variables with respect to time are zero. That is, for any variable x_t,

$$\dot{x}_t = 0$$

In discrete time, the steady state is defined as the point where for any variable x_t

$$x_t = x_{t+1} = \bar{x}$$

The over-bar above is typically used to denote the steady state, though we can also denote the steady state by simply omitting the time subscript on a variable (i.e. simply writing x). Conceptually, we can illustrate the steady state as the rest point of a system, an idea which is simple to illustrate in three dimensions.

Figure 4.1 illustrates a globally stable system with a rest point at (0,0). If you imagine placing a ball on the surface of figure 4.1, the rest point is the point at which it would not move. This system is globally stable because no matter where you initially place the ball, it will always end up at the rest point. The Solow model is an example of a globally stable model in macroeconomics. Figure 4.2 illustrates a saddle path stable
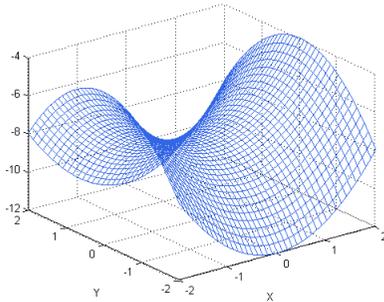

Figure 4.2: A saddle path stable system, $f(x, y) = x^2 - y^2 - 8$

system also with a rest point at (0,0). That is, if you placed a perfectly shaped ball at (0,0) it would remain still. However, placed elsewhere the ball will only end up at the rest point if it starts on the line $y = 0$ (this line denotes the model's saddle path or stable arm). In macroeconomics, we typically deal with saddle path stable systems, though that feature does not come into play until we solve the linearised system of difference or differential equations.

Examples

Equipped with the above definition, finding the deterministic steady state is usually fairly simple. The determanistic steady state is the rest point of a system in which there is no uncertainty

(all variables are known in advance), that is, it assumes $\epsilon_t = 0.$[5]
A few examples follow.

1. Model:

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \ln(c_{t+s})$$

   subject to

$$a_t y_t + (1 + r)b_t = c_t + b_{t+1}$$

   First Order Condition:

$$E_t \frac{c_t}{c_{t+1}} \beta(1 + r) = 1$$

   Deterministic Steady State:

$$\frac{\bar{c}}{\bar{c}} = \beta(1 + r)$$

   thus

$$\beta(1 + r) = 1$$

2. Model:

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{c_{t+s}^{1-\sigma}}{1 - \sigma}$$

   subject to

$$a_t k_t^{\alpha} + (1 - \delta)k_t = c_t + k_{t+1}$$

___

[5]This point is only an approximation of the actual rest point of the model
if there is consumption risk, that is, if consumption fluctuates over time.
The reason is that $u'(\cdot)$ is a convex function. Thus if $u'(c_t) = E_t u'(c_{t+1})$
(which would be the case if $\beta(1 + r) = 1$ as in the first example) then
using Jensen's inequality $c_t < E_t c_{t+1}$ implying consumption will grow over
time, not remain constant as we claim if we ignore uncertainty. This result
underlies the literature on precautionary saving.

First Order Condition:

$$c_t^{-\sigma} = \beta E_t c_{t+1}^{-\sigma} (a_{t+1} \alpha k_{t+1}^{\alpha-1} + 1 - \delta)$$

Deterministic Steady State:

$$1 = \beta \left(\alpha k^{\alpha-1} + 1 - \delta \right)$$

$$\frac{1}{\beta} - (1 - \delta) = \alpha k^{\alpha-1}$$

thus

$$k = \left[\frac{1}{\alpha} \left(\frac{1}{\beta} - (1 - \delta) \right) \right]^{\frac{1}{\alpha-1}}$$

or

$$k = \left[\frac{\alpha\beta}{1 - \beta(1 - \delta)} \right]^{\frac{1}{1-\alpha}}$$

3. Model:

$$\mathcal{U} = \int_{t=0}^{\infty} e^{-\rho t} \ln(c_t) dt$$

subject to

$$\dot{b}_t = y - rb_t - c_t$$

First Order Condition:

$$\dot{c}_t = (r - \rho) c_t$$

Steady State:

$$0 = (r - \rho)\bar{c}$$

thus

$$r = \rho$$

4.3   Log Linearizing

The equations that make up the first order conditions, budget constraints, and shock processes of our economic models form a system of non-linear difference (or differential) equations. In order to solve this system of equations we need to make it linear; if we want to look at a model's behaviour around the deterministic steady state then taking a linear approximation at that point will give us a fairly accurate idea of how all the variables interact.[6] As we move further from the steady state, however, the accuracy of our linear approximation decreases.
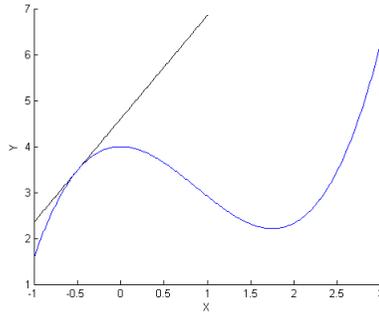

Figure 4.3: A linear approximation of $y = 4 - \frac{7}{4}x^2 + \frac{2}{3}x^3$ at $x = -\frac{1}{2}$

A linear approximation is the line, plane, or hyperplane that best approximates the behaviour of a function at a certain point. Figure 4.3 illustrates a linear approximation in two dimensions, figure 4.4 in three dimensions. As is evident in figure 4.3, the farther we move from the point of approximation, the less accurate

---

[6]However, a linear approximation will not take into account the impact of higher order terms like the variance of the shock process. For a comparison of a linear approximation with the true, non-linear solution to a simple model see chapter 8.

our approximation becomes (both the positions and the slopes diverge as we move away from the point of approximation). In
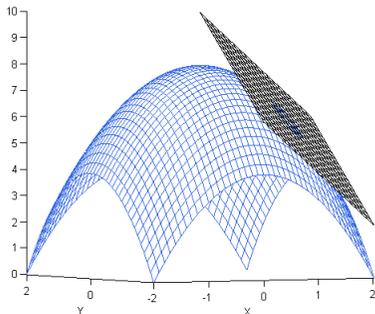

Figure 4.4: A linear approximation of $z = 8 - x^2 - y^2$ at $(1, -1)$

two dimensions the formula describing a linear approximation of some function $f(x)$ at point $\bar{x}$ is

$$(4.1) \qquad f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x})$$

$f(\bar{x})$ is called the zero order term and $f'(\bar{x})(x - \bar{x})$ is the first order term. We can generalize this approximation to an nth order approximation, called a **Taylor expansion**, given by

$$f(x) \approx f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + \frac{1}{2}f''(\bar{x})(x - \bar{x})^2 + \ldots + \frac{1}{n!}f^n(\bar{x})(x - \bar{x})^n$$

though we will typically only be concerned with first order (or linear) expansions. In more than one dimension our first order expansion around $f(\bar{x}_1, \ldots, \bar{x}_n)$ is

(4.2)

$$\begin{aligned} f(x_1, \ldots, x_n) &\approx f(\bar{x}_1, \ldots, \bar{x}_n) \\ &+ \nabla f(\bar{x}_1, \ldots, \bar{x}_n) \cdot \langle x_1 - \bar{x}_1, \ldots, x_n - \bar{x}_n \rangle \end{aligned}$$

or, writing equation 4.2 out explicitly,

$$
\begin{aligned}
f(x_1, \ldots, x_n) \approx\ & f(\bar{x}_1, \ldots, \bar{x}_n) \\
+\ & f_{x_1}(\bar{x}_1, \ldots, \bar{x}_n)(x_1 - \bar{x}_1) + \ldots \\
+\ & f_{x_n}(\bar{x}_1, \ldots, \bar{x}_n)(x_n - \bar{x}_n)
\end{aligned}
$$

This equation also generalizes to an pth order Taylor approximation though one has to be careful to include cross-partials as well.

The only twist on the above in economics is that we usually use log linearization instead of linearization in levels. That is, we often want our system to be linear in $\hat{x} = \ln(x) - \ln(\bar{x})$. Making our model linear in log deviations from the steady state as opposed to absolute deviations from the steady state allows for some non-linearity in the approximation, but more importantly allows us to talk about the model in percent deviations. That is, $\hat{x}$ gives the percent deviation of a variable from its steady state value. This is because

$$
\log(x) - \log(\bar{x}) \approx d\log(\bar{x}) = \frac{d\bar{x}}{\bar{x}} \approx \frac{x - \bar{x}}{\bar{x}}
$$

There are two main approaches to log linearizing; I introduce both here and one can use either as they yield the same result.

Uhlig's Approach to Log Linearizing

Let x_t be our variable of interest around the point x_0. Letting

$$
\hat{x}_t = \log(x_t) - \log(x_0) = \log(\frac{x_t}{x_0})
$$

that is, $\hat{x}_t$ is the log deviation of x_t from x_0, then

$$
x_t = x_0 e^{\hat{x}_t} = x_0 e^{\log(\frac{x_t}{x_0})}
$$

I will denote our point of approximation $\hat{x}_0$. Since we are interested in an approximation around the steady state, our point

of approximation will be $\hat{x}_0 = 0$, that is, $\log(x_t) = \log(x_0)$. We then have (with the corresponding terms from equation 4.1 written below)

$$x_0 e^{\hat{x}_t} \approx \underbrace{x_0 e^{\hat{x}_0}}_{f(\bar{x})} + \underbrace{x_0 e^{\hat{x}_0}}_{f'(\bar{x})} \underbrace{(\hat{x}_t - \hat{x}_0)}_{(x-\bar{x})}$$

Plugging in $\hat{x}_0 = 0$ this simplifies to

$$x_t = x_0 e^{\hat{x}_t} \approx x_0(1 + \hat{x}_t)$$

Products and powers are easily addressed by this approach. For example, linearizing

$$1 = x_t^{\rho} y_t^{1-\sigma}$$

around $\hat{x}_0 = 0, \hat{y}_0 = 0$, or, equivalently, around $\log(x_t) = \log(x_0)$, $\log(y_t) = \log(y_0)$ we have

$$1 = (x_0 e^{\hat{x}_t})^{\rho}(y_0 e^{\hat{y}_t})^{1-\sigma} = x_0^{\rho} y_0^{1-\sigma} e^{\rho \hat{x}_t + (1-\sigma)\hat{y}_t}$$

Taking a first order approximation this becomes (with the corresponding terms form equation 4.2 written below)

$$
\begin{aligned}
1 \quad \approx \quad & \underbrace{x_0^{\rho} y_0^{1-\sigma} e^{\rho \hat{x}_0 + (1-\sigma)\hat{y}_0}}_{f(\bar{x}_1, \bar{x}_2)} + \underbrace{\rho x_0^{\rho} y_0^{1-\sigma} e^{\rho \hat{x}_0 + (1-\sigma)\hat{y}_0}}_{f_x(\bar{x}_1, \bar{x}_2)} \underbrace{(\hat{x}_t - \hat{x}_0)}_{(x_1 - \bar{x}_1)} \\
+ \quad & \underbrace{(1-\sigma)x_0^{\rho} y_0^{1-\sigma} e^{\rho \hat{x}_0 + (1-\sigma)\hat{y}_0}}_{f_y(\bar{x}_1, \bar{x}_2)} \underbrace{(\hat{y}_t - \hat{y}_0)}_{x_2 - \bar{x}_2}
\end{aligned}
$$

Plugging in $\hat{x}_0 = 0$ and $\hat{y}_0 = 0$ we have simply

$$1 \approx x_0^{\rho} y_0^{1-\sigma}(1 + \rho \hat{x}_t + (1-\sigma)\hat{y}_t)$$

It is worth noting here that the zero order terms, $f(\bar{x}_1, \bar{x}_2)$ and 1 in the above example, will always cancel due to the steady state relationship (which in this case would be $1 = x_0^{\rho} y_0^{1-\sigma}$).

Thus in taking our first order approximations we can ignore (drop) zero order terms, and our above approximation becomes

$$0 \approx x_0^\rho y_0^{1-\sigma}(\rho\hat{x}_t + (1-\sigma)\hat{y}_t)$$

or

$$0 \approx \rho\hat{x}_t + (1-\sigma)\hat{y}_t$$

As an aside, we did not really have to approximate this equation. From

$$1 = x_0^\rho y_0^{1-\sigma}e^{\rho\hat{x}_t+(1-\sigma)\hat{y}_t}$$

we can simply take the log of each side of the equation yielding

$$0 = \ln(x_0^\rho y_0^{1-\sigma}) + \rho\hat{x}_t + (1-\sigma)\hat{y}_t$$

holding exactly. Since $x_0^\rho y_0^{1-\sigma} = 1$ we simply have

$$0 = \rho\hat{x}_t + (1-\sigma)\hat{y}_t$$

again holding exactly.

A Very Handy Trick

Notice that beginning with the equation

$$1 = x_t^\rho y_t^{1-\sigma}$$

we ended with

$$0 = \rho\hat{x}_t + (1-\sigma)\hat{y}_t$$

When log linearising it will often be the case that we can write our solution by

1. Ignoring the zero order terms because they will always cancel.

2.
- If the variables are written as products (such as $x_t^3 y_t^3 = x_t^{2\gamma} + 2x_t^\gamma y_t^\sigma + y_t^{2\sigma}$): writing the solution as the zero order term multiplied by the log deviation of each variable times its power. That is, the current example can be log linearized as

$$x_0^3 y_0^3 (3\hat{x}_t + 3\hat{y}_t) = x_0^{2\gamma} 2\gamma\hat{x}_t + 2x_0^\gamma y_0^\sigma (\gamma\hat{x}_t + \sigma\hat{y}_t) + y_0^{2\sigma} 2\sigma\hat{y}_t$$

- If the variables are still inside some "outer" function, for example $(x_t y_t)^3 = (x_t^\gamma + y_t^\sigma)^2$ where the "outer" functions are $(\cdot)^3$ and $(\cdot)^2$: writing the solution as the derivative of the "outer" function evaluated at the points of approximation times the log linearized "inner" functions. That is, the current example can be log linearized as

$$3(x_0 y_0)^2 (x_0 y_0 \hat{x}_t + x_0 y_0 \hat{y}_t) = 2(x_0^\gamma + y_0^\sigma)(x_0^\gamma \gamma\hat{x}_t + y_0^\sigma \sigma\hat{y}_t)$$

Of course, in this example the solutions to the first and second points above are the same.

Log Linearizing Using Total Differentiation

As an alternative approach we can use the fact that

$$\hat{x}_t = \log(x_t) - \log(\bar{x}) \approx d\log(\bar{x}) = \frac{d\bar{x}}{\bar{x}}$$

Thus if we can manipulate our equations into the form $\frac{d\bar{x}}{\bar{x}}$ by totally differentiating each side and potentially dividing by steady state values we can replace that variable with $\hat{x}_t$ to get our log linearized system.

Again using the example

$$1 = x_t^\rho y_t^{1-\sigma}$$

then totally differentiating each side we have

$$0 = \rho x_t^{\rho-1} y_t^{1-\sigma} dx_t + (1-\sigma) x_t^{\rho} y_t^{-\sigma} dy_t$$

$$0 = \rho x_t^{\rho} y_t^{1-\sigma} \frac{dx_t}{x_t} + (1-\sigma) x_t^{\rho} y_t^{1-\sigma} \frac{dy_t}{y_t}$$

and finally, plugging in the steady state values of the variables

$$0 = \rho \bar{x}^{\rho} \bar{y}^{1-\sigma} \frac{d\bar{x}_t}{\bar{x}} + (1-\sigma) \bar{x}^{\rho} \bar{y}^{1-\sigma} \frac{d\bar{y}_t}{\bar{y}}$$

or

$$0 \approx \bar{x}^{\rho} \bar{y}^{1-\sigma} (\rho \hat{x}_t + (1-\sigma) \hat{y}_t)$$

and the steady state relationship $1 = \bar{x}^{\rho} \bar{y}^{1-\sigma}$

$$0 \approx \rho \hat{x}_t + (1-\sigma) \hat{y}_t$$

Examples

At this point it should also be clear why we assumed that the shock process takes the form $\hat{a}_{t+1} = \rho \hat{a}_t + \epsilon_{t+1}$; it is already log linear so we don't need to worry about approximating it.

1. For example 1 in section 4.2 we have two equations to linearize: the FOC and the budget constraint. The shock process, which also forms part of our system, is already log linear. We can write the FOC using Uhlig's technique as

    $$E_t \frac{c_t}{c_{t+1}} \beta(1+r) = 1$$

    $$E_t \frac{\bar{c}}{\bar{c}} e^{\hat{c}_t - \hat{c}_{t+1}} \beta(1+r) = 1$$

    Taking a first order approximation around $\hat{c}_t = 0$ and $\hat{c}_{t+1} = 0$ and using the steady state condition that $\beta(1+r) = 1$ we have

    $$E_t(1 + \hat{c}_t + \hat{c}_{t+1}) = 1$$

or

$$E_t \hat{c}_{t+1} = \hat{c}_t$$

For the budget constraint

$$a_t y_t + (1 + r)b_t = c_t + b_{t+1}$$

we have

$$\bar{a}\bar{y}e^{\hat{a}_t + \hat{y}_t} + (1 + r)\bar{b}e^{\hat{b}_t} = \bar{c}e^{\hat{c}_t} + \bar{b}e^{\hat{b}_{t+1}}$$

Taking a first order approximation around $\hat{x}_t = 0$ for each variable

$$\bar{a}\bar{y}(1 + \hat{a}_t + \hat{y}_t) + (1 + r)\bar{b}(1 + \hat{b}_t) = \bar{c}(1 + \hat{c}_t) + \bar{b}(1 + \hat{b}_{t+1})$$

or, dropping the zero order terms and recalling $\bar{a} = 1$,

$$\bar{y}(\hat{a}_t + \hat{y}_t) + (1 + r)\bar{b}\hat{b}_t = \bar{c}\hat{c}_t + \bar{b}\hat{b}_{t+1}$$

Thus we have a system of three linear difference equations:

$$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

$$E_t\hat{c}_{t+1} = \hat{c}_t$$

and

$$\bar{y}(\hat{a}_t + \hat{y}_t) + (1 + r)\bar{b}\hat{b}_t = \bar{c}\hat{c}_t + \bar{b}\hat{b}_{t+1}$$

2. For example 2 in section 4.2 we also have two equations to linearisze, again the FOC and budget constraint:

$$c_t^{-\sigma} = \beta E_t c_{t+1}^{-\sigma}(a_{t+1}\alpha k_{t+1}^{\alpha-1} + 1 - \delta)$$

and

$$a_t k_t^{\alpha} + (1 - \delta)k_t = c_t + k_{t+1}$$

Using total differentiation and beginning with the FOC we have

$$
\begin{aligned}
-\sigma c_t^{-\sigma-1} dc_t \;=\;& \beta E_t\Big(c_{t+1}^{-\sigma}\alpha k_{t+1}^{\alpha-1} da_{t+1} \\
+\;& c_{t+1}^{-\sigma} a_{t+1}\alpha(\alpha-1)k_{t+1}^{\alpha-2} dk_{t+1} \\
-\;& \sigma c_{t+1}^{-\sigma-1}(a_{t+1}\alpha k_{t+1}^{\alpha-1}+1-\delta)dc_{t+1}\Big)
\end{aligned}
$$

thus

$$
\begin{aligned}
-\sigma c_t^{-\sigma}\tfrac{dc_t}{c_t} \;=\;& \beta E_t\Big(c_{t+1}^{-\sigma}\alpha k_{t+1}^{\alpha-1} a_{t+1}\tfrac{da_{t+1}}{a_{t+1}} \\
+\;& c_{t+1}^{-\sigma} a_{t+1}\alpha(\alpha-1)k_{t+1}^{\alpha-1}\tfrac{dk_{t+1}}{k_{t+1}} \\
-\;& \sigma c_{t+1}^{-\sigma}(a_{t+1}\alpha k_{t+1}^{\alpha-1}+1-\delta)\tfrac{dc_{t+1}}{c_{t+1}}\Big)
\end{aligned}
$$

or

$$
\begin{aligned}
-\sigma\hat{c}_t \;=\;& E_t\Big[\beta\alpha\bar{k}^{\alpha-1}(\hat{a}_{t+1}+(\alpha-1)\hat{k}_{t+1}) \\
-\;& \beta\sigma(\alpha\bar{k}^{\alpha-1}+1-\delta)\hat{c}_{t+1}\Big]
\end{aligned}
$$

For the budget constraint, differentiating yields

$$
k_t^\alpha da_t + a_t\alpha k_t^{\alpha-1} dk_t + (1-\delta)dk_t = dc_t + dk_{t+1}
$$

or

$$
a_t k_t^\alpha \frac{da_t}{a_t} + a_t\alpha k_t^\alpha \frac{dk_t}{k_t} + (1-\delta)k_t\frac{dk_t}{k_t} = c_t\frac{dc_t}{c_t} + k_{t+1}\frac{dk_{t+1}}{k_{t+1}}
$$

which becomes (using the steady state value of $\bar{a}=1$)

$$
\bar{k}^\alpha\hat{a}_t + \left(\alpha\bar{k}^\alpha + (1-\delta)\bar{k}\right)\hat{k}_t = \bar{c}\hat{c}_t + \bar{k}\hat{k}_{t+1}
$$

We now have a system of three linear difference equations

$$
\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}
$$

$$
\begin{aligned}
E_t\Big[-\beta\alpha\bar{k}^{\alpha-1}\hat{a}_{t+1} + \beta\alpha(1-\alpha)\bar{k}^{\alpha-1}\hat{k}_{t+1} \;+\;& \\
\beta\sigma(\alpha\bar{k}^{\alpha-1}+1-\delta)\hat{c}_{t+1}\Big] \;=\;& \sigma\hat{c}_t
\end{aligned}
$$

and

$$
\bar{k}\hat{k}_{t+1} = \bar{k}^\alpha\hat{a}_t + \left(\alpha\bar{k}^\alpha + \bar{k}(1-\delta)\right)\hat{k}_t - \bar{c}\hat{c}_t
$$

Further Reading

Log linearization, being somewhat specific to macroeconomic models, is not something that is particularly well covered elsewhere. Taylor expansions are covered in most good calculus textbooks such as, for example Varberg, Purcell, and Rigdon (2000). Harald Uhlig's (1998) "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily", though somewhat dated now in terms of the coding component, remains an excellent reference for the topic.

Chapter 5

Solving the Stochastic System of Difference Equations

Thus far we've covered finding a deterministic steady state of macro models (our point of approximation) and log linearizing the system of difference equations which describes our model around that point. The next step is to solve that system of difference equations. This is typically considered the hardest part of solving macro models and thus tends to be neglected in the literature. In this chapter I cover three different ways to solve our linearized system of stochastic difference equations. The first approach is called undetermined coefficients. This is the simplest way to solve the system by hand since it doesn't involve matrices (and finding their eigenvalues and eigenvectors) but it gets messy quickly as models get more involved. The second two approaches use matrix methods. The first is my own technique based on what I call the "math class approach" to solving systems of difference equations (since it's based on standard material one might cover in a non-economics course on differential

and difference equations). The second is Blanchard and Kahn's (1980) approach which is more general than my own and is the technique used by software such as Dynare. Before going into each method, however, I discuss how to solve deterministic linear systems of difference equations and illustrate what we're actually trying to find — the saddle path stable solution or stable manifold — using phase diagrams.

5.1   Solving Deterministic Systems of Linear Difference Equations

This section outlines the standard "math class" approach to solving systems of linear difference equations.

Suppose we were looking at a system of linear difference equations

$$
\begin{aligned}
x_t &= a_{11}x_{t-1} + a_{12}y_{t-1} + b_1 \\
y_t &= a_{21}x_{t-1} + a_{22}y_{t-1} + b_2
\end{aligned}
$$

We can write this more generally (where the system can be of n equations) in matrix form as

(5.1) $$Y_t = AY_{t-1} + B$$

We proceed by guessing a solution to the **homogeneous equations** (that is, ignoring for the moment B) of the form

$$Y_t = k_i C_i \lambda_i^t$$

where k_i is a constant determined by the initial (and/or terminal) conditions and we need to solve for the vectors C_i and the scalars λ_i. The subscript i is included since there will be n independent solutions. Iterating our putative solution back one period we have

$$Y_{t-1} = k_i C_i \lambda_i^{t-1}$$

Plugging this into 5.1 yields

$$k_i C_i \lambda_i^t = A k_i C_i \lambda_i^{t-1}$$

or

$$C_i \lambda_i = A C_i$$

which, rearranging, states that

$$0 = (A - \lambda_i I) C_i$$

Thus λ_i are the eigenvalues of the matrix A and C_i the eigenvectors. For the **particular solution** (that is, focusing only on the vector of constants B), we have

$$Y_p = A Y_p + B$$

where Y_p is not a function of time as B does not depend on time. Simple matrix algebra yields

$$Y_p = (I - A)^{-1} B$$

The complete solution to our linear system of difference equations is then

$$Y_t = k_1 C_1 \lambda_1^t + k_2 C_2 \lambda_2^t + \ldots + k_n C_n \lambda_n^t + Y_p$$

In macroeconomics we are typically not concerned with the particular solution to a system of linear difference equations since we usually look at deviations from the steady state (or some other point of approximation), thus in most situations $B = [0]$.

One issue we are usually concerned with, however, is the stability of the system, particularly in the case of saddle path stability. Evidently, if one of the eigenvalues of A, that is, one of the λ_i's, is greater than one in absolute value our system will

diverge as $t \to \infty$. **Saddle path stability** is the case where some of the eigenvalues are greater than one (in magnitude) while others are smaller. In this case the system is only stable where the coefficients (the k_i's above) multiplying the unstable eigenvalues are zero. For example, suppose

$$Y_t = \begin{bmatrix} 1.5 & 1 \\ 0.5 & 1 \end{bmatrix} Y_{t-1} + \begin{bmatrix} -2 \\ -4 \end{bmatrix}$$

The eigenvalues of A satisfy

$$0 = (1.5 - \lambda_i)(1 - \lambda_i) - 0.5 = \lambda_i^2 - 2.5\lambda + 1 = (\lambda_i - 0.5)(\lambda_i - 2)$$

thus $\lambda_1 = 0.5$ and $\lambda_2 = 2$. The eigenvectors of this system satisfy, for $\lambda = 0.5$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} c_{1,1} \\ c_{2,1} \end{bmatrix}$$

and for $\lambda = 2$

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -0.5 & 1 \\ 0.5 & -1 \end{bmatrix} \begin{bmatrix} c_{1,1} \\ c_{2,1} \end{bmatrix}$$

Thus the homogeneous solution to this problem is

$$Y_{t,h} = k_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} 0.5^t + k_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} 2^t$$

The particular solution is

$$Y_p = (I - A)^{-1} B = \begin{bmatrix} 8 \\ -2 \end{bmatrix}$$

thus the complete solution to this problem is

$$Y_t = k_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} 0.5^t + k_2 \begin{bmatrix} 2 \\ -1 \end{bmatrix} 2^t + \begin{bmatrix} 8 \\ -2 \end{bmatrix}$$

Clearly, this system will only be stable when $k_2 = 0$. Stability in this example corresponds to a **transversality condition** in macroeconomics and is the second endpoint in our two point boundary value problem.[1] Supposing the two variables in Y_t to be $\begin{bmatrix} s_t \\ c_t \end{bmatrix}$ (which I will later call state and control) we then have

$$s_t = k_1 \left(\frac{1}{2}\right)^t + 8$$

$$c_t = -k_1 \left(\frac{1}{2}\right)^t - 2$$

thus

$$c_t = 6 - s_t$$

Our first boundary point is an initial condition for s_t, and we have solved a two point boundary value problem.

5.2 A Graphical Representation of the Solution

In chapter 4 we considered the model described by the objective function

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{(c_{t+s})^{1-\sigma}}{1-\sigma}$$

and the budget constraint

$$a_t k_t^{\alpha} = c_t + k_{t+1}$$

---

[1]The tranversality or "no Ponzi scheme" condition is a condition we impose on our models to ensure variables (wealth or capital in particular) do not explode (go to $+\infty$ or $-\infty$).

The first order conditions for an optimum in this model are the Euler equation

$$c_t^{-\sigma} = \beta E_t(a_{t+1}\alpha k_{t+1}^{\alpha-1} + 1 - \delta)c_{t+1}^{-\sigma}$$

and the budget constraint. Before log linearizing and solving the approximated model we can use a **phase diagram** as a graphical representation of the solution to get a feel for how this system will behave. A phase diagram illustrates the laws of motion which govern our system of equations around the deterministic steady state or rest point of the system. Dealing with the budget constraint first, we can find the conditions under which capital (k_t) is constant by re-writing the equation as
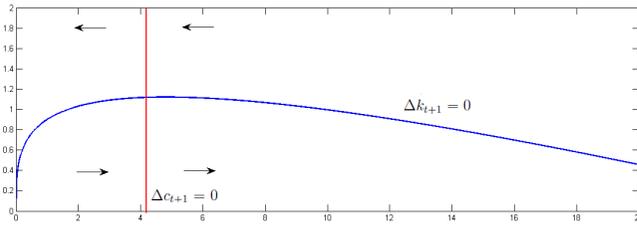

Figure 5.1: Illustrating the steady state with the law of motion for capital illustrated by arrows

$$(5.2) \qquad\qquad k_{t+1} - k_t = a_t k_t^\alpha - c_t - k_t$$

Then k_t is constant if $a_t k_t^\alpha - c_t - k_t = 0$. We can re-write the Euler equation as

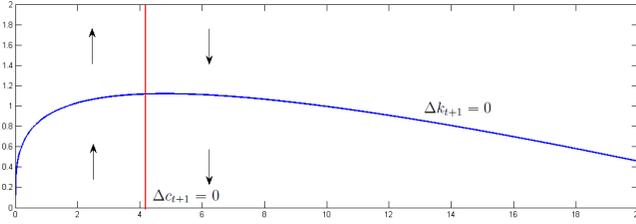$$E_t\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma}\beta(a_{t+1}\alpha k_{t+1}^{\alpha-1} + 1 - \delta) = 1$$

Figure 5.2: Illustrating the steady state with the law of motion for consumption illustrated by arrows

Then, abstracting from uncertainty[2], c_t is constant if

(5.3) $$\beta(\alpha k_{t+1}^{\alpha-1} + 1 - \delta) = 1$$

These two conditions are plotted in figure 5.1; their intersec-
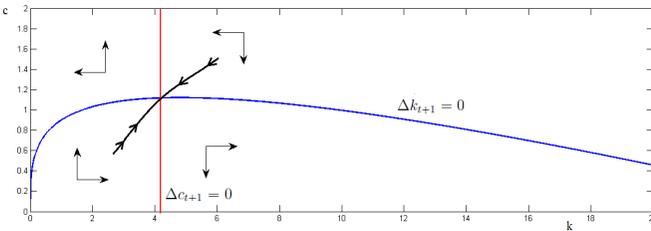

Figure 5.3: The saddle path or stable manifold (with arrows pointing towards the deterministic steady state) illustrates the only points from which the system of equations will converge towards the steady state.

---

[2]Abstracting from uncertainty is of course wrong if c_{t+1} fluctuates. Solving the model including the influence of this uncertainty is however vastly more difficult thus models typically deal with the deterministic steady state - the steady state that would occur if there was in fact no uncertainty.

tion is the deterministic steady state since at that point in the absence of any exogenous shocks both variables will remain constant. The next question is: how will the system behave if we are not at the steady state? To answer this question we can use equations 5.2 and 5.3. If, holding all else constant, we increase consumption in equation 5.2 then $k_{t+1} - k_t$ will be negative implying decreasing capital; this is also illustrated in figure 5.1 by the arrow pointing left above the line $\Delta k_{t+1} = 0$. On the other hand, if we decrease consumption in equation 5.2 then $k_{t+1} - k_t$ will be positive as illustrated by the arrow pointing right below the line $\Delta k_{t+1} = 0$. Similarly, if, holding all else equal (and abstracting from uncertainty), we increase capital in equation 5.3, $k_t^{\alpha-1}$ will decrease (α is less than 1) so that, to ensure the Euler equation is satisfied, $\left(\frac{c_{t+1}}{c_t}\right)^{-\sigma}$ must increase meaning $\left(\frac{c_{t+1}}{c_t}\right)$ must decrease; consumption will be decreasing. This is illustrated by the arrows pointing down to the right of $\Delta c_{t+1} = 0$. Conversely, if we increase capital in equation 5.3 holding all else equal then consumption will be increasing as illustrated by the arrows pointing up to the left of the line $\Delta c_{t+1} = 0$. These dynamics imply there is only one line along which the system will converge to the steady state. This line is the **saddle path** or **stable manifold** of our system of equations and is the solution that we seek.

5.3   Systems of Difference Equations in Macroeconomics Using Undetermined Coefficients

Undetermined coefficients is the easiest way to solve our systems of difference equations by hand. The advantage of matrix methods, discussed below, is that we can let a computer manipulate the relevant matrices quickly and easily. Undetermined coeffi-

cients proceeds by assuming a recursive law of motion (that is, a solution to our system of equations) and then plugging this assumed solution into our model. I will illustrate the approach with an example.

Suppose we are interested in solving a model in which consumers maximize

$$\mathcal{U} = E_t \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$

subject to the budget constraint

$$a_t k_t^{\alpha} = c_t + k_{t+1}$$

where a_t is a stochastic process governed by $\hat{a}_{t+1} = \rho \hat{a}_t + \epsilon_{t+1}$ where hatted variables denote log deviations from the steady state. The first order conditions for this problem are

$$1 = E_t \beta \frac{c_t}{c_{t+1}} a_{t+1} \alpha k_{t+1}^{\alpha-1}$$

and the budget constraint. Log linearising these conditions we have

$$0 = E_t \{ -\hat{c}_{t+1} + \hat{c}_t + \hat{a}_{t+1} + (\alpha - 1)\hat{k}_{t+1} \}$$

$$k_0^{\alpha}(\hat{a}_t + \alpha \hat{k}_t) = c_0 \hat{c}_t + k_0 \hat{k}_{t+1}$$

and

$$\hat{a}_{t+1} = \rho \hat{a}_t + \epsilon_{t+1}$$

where k_0 and c_0 are the points at which we approximate the budget constraint (these points are the steady state values of each variable; recall $a_0 = 1$ so we can drop it from the budget constraint).

The first step in using undetermined coefficients is to write down our assumed recursive law of motion for each variable. To do so we need to identify the state variables in the model, in

this case a_t and k_t. If a linear solution to our model exists, it will then have the form

$$\hat{c}_t = \gamma_{ca}\hat{a}_t + \gamma_{ck}\hat{k}_t$$

$$\hat{k}_{t+1} = \gamma_{ka}\hat{a}_t + \gamma_{kk}\hat{k}_t$$

$$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

Next, we plug these assumed laws of motion into our linearized first order conditions. Thus for the budget constraint

$$k_0^\alpha\hat{a}_t + \alpha k_0^\alpha\hat{k}_t = c_0(\gamma_{ca}\hat{a}_t + \gamma_{ck}\hat{k}_t) + k_0(\gamma_{ka}\hat{a}_t + \gamma_{kk}\hat{k}_t)$$

Since state variables can take any value, it must be the case that coefficients multiplying each state variable add up. Thus for $\hat{a}_t$

$$k_0^\alpha = c_0\gamma_{ca} + k_0\gamma_{ka}$$

and for $\hat{k}_t$

$$\alpha k_0^\alpha = c_0\gamma_{ck} + k_0\gamma_{kk}$$

or

(5.4)
$$\gamma_{ck} = \frac{1}{c_0}(\alpha k_0^\alpha - k_0\gamma_{kk})$$

The Euler equation is somewhat messier. We have

$$0 = E_t\left\{-(\gamma_{ca}\hat{a}_{t+1} + \gamma_{ck}\hat{k}_{t+1}) + \gamma_{ca}\hat{a}_t + \gamma_{ck}\hat{k}_t + \hat{a}_{t+1}\right.$$

$$\left. + (\alpha-1)(\gamma_{ka}\hat{a}_t + \gamma_{kk}\hat{k}_t)\right\}$$

As the whole equation is in expectations we can replace $\hat{a}_{t+1}$ with $\rho\hat{a}_t$ because $E_t(\hat{a}_{t+1}) = E_t(\rho\hat{a}_t + \epsilon_{t+1}) = \rho\hat{a}_t$ since the shock ϵ_{t+1} is zero in expectations. However, we don't yet have

the equation strictly in terms of state variables; to do so we need
to plug in our assumed laws of motion a second time to get rid
of the term $\hat{k}_{t+1}$. Thus the Euler becomes

$$\gamma_{ca}\rho\hat{a}_t+\gamma_{ck}(\gamma_{ka}\hat{a}_t+\gamma_{kk}\hat{k}_t) = \gamma_{ca}\hat{a}_t+\gamma_{ck}\hat{k}_t+\hat{a}_{t+1}+(\alpha-1)(\gamma_{ka}\hat{a}_t+\gamma_{kk}\hat{k}_t)$$

Equating the terms that multiply $\hat{a}_t$ we have

(5.5) $$\gamma_{ca}\rho + \gamma_{ck}\gamma_{ka} = \gamma_{ca} + \rho + (\alpha - 1)\gamma_{ka}$$

Equating the terms that multiply $\hat{k}_t$ we have

(5.6) $$\gamma_{ck}\gamma_{kk} = \gamma_{ck} + (\alpha - 1)\gamma_{kk}$$

Plugging equation 5.4 into equation 5.6 we then have

$$\left(\frac{1}{c_0}(\alpha k_0^\alpha - k_0\gamma_{kk})\right)\gamma_{kk} = \frac{1}{c_0}(\alpha k_0^\alpha - k_0\gamma_{kk}) + (\alpha - 1)\gamma_{kk}$$

which is a quadratic equation

$$\frac{k_0}{c_0}\gamma_{kk}^2 + \left((\alpha - 1) - \frac{k_0}{c_0} - \frac{\alpha k_0^\alpha}{c_0}\right)\gamma_{kk} + \frac{\alpha k_0^\alpha}{c_0} = 0$$

Once we have γ_{kk} we can plug back in to find the other coeffi-
cients of interest. For example, letting $\alpha = 0.35$ and $\beta = 0.99$
we have

$$k_0 = 0.20$$

$$c_0 = 0.37$$

From our above quadratic equation we then have two solutions:

$$\gamma_{kk} = 0.35$$

or

$$\gamma_{kk} = 2.89$$

To eliminate one of these solutions we use the fact that we are looking for the saddle path. This means we enforce stability on the model (again, often called the transversality of "no ponzi" condition). In this case stability requires that $|\gamma_{kk}| < 1$ (if γ_{kk} was greater than one in magnitude a shock would cause capital to grow in magnitude without bound), leaving us with the solution $\gamma_{kk} = 0.35$. With this solution we additionally have $\gamma_{ca} = 1$ and $\gamma_{ka} = 1$. Thus the solution to this macro model is

$$
\begin{aligned}
\hat{a}_{t+1} &= 0.7\hat{a}_t + 0\hat{k}_t + \epsilon_{t+1} \\
\hat{k}_{t+1} &= \hat{a}_t + 0.35\hat{k}_t \\
\hat{c}_t &= \hat{a}_t + 0.35\hat{k}_t
\end{aligned}
$$

(5.7)

Figure 5.4 illustrates this result by plotting a one time shock of $\epsilon_0 = 1$ and tracing the response of capital and consumption over time; these responses are called **impulse response functions** and are usually the solution we seek to a macro model.

5.4 Systems of Difference Equations in Macroeconomics Using the "Math Class" Approach

In this section I propose a technique that follows the "math class" approach to solving systems of linear difference equations as closely as possible which will, I hope, illustrate the link between section 5.1 and the solution to macro models.

Deriving the Procedure

The first step in using matrix methods to solve a macro model is to put the model into the form

$$
E_t Y_{t+1} = A Y_t + B
$$

or, assuming $B = [0]$, which is usually the case since our lin-earized system of equations is in log deviations from the steady state,

$$(5.8) \qquad\qquad E_t Y_{t+1} = A Y_t$$

Instead of considering the entire span of t for the model this derivation is based on a two period question: Given our initial point, the state variables (which include any exogenous shock processes) in the current period, can we determine the value of the state variables next period and the value of other variables (choice and co-state) in the current period in a way which will not violate our terminal point, the transversally condition? That is, we seek an answer of the form

$$(5.9) \qquad\qquad Y_{s,t+1} = W Y_{s,t} + \epsilon_{t+1}$$

$$(5.10) \qquad\qquad Y_{c,t} = V Y_{s,t}$$

where ϵ_{t+1} is any exogenous shock to the state variables, $Y_{s,t}$ is the submatrix of Y_t that contains only state variables, and $Y_{c,t}$ is the submatrix of Y_t that contains other variables (those for which we do not have initial conditions). If so, we have solved the whole problem as we can iterate this solution as far forward as we like. Suppose that the first m variables of the DSGE model represented in equation 5.8 are state variables and that the remaining $n - m$ variables are jump variables where A is an $n \times n$ matrix. Then the general "math class" solution to this system of difference equation is

$$(5.11) \quad E_t Y_{t+l} = k_1 C_1 \lambda_1^l + k_2 C_2 \lambda_2^l + ... + k_i C_i \lambda_i^l + ... + k_n C_n \lambda_n^l$$

If we were only interested in the values of state variables in future periods, we could re-write this solution as

$$(5.12)$$
$$E_t Y_{s,t+l} = k_1 C_{s,1} \lambda_1^l + k_2 C_{s,2} \lambda_2^l + ... + k_i C_{s,i} \lambda_i^l + ... + k_n C_{s,n} \lambda_n^l$$

where $C_{s,i}$ is the vector of the first m elements of the eigenvector C_i.

A unique solution in the form of equations 5.9 and 5.10 to this model requires that A have m stable eigenvalues (eigenvalues that are less than one in absolute value) and $n - m$ unstable eigenvalues (eigenvalues that are greater than one in absolute value). The saddle path, the points along which the model does not diverge, is given by setting the constants k in equation 5.11 multiplying the unstable eigenvalues to zero. This rules out explosive solutions enforcing the terminal (no Ponzi) condition. Then in the current period, $l = 0$, the unique (saddle path) solution is given by

$$(5.13) \qquad Y_t = k_1 C_1 + k_2 C_2 + ... + k_i C_i + ... + k_m C_m$$

where I have ordered $k_i C_i$ in ascending order according to the associated eigenvalues λ_i. We can write equation 5.13 in matrix notation as

$$
\underbrace{\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ \vdots \\ y_{m,t} \\ \vdots \\ y_{n,t} \end{bmatrix}}_{n \times 1}
=
\underbrace{\begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n,1} & c_{n,2} & \cdots & c_{n,m} \end{bmatrix}}_{n \times m}
\underbrace{\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix}}_{m \times 1}
$$

or

$$(5.14) \qquad \begin{bmatrix} Y_{s,t} \\ Y_{c,t} \end{bmatrix} = \begin{bmatrix} C_s \\ C_c \end{bmatrix} K$$

where $Y_{s,t}$ is the submatrix of state variables, $Y_{c,t}$ is the sub-matrix of control or co-state variables (those for which we do not have initial conditions), C_s is the top $m \times m$ submatrix of

C, and C_c is the bottom $(m - n) \times m$ submatrix of C. Note that from this partitioning

(5.15) $$Y_{s,t} = C_s K$$

The conditions for K to be unique, that the matrix A have m stable eigenvalues and $n - m$ unstable eigenvalues, are evident from equation 5.15. Denoting the actual number of stable eigenvalues as j, the unique solution for K, $K = C_s^{-1} Y_{s,t}$, requires C_s is invertible in which case the number of stable eigenvalues j must equal the number of state variables m. Assuming that this is the case, in period $t+1$ the solution for next period state variables will be

(5.16)
$$Y_{s,t+1} = k_1 C_{s,1} \lambda_1 + k_2 C_{s,2} \lambda_2 + ... + k_i C_{s,i} \lambda_i + ... + k_m C_{s,m} \lambda_m + \epsilon_{t+1}$$

which we can write in matrix notation as

$$\underbrace{\begin{bmatrix} y_{1,t+1} \\ y_{2,t+1} \\ \vdots \\ y_{m,t+1} \end{bmatrix}}_{m \times 1} = \underbrace{\begin{bmatrix} c_{1,1} & c_{1,2} & \cdots & c_{1,m} \\ c_{2,1} & c_{2,2} & \cdots & c_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m,1} & c_{m,2} & \cdots & c_{m,m} \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_m \end{bmatrix}}_{m \times m} \underbrace{\begin{bmatrix} k_1 \\ k_2 \\ \vdots \\ k_m \end{bmatrix}}_{m \times 1}$$

$$+ \underbrace{\begin{bmatrix} \epsilon_{1,t+1} \\ \epsilon_{2,t+1} \\ \vdots \\ \epsilon_{m,t+1} \end{bmatrix}}_{m \times 1}$$

or in more compact notation

(5.17) $$Y_{s,t+1} = C_s \Lambda K + \epsilon_{t+1}$$

The innovation in this approach is to note that equation 5.17 can be written as

$$Y_{t+1} = C_s \Lambda I K + \epsilon_{t+1}$$

or

$$Y_{s,t+1} = C_s \Lambda C_s^{-1} C_s K + \epsilon_{t+1}$$

thus, since $Y_{s,t} = C_s K$

$$Y_{s,t+1} = C_s \Lambda C_s^{-1} Y_{s,t} + \epsilon_{t+1}$$

Therefore

(5.18) $$Y_{s,t+1} = W Y_{s,t} + \epsilon_{t+1}$$

where

$$W = C_s \Lambda C_s^{-1}$$

and we have next period state variables as a function of current state variables and the shock ϵ. Variables for which we do not have initial conditions (choice and co-state) are determined by the saddle path of the system. We can find the current period choice and co-state variables as a function of the state variables using equation 5.14 from which we have equation 5.15,

$$Y_{s,t} = C_s K$$

and

$$Y_{c,t} = C_c K$$

thus

$$C_s^{-1} Y_{s,t} = K$$

and

$$Y_{c,t} = C_c C_s^{-1} Y_{s,t}$$

Given the above solutions we can write the complete solution to our DSGE model as

(5.19) $$Y_{s,t+1} = W Y_{s,t} + \epsilon_{t+1}$$

and [3]

(5.20) $$Y_{c,t} = C_c C_s^{-1} Y_{s,t}$$

Summing up the procedure

Linearised DSGE models typically give rise to a two point boundary value problem in which our first point is a set of initial conditions for state variables and our second is a transversality or stability condition. Dividing Y_t into m state variables (those for which we have initial conditions, our first point) and $n - m$ control or co-state variables (those for which we do not have initial conditions) such that

$$Y_t = \begin{bmatrix} Y_s \\ Y_c \end{bmatrix}$$

we can write our linearised macroeconomic model as

$$E_t Y_{t+1} = A Y_t$$

Solving for the recursive law of motion of this system using the "math class" approach is fairly simple. Denote the stable eigenvalues of A as Λ and the stable eigenvectors of A (those associated with Λ) as $C = \begin{bmatrix} C_s \\ C_c \end{bmatrix}$ where C_s is the upper (assuming the number of stable eigenvalues j equals the number of state variables m) $m \times m$ square submatirx of C. Then the solution to our macroeconomic model is

$$Y_{s,t+1} = W Y_{s,t} + \epsilon_{s,t+1}$$

---

[3]If equation 5.9 was not homogeneous, but rather $Y_{t+1} = AY_t + B + \epsilon_{t+1}$, then the solution to our DSGE model would include the particular solution to this new system of difference equations, call it Y_p. In that case our new solution would be $Y_{s,t+1} = W_s Y_{s,t} + (Y_{s,p} - W_s Y_{s,p}) + \epsilon_{s,t+1}$ and $Y_{c,t} = VY_{s,t} - VY_{s,p} + Y_{c,p}$ where $Y_p = (I - A)^{-1}B$, $Y_{s,p}$ includes the first m elements of Y_p, and $Y_{c,p}$ the last $m - n$ elements of Y_p.

and

$$Y_{c,t} = V Y_{s,t}$$

where

$$W = C_s \Lambda C_s^{-1}$$

$$V = C_c C_s^{-1}$$

Examples

A Simple Example

Suppose we are interested in solving the model of section 5.3 in which consumers maximize

$$\mathcal{U} = E_t \sum_{t=0}^{\infty} \beta^t \ln(c_t)$$

subject to the budget constraint

$$a_t k_t^\alpha = c_t + k_{t+1}$$

Our long linearised first order conditions are again

$$0 = E_t\{-\hat{c}_{t+1} + \hat{c}_t + \hat{a}_{t+1} + (\alpha - 1)\hat{k}_{t+1}\}$$

$$k_0^\alpha(\hat{a}_t + \alpha\hat{k}_t) = c_0\hat{c}_t + k_0\hat{k}_{t+1}$$

and

$$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

To use matrix methods we need to put the model into matrix form. Rearranging the above equations we can write our model as

$$\hat{a}_{t+1} = \rho\hat{a}_t$$

$$\hat{k}_{t+1} = k_0^{\alpha-1}\hat{a}_t + \alpha k_0^{\alpha-1}\hat{k}_t - \frac{c_0}{k_0}\hat{c}_t$$

$$\hat{c}_{t+1} = \left(\rho + (\alpha - 1)k_0^{\alpha-1}\right)\hat{a}_t + \left((\alpha - 1)\alpha k_0^{\alpha-1}\right)\hat{k}_t + \left(\frac{1-\alpha}{k_0}c_0 + 1\right)\hat{c}_t$$

Note that above I have written the expressions for variables which are state variables in period t, that is, $\hat{a}_t$ and $\hat{k}_t$, first. We can put these equations in matrix form as

$$E_t \begin{bmatrix} \hat{a}_{t+1} \\ \hat{k}_{t+1} \\ \hat{c}_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ k_0^{\alpha-1} & \alpha k_0^{\alpha-1} & -\frac{c_0}{k_0} \\ \rho + (\alpha-1)k_0^{\alpha-1} & (\alpha-1)\alpha k_0^{\alpha-1} & \frac{1-\alpha}{k_0}c_0 + 1 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{c}_t \end{bmatrix}$$

or

$$E_t Y_{t+1} = A Y_t$$

where

$$A = \begin{bmatrix} \rho & 0 & 0 \\ k_0^{\alpha-1} & \alpha k_0^{\alpha-1} & -\frac{c_0}{k_0} \\ \rho + (\alpha-1)k_0^{\alpha-1} & (\alpha-1)\alpha k_0^{\alpha-1} & \frac{1-\alpha}{k_0}c_0 + 1 \end{bmatrix}$$

and

$$Y_t = \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{c}_t \end{bmatrix}$$

Plugging in[4] the same values as in section 5.3 ($\beta = 0.99$, $\alpha = 0.35$, and $\rho = 0.7$) we have

$$A = \begin{bmatrix} 0.70 & 0 & 0 \\ 2.89 & 1.01 & -1.89 \\ -1.18 & -0.66 & 2.23 \end{bmatrix}$$

with eigenvalues

$$\begin{bmatrix} 0.7000 & 0 & 0 \\ 0 & 0.3500 & 0 \\ 0 & 0 & 2.8860 \end{bmatrix}$$

---

[4]Note that if we need a closed form solution the Matlab command **eig(A)** works with the symbolic math toolbox to find the eigenvalues and eigenvectors of our system in terms of the original variables. Very handy.
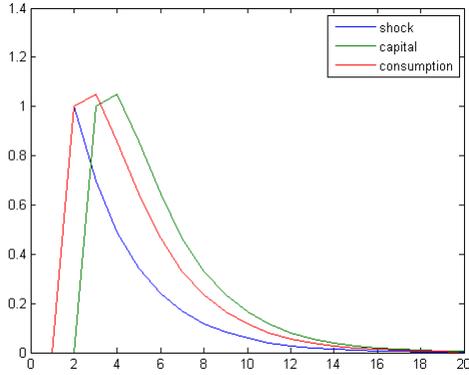
Figure 5.4: Impulse responses for a simple model

and eigenvectors

$$\begin{bmatrix} 0.2756 & 0 & 0 \\ 0.7875 & 0.9439 & -0.7090 \\ 0.5512 & 0.3304 & 0.7052 \end{bmatrix}$$

The matrices we need for a solution are

$$\Lambda = \begin{bmatrix} 0.7000 & 0 \\ 0 & 0.3500 \end{bmatrix}$$

$$C_s = \begin{bmatrix} 0.2756 & 0 \\ 0.7875 & 0.9439 \end{bmatrix}$$

and

$$C_c = \begin{bmatrix} 0.5512 & 0.3304 \end{bmatrix}$$

thus our solution is

$$W = C_s \Lambda C_s^{-1} = \begin{bmatrix} 0.7000 & 0 \\ 1.0000 & 0.3500 \end{bmatrix}$$

and

$$V = C_c C_s^{-1} = \begin{bmatrix} 1.0000 & 0.3500 \end{bmatrix}$$

that is,

$$\begin{bmatrix} \hat{a}_{t+1} \\ \hat{k}_{t+1} \end{bmatrix} = \begin{bmatrix} 0.7000 & 0 \\ 1.0000 & 0.3500 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

and

$$\hat{c}_t = \begin{bmatrix} 1.0000 & 0.3500 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \end{bmatrix}$$

Figure 5.4 again plots the impulse response functions for this model around the steady state.

A More Involved Example

As a more challenging example suppose we were interested in modelling an economy in which consumers maximize lifetime utility defined as

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \left(\frac{c_{t+s}^{1-\sigma}}{1-\sigma} + \xi \frac{l_{t+s}^{1-\gamma}}{1-\gamma} \right)$$

subject to

$$1 = l_t + n_t$$

$$a_t k_t^\alpha n_t^{1-\alpha} = c_t + i_t + \frac{\phi}{2} i_t^2$$

$$i_t = k_{t+1} - (1-\delta)k_t$$

$$\hat{a}_{t+1} = \rho \hat{a}_t + \epsilon_{t+1}$$

where c_t is consumption, l_t is leisure, n_t is labor, k_t is capital, and i_t is investment. In this problem we must contend with both

the agent's labor versus leisure choice as well as a cost to capital adjustment. Writing our constrained maximization problem as

$$
\begin{aligned}
\mathcal{L} \;=\;& E_t \sum_{s=0}^{t} \beta^s \Bigg[\left(\frac{c_{t+s}^{1-\sigma}}{1-\sigma} + \xi \frac{(1-n_{t+s})^{1-\gamma}}{1-\gamma} \right) \\
+\;& \lambda_{t+s} \left(a_{t+s} k_{t+s}^{\alpha} n_{t+s}^{1-\alpha} - c_{t+s} - i_{t+s} - \tfrac{\phi}{2} i_{t+s}^2 \right) \\
+\;& \mu_{t+s} \left(i_{t+s} + (1-\delta) k_{t+s} - k_{t+s+1} \right) \Bigg]
\end{aligned}
$$

the first order conditions are

$$
c_t^{-\sigma} = \lambda_t
$$

$$
\xi(1-n_t)^{-\gamma} = \lambda_t(1-\alpha)a_t k_t^{\alpha} n_t^{-\alpha}
$$

$$
\mu_t = \beta E_t\big(\alpha\lambda_{t+1}a_{t+1}k_{t+1}^{\alpha-1}n_{t+1}^{1-\alpha} + \mu_{t+1}(1-\delta)\big)
$$

$$
\lambda_t(1+\phi i_t) = \mu_t
$$

and the budget constraints. Our linearised system of difference equations thus includes the contemporaneous equations (equations which only have variables in period t)

$$
\hat{\lambda}_t = -\sigma\hat{c}_t
$$

$$
\hat{\lambda}_t + \hat{a}_t + \alpha\hat{k}_t = \left(\frac{\gamma n_0}{1 - n_0} + \alpha \right) \hat{n}_t
$$

$$
\hat{\lambda}_t \lambda_0(1+\phi i_0) = \mu_0\hat{\mu}_t - \phi\lambda_0 i_0\hat{i}_t
$$

$$
\kappa_1\hat{a}_t + \alpha\kappa_1\hat{k}_t = c_0\hat{c}_t + i_0(1+\phi i_0)\hat{i}_t - (1-\alpha)\kappa_1\hat{n}_t
$$

where $\kappa_1 = a_0 k_0^{\alpha} n_0^{1-\alpha}$. The dynamic equations (equations with variables in period t and $t+1$) in our linearised system are

$$
\mu_0\hat{\mu}_t = E_t\left(\kappa_2(\hat{\lambda}_{t+1}+\hat{a}_{t+1}+(\alpha-1)\hat{k}_{t+1}+(1-\alpha)\hat{n}_{t+1})+\beta\mu_1(1-\delta)\hat{\mu}_{t+1} \right)
$$

$$
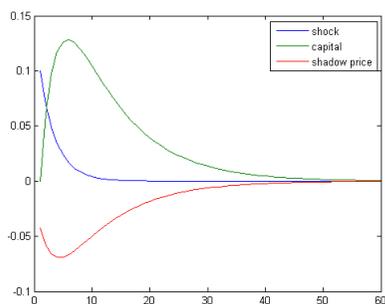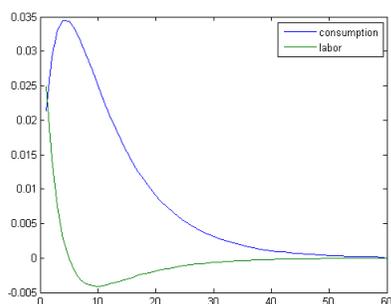k_1\hat{k}_{t+1} = (1-\delta)k_0\hat{k}_t + i_0\hat{i}_t
$$

Figure 5.5: Y variable IRFs Around the Steady State

Figure 5.6: Consumption and Investment IRFs Around the Steady State

$$E_t \hat{a}_{t+1} = \rho \hat{a}_t$$

where $\kappa_2 = \alpha \beta \lambda_1 a_1 k^{\alpha-1} n_1^{1-\alpha}$

For the parameters $\rho = 0.7$, $\beta = .99$, $\alpha = .35$, $\xi = 1$, $\sigma = 2$, $\gamma = 1.5$, $\phi = 0.2$, and $\delta = 0.2$ the steady state values of our relevant variables are (found numerically in Matlab) $k_0 = 1.16$, $c_0 = 0.49$, $n_0 = 0.564$, $\lambda_0 = 4.17$, $i_0 = 0.23$, and $\mu_0 = 4.37$. After a little bit of arduous algebra using the linearized first

order conditions we can write our model as

$$
E_t \underbrace{\begin{bmatrix} a_{t+1} \\ k_{t+1} \\ \lambda_{t+1} \end{bmatrix}}_{Y_{t+1}} = \underbrace{\begin{bmatrix} 0.70 & 0 & 0 \\ 0.77 & 1.07 & 0.37 \\ -0.06 & 0.14 & 1.21 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} a_t \\ k_t \\ \lambda_t \end{bmatrix}}_{Y_t}
$$

Note that I have solved the model here in terms of $\hat{\lambda}_t$ instead of $\hat{c}_t$; either one is fine since their relationship is given by $\hat{\lambda}_t = -\sigma \hat{c}_t$. The eigenvalues of A are $\lambda_1 = 0.70$, $\lambda_2 = 0.90$, and $\lambda_3 = 1.38$ with the associated eigenvectors

$$
C = \begin{bmatrix} 0.29 & 0 & 0 \\ -0.91 & 0.91 & 0.78 \\ 0.29 & -0.42 & 0.64 \end{bmatrix}
$$

so that the matrices we need for a solution are

$$
\Lambda = \begin{bmatrix} 0.70 & 0 \\ 0 & 0.90 \end{bmatrix}
$$

and

$$
C_s = \begin{bmatrix} 0.29 & 0 \\ -0.91 & 0.91 \end{bmatrix}, \qquad C_c = \begin{bmatrix} 0.29 & -0.42 \end{bmatrix}
$$

Our solution matrices are

$$
W = C_s \Lambda C_s^{-1} = \begin{bmatrix} 0.70 & 0 \\ 0.61 & 0.90 \end{bmatrix}, \qquad V = C_s C_c = \begin{bmatrix} -0.43 & -0.46 \end{bmatrix}
$$

thus the solution to our model is

$$
\begin{bmatrix} a_{t+1} \\ k_{t+1} \end{bmatrix} = \begin{bmatrix} 0.70 & 0 \\ 0.61 & 0.90 \end{bmatrix} \begin{bmatrix} a_t \\ k_t \end{bmatrix} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}, \qquad \lambda_t = \begin{bmatrix} -0.43 & -0.46 \end{bmatrix} \begin{bmatrix} a_t \\ k_t \end{bmatrix}
$$

Figure 5.5 plots the impulse responses for the Y variables in this model, figure 5.6 plots the impulse responses for consumption and labor, and figure 5.7 plots log deviations in investment and Tobin's Q, defined here as $Q_t = \frac{\mu_t}{\lambda_t}$ (and thus $\hat{Q}_t = \hat{\mu}_t - \hat{\lambda}_t$), all around the steady state.
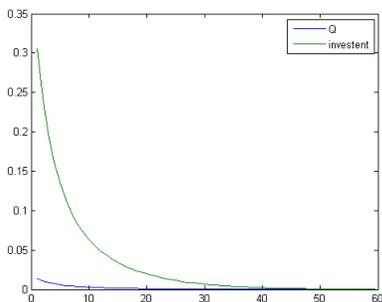
Figure 5.7: IRF for $\hat{Q}_t$ and $\hat{i}_t$ Around the Steady State

5.5 Systems of Difference Equations in Macroeconomics using Blanchard and Kahn (1980)

Blanchard and Kahn (1980) (from here on BK (1980)) describe our linearized macro model as

(5.21)
$$\begin{bmatrix} X_{t+1} \\ {}_tP_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ P_{t+1} \end{bmatrix} + \gamma Z_t$$

where, in BK (1980) notation, X_t is a vector of predetermined state variables (my $Y_{s,t}$), P_t is a vector of jump variables (my $Y_{c,t}$), and Z_t is a vector of exogenous forcing variables (the shock process, which I include in the vector $Y_{s,t}$). The notation ${}_tP_{t+1}$ which BK (1980) use means E_tP_{t+1}. Note that equation 5.21 could also be written as

$$E_t \begin{bmatrix} X_{t+1} \\ P_{t+1} \end{bmatrix} = A \begin{bmatrix} X_t \\ P_{t+1} \end{bmatrix} + \gamma Z_t$$

since taking the expectation of X_{t+1}, the predetermined variables, is X_{t+1} itself. The difference between my formulation of

the problem and BK (1980) is that I consider the case where all forcing variables (the Z_t vector, which is a little confusing since this is my ϵ_{t+1} vector, that is, I explicitly include exogenous orthogonal i.i.d. shocks) are zero in expectation. That is, I include the shock process (an AR(1) process which fits nicely into the linear system of equations) in the X vectors and the A matrix to eliminate the matrix γ from equation 5.21; BK (1980) include the shock process in the vector Z so that it is typically not zero in expectation. Comparing the two approaches requires me to rewrite BK (1980) to eliminate γ. To illustrate this difference, consider the example from section 5.4. BK (1980) might write this problem as

(5.22)

$$
E_t \begin{bmatrix} \hat{k}_{t+1} \\ \hat{\lambda}_{t+1} \end{bmatrix} = \begin{bmatrix} \frac{\alpha a_0 k_0^\alpha}{k_1} & \frac{c_0}{k_1} \\ \frac{1-\alpha}{k_1}\alpha a_0 k_0^\alpha & \frac{1-\alpha}{k_1}c_0 + 1 \end{bmatrix} \begin{bmatrix} \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix} + \begin{bmatrix} \frac{a_0 k_0^\alpha}{k_1} \\ -\rho + \frac{1-\alpha}{k_1}a_0 k_0^\alpha \end{bmatrix} \hat{a}_t
$$

where re-writing it as

(5.23)

$$
E_t \begin{bmatrix} \hat{a}_{t+1} \\ \hat{k}_{t+1} \\ \hat{\lambda}_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 \\ \frac{a_0 k_0^\alpha}{k_1} & \frac{\alpha a_0 k_0^\alpha}{k_1} & \frac{c_0}{k_1} \\ -\rho + \frac{1-\alpha}{k_1}a_0 k_0^\alpha & \frac{1-\alpha}{k_1}\alpha a_0 k_0^\alpha & \frac{1-\alpha}{k_1}c_0 + 1 \end{bmatrix} \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix}
$$

eliminates γ and ensures that Z is zero in expectation (since by taking expectations of both sides of equation 5.23 eliminates ϵ_{t+1}). Following BK (1980) we can rewrite our model in the form of equation 5.23 (that is, where γ is the identity matrix and Z_t zero in expectations) as

$$
E_t \begin{bmatrix} X_{t+1} \\ P_{t+1} \end{bmatrix} = C^{-1}\Lambda C \begin{bmatrix} X_t \\ P_{t+1} \end{bmatrix}
$$

where $C^{-1}\Lambda C$ is the eigenvector-eigenvalue decomposition of A. Using the substitution

$$
C \begin{bmatrix} X_t \\ P_t \end{bmatrix} = \begin{bmatrix} Y_t \\ Q_t \end{bmatrix}
$$

or more explicitly

(5.24)
$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} \begin{bmatrix} X_t \\ P_t \end{bmatrix} = \begin{bmatrix} Y_t \\ Q_t \end{bmatrix}$$

we have

(5.25)
$$E_t \begin{bmatrix} Y_{t+1} \\ Q_{t+1} \end{bmatrix} = \Lambda \begin{bmatrix} X_t \\ P_{t+1} \end{bmatrix}$$

Equation 5.25 is our model in "canonical form". Λ is a diagonal matrix,

$$\Lambda = \begin{bmatrix} J_1 & 0 \\ 0 & J_2 \end{bmatrix}$$

where we organize the eigenvalues in ascending order from element $1, 1$ to element n, n. The submatrix J_1 has m elements where m is the number of state variables and the submatrix J_2 has $n - m$ elements, the number of jump variables. As in the previous, a unique solution requires that the elements of J_1 are less than one in absolute value and the elements of J_2 greater than one in absolute value. Because Λ is a diagonal matrix we can rewrite the bottom elements of equation 5.25 as

$$E_t Q_{t+1} = J_2 Q_t$$

or

$$Q_t = J_2^{-1} E_t Q_{t+1}$$

Iterating this equation forward

$$Q_t = J_2^{-(T-t+1)} E_t Q_T$$

which, taking the limit as $T \to \infty$ and using our transversality condition (requiring stability) says that

$$Q_t = 0$$

Using equation 5.24 this says that

$$C_{21}X_t + C_{22}P_t = 0$$

or

$$P_t = -C_{22}^{-1}C_{21}X_t$$

which gives our jump variables as a function of our choice variables. The matrix $-C_{22}^{-1}C_{21}$ is the matrix V in my approach. Letting

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1} = \begin{bmatrix} C_s & C_{s2} \\ C_c & C_{c2} \end{bmatrix}$$

then $V = C_c C_s^{-1} = -C_{22}^{-1}C_{21}$ where the left hand side is my solution and the right hand side is the BK (1980) solution.

The solution for the next period state variables using the BK (1980) solution technique from equation 5.25 is

$$E_t Y_{t+1} = J_1 Y_t$$

or, using equation 5.24

$$C_{11}E_tX_{t+1} + C_{12}E_tP_{t+1} = J_1(C_{11}E_tX_t + C_{12}E_tP_t)$$

or

$$E_tX_{t+1} = C_{11}^{-1}J_1C_{11}X_t + C_{11}J_1C_{12}P_t - C_{11}^{-1}C_{12}E_tP_{t+1}$$

Substituting the solution $P_t = -C_{22}^{-1}C_{21}X_t$ we can rewrite this equation as

$$(I_m - C_{11}^{-1}C_{12}C_{22}^{-1}C_{21})E_tX_{t+1} = C_{11}^{-1}J_1(C_{11} - C_{12}C_{22}^{-1}C_{21})X_t$$

or

$$(C_{11} - C_{12}C_{22}^{-1}C_{21})E_tX_{t+1} = J_1(C_{11} - C_{12}C_{22}^{-1}C_{21})X_t$$

thus

(5.26) $E_t X_{t+1} = (C_{11} - C_{12}C_{22}^{-1}C_{21})^{-1}J_1(C_{11} - C_{12}C_{22}^{-1}C_{21})X_t$

which is the solution for next period state variables as a function of current period state variables. Again letting

$$\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}^{-1} = \begin{bmatrix} C_s & C_{s2} \\ C_c & C_{c2} \end{bmatrix} = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

and $\Lambda_{my} = J_1$ where the matrix Λ_{my} from my approach is the same matrix as J_1 in the BK (1980) approach then

$$\begin{aligned} W &= C_s \Lambda_{my} C_s^{-1} = (C_{11} - C_{12}C_{22}^{-1}C_{21})^{-1}J_1(C_{11} - C_{12}C_{22}^{-1}C_{21}) \\ &= B_{11}J_1 B_{11}^{-1} \end{aligned}$$

where the left hand side is my solution and the right hand side the BK (1980) solution (which are obviously the same). This comparison of course assumes that $\gamma = I_n$ and $E_t Z_t = 0$. If that is not the case than following the same approach as above we get the BK (1980) solution of

$$P_t = -C_{22}^{-1}C_{21}X_t - C_{22}^{-1}\sum_{i=0}^{\infty} J_2^{-i-1}(C_{21}\gamma_1 + C_{22}\gamma_2)E_t Z_{t+i}$$

and

$$\begin{aligned} E_t X_{t+1} &= B_{11}J_1 B_{11}^{-1}X_t + \gamma_1 Z_t \\ &- (B_{11}J_1 C_{12} + B_{12}J_2 C_{22})C_{22}^{-1}\sum_{i=0}^{\infty} J_2^{-i-1}(C_{21}\gamma_1 + C_{22}\gamma_2)E_t Z_{t+i} \end{aligned}$$

Further Reading

Solving the stochastic system of linear difference equations arising from a linearized macroeconomic model is not something that is particularly well covered in the literature. One of the

best references remains Blanchard and Kahn's original 1980 article. Harald Uhlig's (1998) "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily" also covers solving the system of equations. Dowling (2011) and Chaing and Wainwright (2005) both cover the simpler case of deterministic difference equations, and for the self-taught student Bronson and Costa (2006) is a particularly good introduction to differential and, to a lesser extent, difference equations. Finally, the documentation for Dynare, the Matlab based program written to solve a wide range of macroeconomic models, is an excellent resource. It is available online at www.dynare.org.

Chapter 6

Solving a Simple DSGE Model: A Model with Cost of Capital Adjustment

This chapter works through a simple DSGE model from start to finish with an emphasis on solving the stochastic system of difference equations arising from the model (repeating some of the results from chapter 5 for clarity) as this is often the most difficult part of a problem. The particular model we'll solve includes a cost of capital adjustment as in section 2.4. We'll begin with our representative agent's objective function

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{c_{t+s}^{1-\sigma} - 1}{1 - \sigma}$$

subject to

(6.1) $$a_t k_t^\alpha = c_t + i_t + \frac{\phi}{2}(i_t - \delta k_t)^2$$

(6.2) $$i_t = k_{t+1} - (1 - \delta)k_t$$

(6.3) $$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

Thus the cost of investing (that is, of adjusting capital) is $\frac{\phi}{2}(i_t - \delta k_t)^2$. Note, importantly, that this function is convex: the bigger the investment (once it exceeds the replacement cost of capital), the bigger the marginal cost of adjustment. Thus agents have an incentive to spread investments out over time.

Our goal in looking at this problem is to figure out how variables of our model economy (consumption, capital, investment, asset prices) respond to exogenous shocks, in this case an exogenous shock to productivity (a_t). That is, we seek a set of impulse response functions describing the dynamics of the system. One way we might describe our system (in log deviations as opposed to levels, denoted by $\hat{x}_t$) in matrix notation is, letting

$$Y_t = \begin{bmatrix} \hat{a}_t & \hat{k}_t & \hat{\lambda}_t \end{bmatrix}$$

and recalling $\hat{\lambda}_t = -\sigma\hat{c}_t$, an equation that will come out of our linearised first order conditions below,

$$E_t Y_{t+1} = AY_t$$

where A describes how our system transitions from one period to the next. While this may initially look good, there is a problem. Consumption is a jump variable and thus $\hat{\lambda}_t$ can change within the period. That is, its value is not predetermined. More generally, if

$$Y_t = \begin{bmatrix} Y_{s,t} \\ Y_{c,t} \end{bmatrix}$$

where $Y_{s,t}$ contains our state variables and $Y_{c,t}$ contains our control and other jump variables (λ_t is called a co-state variable) then what we need is a matrix W such that

$$E_t Y_{s,t+1} = WY_{s,t}$$

and a matrix V such that

$$Y_{c,t} = VY_{s,t}$$

That is, our goal, as described in chapter 5, is to find our next period state variables and our current period control or co-state variables as a function of our current period state variables.

Our first step towards this goal is to set up the Lagrangian for this problem as

$$
\begin{aligned}
\mathcal{L} &= E_t \sum_{s=0}^{t} \beta^s \left[\frac{c_{t+s}^{1-\sigma}}{1-\sigma} \right. \\
&+ \lambda_{t+s} \left(a_{t+s} k_{t+s}^{\alpha} - c_{t+s} - i_{t+s} - \frac{\phi}{2}(i_{t+s} - \delta k_{t+s})^2 \right) \\
&+ \mu_{t+s} \left(i_{t+s} + (1-\delta)k_{t+s} - k_{t+s+1} \right) \Big]
\end{aligned}
$$

thus the first order conditions are

(6.4)
$$c_t^{-\sigma} = \lambda_t$$

(6.5)
$$\mu_t = \beta E_t \left[\lambda_{t+1} \left(\alpha a_{t+1} k_{t+1}^{\alpha-1} + \phi(i_{t+1} - \delta k_{t+1})\delta \right) + \mu_{t+1}(1-\delta) \right]$$

(6.6)
$$\lambda_t (1 + \phi(i_t - \delta k_t)) = \mu_t$$

and the budget constraints. If we want asset prices to be part of our model we need a way to measure the difference in value (in terms of consumer utility) between investment and output. Looking at equation 6.6, whenever $\mu_t > \lambda_t$, investment must be more than that which simply replaces depreciated capital. Thus the ratio of μ_t to λ_t seems like a reasonable way to measure the value of investment: whenever this ratio is greater than one, our agent should invest (since the return to capital is high), and

whenever it is less than one, our agent should consume (i.e. allow capital to depreciate). Thus we can define our price of capital as $Q \equiv \frac{\mu_t}{\lambda_t}$; this measure of the value of investing is called **Tobin's Q**.

The next step in solving this problem is to find the steady state values of each variable. From equation 6.2 we have (where the subscript 0 denotes steady state values)

$$i_0 = \delta k_0$$

and thus from equation 6.6

$$\lambda_0 = \mu_0$$

Plugging this into equation 6.5 we have (recall $a_0 = 1$)

$$1 = \beta(\alpha k_0^{\alpha-1} + 1 - \delta)$$

thus

$$k_0 = \left(\frac{\alpha}{\frac{1}{\beta} - (1-\delta)} \right)^{\frac{1}{1-\alpha}}$$

Plugging this back into equation 6.1

$$c_0 = k_0^\alpha - \delta k_0$$

From here we can proceed in two ways. We could substitute out as many variables as possible, linearise those relationships, and than use the method of undetermined coefficients to solve for the linearised recursive law of motion that governs this system (that is, the impulse response functions we seek as the solution to this problem). Alternatively, we could linearise the system as it is and use matrix techniques to solve for the impulse response functions. This chapter uses the second approach.

To use a matrix based approach it is useful to sort our equations into static equations, those which deal only with period t,

and dynamic equations, those which have variables in period t and in period $t+1$. The linearised static equations are:
For equation 6.4

$$\hat{\lambda}_t = -\sigma \hat{c}_t$$

For equation 6.6, using the fact that $\lambda_0 = \mu_0$

$$(1 + \phi i_0 - \delta k_0)\hat{\lambda}_t - \delta k_0 \hat{k}_t = \hat{\mu}_t - \phi i_0 \hat{i}_t$$

Again using steady state relationships, this time that $i_0 = \delta k_0$, we have

$$\hat{\mu}_t = \hat{\lambda}_t + \phi \delta k_0 (\hat{i}_t - \hat{k}_t)$$

For equation 6.1

$$k_0^\alpha \hat{a}_t + (\alpha k_0^\alpha + \frac{\phi}{2}\delta i_0 k_0 - \phi \delta^2 k_0^2)\hat{k}_t = c_0 \hat{c}_t + i_0(1 + \phi i_0 - \frac{\phi}{2}\delta k_0)\hat{i}_t$$

Again using the steady state value for investment

$$k_0^\alpha \hat{a}_t + (\alpha k_0^\alpha - \frac{\phi}{2}i_0^2)\hat{k}_t = c_0 \hat{c}_t + i_0(1 + \frac{\phi}{2}i_0)\hat{i}_t$$

The linearised dynamic equations in our linearised system are:
For equation 6.5, using the fact that $\mu_0 = \lambda_0$ and $a_0 = 1$

$$
\begin{aligned}
\hat{\mu}_t \;\;=\;\; & E_t \Big[\beta \alpha k_0^{\alpha-1} \hat{a}_{t+1} + \beta(\alpha k_0^{\alpha-1}(\alpha - 1) - \phi \delta^2 k_0)\hat{k}_{t+1} \\
& + \;\; \beta(\alpha k_0^{\alpha-1} + \delta \phi i_0 - \phi \delta^2 k_0)\hat{\lambda}_{t+1} + \beta \delta \phi i_0 \hat{i}_{t+1} + \beta(1 - \delta)\hat{\mu}_{t+1} \Big]
\end{aligned}
$$

Using $i_0 = \delta k_0$ we have

$$
\begin{aligned}
\hat{\mu}_t \;\;=\;\; & E_t \Big[\beta \alpha k_0^{\alpha-1} \hat{a}_{t+1} + \beta(\alpha k_0^{\alpha-1}(\alpha - 1) - \phi \delta^2 k_0)\hat{k}_{t+1} \\
& + \;\; \beta(\alpha k_0^{\alpha-1})\hat{\lambda}_{t+1} + \beta \delta \phi i_0 \hat{i}_{t+1} + \beta(1 - \delta)\hat{\mu}_{t+1} \Big]
\end{aligned}
$$

For equation 6.2

$$k_0 \hat{k}_{t+1} = (1 - \delta) k_0 \hat{k}_t + i_0 \hat{i}_t$$

or

$$\hat{k}_{t+1} = \hat{k}_t + \delta(\hat{i}_t - \hat{k}_t)$$

And equation 6.3, which is already log linear

$$E_t \hat{a}_{t+1} = \rho \hat{a}_t$$

The linearised equations 6.6 and 6.3 are useful in looking at the relationship between $\hat{Q}_t = \hat{\mu}_t - \hat{\lambda}_t$, investment, and capital. The linearised equation 6.6 says that whenever $\hat{Q}_t$ is positive, then the log deviation of investment in the steady state is greater than the log deviation in capital. The linearised equation 6.2 lets us translate this into something more meaningful: whenever $\hat{Q}$ is positive $\hat{k}$ is increasing, and whenever $\hat{Q}$ is negative $\hat{k}$ is decreasing. In levels this means that when Q_t is greater than one investment exceeds the replacement cost of capital, and the reverse when Q_t is less than one. If capital is more valuable than output (which the agent uses to purchase consumption), our agent should increase his or her capital stock by investing more than δk_t.

Once the equations describing our system are linear we can put them into matrix form; the trick is to organize them in such a way that we can use standard math to solve the system. That is, we want to put our system in the form

(6.7)
$$E_t \begin{bmatrix} Y_{s,t+1} \\ Y_{c,t+1} \end{bmatrix} = A \begin{bmatrix} Y_{s,t} \\ Y_{c,t} \end{bmatrix}$$

where $Y_{s,t}$ is the sub-vector of Y_t that includes only state variables (variables for which we have initial conditions) and $Y_{c,t}$ is the sub-vector of Y_t that includes everything else. One approach to getting our system into this form is to start substituting out

as many variables as possible as we might have done had we
wanted to use undetermined coefficients. Alternatively, we could
use matrix algebra to eliminate unwanted variables. The trou-
ble with substituting things out (in this case we would have to
substitute both $\hat{\mu}$ and $\hat{i}$ out of the linearised equation 6.5 and $\hat{i}$
our of the linearised equation 6.6) is that our equations become
even bigger and messier, so I will use matrix algebra.

I begin by defining the matrices

$$Y_t = \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix}$$

$$\tilde{Y}_t = \begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}$$

and

$$X_t = \begin{bmatrix} \hat{c}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}$$

Here $\tilde{Y}_t$ is my augmented vector of dynamic variables which
I would like to slim down into Y_t. With these definitions I can
write my system in matrix form as

$$E_t \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ \beta\alpha k_0^{\alpha-1} & \tilde{h}_{2,2} & \beta\alpha k_0^{\alpha-1} & \beta(1-\delta) & \beta\delta\phi i_0 \\ 0 & k_0 & 0 & 0 & 0 \end{bmatrix}}_{\tilde{H}} \underbrace{\begin{bmatrix} \hat{a}_{t+1} \\ \hat{k}_{t+1} \\ \hat{\lambda}_{t+1} \\ \hat{\mu}_{t+1} \\ \hat{i}_{t+1} \end{bmatrix}}_{\tilde{Y}_{t+1}} =$$

$$
\underbrace{\begin{bmatrix} \rho & 0 & 0 \\ 0 & 0 & 0 \\ 0 & (1-\delta)k_0 & 0 \end{bmatrix}}_{N} \underbrace{\begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix}}_{Y_t} + \underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & i_0 \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} \hat{c}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}}_{X_t}
$$

where

$$
\tilde{h}_{2,2} = \beta(\alpha(\alpha - 1)k_0^{\alpha-1} - \phi\delta^2 k_0)
$$

Furthermore, we have

$$
\underbrace{\begin{bmatrix} 0 & 0 & 1 \\ 0 & -\delta\phi k_0 & 1 \\ k_0^{\alpha} & (\alpha k_0^{\alpha} - \frac{\phi}{2}i_0^2) & 0 \end{bmatrix}}_{M_y} \underbrace{\begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix}}_{Y_t} =
$$

$$
\underbrace{\begin{bmatrix} -\sigma & 0 & 0 \\ 0 & 1 & -\phi i_0 \\ c_0 & 0 & i_0(1 + \frac{\phi}{2}i_0) \end{bmatrix}}_{M_x} \underbrace{\begin{bmatrix} \hat{c}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}}_{X_t}
$$

And finally:

$$
\underbrace{\begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}}_{\tilde{Y}_t} =
$$

$$
\underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{D_y} \underbrace{\begin{bmatrix} \hat{a}_t \\ \hat{k}_t \\ \hat{\lambda}_t \end{bmatrix}}_{Y_t} +
$$

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{D_x} \underbrace{\begin{bmatrix} \hat{c}_t \\ \hat{\mu}_t \\ \hat{i}_t \end{bmatrix}}_{X_t}$$

Phew. Rewriting the above matrix equations we have

$$(6.8) \qquad E_t \tilde{H} \tilde{Y}_{t+1} = NY_t + QX_t$$

$$(6.9) \qquad M_y Y_t = M_x X_t$$

and

$$(6.10) \qquad \tilde{Y}_t = D_y Y_t + D_x X_t$$

Using these matrices, we can now manipulate our system into a standard stochastic system of difference equations. Noting that

$$(6.11) \qquad X_t = M_x^{-1} M_y Y_t$$

then we have

$$(6.12) \qquad \tilde{Y}_t = \left(D_y + D_x M_x^{-1} M_y \right) Y_t$$

We can next plug this equation into equation 6.8 yielding

$$(6.13) \qquad \underbrace{\tilde{H} \left(D_y + D_x M_x^{-1} M_y \right)}_{H} E_t Y_{t+1} = NY_t + QX_t$$

We can also use equation 6.11 in equation 6.13 so that we have

$$(6.14) \qquad H E_t Y_{t+1} = (N + Q M_x^{-1} M_y) Y_t$$

which finally gets us to

$$(6.15) \qquad E_t Y_{t+1} = \underbrace{H^{-1} (N + Q M_x^{-1} M_y)}_{A} Y_t$$

our standard system of difference equations in expectation. The
solution to the model is determined by the eigenvalues and eigen-
vectors of A. Specifically, the solution to this system of difference
equations will be

$$(6.16) \qquad E_t Y_{t+l} = k_1 C_1 \lambda_1^l + k_2 C_2 \lambda_2^l + k_3 C_3 \lambda_3^l$$

where λ_i are the eigenvalues of A, C_i are the eigenvectors of
A, k_i are constants to be determined by the initial values of
the state variables and our transversality or no-Ponzi condition,
and l measures the number of periods forward from the current
period. Some of the eigenvalues (two of them in this case - equal
the number of state variables) will be less than one in absolute
value while others (one of them in this case) will be greater. If
the no-Ponzi condition holds, then the constant multiplying the
eigenvalue that is greater than one in magnitude must be zero.
Suppose the eigenvalue with magnitude greater than one is λ_3.
Then the current period value of our model is

$$Y_t = k_1 C_1 + k_2 C_2$$

or in matrix form

$$\begin{bmatrix} Y_{s,t} \\ Y_{c,t} \end{bmatrix} = \begin{bmatrix} C_s \\ C_c \end{bmatrix} K$$

Note that this implies

$$(6.17) \qquad\qquad\qquad Y_{s,t} = C_s K$$

and

$$(6.18) \qquad\qquad\qquad Y_{c,t} = C_c K$$

From this we immediately have the solution for current period
choice or co-state variables as a function of the current state:
since

$$C_s^{-1} Y_{s,t} = K$$

from equation 6.18

$$Y_{c,t} = VY_{s,t}$$

where $V = C_c C_s^{-1}$. The one period ahead expected solution to our system of difference equations is

$$E_t Y_{t+1} = k_1 C_1 \lambda_1 + k_2 C_2 \lambda_2$$

or in matrix form

(6.19) $$E_t Y_{t+1} = C \Lambda K$$

If we are only interested in the next period state variables (since control variables can change within the period anyhow) we could write the first two rows of the matrices in equation 6.19 as

$$Y_{s,t+1} = C_s \Lambda K + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

The trick here is noticing we can write this equation as

$$Y_{s,t+1} = C_s \Lambda I K + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

or

$$Y_{s,t+1} = C_s \Lambda C_s^{-1} C_s K + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

thus, since $Y_{s,t} = C_s K$

$$Y_{s,t+1} = C_s \Lambda C_s^{-1} Y_{s,t} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$

Therefore the solution to our DSGE model will be

(6.20) $$Y_{s,t+1} = W Y_{s,t} + \begin{bmatrix} \epsilon_{t+1} \\ 0 \end{bmatrix}$$
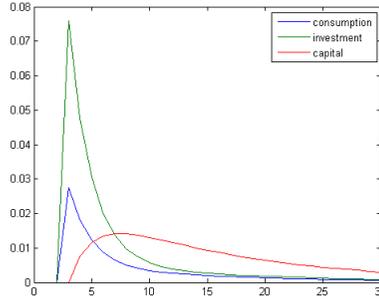
Figure 6.1: Impulse Responses for Capital, Investment, and Consumption

and

$$Y_{c,t} = VY_{s,t}$$

where $W = C_s \Lambda C_s^{-1}$ and $V = C_c C_s^{-1}$, which is the desired form: next period variables are a function of current state variables and the shock ϵ. Having finally solved the model, we can plug in some numbers to see what it looks like. In the following example, I use the parameters $\rho = .6$, $\beta = .99$, $\alpha = .35$, $\sigma = 2$, $\phi = 1$, $\delta = 0.1$. Both the cost of adjustment and rate of depreciation are high in this example to highlight their effect on the model. Plugging these parameters into our matrices, we have

$$A = \begin{bmatrix} 0.6000 & 0 & 0 \\ 0.2427 & 0.9621 & 0.0828 \\ 0.3969 & 0.0387 & 1.0196 \end{bmatrix}$$

with eigenvalues

$$\begin{bmatrix} 1.0544 & 0 & 0 \\ 0 & 0.9274 & 0 \\ 0 & 0 & 0.6000 \end{bmatrix}$$

and eigenvectors

$$
\begin{bmatrix}
0 & 0 & 0.7018 \\
0.6677 & 0.9222 & -0.3255 \\
0.7444 & -0.3867 & -0.6337
\end{bmatrix}
$$

Clearly, the first eigenvalue, 1.05, is greater than one. If this were not the case then we would not have a unique solution for our model. Instead of being saddle path stable it would be globally stable, and the jump variables as a function of the state variables would be indeterminate (if the model is globally stable consumption, for example, could be anything and our model would never violate the trasversality condition). As it is, we can get rid of the eigenvalue 1.05 and its associated eigenvectors to force our model onto the saddle path. Thus we have

$$
\Lambda = \begin{bmatrix}
0.9274 & 0 \\
0 & 0.6000
\end{bmatrix}
$$

$$
C_c = \begin{bmatrix} -0.3867 & -0.6337 \end{bmatrix}
$$

and

$$
C_s = \begin{bmatrix}
0 & 0.7018 \\
0.9222 & -0.3255
\end{bmatrix}
$$

and the solution to our model is

$$
C_s \Lambda C_s^{-1} = W = \begin{bmatrix}
0.6000 & 0 \\
0.1518 & 0.9274
\end{bmatrix}
$$

and

$$
C_c C_s^{-1} = V = \begin{bmatrix} -1.0975 & -0.4194 \end{bmatrix}
$$

What this says is that, for example, the percent deviation from the steady state of next period capital is 0.15 times the current deviation in productivity plus 0.93 times the current deviation in capital. The coefficient 0.93 is of particular interest -
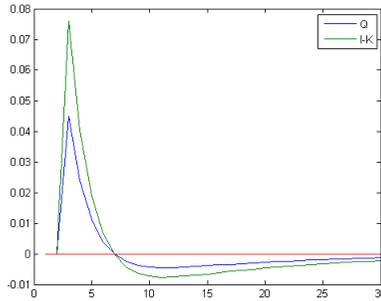
Figure 6.2: Investment Behaviour and Tobin's Q

this tells us that changes in capital are persistent (if this were 1 capital would never return to its steady state level). Figure 6.1 plots the impulse response functions for several variables of interest. Figure 6.2 looks at Tobin's Q (our asset price) and investment behaviour. When $\hat{Q}_t$ is less than zero, that is, Q_t is less than one, then the deviation in investment from the steady state is less than the deviation in capital. What this means is that when Q_t falls our agent allows his capital stock to depreciate. Specifically, if $Q_t < 1$ then capital is less valuable than consumption thus our agent consumes the output he would have otherwise spent on maintaining his capital stock at the same level. Note that if we let $\phi = 0$, that is, there is no cost to capital adjustment, then $Q = 1$ in every period.

Further Reading

Examples of simple macro models worked through from start to finish in a way that makes intermediate steps clear to the reader tend to be hard to find. Again, Harald Uhlig's (1998) "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily" is

a good reference (and is how I learned to work through macro models by hand). Once one has mastered the basics presented here it is good practice to try and replicate results in classic economics papers (more recent papers tend to be harder to follow or use more sophisticated techniques) such as Backus, Kehoe, and Kydland (1992) or King and Rebelo (1999). Another great way to practice working through macro models is to familiarize yourself with Dynare (which is very well documented) and compare your own results to those computed by the program.

Chapter 7

Growth

Our analysis thus far has missed an important, indeed central, feature of most economies: they grow over time. There are at least two major reasons for the bias towards short run deviations from steady states — the impulse response functions of chapters 5 and 6 — in most macroeconomics courses. First, long run growth, at least in the United States, tends to be very stable. It is short run deviations from this path, booms and busts, that capture the attention of both the economics profession and the general public. Second, the math upon which the theoretical models of the previous chapters are built require a stationary solution. That is, a unique solution to the model is one in which all variables revert to their steady state values in the absence of shocks, values which are constant over time. Yet, as the second model in this chapter shows, growth can have interesting implications for the response of capital and consumption to exogenous shocks. Perhaps more importantly, theoretical models of economic growth can inform how we should think about the economic indicators we observe in practice: industrial production, asset prices, or consumption for example. A coherent theoretical model of how economic variables grow over time provides a

solid foundation for how we might construct empirical models to make sense of observed economic data.

This chapter begins with a brief look at the Solow growth model, a point at which many macroeconomics courses begin. It then moves on to a simple example of a more satisfying growth model in the style of what we have already seen, one in which dynamics are driven by rational agents with a quantifiable objective function. I end with a brief comment on the implications for empirical macroeconomics and, as this is strictly a macro theory text, a few notes on where to find introductory macroeconometrics material.

7.1   The Solow Growth Model

The Solow growth model is not micro founded. Instead of modeling optimizing consumers, we simply suppose that the economy saves (and thus invests since in this closed economy the only thing to do with savings is invest it in capital) a fixed fraction of its income. Thus we have

$$(7.1) \qquad\qquad\qquad I_t = sY_t$$

where I_t is investment, s is some exogenous saving rate, and Y_t is production. The other two equations governing this model are the production function,

$$(7.2) \qquad\qquad\qquad Y_t = K_t^\alpha \left(A_t L_t\right)^{1-\alpha}$$

where K_t is capital, A_t productivity, and L_t labor, and the budget constraint

$$(7.3) \qquad\qquad\qquad \dot{K_t} = I_t + \delta K_t$$

where $\dot{K_t}$ is change in capital with respect to time (the model is typically written in continuous time) and δ the rate at which

capital depreciates. We also assume some exogenous rate of productivity growth,

$$(7.4) \qquad\qquad a = \frac{\dot{A}_t}{A_t}$$

and population growth

$$(7.5) \qquad\qquad n = \frac{\dot{L}_t}{L_t}$$

The trick to solving this model, one that carries over into our micro-founded growth model as well, is to scale everything in a way that makes variables stationary. Upper case letters above denote aggregates, that is, K_t is the aggregate level of capital in the economy. The model is written in such a way that if we scale by productive labor and define lower case letters as scaled variables, for example, $k_t = \frac{K_t}{A_t L_t}$, we'll be back in a situation in which we can solve for a steady state. The fact that the production function is homogeneous of degree one[1] makes this particularly convenient. Thus, for example, $y_t = k_t^\alpha$.

The only tricky part of solving the model is figuring out the budget constrained for the scaled variables. That is, we need to solve for $\dot{k}_t$. To do so we have (using either the product of quotient rule for derivatives)

$$\frac{d\left(\frac{K_t}{A_t L_t}\right)}{dt} = \frac{\dot{K}_t}{A_t L_t} - \frac{K}{AL}\left(\frac{\dot{A}_t}{A_t} + \frac{\dot{L}_t}{L_t}\right)$$

or

$$\dot{k}_t = sk_t^\alpha - (\delta + a + n)k_t$$

---

[1]This means that multiplying all the arguments by a constant is equivalent to multiplying the result by the same constant. That is, for some constant m, $mY_t = (mK_t)^\alpha (mA_t m L_t)^{1-\alpha}$.

As in previous chapters we find the steady state by setting $\dot{k}_t$ to zero, which gives us the solution

$$\bar{k} = \left(\frac{s}{\delta + a + n} \right)^{\frac{1}{1-\alpha}}$$

We could even, if we were so inclined, show that the level of savings that maximizes per unit of productive labor steady state consumption is $s = \alpha$.[2] However, from a methodological standpoint what is perhaps more interesting are the tricks we used to solve the model: scaling by productive labor and a production function that is homogeneous of degree one.

7.2   A Simple Micro-Founded Stochastic Growth Model

A micro founded growth model requires a few more ingredients than the Solow model. In particular, we will want utility maximizing agents to determine what fraction of output is re-invested (the s we simply assumed was fixed in section 7.1). However, it's easiest to begin with similarities. As before, let upper case

---

[2] To show this we want to maximize steady state consumption, which is

$$\begin{aligned} c &= (1-s)y \\ &= (1-s)k^{\alpha} \\ &= k^{\alpha} - (\delta + a + n)k \end{aligned}$$

The first order condition with respect to k is

$$0 = \alpha k^{\alpha - 1} - (\delta + a + n)$$

or

$$k = \left(\frac{\alpha}{\delta + a + n} \right)^{\frac{1}{1-\alpha}}$$

implying that $s^* = a$.

letters denote aggregates and lower case letters denote variables per unit of productive labor. Production is

$$(7.6) \qquad Y_t = K_t^\alpha L_t^{1-\alpha}$$

where for simplicity L_t is productive labor (which was $A_t L_t$ previously) and I assume population growth is zero. The aggregate budget constraint is

$$(7.7) \qquad a_t K_t^\alpha L_t^{1-\alpha} + (1-\delta)K_t = C_t + K_{t+1}$$

where a_t is productivity, K_t is physical capital, H_t is human capital, L_t productive labor, and C_t is consumption. Instead of assuming productive labor grows at a fixed rate, I'll introduce the possibility of growth shocks so that $L_{t+1} = \tau_{t+1} L_t$ where τ is labor productivity growth. τ itself in this model remains exogenous (this is not an endogenous growth model) and is governed by

$$(7.8) \qquad \hat{\tau}_{t+1} = \rho_\tau \hat{\tau}_t + \varepsilon_{t+1}$$

where hats denote log deviations from some steady state growth level. Total factor productivity a_t is also subject to shocks and governed by the usual AR(1) process

$$(7.9) \qquad \hat{a}_{t+1} = \rho_a \hat{a}_t + \epsilon_{t+1}$$

Note that both the growth rate and total factor productivity are stationary, but productive labor, capital, and consumption will grow over time. Finally, unlike the Solow model we have a representative agent who maximizes the utility function

$$(7.10) \qquad \mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{C_{t+s}^{1-\sigma}}{1-\sigma}$$

Note that because we have not included population growth in this model we need not distinguish between aggregate and per

capita variables. Additionally, it is worth pointing out that our consumer wants to maximize total consumption, not consumption scaled by labor productivity.

To solve this model we need to combine the tools of the previous chapter with the tricks from section 7.1. We begin by finding the first order conditions for the unscaled choice variables, consumption and next period capital. These combine to yield

$$(7.11) \qquad C_t^{-\sigma} = \beta E_t \left[C_{t+1}^{-\sigma} \left(\alpha a_{t+1} K_{t+1}^{\alpha-1} L_{t+1}^{1-\alpha} + 1 - \delta \right) \right]$$

Equations 7.7, 7.8, 7.9, and 7.11 form the system of difference equations that describe our model. However, only the two that govern exogenous shocks, equations 7.8 and 7.9, are log linear and stationary. The first step in working equations 7.7 and 7.11 towards this goal is to scale by productive labor. For the first order condition, equation 7.11, the fact that the marginal product of capital is homogeneous of degree zero[3] makes this simple. Dividing 7.11 by $L_t^{-\sigma} = \left(\frac{L_{t+1}}{\tau_{t+1}} \right)^{-\sigma}$ we have

$$(7.12) \qquad c_t^{-\sigma} = \beta E_t \left[c_{t+1}^{-\sigma} \tau_{t+1}^{-\sigma} \left(\alpha a_{t+1} k_{t+1}^{\alpha-1} + 1 - \delta \right) \right]$$

Doing the same for the budget constraint yields

$$(7.13) \qquad a_t k_t^\alpha + (1 - \delta)k_t = c_t + \tau_{t+1} k_{t+1}$$

The first order condition in 7.12 gives us a steady state for capital per unit of productive labor[4] which, through the budget constraint, gives a steady state value for consumption per unit of productive labor.

---

[3]That is, if we multiply each argument of $\alpha K_{t+1}^{\alpha-1} L_{t+1}^{1-\alpha}$ by a constant the result is unchanged: $\alpha(mK_{t+1})^{\alpha-1}(mL_{t+1})^{1-\alpha} = \alpha K_{t+1}^{\alpha-1} L_{t+1}^{1-\alpha}$.

[4]Steady state capital is $k = \left(\frac{\alpha\beta\tau^{-\sigma}}{1-\beta\tau^{-\sigma}(1-\delta)} \right)^{\frac{1}{1-\alpha}}$.

The final step before solving is to log linearize these two equations around the steady state values of variables. For the first order condition, equation 7.12 we have

$$(7.14) \quad \begin{aligned} -\sigma\hat{c}_t &= E_t\big[-\sigma\hat{c}_{t+1} - \sigma\hat{\tau}_{t+1} + \beta\tau^{-\sigma}\alpha k^{\alpha-1}\hat{a}_{t+1} \\ &\quad + (\alpha-1)\beta\tau^{-\sigma}\alpha k^{\alpha-1}\hat{k}_{t+1}\big] \end{aligned}$$

where variables without time subscripts denote steady state values and variables with hats denote log deviations from steady state values. For equation 7.13, the budget constraint, we have

$$(7.15) \quad k^\alpha\hat{a}_t + \big(\alpha k^\alpha + (1-\delta)k\big)\hat{k}_t = c\hat{c}_t + \tau k\hat{\tau}_{t+1} + \tau k\hat{k}_{t+1}$$

These two equations, along with the laws of motion for the total factor productivity (TFP) shock (equation 7.9) and the technology shock (equation 7.8) complete the linear system of difference equations that describe this model.

We are now at a point where we can plug in values and solve for impulse response functions. Using the parameters specified

Table 7.1: Parameter Values

Parameter	Value
ρ_a	.7
ρ_τ	.7
β	.99
σ	2
α	1/3
δ	.1
$\bar{\tau}$	1.01

in table 7.1, we can write the model in matrix form as

$$\underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 4.838 & 4.838 & 0 \\ 0.115 & -2 & -0.077 & -2 \end{bmatrix}}_{H} E_t \begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{c}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 1.689 & 0 & 4.903 & -2.188 \\ 0 & 0 & 0 & -2 \end{bmatrix}}_{N} \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_t \\ \hat{k}_t \\ \hat{c}_t \end{bmatrix}$$

where the first equation is the law of motion for $\hat{a}_t$, the second the law of motion for $\hat{\tau}_t$, the third the log linearized budget constraint, and the last the log linearized first order condition. We can re-write this system of equations as

$$
E_t \begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \\ \hat{c}_{t+1} \end{bmatrix} = \underbrace{\begin{bmatrix} 0.7 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 \\ 0.349 & -0.7 & 1.013 & -0.452 \\ 0.027 & -0.673 & -0.039 & 1.017 \end{bmatrix}}_{A} \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_t \\ \hat{k}_t \\ \hat{c}_t \end{bmatrix}
$$

where $A = H^{-1}N$. The eigenvalues of A are 1.15, 0.883, 0.7, and 0.7. Eliminating the unstable eigenvalue and associated eigenvector we then have, for the state variables,

$$
\underbrace{\begin{bmatrix} 0.7 & 0 & 0 \\ 0 & 0.7 & 0 \\ 0.274 & -1.176 & 0.883 \end{bmatrix}}_{W_s} =
$$

$$
\underbrace{\begin{bmatrix} 0 & 0.548 & 0 \\ 0 & 0 & 0.140 \\ -0.961 & -0.823 & 0.902 \end{bmatrix}}_{C_s} \underbrace{\begin{bmatrix} 0.883 & 0 & 0 \\ 0 & .7 & 0 \\ 0 & 0 & .7 \end{bmatrix}}_{\Lambda} \underbrace{\begin{bmatrix} -1.562 & 6.690 & -1.041 \\ 1.823 & 0 & 0 \\ 0 & 7.1251 & 0 \end{bmatrix}}_{C_s^{-1}}
$$

and for the choice variable, consumption,

$$
\underbrace{\begin{bmatrix} 0.165 & 1.052 & 0.289 \end{bmatrix}}_{W_c} = \underbrace{\begin{bmatrix} -0.277 & -0.147 & 0.408 \end{bmatrix}}_{C_c} \underbrace{\begin{bmatrix} -1.562 & 6.690 & -1.041 \\ 1.823 & 0 & 0 \\ 0 & 7.1251 & 0 \end{bmatrix}}_{C_s^{-1}}
$$

That is, our solution is described by

$$
E_t \begin{bmatrix} \hat{a}_{t+1} \\ \hat{\tau}_{t+1} \\ \hat{k}_{t+1} \end{bmatrix} = W_s \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_t \\ \hat{k}_t \end{bmatrix}, \qquad \hat{c}_t = W_c \begin{bmatrix} \hat{a}_t \\ \hat{\tau}_t \\ \hat{k}_t \end{bmatrix}
$$

With this solution we can now look at the impulse response functions that describe how our model responds to shocks. Fig-
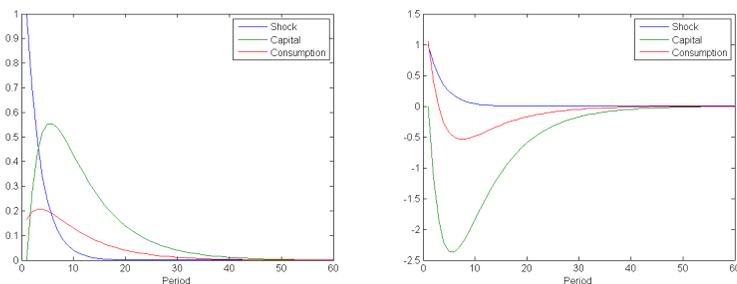
Figure 7.1: IRFs for TFP shock (left) and Labor Productivity shock (right)

ure 7.1 illustrates these impulse responses for a TFP shock (left panel) and a labor productivity shock (right panel). In both cases the capital and consumption response is more persistent then the shock itself — this was in fact already evident as the largest eigenvalue in Λ, 0.88, is larger than the persistence of either shock. Note, however, that while the capital response is less, in terms of percent deviation from steady state values, than a TFP shock itself, the opposite is true for a labor productivity shock. That is, when growth is part of our model not only is the response of variables to shocks more persistent than the shocks themselves, responses of variables are also larger: stochastic growth potentially augments the impact of exogenous shocks.

The fact that the response of capital to a positive labor productivity shock is negative may initially seem counter-intuitive, but recall that these response are scaled by labor productivity. To get a better feel for what actually goes on in the model, it is helpful to look at impulse responses for level variables as well. These level IRFs are depicted in figure 7.2. For the TFP shock (left), the IRFs are not very different from the previous figure.
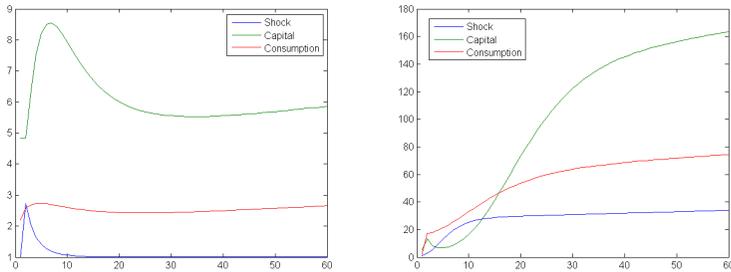
Figure 7.2: IRFs in Levels for TFP shock (left) and Labor Productivity shock (right)

For the labor productivity shock (right), however, the level IRFs look very different. While capital goes up the period after the shock is realized, it actually declines again for a few periods as agents consume more before increasing again. As we might expect given agents' preferences, consumption jumps up initially but then increases smoothly. Note also that the TFP and labor productivity shocks are of the same magnitude — one — and follow the same autoregressive process (though labor productivity itself grows exponentially), yet the response to a labor productivity shock in terms of consumption is initially around ten times larger and this difference only increases as time progresses.

7.3   Why Growth Matters

Clearly, the results of our model when shocks can impact growth are potentially very different from shocks to a stationary productivity process. Not only might impulse responses be more dramatic in terms of log deviations from the steady states of scaled variables, level impulse responses will be more dramatic still as the growth rate itself — the denominator for our scaled

variables — changes over time too. This is made clear in figure 7.2. However, there is another reason why theoretical models of growth are important: they can inform how we think about the data we observe in practice. As this is a text on macro theory I will not get into a discussion of time series econometrics here, but it is worth noting that what we created in section 7.2 is a non-stationary state space model. The insights from this framework highlight the importance of differentiating between state and jump variables and suggest ways of estimating series which may be covariance stationary without having to difference variables — a procedure that remains the standard practice in econometrics. I have included some references to relevant econometric material in the further reading section.

7.4 Further Reading

Probably the most authoritative source on growth theory is "Introduction to Modern Economic Growth" by Daron Acemoglu. For a less daunting approach chapters 1 and 3 of Romer's "Advanced Macroeconomics" also deal with growth. While there is a large literature on stochastic growth, I am not aware of any other work that deals with it in the sense of section 7.2 in which exogenous shocks impact labor productivity.

On the econometric side, beginning students might want to look at the notes I have posted on-line at srlanalytics.com. Much of my own research on that web page deals specifically with factor models including factor models for non-stationary series. An excellent and accessible resource is Stock and Watson's *chapter Factor Models and Structural Vector Autoregressions in Macroeconomics* for the Handbook of Macroeconomics available on professor Watson's Princeton web page.[5] Finally, "Non-Stationary

---

[5]https://www.princeton.edu/~mwatson/

Dynamic Factor Models for Large Datasets" by Barigozzi, Lippi, and Luciani (2016) presents a novel and interesting approach to estimating factor models for non-stationary data.

Chapter 8

The Shortcomings of Locally Approximated Macro Models

Log linear solutions to stochastic systems of difference equations (i.e. the conditions arising from a macroeconomic model) remain popular for a number of good reasons: the solution technique is relatively simple, can often be done with paper and pencil, and yields insight into the workings of a model in intermediate steps. However, because by construction linear approximations ignore higher moments of variables results thus obtained may be inaccurate. This is not due to the fact that such solutions are approximations per say, but rather due to the fact that, due to the precautionary motive, both the zero order terms (the point of approximation) and the first order terms (the impulse responses) may be wrong. A linearly approximated solution is thus not the rational expectations solution, even to a first order approximation, when the precautionary motive exists.

This chapter examines the issue in the context of a simple RBC model with CRRA utility and capital as the only produc-

tion input; in this setting the log linear solution is accurate in percent change terms, though not in levels. The shortcomings of the log linear solution are (1) failing to identify the correct steady state and thus the response to productivity shocks in levels and (2) failing to account for changes in risk (which is natural as linear solutions ignore higher moments of the variables).

8.1   The Model

Consumers maximize

$$\mathcal{U} = E_t \sum_{s=0}^{\infty} \beta^s \frac{c_{t+s}^{1-\sigma}}{1-\sigma}$$

subject to

$$a_t k_t^{\alpha} + (1-\delta)k_t = c_t + k_{t+1}$$

and an exogenous productivity shock so that the system of difference equations which describes this model are the first order condition

$$(8.1) \qquad c_t^{-\sigma} = \beta E_t c_{t+1}^{-\sigma}\left(\alpha a_{t+1} k_{t+1}^{\alpha-1} + (1-\delta)\right)$$

the budget constraint

$$(8.2) \qquad a_t k_t^{\alpha} + (1-\delta)k_t = c_t + k_{t+1}$$

and the exogenous productivity process

$$(8.3) \qquad \log(a_{t+1}) = \rho \log(a_t) + \varepsilon_{t+1}$$

The deterministic steady state is the rest point of the model that will arise if $\varepsilon_{t+1} = 0$ $\forall t$, that is, if the model is in fact deterministic. It is not the steady state, the rest point of the model in the absence of shocks, that will arise if agents form expectations that incorporate some stochastic process for ε_{t+1}. Using

the deterministic steady state as the point of approximation for a stochastic model thus contradicts the idea that agents form rational expectations. The true (stochastic) steady state of the model is increasing in risk due to the **precautionary motive**. Larger shocks, when negative, correspond to larger falls in consumption. To insure against this risk agents increase the amount of capital they hold when negative shocks are not realized (that is, in "normal" times), so that when a negative shock is realized it is not so painful.

The deterministic steady state for this model is given by

$$\bar{a} = 1$$

$$\bar{k} = \left[\frac{\beta\alpha}{1 - \beta(1 - \delta)} \right]^{\frac{1}{1-\alpha}}$$

$$\bar{c} = \bar{k}^\alpha - \delta\bar{k}$$

Log linearising conditions 8.1 - 8.2 we have

$$\hat{k}_{t+1} = \bar{k}^{\alpha-1}\hat{a}_t + (\alpha\bar{k}^{\alpha-1} + 1 - \delta)\hat{k}_t - \frac{\bar{c}}{\bar{k}}\hat{c}_t$$

$$E_t\left[-\beta\alpha\bar{k}^{\alpha-1}\hat{a}_{t+1} + \beta\alpha(1-\alpha)\bar{k}^{\alpha-1}\hat{k}_{t+1} + \beta\sigma(\alpha\bar{k}^{\alpha-1} + 1 - \delta)\hat{c}_{t+1} \right] = \sigma\hat{c}_t$$

and

$$\hat{a}_{t+1} = \rho\hat{a}_t + \epsilon_{t+1}$$

where $\hat{x}_t = \log(x_t) - \log(\bar{x})$ is approximately the percent deviation of variable x_t from its steady state value.

8.2 Impulse Responses

Figure 8.1 plots the impulse response functions for the simple case of full depreciation, that is, $\delta = 1$. For a simplified solution to this model when consumers have log utility see section **??**.

Other parameter values for this problem are the coefficient of relative risk aversion $\sigma = 2$, the time preference $\beta = 0.96$, and the capital intensity parameter $\alpha = 1/3$. Impulse response functions are in levels. The figure contrasts the solution found using log linearization around the deterministic steady state against the solution for the same model obtained through policy function iteration in which the exogenous shock has a standard deviation of 0.2 or 0.4.[1] In this case the log linear approach does very well
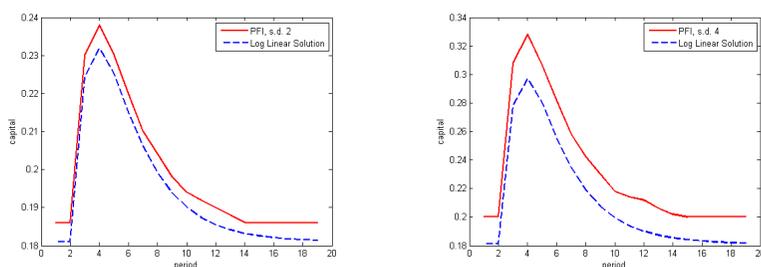
Figure 8.1: Capital IRF to a Productivity Shock with Full Depreciation

While the dynamics of the capital response to an exogenous productivity shock from the log linear solution are quite good in percent deviations, the results in levels are not.

in terms of percent deviation from the initial value. In every case capital increases to about 28% of its initial value by period 4 (the shock occurs in period 2) before falling back to that initial value. Where the log linear solution does not do well is in determining the steady state and the dynamics of the model in levels.

---

[1]This requires computing Markov transition matrices that simulate an AR(1) process; see Tauchen (1986). For this exercise I have used such a transition matrix for 100 productivity states and a capital grid from 0.001 to 1 at intervals of 0.002.

Allowing for partial depreciation increases the steady state level of capital both in the linear approximation and the numerical solution. The result is that the precautionary motive becomes less important. With full depreciation steady state capital is around 0.2 — close to the natural lower bound on capital of zero. As the degree of capital depreciation δ falls both the stochastic and the deterministic steady states increase, moving away from this lower bound. Surprisingly (at least to me), if the depreciation rate is sufficiently low the stochastic steady state for capital may in fact be less than the deterministic steady state due to the high correlation between the marginal product of capital and the marginal utility of consumption (which makes capital a risky asset and thus less desirable).

8.3   Risk Shocks

Aside from problems with impulse responses to productivity shocks, a second and perhaps more important shortcoming of log linear solutions is their failure to take into account shocks that impact higher moments of variables. This is natural in a first order approximation, but by construction ignores risk shocks. As an example, figure 8.2 plots the true impulse response to an increase in the standard deviation of the productivity shock from 0.2 to 0.4 for the model with partial capital depreciation of $\delta = 0.5$. In the log linear solution there is no response as higher moments of the productivity shock do not enter, while the numerical solution shows a 13 percent increase in capital and corresponding fall in consumption via the budget constraint. This is why, as pointed out by Leduc and Liu (2014), uncertainty shocks are aggregate demand shocks. Risk shocks need not only be uncertainty shocks however. For example, while higher asset price volatility constitutes a risk shock, a fall in asset price levels, despite being first order, acts like a risk shock: lower as-
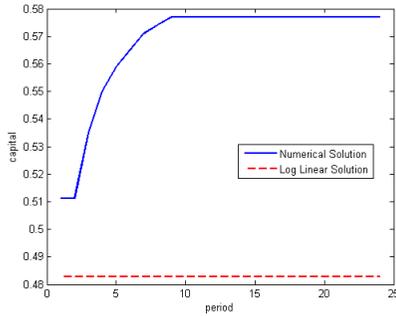
Figure 8.2: Impulse Responses to a Risk Shock

An unexpected increase in the standard deviation of the productivity
shock induces agents to accumulate more capital to insure against the
increased likelihood of a large negatvie shock.

set prices reduce the value of precautionary wealth increasing
agents' exposure to other negative shocks. Log linear models do
not capture these dynamics though they likely play a large role
in business cycle volatility.

Bibliography

[1] Daron Acemoglu. *Introduction to Modern Economic Growth*. Princeton University Press, Princeton, NJ, 2009.

[2] Jérôme Adda and Russell Cooper. *Dynamic Economics*. MIT Press, Cambridge, MA, 2003.

[3] S Rao Aiyagari. Uninsured Idiosyncratic Risk and Aggregate Saving. *The Quarterly Journal of Economics*, 109(3):659–84, August 1994.

[4] David K Backus, Patrick J Kehoe, and Finn E Kydland. International Real Business Cycles. *Journal of Political Economy*, 100(4):745–75, August 1992.

[5] Matteo Barigozzi, Macro Lippi, and Matteo Luciani. Non-Stationary Dynamic Factor Models for Large Datasets. Finance and economics discussion series 2016-024, Federal Reserve Board, 2016.

[6] Dimitri Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific, Cambridge, MA, 2007.

[7] Olivier Jean Blanchard and Charles M Kahn. The Solution of Linear Difference Models under Rational Expectations. *Econometrica*, 48(5):1305–11, July 1980.

[8]   Richard Bronson and Gabriel Costa. *Shaum's Outlines, Differential Equations.* McGraw Hill, New York, 2006.

[9]   Alpha Chaing and Kevin Wainright. *Fundamental Methods of Mathematical Economics.* McGraw Hill, Boston, MA, 2004.

[10] Edwin Purcell Dale Varberg and Steven Rigdon. *Calculus, Eighth Edition.* Prentice Hall, Upper Saddle River, NJ, 2000.

[11] Edward Dowling. *Schaum's Outlines, Introduction to Mathematical Economics.* McGraw Hill, New York, NY, 2011.

[12] Jordi Gali. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications.* Princeton University Press, Princeton, NJ, 2015.

[13] James Hamilton. *Time Series Analysis.* Princeton University Press, Princeton, NJ, 1989.

[14] Hwei Hsu. *Schaum's Outlines, Probability, Random Variables, and Random Processes.* McGraw Hill, New York, NY, 2011.

[15] M.I. Kamien and N.L. Schwartz. *Dynamic Optimization: The Calculus of Variations and Optimal Control in Economics and Management.* Dover books on mathematics. Dover Publications, 2012.

[16] Robert G. King, Charles I. Plosser, and Sergio T. Rebelo. Production, growth and business cycles : I. The basic neoclassical model. *Journal of Monetary Economics*, 21(2-3):195–232, 1988.

[17] Robert G. King and Sergio T. Rebelo. Resuscitating Real Business Cycles. RCER Working Papers 467, University of Rochester - Center for Economic Research (RCER), January 2000.

[18] Per Krusell, Anthony A. Smith, and Jr. Income and Wealth Heterogeneity in the Macroeconomy. *Journal of Political Economy*, 106(5):867–896, October 1998.

[19] Sylvain Leduc and Zheng Liu. Uncertainty Shocks are Aggregate Demand Shocks. Working paper series, Federal Reserve Bank of San Francicso, 2014.

[20] Helmut Lutkepohl. *New Introduction to Multiple Time Series Analysis*. Springer, Berlin, 2006.

[21] Tomasso Mancini Griffoli. Dynare User Guide: An introduction to the solution & estimation of DSGE models. User manual, 2013.

[22] Maurice Obstfeld and Kenneth Rogoff. *Foundations of International Macroeconomics*. MIT Press, Cambridge, MA, 1996.

[23] David Romer. *Advanced Macroeconomics*. McGraw Hill, Boston, MA, 2011.

[24] Thomas Sargent and Lars Ljungqvist. *Recursive Macroeconomic Theory*. MIT Press, Cambridge, MA, 2012.

[25] Christopher A Sims. Macroeconomics and Reality. *Econometrica*, 48(1):1–48, January 1980.

[26] James H Stock and Mark W Watson. *Introduction to Econometrics*. Prentice Hall, Upper Saddle River, NJ, 2011.

[27] James H Stock and Mark W Watson. Factor Models and
Structural Vector Autoregressions in Macroeconomics. In
Michael Woodford, editor, *Handbook of Macroeconomics*.
Oxford, Oxford Handbooks, Forthcoming.

[28] Harald Uhlig. A Toolkit for Analysing Nonlinear Dynamic
Stochastic Models Easily. QM&RBC Codes, Quantitative
Macroeconomics & Real Business Cycles, 1998.

[29] Thomas Weber. *Optimal Control Theory with Applications
in Economics*. MIT Press, Cambridge, MA, 2011.

[30] Jeffery Wooldridge. *Econometric Analysis of Cross Sec-
tional and Panel Data*. MIT Press, Cambridge, MA, 2002.

Index