Methods in Macroeconomic Dynamics

Ambrogio Cesa Bianchi[1]

November 24, 2012

[1] *Disclaimer*. These explanatory notes result from the elaboration of text and figures found in published papers, unpublished manuscripts, personal notes, and a lot of material found on the web that I collected over a long time. If you find any lack of attribution or error in quoting the source of the material in these notes —as well as if you find any errors or you would like to give me comments— please email me at ambrogio.cesabianchi@gmail.com.

Contents

Chapter 1

DIFFERENTIAL CALCULUS AND LINEARIZATION

1.1 Linear Differential Equations With Constant Coefficients

We will frequently want to solve a differential equation of the form

$$\dot{x}_t - ax_t = b, \qquad t \geq 0 \tag{1.1}$$

where the term $\dot{x}_t = \frac{dx_t}{dt}$, and the scalars a, b, and an initial condition x_0 are given. If the coefficient $b = 0$, the equation is called homogeneous, otherwise it is called inhomogenous.

1.1.1 Homogenous equations

If $b = 0$, we have the homogenous equation

$$\dot{x}_t - ax_t = 0, \qquad t \geq 0 \tag{1.2}$$

Introspection tells us that the solution is a function x_t that grows or decays exponentially at rate a (growing if $a > 0$, decaying if $a < 0$). That is, we expect the solution to be

$$x_t = e^{at}x_0, \qquad t \geq 0 \tag{1.3}$$

To see why this is the solution, write the problem as

$$\frac{dx_t}{x_t} = a \cdot dt, \qquad t \geq 0$$

and integrate both sides over the interval $[0, \tau)$. This gives

$$\begin{aligned}\log(x_\tau) - \log(x_0) &= \int_0^\tau \frac{dx_t}{x_t} = \int_0^\tau a \cdot dt = a\tau - 0 \\ \log(\frac{x_\tau}{x_0}) &= a\tau \\ x_\tau &= e^{a\tau} x_0, \qquad \tau \geq 0\end{aligned}$$

This solution diverges to $\pm\infty$ (depending on the sign of x_0) if $a > 0$ but instead decays to zero if $a < 0$. Put differently, the differential equation is unstable if $a > 0$ but stable if $a < 0$.

1.1.2 Inhomogenous equations

Otherwise, if $b \neq 0$, we have to do a bit more work. The trick is to transform the inhomogenous equation into a homogenous equation by a change of variables. We introduce here the definition of steady state

Definition. *Denote by $\bar{x}$ that unique value of x_t such that $\dot{x}_t = 0$. This is the steady state of x_t.*

Clearly,

$$\bar{x} = -\frac{b}{a} \tag{1.4}$$

which is well defined so long as $a = 0$. Now introduce a new variable

$$y_t = x_t - \bar{x}$$

(i.e., the difference between the actual and steady state value of x_t). Note that this change of variables implies

$$\dot{y}_t = \frac{d(x_t - \bar{x})}{dt} = \dot{x}_t$$

and therefore, substituting x_t for $y_t + \bar{x}$,

$$\dot{y}_t = a(y_t + \bar{x}) + b = ay_t + a\bar{x} + b$$

but from the steady state defined by 1.4 $a\bar{x} + b = 0$, then

$$\dot{y}_t = ay_t$$

So the new variable obeys a homogenous differential equation and therefore has the solution

$$y_t = e^{at} y_0, \qquad t \geq 0$$

Plugging back in the definition $y_t = x_t - \bar{x}$ gives

$$x_t - \bar{x} = e^{at}(x_0 - \bar{x}), \qquad t \geq 0$$

which is sometimes re-written

$$x_t = (1 - e^{at})\bar{x} + e^{at} x_0, \qquad t \geq 0 \tag{1.5}$$

Inhomogeneous equation can also take the form of the trivial differential equation

$$\dot{x}_t = b, \qquad t \geq 0$$

has the solution

$$x_t = bt + x_0, \qquad t \geq 0$$

1.1.3 Stability

The stability properties of this solution are easy. If $a < 0$, then

$$\lim_{t\to\infty} e^{at} = 0$$

which implies that

$$\lim_{t\to\infty} x_t = \bar{x}$$

irrespective of the value of the initial condition. That is, if $a < 0$, the steady state $\bar{x}$ is **globally stable**. If $a > 0$, then

$$\lim_{t\to\infty} e^{at} = \pm\infty$$

depending on the sign of $x_0 - \bar{x}$, i.e., depending on whether the variable starts above or below its steady state value. If the initial condition happens to be $x_0 = \bar{x}$, the system stays there irrespective of the value of a.

1.2 Linear difference equations

Similarly, if we have the linear difference equation

$$x_{t+1} - ax_t = b, \qquad t = 0, 1, 2, ... \tag{1.6}$$

given scalars a, b and an initial condition x_0, then the **homogenous solution** (when $b = 0$) is

$$x_t = a^t x_0, \qquad t = 0, 1, 2, ... \tag{1.7}$$

You can verify this by iterating as follows

$$\begin{aligned} x_1 &= a^1 x_0 = \\ x_2 &= a^1 x_1 = a^2 x_0 \\ &... \\ x_t &= a^1 x_{t-1} = a^t x_0 \end{aligned}$$

The general solution when $b \neq 0$ can be obtained as for the differential equations. First define the steady state as

Definition. *In discrete time we denote by $\bar{x}$ that unique value of x_t such that $x_{t+1} = x_t$. This is the steady state of x_t.*

The steady steate is then

$$\bar{x} = \frac{b}{1-a} \tag{1.8}$$

Now introducing a new variable as before

$$y_t = x_t - \bar{x}$$

and substituting x_t for $y_t + \bar{x}$,

$$\begin{aligned} y_{t+1} + \bar{x} &= a(y_t + \bar{x}) + b \\ y_{t+1} &= ay_t - (1-a)\bar{x} + b \end{aligned}$$

but from the steady state defined by 1.8, we have that $(1-a)\bar{x} = b$, then

$$y_{t+1} = ay_t$$

So the new variable obeys a homogenous difrference equation and therefore has the solution

$$y_t = a^t y_0, \qquad t \geq 0$$

Plugging back in the definition $y_t = x_t - \bar{x}$ gives

$$x_t - \bar{x} = a^t(x_0 - \bar{x}), \qquad t \geq 0$$

which is sometimes re-written

$$x_t = (1-a^t)\bar{x} + a^t x_0, \qquad t \geq 0 \tag{1.9}$$

In discrete time, stability properties are determined by whether $|a| < 1$ or not. If $|a| < 1$, then

$$\lim_{t\to\infty} a^t = 0$$

so that $x_t \to \bar{x}$ irrespective of the value of the initial condition. If $|a| > 1$, then

$$\lim_{t\to\infty} a^t = \pm\infty$$

so the steady state is not stable. Notice that if $a > 0$, the motion of x_t is monotonic but if $a < 0$, the motion of x_t is oscillatory. Again, if the initial condition happens to be $x_0 = \bar{x}$, the system stays there irrespective of the value of a.

1.3 Systems of Linear Difference Equations

We will frequently want to solve a two-dimensional system of difference equations

$$x_{t+1} - Ax_t = b, \qquad t \geq 0 \tag{1.10}$$

given a 2-by-2 matrix A, a 2-by-1 vector b and a 2-by-1 vector of initial conditions x_0. (Higher dimensional systems generalize in an obvious way, so I will do almost everyting here for the two dimensional case).

1.3.1 Uncoupled equations

If the matrix A is diagonal, i.e., of the form

$$\begin{pmatrix} a_{11} & 0 \\ 0 & a_{22} \end{pmatrix}$$

then there is no interaction between the two equations (they are independent) and can be written

$$\begin{aligned} x_{1,t+1} - a_{11}x_{1,t} &= b_1 \\ x_{2,t+1} - a_{22}x_{2,t} &= b_2 \end{aligned}$$

These have the obvious solutions

$$\begin{aligned} x_{1t} &= (1 - a_{11}^t)\bar{x}_1 + a_{11}^t x_{1,0} \\ x_{2t} &= (1 - a_{22}^t)\bar{x}_2 + a_{22}^t x_{2,0} \end{aligned}$$

or in matrix form

$$x_t = (I_2 - A^t)\bar{x} + A^t x_0$$

where I is a conformable identity matrix. In this expression, the matrix power A^t is justified because A is diagonal so that

$$A^t = \begin{pmatrix} a_{11}^t & 0 \\ 0 & a_{22}^t \end{pmatrix}$$

Remark. *A^t does not equal each of its elements to the power t unless A is diagonal.*

The steady state vector $\bar{x}$ is the natural analogue to the scalar case,

$$\bar{x} = (I - A)^{-1} b$$

1.3.2 Coupled equations

If the square matrix A is not diagonal, there is feedback between the two equations and we can no longer solve them independently. In this case, the equations are said to be **coupled**. Since solving diagonal system is easy, it would be nice if there was a change of variables that allowed us to ”uncouple” or ”diagonalize” the system. In fact, for an important class of matrices, we can make this change of variables quite easily.

Definition. *A square matrix A can be diagonalized if there exists an invertible matrix Q such that the matrix D defined by*

$$D = Q^{-1} A Q$$

is diagonal. The necessary and sufficient condition for A to be diagonalized is that its eigenvalues are distinct (e.g, nonzero nilpotent matrices are not diagonalizable).

To know if a matrix is diagonalizable, we need to answer the following questions: What are the eigenvalues? How can we know if the eigenvalues are distinct? and also we could be interested in which is the nature of the matrix Q? Let's brush up on some important concepts from linear algebra that will actually help us do diagonalizations. For the following examples I will consider a 2-by-2 matrix, namely

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

Definition. *If A is a square matrix, a scalar λ is an eigenvalue of A if and only if $B = A - \lambda I$ is a singular matrix.*

Remember that a square matrix is singular if and only if its determinant is zero. Therefore we are looking for the values od λ such that B is singular, i.e., such that $\det(B) = \det(A - \lambda I) = 0$. In the example we are considering

$$\begin{aligned}
\det(A - \lambda I) &= \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} = \\
&= (a_{11} - \lambda)(a_{22} - \lambda) - a_{21} a_{12} = \\
&= \lambda^2 - (a11 + a22)\lambda + (a11a22 - a12a21) = \\
&= p(\lambda) = 0
\end{aligned}$$

We define the polynomial $p(\lambda)$ as the **characteristic polynomial** in λ. For a 2-by-2 system, this is just a quadratic equation (higher dimensional A lead to higher dimensional polynomials, but that need not bother us here). That is, finding the eigenvalues of A just means finding the roots of

$$\lambda^2 - (a11 + a22)\lambda + (a11a22 - a12a21) = 0$$

or by the quadratic formula

$$\lambda_1, \lambda_2 = \frac{-\beta \pm \sqrt{\beta^2 - 4\alpha\gamma}}{}$$

where

$$\begin{aligned} \alpha &= 1 \\ \beta &= -(a11 + a22) \\ \gamma &= a11a22 - a12a21 \end{aligned}$$

Of course the roots may be real or complex. Generally, a square matrix has as many eigenvalues as its dimension, but some of these values may be repeated. The discriminant of a polynomial is an expression which gives information about the nature of the polynomial's roots. In particular a polynomial has a multiple root (i.e. a root with multiplicity greater than one) in the complex numbers if and only if its discriminant is zero. For example, in the 2-by-2 case, if the discriminant $\beta 2 - 4\alpha\gamma = 0$, both roots are the same.

Definition. *For a diagonal matrix the eigenvalues are equal to the diagonal elements*

To prove this simply substitute $a_{21} = a_{12} = 0$ in the formula to compute the eigenvales and see that,

$$\begin{aligned} \lambda_1, \lambda_2 &= \frac{a11 + a22 \pm \sqrt{(a11 + a22)^2 - 4a11a22}}{2} = \\ &= \frac{a11 + a22 \pm \sqrt{(a11 - a22)^2}}{2} = a_{11}, a_{22} \end{aligned}$$

In this case, $\lambda_1 = a_{11}$ and $\lambda_2 = a_{22}$.

We just showed that the eigenvalues are scalars such that the matrix $B = A - I\lambda$ is singular, i.e., $\det(B) = 0$. Notice that, according to another definitiion of singularity, the matrix B is singular if and only if the equation $Bx = 0$ has solutions other than $x = 0$. Then

Definition. *If λ is such that $B = A - I\lambda$ is singular (i.e., is an eigenvalue of A), there must be a vector $x \neq 0$ such that $Bx = 0$, or equivalently*

$$Ax = \lambda x, \qquad \text{for } x \neq 0.$$

Such vector x is called **eigenvector**

It can be shown that the remarkable relationship between a diagonalized matrix, eigenvalues, and eigenvectors follows from the beautiful mathematical identity (the eigen decomposition) that a square matrix can be decomposed into the very special form

$$A = PDP^{-1}$$

where P is a matrix composed of the eigenvectors of A, and D is the diagonal matrix constructed from the corresponding eigenvalues.

To sum up, if the matrix A is well-behaved (having distinct eigenvalues) we can diagonalize it by $A = QDQ^{-1}$ where D and Q are defined above. Then if we have the system of difference equations as in 1.10

$$\begin{aligned} x_{1,t+1} - a_{11}x_{1,t} - a_{12}x_{2,t} &= b_1 \\ x_{2,t+1} - a_{21}x_{1,t} - a_{22}x_{2,t} &= b_2 \end{aligned}$$

we have the following procedure.

1. Compute the steady state: $\bar{x} = (I - A)^{-1}b$
2. Set up the equivalent homogeneous system by defining $y_t = x_t - \bar{x}$ so that $y_{t+1} = Ay_t$ (as in the one-equation setting)
3. If A has all distinct eigenvalues, use the change of variables $y_t = Qz_t$ with Q matrix of the eigenvector of A to get

$$\begin{aligned} Qz_{t+1} &= AQz_t \\ z_{t+1} &= Q^{-1}AQz_t \\ z_{t+1} &= Dz_t \end{aligned}$$

   which is an uncoupled system with solution

$$z_t = D^t z_0$$

4. Use the fact that $z_t = Q^{-1}y_t$ to get

$$y_t = QD^tQ^{-1}y_0$$

   And finally use $y_t = x_t - \bar{x}$ to get

$$\begin{aligned} x_t - \bar{x} &= QD^tQ^{-1}(x_0 - \bar{x}) \\ x_t &= (I - QD^tQ^{-1})\bar{x} + QD^tQ^{-1}x_0 \end{aligned}$$

The stability properties are now easy to determine. Thanks to the fact that D is diagonal I can write

$$D^t = \begin{pmatrix} \lambda_1^t & 0 \\ 0 & \lambda_2^t \end{pmatrix}$$

Then, if all of the eigenvalues of A are less than one in absolute value, the matrix $D^t \rightarrow 0$ as $t \rightarrow \infty$, so that $x_t \rightarrow \bar{x}$. In this case, the system is **(globally) stable**. Otherwise, if any eigenvalues of A are larger than one in absolute value, the system is unstable.

Finally, it's instructive to write the solution of the homogeneous equation as (from point 3)

$$\begin{aligned} y_t &= Qz_t = \\ &= QD^t z_0 \end{aligned}$$

which can be re-written as

$$y_t = \sum_{i=1}^{n} q_i \lambda_i^t z_{i,0}$$

That is the solution y_t (the distance from the steady state) is a weighted average of the eigenvalues λ_i with weights corresponding to the eigenvectors q_i times the corresponding initial condition $z_{i,0}$. Here we see explicitly that if any λ_i is bigger than one in absolute value, the vector y_t will diverge unless the "weight" given to that eigenvalue is zero.

1.4 Linearization

In mathematics and its applications, linearization refers to finding the linear approximation to a function at a given point. In the study of dynamical systems, linearization is a method for assessing the local stability of an equilibrium point of a system of nonlinear differential equations or discrete dynamical systems.

In fact will frequently want to linearize a difference equation of the form

$$x_{t+1} = \Psi(x_t) \tag{1.11}$$

given an initial condition x_0. Equation 1.11 can be approximated by a linear Taylor expansion around a given point $(\bar{x})$

$$x_{t+1} = \Psi(\bar{x}) + \Psi'(\bar{x})(x_t - \bar{x}). \tag{1.12}$$

The right hand side of this expression is a linear equation in x_t with slope $\Psi'(\bar{x})$ equal to the derivative of $\Psi(\cdot)$ evaluated at the point $\bar{x}$. Typically, we will want to linearize around a steady state, $\bar{x}$, which also has the property $\Psi(\bar{x}) = \bar{x}$ so that

$$x_{t+1} = \bar{x} + \Psi'(\bar{x})(x_t - \bar{x}).$$

If we treat this approximation as exact, we have an inhomogenous difference equation with constant coefficients of the form

$$x_{t+1} - ax_t = b, \qquad t = 0, 1, 2, ...$$

where

$$\begin{aligned} a &= \Psi'(\bar{x}) \\ b &= (1-a)\bar{x} \end{aligned}$$

And we know that the solution to this difference equation is

$$x_t = (1-a^t)\bar{x} + a^t x_0$$

The stability properties depend on the coefficient a, as usual. If $|a| < 1$, then $\lim_{t\to\infty} a^t = 0$ so that $x_t \to \bar{x}$ irrespective of the value of the initial condition. If $|a| > 1$, then $\lim_{t\to\infty} a^t = \pm\infty$ so the steady state is not stable. Notice that these are conditions for the **local stability** of the difference equation (since we have had to take a local approximation to the function $\Psi()$): we cannot use this method to determine global properties of the solution. In a linear model, local and global stability are the same thing, but that is not true here. A lot of applied work in economics works by using linear approximations to non-linear models.

Finally, let's consider tha case in which we wan to take an approximation of a two-variable function $f(x,y)$ around a given point (x_0, y_0). This is given by is given by

$$f(x,y) \simeq f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0)$$

The Taylor series may also be generalized to functions of more than two variables.

1.5 Log-Linearization

A popular alternative to linearizing a model is to log-linearize it. To do this, first notice that the fact that the percentage deviation between x_t and $\bar{x}$ can be approximated by

$$\log\left(\frac{x_t}{\bar{x}}\right) \simeq \frac{x_t - \bar{x}}{\bar{x}}$$

To see it first notice that

$$\log\left(\frac{x_t}{\bar{x}}\right) \simeq \log\left(1 + \frac{x_t - \bar{x}}{\bar{x}}\right)$$

Then define $z_t = \frac{x_t - \bar{x}}{\bar{x}}$ take a Taylor approximation of $\log(1+z_t)$ around the point $z_0 = 0$.

$$\begin{aligned} \log(1+z_t) &\simeq \log(1+z_0) + \frac{1}{1+z_0}(z_t - z_0) \\ \log(1+z_t) &\simeq z_t \end{aligned}$$

Now define a new variable as the deviation of the variable of interest from its steady state

$$\hat{x}_t = \log\left(\frac{x_t}{\bar{x}}\right)$$

and notice that we can write

$$x_t = \bar{x}e^{\log\left(\frac{x_t}{\bar{x}}\right)} = \bar{x}e^{\hat{x}_t}.$$

We can substitute now the new defined x_t in equation 1.11 and we get

$$\bar{x}e^{\hat{x}_{t+1}} = \Psi(\bar{x}e^{\hat{x}_t})$$

Taking an approximation of this function around the steady state, means computing a Taylor expansion around the point $\hat{x}_t = 0$ (the deviation from the steady state in the steady state is obviously null)

$$\begin{aligned} \bar{x}e^0 + \bar{x}e^0(\hat{x}_{t+1} - 0) &= \Psi(\bar{x}e^0) + [\Psi'(\bar{x}e^0) \cdot \bar{x}e^0](\hat{x}_t - 0) \\ \bar{x} + \bar{x}\hat{x}_{t+1} &= \Psi(\bar{x}) + \Psi'(\bar{x})\bar{x}\hat{x}_t \end{aligned}$$

Moreover, remember that the steady state is given by a set of $\bar{x}$ such that

$$x_{t+1} = \Psi(x_t) \Leftrightarrow \bar{x} = \Psi(\bar{x}).$$

Finally we can re-write the general dynamic system expressed by 1.11 in its approximate form as

$$\hat{x}_{t+1} = \Psi'(\bar{x})\hat{x}_t \tag{1.13}$$

which is a homogeneous difference equation in log deviations from the steady state. Local stability again crucially depends on the absolute magnitude of $\Psi'(\bar{x})$, with $\Psi'(\bar{x}) < 1$ giving a locally stable steady state.

Similar derivations allow us to do more complicated log-linearizations (this is sometimes known as the "Campbell calculus"). The following basic rule often helps immensely. Suppose that we have a differentiable function $z_t = f(x_t, y_t)$ (more arguments will generalize in an obvious fashion). Then the log-linear approximation is

$$\bar{z}_t\hat{z}_t \simeq f_x(\bar{x},\bar{y})\bar{x}\hat{x}_t + f_y(\bar{x},\bar{y})\bar{y}\hat{y}_t$$

Obviously the function z_t can be one of the two independent variables forwarded by one. This is a typical first order condition in dynamic optimization problems

$$\begin{aligned} \bar{x}\hat{x}_{t+1} &\simeq f_x(\bar{x},\bar{y})\bar{x}\hat{x}_t + f_y(\bar{x},\bar{y})\bar{y}\hat{y}_t \\ \hat{x}_{t+1} &\simeq f_x(\bar{x},\bar{y})\hat{x}_t + f_y(\bar{x},\bar{y})\frac{\bar{y}}{\bar{x}}\hat{y}_t \end{aligned}$$

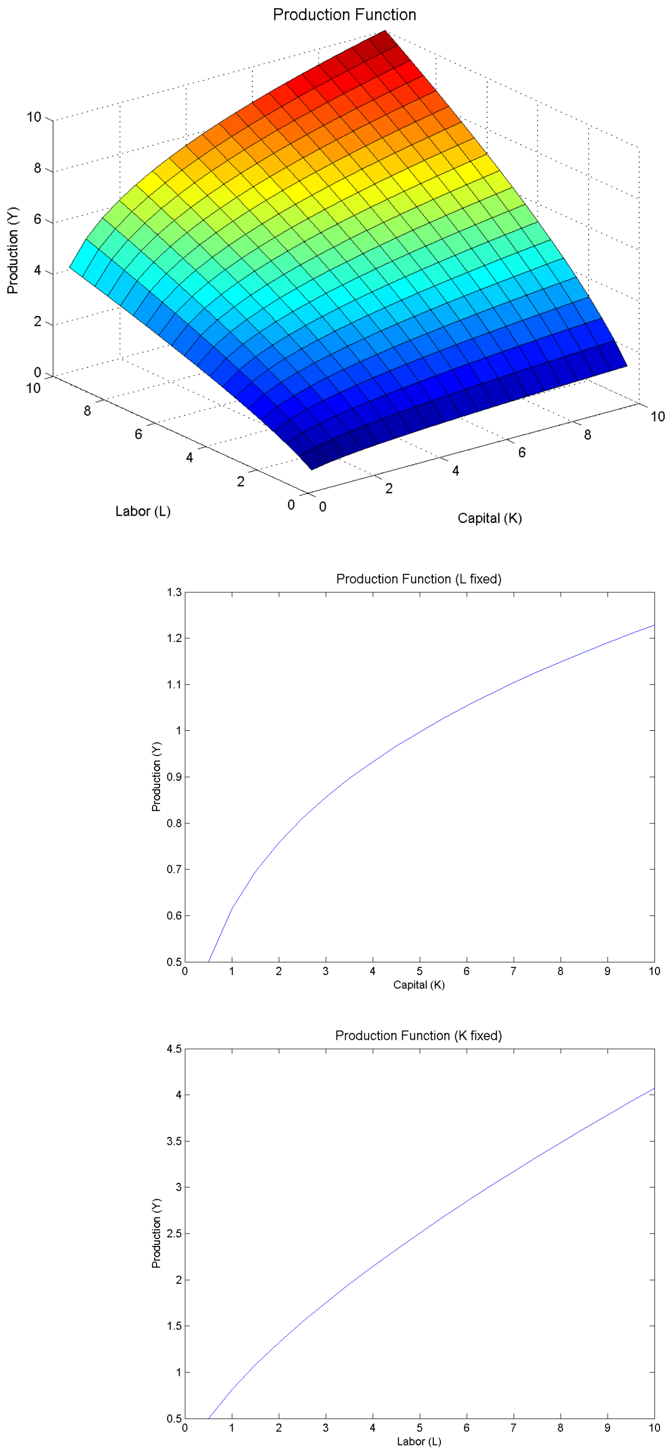
Chapter 2

THE NEOCLASSICAL FIRM AND THE PRODUCTION FUNCTION

2.1 The Production function

The production, Y, is a function of two factors, capital and labor $F(K, L)$. Firms produce investment and consumption goods using the following technology: $Y_t = F(K_t, L_t)$.

Figure 2.1 The Production Function - A graphical illustration of $Y = K^{0.3}L^{0.7}$



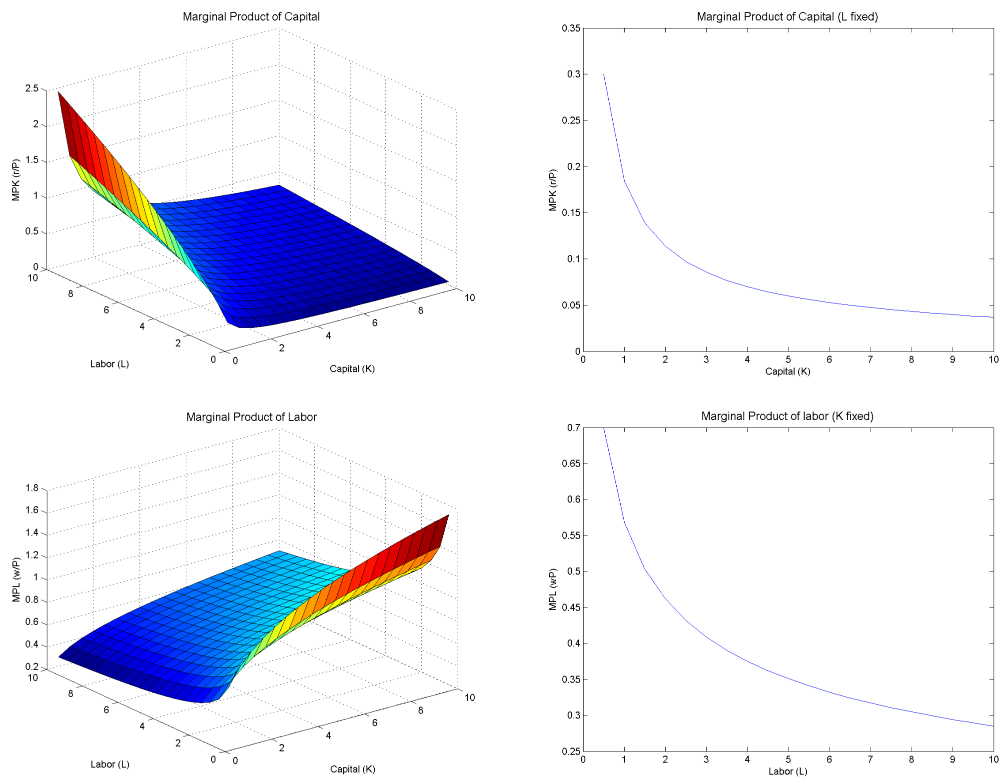
Assumption. *Constant returns to scale (CRTS) to the inputs K and L. For any* $b > 0$, $bY_t = F([bK_t], A_t[bL_t])$. *Some implications:*

- *Output per worker only depends on capital per worker. To see this take* $b = 1/L$ *and notice that* $y = \frac{Y}{L} = F(\frac{K}{L}, 1) = F(k, 1) = f(k)$.
- *Then CRTS implies that small firms produce as much output per worker as small firms if they have the same K/L (level does not matter, the only thing is the ratio. Is this credible?)*

Assumption. *Positive but decreasing marginal products (MP) to capital and labor taken separately:* $\frac{\partial F(K,L)}{\partial K} > 0$ *and* $\frac{\partial^2 F(K,L)}{\partial^2 K} < 0$.

Assumption. *Inada conditions:* $f\prime(0) = \infty$ *and* $f\prime(\infty) = 0$

Figure 2.2 Marginal Productivities - A graphical illustration of $Y = K^{0.3}L^{0.7}$



Small Digression: Derivatives
Given a function $y = f(x)$ first define:

Differential: $\Delta y = f(x + \Delta x) - f(x)$
Derivative: $f'(x) = lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$

Or: $\Delta y = f'(x)\Delta x + \epsilon$

Then the derivative is the ratio $\frac{\Delta y}{\Delta x}$ when Δx is sufficiently small. Its geometrical interpretation is the amount by which y grows (in level) when x changes by a given amount. For example, consider $y = x^2$, $f'(x) = 2x$, where x are the hours I work and y the dollars I am paid per hour. Let's assume I work ten hours per week, then my wage is 100$ per week. I want to know how by how much my wage increases if I work an additional hour. The derivative is the answer: $\Delta y = f'(x)\Delta x + \epsilon = (2 \cdot 11) \cdot 1 = 22$. In reality, I know that if I work 11 hours per week I'll get an increase of wage of 21. The derivative is a good approximation ($\epsilon = 1$ in this case) and it is more precise the smaller is the increment Δx.

When considering functions of several variables, we can introduce the partial differentiation, i.e. the process of calculating the derivative function of the dependent variable when only one independent variable is undergoing a small change and the others remain constant. For each bi-variate function there are therefore two partial derivative functions. $\frac{\partial F(K,L)}{\partial K}$ measures the change of the dependent variable $Y = F(K, L)$ relative to a change in K keeping L constant. It is the margin of Y with respect to K holding L constant. Consequently, $\frac{\partial F(K,L)}{\partial L}$ measures the change of Y when L changes and K is fixed. For a graphical representation of partial derivatives see this page.
In the univariate case, consider the function $y = f(x) = x^2$ and its derivative, $f'(x) = 2x$. The derivative is a measure of how a function changes as its input changes. The numerical interpretation of the derivative is the following: in any point in the space (x, y) i can evaluate by how much the $f(x)$ changes if x increases by an infinitesimal quantity ($dx = 0.01$). Then we get:

x	$f(x)$	$f'(x)$	$f(x) + f'(x)dx$	$f(x + dx)$
	x^2	$2x$	$x^2 + 2xdx$	$(x + dx)^2$
1	1	2	1.020	1.020
2	4	4	4.040	4.040
3	9	6	9.060	9.060
4	16	8	16.080	16.080
5	25	10	25.100	25.100
6	36	12	36.120	36.120

In the multivariate case, consider the function $f(x, y) = x^2y^2$ and its partial derivative with respect to the x, $\partial f(x, y)/\partial x = 2xy^2$. The numerical interpretation of the partial derivative is the following: in any point in the space (x, y) i can evaluate by how much the $f(x)$ changes if I increase x by an infinitesimal quantity (say 0.01) and keep fixed the y to an arbitrary value. Then we get:

$x, y = 5$	$f(x, y = 5)$	$f_x(x, y = 5)$	$f(x, y = 5) + f_x(x, y = 5)dx$	$f(x + dx, y = 5)$
	x^2y^2	$2xy^2$	$x^2y^2 + 2xy^2dx$	$(x + dx)^2y^2$
1	25	50.0	25.50	25.50
2	100	100.0	101.00	101.00
3	225	150.0	226.50	226.50
4	400	200.0	402.00	402.00
5	625	250.0	627.50	627.50

2.2 Production Function and Technological Change

We can think on technological change as a process that enhance the efficiency of the factors of production. In this sense, technology (e.g. industrial revolution, Ford T,...) increases much more the efficiency of labor than the efficiency of capital. Therefore we consider a labor-augmenting technology in the production function, $F(K_t, A_t L_t)$.

- A is a factor that will capture technological progress or improvements in efficiency. The greater A, the more the firm produces with a given amount of resources K and L (the "more efficient" is the firm).
- Technological progress is taken as given; the firm cannot control it.
- Technological progress multiplies labor. This is called labor-augmenting technological progress.

2.2.1 The Static Market Equilibrium

There are two markets, one for each factor and every factor is determined in its market.

Assumption. *All markets are assumed to be perfectly competitive, which means Demand=Supply in all markets*

Assumption. *Supply of capital and labor are fixed*

Assumption. *The representative neoclassical firm, chooses K_t and L_t in order to maximize profits. The price of each factor is determined by firms demand.*

The profits of the representative firm are given by

$$\begin{aligned} \pi &= PY - rK - wL && (2.1) \\ s.t.\ Y &= F(K, AL), && (2.2) \end{aligned}$$

where P is the price of the good produced, r and w are the price of capital and labor respectively. In any competitive equilibrium, the optimal choice of the firm is given by

$$\begin{aligned} \frac{r_t}{P} &= F_K(K_t, A_t L_t) = MPK && (2.3) \\ \frac{w_t}{P} &= F_L(K_t, A_t L_t) = MPL. \end{aligned}$$

Summarizing:

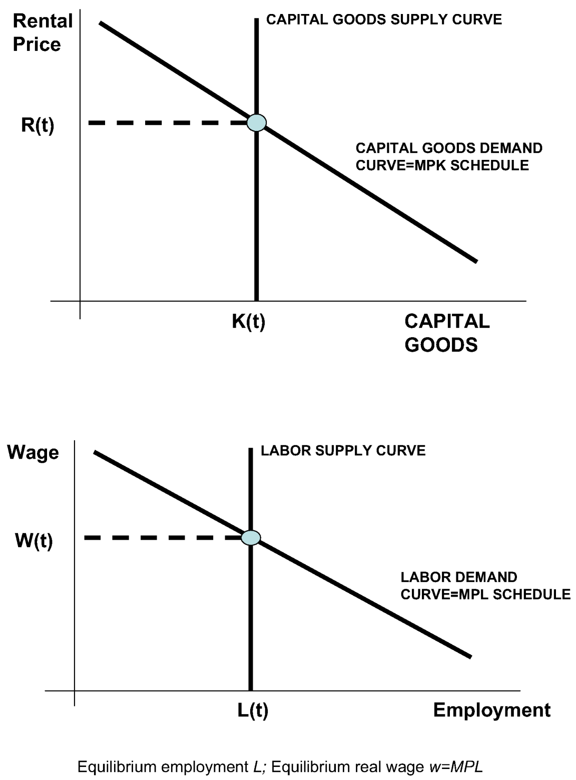
1. Labor market

- Supply and demand: Labor supply is inelastic, Labor demand by firms
- For any given capital stock firms hire labor to maximize profits[1]: $\max_L F(K, AL) - wL \Longrightarrow F_L(K, AL) - w = 0 \Longrightarrow MPL = w$. Which means that firms hire additional workers until this condition is verified

2. Capital Market

- Supply and demand: Capital supply is inelastic, firms own capital but they can also rent it (if they think is cheaper to do so)
- For any given level of employment firms rent capital to maximize profits: $\max_K F(K, AL) - rK \Longrightarrow F_K(K, AL) - r = 0 \Longrightarrow MPK = r$. Which means that firms rent additional units of capital until this condition is met

Figure 2.3 Market Clearing Conditions



Conclusion. *By maximizing profits, firms find a an optimal relation between real wage and quantity of labor. For any given quantity of labor supplied by households, firms will choose that amount of labor and will offer a real wage such that* $MPL = w/P$

[1] Price here is normalized to one.

Conclusion. *By maximizing profits, firms find a an optimal relation between real interest rate and quantity of capital. For any given quantity of capital supplied by households, firms will choose that amount of capital and will offer a real rate such that* $MPK = r/P$

Chapter 3

THE SOLOW MODEL

3.1 Description of the model

Accumulation of physical capital cannot account for 1) the dramatic growth over time of output per capita 2) the vast geographic differences in output per person. Exogenous technological progress can be one source of differences in output per capita. The model is first presented in continuous time; in Section XX explains a discrete time version of the model.

Assumption: *Neoclassical firm with labor-augmenting technology in the production function, constant return to scale, and usual concavity properties:* $Y_t = F(K_t, A_t L_t)$

Assumption: *The terms* L_t *and* A_t *grow exponentially at constant rates,* n *and* g*, respectively:* $A_t = e^{gt} A_0$ *and* $L_t = e^{nt} L_0$

Assumption: *Capital depreciates at rate* δ*. To keep capital (K) at the desired level, firms have to invest*

Assumption: *Closed economy, no government:* $Y = C + I$ *and therefore* $S = Y - C = I$*. Savings are assumed to be a constant fraction,* s*, of income*

In continuous time, the change in stock of capital evolves according to

$$\begin{aligned} \dot{K}_t &= I_t - \delta K_t = \\ &= sY_t - \delta K_t \end{aligned} \tag{3.1}$$

To solve the model rewrite all the equations in terms of output per efficiency worker. First the production function

$$y = f(\tilde{k}_t)$$

and then the equation of the evolution of the capital stock, also called law of motion of capital

$$\dot{\tilde{k}}_t = sf(\tilde{k}_t) - (\delta + n + g)\tilde{k}_t \tag{3.2}$$

Steady State: a value of $\tilde{k}_t$ which solves the law of motion (i.e. such that $\dot{\tilde{k}}_t = 0$)

Small Digression: Growth rates

Derivative of X with respect to time: $\frac{\partial X}{\partial t} = \dot{X}$

The growth rate of X with respect to time: $\hat{X} = \frac{X_{t+1} - X_t}{X_t}$

If Δt is sufficiently small: $\hat{X} = \frac{\dot{X}}{X}$

Some tricks:

- $z = xy \Rightarrow \hat{z} = \hat{x} + \hat{y}$
- $z = x = y \Rightarrow \hat{z} = \hat{x} - \hat{y}$
- $z = ax \Rightarrow \hat{z} = \hat{x}$
- $z = x^a \Rightarrow \hat{z} = a\hat{x}$
- $z = x + y \Rightarrow \hat{z} = \frac{\bar{x}}{\bar{z}}\hat{x} + \frac{\bar{y}}{\bar{z}}\hat{y}$

where bar variables are steady state values. For a formal description of log-linear approximation of growth rates see Chris Edmond notes

3.1.1 Output per efficiency worker

In this section I'll show in a few bullets how to derive the production function and the law of motion of capital in terms of capital per efficiency worker.

Production Function: $Y_t = F(K_t, A_t L_t)$

- Given CRTS, choose $b = \frac{1}{A_t L_t}$
- And get: $\frac{Y_t}{A_t L_t} = F(\frac{K_t}{A_t L_t}, 1)$
- Then define: $\tilde{y} = \frac{Y_t}{A_t L_t}$ and $\tilde{k} = \frac{K_t}{A_t L_t}$
- And finally: $\mathbf{\tilde{y}} = \mathbf{f}(\mathbf{\tilde{k}})$

Evolution of Capital stock: $\dot{K}_t = sF(K_t, A_t L_t) - \delta K_t$

- Premultiply by $A_t L_t$ the RHS: $\dot{K}_t = A_t L_t [sF(\frac{K_t}{A_t L_t}, 1) - \delta \frac{K_t}{A_t L_t}]$
- to get: $\dot{K}_t = A_t L_t [sf(\tilde{k}) - \delta \tilde{k}]$
- then premultiply by $\frac{K}{K}$ the LHS: $\frac{\dot{K}_t}{K_t} \frac{K_t}{A_t L_t} = sf(\tilde{k}_t) - \delta \tilde{k}_t$
- Notice that: $\hat{\tilde{k}} = \frac{\dot{\tilde{k}}_t}{\tilde{k}_t} = \frac{\dot{K}_t}{K_t} - (n + g)$

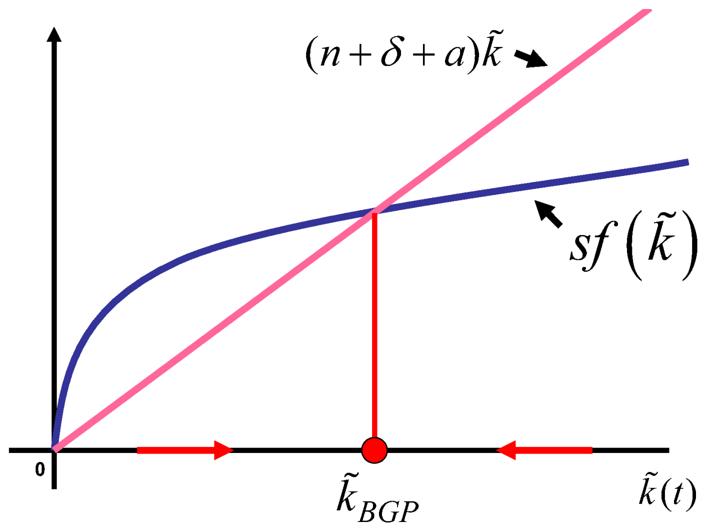
- multiply by $\tilde{k}_t$ both sides: $\dot{\tilde{k}}_t = \frac{\dot{K}_t}{K_t}\tilde{k}_t - (n+g)\tilde{k}_t$
- Plug last expression in the evolution of capital stock: $\dot{\tilde{k}}_t + (n+g)\tilde{k}_t = sf(\tilde{k}) - \delta\tilde{k}$
- and finally $\mathbf{\dot{\tilde{k}}_t = sf(\tilde{k}) - (\delta + n + g)\tilde{k}_t}$

Then the solution of the Solow model is given by the unique $\tilde{k}_t^*$ such that

$$sf(\tilde{k}_t) = (\delta + n + g)\tilde{k}_t \tag{3.3}$$

Figure 3.1 shows the Balanced Growth Path, an equilibrium where all variables grow at constant rates (this growth rate can be 0). Notice that the BGP is globally stable, which means that the economy ends up in the BGP in the long run no matter where the economy starts.[1] To see this notice that if $\tilde{k}_t < \tilde{k}_t^*$, the derivative of capital (per efficiency worker), $\dot{\tilde{k}}_t$, is positive. Then $\tilde{k}_t$ grows until the condition is met. The opposite happens when $\tilde{k}_t > \tilde{k}_t^*$.

Figure 3.1 Solution of the Solow Model



Conclusion: *The economy will end up at the same level of capital per efficiency worker, no matter what the initial values for A, K, L.*

Conclusion: *In the balanced growth path, the change in capital per efficiency worker is ZERO.*

Another important implication of the Solow model concerns growth rates in the BGP. Remembering that the growth rate of a variable X is defined as $\frac{\dot{X}}{X}$, we can divide the law of motion of capital (equation 3.3) by $\tilde{k}_t$ to get

$$\frac{\dot{\tilde{k}}_t}{\tilde{k}_t} = \hat{\tilde{k}}_t = s\frac{f(\tilde{k}_t)}{\tilde{k}_t} - (\delta + n + g).$$

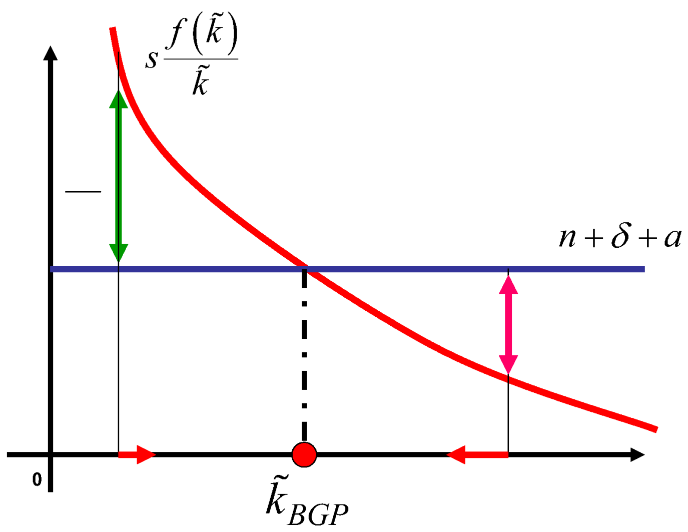
[1]Clearly, ruling out the possibility of $K_0 = 0$ and $L_0 = 0$.

In the BGP $\dot{\tilde{k}}_t = 0$ then also $\hat{\tilde{k}}_t = 0$. Therefore the following relation has to hold: $s\frac{f(\tilde{k}_t)}{\tilde{k}_t} = (\delta + n + g)$. Notice that the term $\frac{f(\tilde{k}_t)}{\tilde{k}_t}$ is the Average Product of Capital (APK) and it is decreasing over time. Figure 3.2 shows that:

Conclusion: *The closer capital per efficiency worker to its BGP value, the lower its growth rate. In the absence of shocks to preferences or technology, the growth rate of capital per efficiency workers is therefore falling over time*

Conclusion: *Once the optimal level of capital is known, real wage and real interest rate in the steady state can be pin down from the optimality conditions of the firm, given by equation 2.3.*

Figure 3.2 The Average Product of Capital



3.2 Golden Rule

Population growth and technological progress are clearly not policy variables. Nonetheless savings can be affected by policy and figure 3.1 shows that different saving ratios lead to different BGP. The policy question then is, which is the optimal saving rate (optimal BGP) in a growing economy? The best path is, by definition, the one which maximizes consumption in the steady state. Remember that consumption is given by

$$C_t = Y_t(1-s) = Y_t - sY_t,$$

and in the BGP

$$\tilde{c}^* = f(\tilde{k}_t^*) - sf(\tilde{k}_t^*).$$

But also remember that in the BGP $sf(\tilde{k}_t) = (\delta + n + g)\tilde{k}_t^*$, then we can re write

$$\tilde{c}^* = f(\tilde{k}_t^*) - (\delta + n + g)\tilde{k}_t.$$

It is possible to compute now the $\tilde{k}_t^*$ which maximizes $\tilde{c}^*$ as

$$\frac{\partial \tilde{c}^*}{\partial \tilde{k}_t^*} = f_{\tilde{k}_t^*}(\tilde{k}_t^*) - (\delta + n + g).$$

Then there exist a value of capital per efficiency worker, $\tilde{k}_t^*$, which maximizes the consumption in the BGP.

3.3 Real Wage and Real Interest Rate

The determination of real wage and interest rate, is a bit different from the neoclassical one presented in section 2.1 because of the assumption of depreciation of capital. In fact, to maximize profits firms have to take investment decisions, which means that firms have to decide about the optimal level of future capital. The inter temporal problem that firms face in *every period* is the following. Assume that at period 0 firms buy capital, and in period 1 sell it (net to depreciation); assume further that the price of the good is one. The maximizing problem then is:

$$\begin{aligned} V &= \max_{K,L,L',I} F(K, AL) - wL - I + \frac{F(K', A'L') - w'L' + (1-\delta)K'}{1+r} \\ s.t. &: K' = K(1-\delta) + I. \end{aligned}$$

The Lagrangean is given by

$$\mathcal{L} = F(K, AL) - wL - I + \frac{F(K', A'L') - w'L' + (1-\delta)K'}{1+r} - \lambda(K' - K(1-\delta) - I)$$

,

$$\begin{aligned} FOC(L) &: F_L(K, AL) - w = 0 \\ FOC(L') &: \frac{F_L(K', A'L') - w'}{1+r} = 0 \\ FOC(K') &: \frac{F_K(K', A'L') + (1-\delta)}{1+r} - \lambda = 0 \\ FOC(I) &: -1 + \lambda = 0. \end{aligned}$$

which tell us that the optimal real wage determination is static, i.e. does not depend on the constraint. On the contrary the optimal capital determination is dynamic and so is the determination of the real interest rate. The optimality conditions yielded by the FOCs are

$$\begin{aligned} w &= F_L(K, AL) \\ r &= F_K(K, AL) - \delta. \end{aligned}$$

Conclusion: *When capital depreciates the equilibrium real interest rate is equal to the marginal productivity of capital adjusted for the depreciation rate, also called net marginal productivity of capital*

3.4 Closed-form Solution with Cobb-Douglas Production Function

Let's consider a Cobb-Douglas production function: $Y_t = K_t^\alpha (A_t L_t)^{1-\alpha}$. First notice that we can easily compute the elasticity of output relative to its factors

$$E_{Y,K} = \frac{\partial F(K_t, A_t L_t)}{\partial K_t}\frac{K_t}{Y_t} = \alpha K_t^{\alpha-1}(A_t L_t)^{1-\alpha}\frac{K}{K_t^\alpha (A_t L_t)^{1-\alpha}} = \alpha,$$

which means that if K increases by 1%, Y increases by $\alpha \cdot 1\%$. Similarly we obtain that $E_{Y,K} = 1 - \alpha$. In terms of efficiency workers,

$$\begin{aligned} \frac{Y_t}{A_t L_t} &= \frac{K_t^\alpha (A_t L_t)^{1-\alpha}}{A_t L_t} \\ &= K_t^\alpha (A_t L_t)^{-\alpha} = \left(\frac{K_t}{A_t L_t}\right)^\alpha. \end{aligned}$$

The Solow model is then summarized by the following equations

$$\begin{aligned} \tilde{y}_t &= \tilde{k}_t^\alpha, \\ \dot{\tilde{k}}_t &= s\tilde{k}_t^\alpha - (\delta + n + g)\tilde{k}_t. \end{aligned}$$

The non-trivial steady state level of capital per efficiency worker is given by

$$\tilde{k}_t^* = \left(\frac{s}{\delta + n + g}\right)^{\frac{1}{1-\alpha}}.$$

or in logarithm

$$\log(\tilde{k}_t^*) = \frac{1}{1-\alpha}\log\left(\frac{s}{\delta + n + g}\right)$$

We can also compute the optimal level of saving given the golden rule

$$\begin{aligned} f_{\tilde{k}_t^*}(\tilde{k}_t^*) - (\delta + n + g) &= 0 \\ \alpha \tilde{k}_t^{*\alpha-1} - (\delta + n + g) &= 0. \end{aligned}$$

Now substituting the steady state level of capital, $\tilde{k}_t^*$, we get

$$\begin{aligned}
\alpha \left[\left(\frac{s}{(\delta+n+g)} \right)^{\frac{1}{1-\alpha}} \right]^{\alpha-1} &= \delta+n+g \\
\alpha \left(\frac{s}{(\delta+n+g)} \right)^{-\frac{1-\alpha}{1-\alpha}} &= \delta+n+g \\
s^* &= \alpha
\end{aligned}$$

Then the optimal saving rate (the one which maximizes consumption) is equal to the elasticity of output to capital.

Finally we can compute the real wage

$$\begin{aligned}
w &= F_L(K, AL) = (1-\alpha) K_t^{\alpha} (A_t L_t)^{-\alpha} A_t = (1-\alpha) K_t^{\alpha} (A_t L_t)^{-\alpha} A_t \frac{L_t}{L_t} = \\
&= (1-\alpha) \frac{K_t^{\alpha} (A_t L_t)^{1-\alpha}}{L_t} = (1-\alpha) \frac{Y_t}{L_t}
\end{aligned}$$

and the real interest rate

$$r = F_K(K, AL) - \delta = \alpha K_t^{\alpha-1} (A_t L_t)^{1-\alpha} - \delta = \alpha \tilde{k}_t^{\alpha-1} - \delta.$$

Notice that the real wage is the real wage is simply a constant fraction of income per capita.

Small Digression: Elasticity
Elasticity is a measure of the sensitivity (or responsiveness) of a dependent variable $y = f(x)$ to changes in the independent variable x. Its concept is similar to the one of derivative: the difference lays in the fact that the elasticity is computed in percentage terms. Remember (as in the derivatives box) that a change in x is defined as Δx, and the correspondent change in y is defined by $\Delta y = f(x + \Delta x) - f(x)$. The (very) basic formula to compute the coefficient of elasticity of y to changes in x is

$$E_{y,x} = \frac{\Delta_{\%} y}{\Delta_{\%} x} = \frac{\frac{f(x+\Delta x)-f(x)}{f(x)}}{\frac{\Delta x}{x}} = \frac{\frac{\Delta y}{y}}{\frac{\Delta x}{x}}$$

Nonetheless the accuracy of the elasticity given by the formula above decreases for two reasons. First, the elasticity for a good is not necessarily constant. In fact, the elasticity can vary for different levels of x and y due to its percentage nature. Second, percentage changes are not symmetric: the percentage change between any two values depends on which one is chosen as the starting value and which as the ending value (for example, if quantity demanded increases from 10 units to 15 units, the percentage change is 50%; but if quantity demanded decreases from 15 units to 10 units, the percentage change is -33.3%).
One way to avoid the accuracy problem described above is to minimize the difference between the starting and ending values of x and y. This is the approach taken in the definition of point-price elasticity, which uses differential calculus to calculate the elasticity for an infinitesimal change in price and quantity at any given value of the variables considered. First consider the definition of differential[a]

$$\begin{aligned} \Delta y &= f(x + \Delta x) - f(x) \\ &= f'(x)\Delta x. \end{aligned}$$

Now simply substitute this last expression in the basic elasticity formula to get

$$E_{y,x} = \frac{\frac{f'(x)\Delta x}{y}}{\frac{\Delta x}{x}} = f'(x)\frac{x}{y}$$

Finally, we can extend this approach to functions with several variables. Take for example the production function $Y = F(K, L)$, its elasticity with respect to K is given by

$$E_{Y,K} = \frac{\partial F(K, L)}{\partial K}\frac{K}{Y}$$

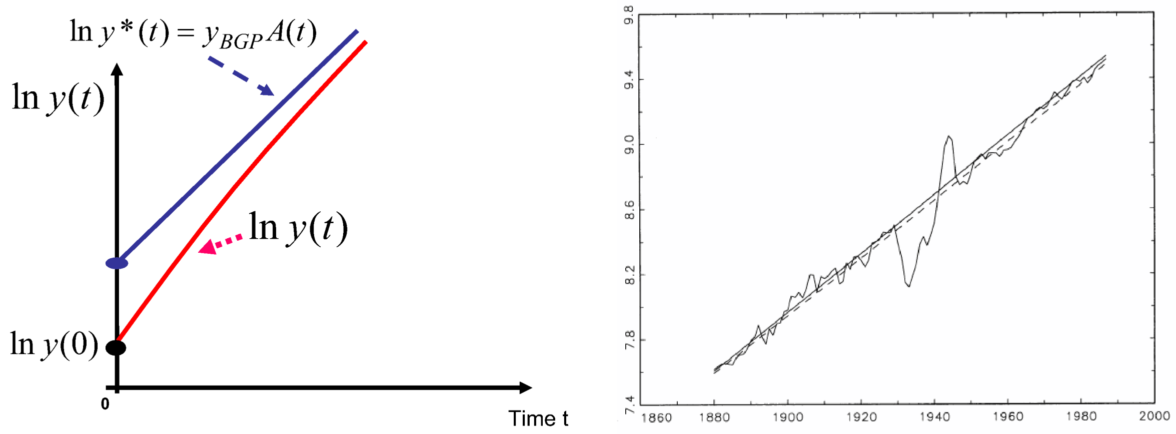
[a] For this expression to be the definition of differential the quantities Δx and Δy should tend to zero. For sake of clarity I keep the Δ notation here.

The notion of output per efficiency worker is not very intuitive. Let's think now in terms of per worker variables. For example output per worker is equal to $y_t = \frac{Y_t}{L_t} = \tilde{y}_t A_t$, and the same can be done with capital per worker. We are now interested on the growth rate of these variables (see digression on growth rates) over time and, in particular, in their BGP.

$$\begin{aligned} \hat{y}_t &= \hat{\tilde{y}}_t + \hat{A}_t = \alpha\hat{\tilde{k}}_t + g = g \\ \hat{k}_t &= \hat{\tilde{k}}_t + \hat{A}_t = g \end{aligned}$$

Then output per worker and capital per worker grow at constant rate g, the same rate of technological progress. A consequence of this result, is that the growth rate of the ratio output over capital is equal to zero.

Figure 3.3 Growth rate of output per capita: Solow prediction and US data



Finally we can compute the growth rate of real wage and real interest rate in the BGP:

$$\begin{aligned} \hat{w}_t &= \hat{y}_t \\ r_t &= (\alpha - 1)\hat{\tilde{k}}_t = 0 \end{aligned}$$

which means that, in the BGP, while real wage is growing over time at constant rate (equal to output per capita growth rate), the real rate growth rate is equal to zero. Let's draw some additional conclusions:

Conclusion: *The long-run growth rate output per worker of a country is determined by the GROWTH RATE OF LABOR EFFICIENCY ONLY. In particular, the long-run growth rate of output per worker does NOT depend on the SAVINGS RATE at all. This is because of decreasing return to capital in production. The savings rate does, however, affect the LEVEL of output per worker.*

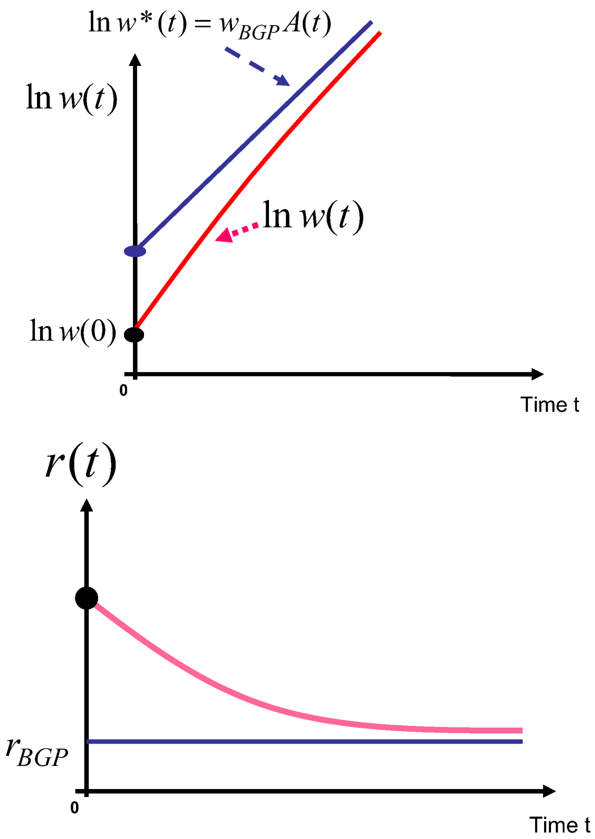
Conclusion: *The CAPITAL-OUTPUT (k/y) ratio is constant in time.*

3.5 Speed of Convergence

Given a certain shock (say a change in the saving rate) we may be interested in how rapidly the system would come back to the equilibrium. The speed of convergence to the steady state measures exactly how rapidly the system described by 3.2 comes back to the equilibrium. To compute it we have to solve the non-linear differential equation of the first order

$$\begin{aligned} \dot{\tilde{k}}_t &= \psi(\tilde{k}_t) = \\ &= s\tilde{k}_t^{\alpha} - (\delta + n + g)\tilde{k}_t. \end{aligned}$$

Figure 3.4 Growth rate of real wage and real interest rate



There are two methods to solve it

1. Transform it analytically to obtian a linear differential equation (normally is pretty complicated)
2. Linearize (or log-linearize) it around the steady state.

Once we obtain a linear differential equation, normal solutions methods can be applied. In the following I give an example of both methods

3.5.1 Analytical Solution

Notice that defining a new variable capital / output ratio

$$x_t = \frac{\tilde{k}_t}{\tilde{y}_t} = \frac{\tilde{k}_t}{\tilde{k}_t^{\alpha}} = \tilde{k}_t^{1-\alpha}$$

if we write the growth rate of x_t we get

$$\begin{aligned} \frac{\dot{x}_t}{x_t} &= \frac{(1-\alpha)\tilde{k}_t^{-\alpha} \cdot \dot{\tilde{k}}_t}{\tilde{k}_t^{1-\alpha}} = (1-\alpha)\frac{\dot{\tilde{k}}_t}{\tilde{k}_t} = \\ &= (1-\alpha)\left[s\frac{\tilde{k}_t^{\alpha}}{\tilde{k}_t} - (\delta+n+g)\right] = \\ &= (1-\alpha)\left[s\frac{1}{x_t} - (\delta+n+g)\right]. \end{aligned}$$

We now multiply both sides by x_t and we get

$$\dot{x}_t = s(1-\alpha) - (1-\alpha)(\delta+n+g)x_t$$

This is clearly a non-homogenous differential equation of the first order, where $a = (1-\alpha)(\delta + n+g)$ and $b = (1-\alpha)$. To solve it first compute the non-trivial steady state, i.e. the value of x_t such that $\dot{x}_t = 0$

$$\bar{x} = \frac{s}{(\delta+n+g)}.$$

Then define a new variable as the difference between original variable and its steady state value, $z_t = x_t - \bar{x}$. Notice that a nice feature of z_t is that $\dot{z}_t = \dot{x}_t$ which implies

$$\begin{aligned} \dot{z}_t &= \dot{x}_t = s(1-\alpha) - (1-\alpha)(\delta+n+g)x_t = \\ &= s(1-\alpha) - (1-\alpha)(\delta+n+g)(z_t+\bar{x}) = \\ &= s(1-\alpha) - (1-\alpha)(\delta+n+g)(z_t + \frac{s}{(\delta+n+g)}) = \\ &= -(1-\alpha)(\delta+n+g)z_t = \\ &= -\lambda z_t \end{aligned}$$

where $\lambda = (1-\alpha)(\delta+n+g)$. The differential equation is now homogenous, and the solution is given by

$$\begin{aligned} z_t &= e^{-\lambda t} z_0 \\ x_t - \bar{x} &= e^{-\lambda t}(x_0 - \bar{x}) \\ x_t &= e^{-\lambda t}x_0 + (1-e^{-\lambda t})\bar{x} \end{aligned}$$

The capital/output ratio is a weighted average of its steady state value and its initial value with the weight on the initial value vanishing asymptotically as $t \to \infty$. The speed of convergence to the steady state is measured by the coefficient $\lambda = -(1-\alpha)(\delta+n+g)$. The speed of convergence is faster the more concave in the production function (the closer α is to zero), or the faster capital wears out, or the faster population or technology grows.

Given that $x_t = \tilde{k}_t^{1-\alpha}$, the final solution for the capital stock is then||

$$\begin{aligned}\tilde{k}_t^{1-\alpha} &= e^{-\lambda t}x_0 + (1-e^{-\lambda t})\frac{s}{(\delta+n+g)} \\ \tilde{k}_t &= \left[e^{-\lambda t}\tilde{k}_0^{1-\alpha} + (1-e^{-\lambda t})\frac{s}{\delta+n+g}\right]^{\frac{1}{1-\alpha}}\end{aligned}$$

which inherits all the stability properties of x_t.

3.5.2 Linearization

The alternative solution is obtained by linearization around the steady state. We want to take a linear approximation of the change of capital stock, $\dot{\tilde{k}}_t$, araound its steady state which is zero. The Taylor expansion is

$$\dot{\tilde{k}}_t \simeq \dot{\tilde{k}}_{t(k=\tilde{k}_t^*)} + f'(\dot{\tilde{k}}_{t(k=\tilde{k}_t^*)})(k - \tilde{k}_t^*)$$

The first term is clearly equal to zero, $\dot{\tilde{k}}_{t(k=\tilde{k}_t^*)} = 0$. Concerning the second term we have to compute the first derivative of $\dot{\tilde{k}}_t$ and evaluate it in $\tilde{k}_t^*$

$$\begin{aligned}f'(\dot{\tilde{k}}_{t(k=\tilde{k}_t^*)}) &= \frac{\partial(sf(\tilde{k}_t^*) - (\delta+n+g)\tilde{k}_t^*)}{\partial k_t} = \\ &= sf'(\tilde{k}_t^*) - (\delta+n+g)\end{aligned}$$

and remember that in the steady state $sf(\tilde{k}_t^*) = (\delta+n+g)\tilde{k}_t^*$, which means that I can substitute for s

$$f'(\dot{\tilde{k}}_{t(k=\tilde{k}_t^*)}) = \frac{(\delta+n+g)\tilde{k}_t^*}{f(\tilde{k}_t^*)}f'(\tilde{k}_t^*) - (\delta+n+g)$$

and finally notice that $f'(\tilde{k}_t^*)\frac{\tilde{k}_t^*}{f(\tilde{k}_t^*)} = \frac{\partial f'(\tilde{k}_t)}{\partial \tilde{k}_t}\frac{\tilde{k}_t}{f'(\tilde{k}_t)}$ is the elasticity of output with respect to capital, whichfor a constant return to scale Cobb-Douglas is given by α. Then we have

$$\begin{aligned}f'(\dot{\tilde{k}}_{t(k=\tilde{k}_t^*)}) &= (\delta+n+g)\alpha - (\delta+n+g) = \\ &= -(1-\alpha)(\delta+n+g)\end{aligned}$$

We can now compute the Taylor expansion

$$\dot{\tilde{k}}_t \simeq -(1-\alpha)(\delta+n+g)(\tilde{k}_t - \tilde{k}_t^*)$$

Moreover if we defne a new variable $z_t = (\tilde{k}_t - \tilde{k}_t^*)$ with the nice property that $\dot{z}_t = \dot{\tilde{k}}_t$, then we can re-write

$$\begin{aligned}\dot{z}_t &= -(1-\alpha)(\delta+n+g)z_t = \\ &= -\lambda z_t\end{aligned}$$

which is exactly the same differential equation we had in the analytical solution.

3.6 Growth Accounting

The aggregate production function makes clear that growth in output can be written in terms of growth in inputs plus growth of efficiency

$$Y_t = K_t^{\alpha}(A_t L_t)^{1-\alpha} = L_t{}^{1-\alpha}(K_t^{\alpha} L_t{}^{1-\alpha}) = TFP_t(K_t^{\alpha} L_t{}^{1-\alpha})$$

where the efficiency factor multiplying all inputs, $L_t{}^{1-\alpha}$, is called **Total Factor Productivity (TFP)**. We can re write the above equation in terms of growth rates, to get the growth rate of the TFP

$$T\hat{F}P_t = \hat{Y}_t - \alpha\hat{K}_t - (1-\alpha)\hat{L}_t$$

Capital (K) and employment (L) are easy data to gather. To compute the TFP we only need to estimate α, the elasticity of output to capital. Remember 1) that $\alpha = \frac{\partial Y}{\partial K_t}\frac{K_t}{Y_t} = MPK\frac{K_t}{Y_t}$ and 2) that in equilibrium the marginal productivity represents the price the factor. The expression $MPK\frac{K_t}{Y_t}$ therefore represent the capital share of income, while $MPK\frac{K_t}{Y_t}$ the labor share of income.

Given that under CRTS we have the shares of labor and income sum up to one (i.e., $\alpha + (1-\alpha) = 1$), we can also write

$$\begin{aligned} T\hat{F}P_t &= \hat{Y}_t - \hat{L}_t + \alpha(\hat{K}_t - \hat{L}_t) = \\ &= \hat{y}_t + \alpha\hat{k}_t \end{aligned}$$

The TFP is also called the Solow residual.

Chapter 4

OPTIMAL GROWTH MODEL

4.1 Description of the Model

The key difference between the Solow growth model and the optimal or Ramsey-Cass-Koopmans growth model is that savings behavior is endogenized. Consider a single representative consumer that has preferences over an infinite stream of consumption $c = \{ct\}_{t=0}^{\infty}$ given by a time-separable utility function of the form

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

The number β is known as the **time discount factor** and is usually assumed to be $0 < \beta < 1$. The period utility function $U(c_t)$ is assumed to be strictly increasing and concave. An important implication of time-separability is that the the marginal utility of consumption at date t

$$\frac{\partial u(c)}{\partial ct} = \beta^t U'(c_t)$$

is independent of the level of consumption at any other date.

Remark. *In the formulation of the model presented here, we abstract from population growth and technological progress (these are easy to reinstate)*

The resource constraint on the aggregate economy is given by

$$Y = C + I \tag{4.1}$$

which means that the resource constraints facing the representative consumer are for each t

$$f(k_t) = c_t + k_{t+1} - k_t, \quad k_0 \text{ given}$$

This means that if output is not consumed, it is invested. Of course we maintain the usual constant returns and concavity assumptions for the intensive production function $f(\cdot)$. Moreover we assume the rate of physical depreciation being constant at $0 < \delta < 1$. Then the resource constraint becomes

$$f(k_t) = c_t + k_{t+1} - (1-\delta)k_t, \quad k_0 \text{ given}$$

or, more commonly,

$$c_t + k_{t+1} = f(k_t) + (1-\delta)k_t, \quad k_0 \text{ given} \tag{4.2}$$

4.2 Optimization

Let us now imagine that our agent is going to live forever. At first sight, this might seem crazy. However, we may conceive our agent as a person who cares about her future generations. An alternative interpretation of the model we are about to present, is that the optimizing agent actually is a social planner who aims at maximizing a social welfare function whose arguments are the discounted utilities of the agents who are alive now and in any possible future date.

The optimization problem is

$$\begin{array}{l} \max_{\{ct\}_{t=0}^{\infty}} u(c) \\ \text{s.t.} \quad c_t + k_{t+1} = f(k_t) + (1-\delta)k_t, \quad k_0 \text{ given.} \end{array} \tag{4.3}$$

There are plenty of methods to solve this problem. A first way to approach the consumer's inter-temporal problem is by forming a "present value" Lagrangian (method of **Lagrange multipliers)**. For each date t, let $\mu_t \geq 0$ denote the multiplier on the resource constraint. Then the Lagrangian is

$$\begin{aligned} \mathcal{L} &= \sum_{t=0}^{\infty} \beta^t U(c_t) + \sum_{t=0}^{\infty} \mu_t [f(k_t) + (1-\delta)k_t - c_t - k_{t+1}] = \qquad (4.4) \\ &= U(c_0) + \mu_0[f(k_0) + (1-\delta)k_0 - c_0 - k_1] + \beta U(c_1) + \mu_1[f(k_1) + (1-\delta)k_1 - c_1 - k_2] + ... \end{aligned}$$

To solve the problem, we must differentiate 4.4 with respect to c_t, k_{t+1}, μ_t (for $\forall t$). The first order conditions are

$$\begin{array}{ll} \text{FOC}(c_t): & \beta^t U'(c_t) - \mu_t = 0 \\ \text{FOC}(k_{t+1}): & -\mu_t + \mu_{t+1}[1 - \delta + f'(k_{t+1})] = 0 \\ \text{FOC}(\mu_t): & f(k_t) + (1-\delta)k_t - c_t - k_{t+1} = 0 \end{array} \tag{4.5}$$

The optimization of 4.3 involves an infinite number of discounted utility terms, and an infinite number of dynamic constraints; moreover the constraint concerning the final level for the stock variable is missing. The key understanding point is that the optimization between t and $t+1$ is the the same optimization between $t+s$ and $t+1+s$. These conditions are necessary, but they are not sufficient: a "final condition" is missing. In our infinite horizon model, the role of this final condition is played by the so-called **transversality condition** (henceforth tvc), which in this set up is:

$$\lim_{t\to\infty} \mu_t k_{t+1} = \lim_{t\to\infty} \beta^t U'(c_t) k_{t+1} = 0 \tag{4.6}$$

which requires that asymptotically the shadow value of more capital is zero. This is the natural infinite-horizon equivalent of the requirement that $k_T + 1 = 0$ in a model with a finite horizon T. The tvc has a clear economic interpretation: it rules out policies implying a "too fast" capital accumulation in the long run.

The FOCs can be manipulated to gain intuition of the result. Plugging the consumption condition in the capital one yields

$$\beta^t U'(c_t) = \mu_{t+1}[1 - \delta + f'(k_{t+1})]$$

and then plugging it again forwarded by one period yields

$$U'(c_t) = \beta U'(c_{t+1})[1 - \delta + f'(k_{t+1})] \tag{4.7}$$

This condition is known as the Euler equation, which is of remarkable importance not only to understand many growth models but also in consumption theory. The Euler equation tells us that an optimal consumption path must be such that -in any period- the marginal utility for consumption is equal to the following period marginal utility, discounted by and capitalized by means of the net marginal productivity of capital. To gain some intuition about the economic meaning of the Euler equation, consider the following manipulation

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

that can be interpreted as prescribing the equality between the marginal rate of substitution between period t and period $t + 1$ consumptions and the physical marginal rate of transformation.

An **equilibrium** is consumption c_t and capital k_t that solve the coupled system of non-linear difference equations

$$\begin{aligned} U'(c_t) &= \beta U'(c_{t+1})[1 - \delta + f'(k_{t+1})] \\ c_t + k_{t+1} &= f(k_t) + (1 - \delta)k_t, \end{aligned} \tag{4.8}$$

with two boundary conditions, the given initial condition k_0 and the transversality condition.

4.3 The Steady State

The steady state is given by numbers $\bar{c} = c_t = c_{t+1}$ and $\bar{k} = k_t = k_{t+1}$. Evidently, these solve

$$\begin{aligned} 1 &= \beta[1 - \delta + f'(\bar{k})] \\ \bar{c} + \bar{k} &= f(\bar{k}) + (1 - \delta)\bar{k} \end{aligned} \tag{4.9}$$

Notice from equation 4.9 that we can solve the steady state value of capital stock independently from consumption (this is an artifact of the time-separable preferences). Specifically,

$$f'(\bar{k}) = \frac{1}{\beta} - 1 + \delta = \rho + \delta$$

where the parameter $\rho = \frac{1}{\beta} - 1$ is known as the time discount rate or **real interest rate in the steady state**. Notice that ρ is such that $\beta = \frac{1}{1+\rho}$.

This condition corresponds to the one found in the real wage and real interest rate determination in the Solow model, where the real interest rate is $r = MPK - \delta$. It is important to notice that the real interest rate, which varies along with the stock of capital, may differ from the time discount rate which, instead, is fixed. Only at steady state, the capital stock is such that the **net marginal product of capital**, $f'(\bar{k}) - \delta$, is equal to the discount rate, ρ. Clearly, more patience (lower ρ) tends to increase capital accumulation and so increase $\bar{k}$. Similarly, lower depreciation δ or more capital intensity in production (higher α) raise $\bar{k}$.

Once the steady state capital stock is computed, the associated consumption level can be backed out from the resource constraint

$$\bar{c} = f(\bar{k}) - \delta\bar{k}.$$

Notice that given the resource constraint (4.1), the term $\delta\bar{k}$ corresponds to the steady state investment.

4.4 Qualitative dynamics

In this model, as it happens in many infinite horizon frameworks, it is useful to draw the phase diagram. To do this, we first consider the stability loci for each of the two variables (i.e. we compute where $\Delta c_{t+1} = c_{t+1} - c_t = 0$ and $\Delta k_{t+1} = k_{t+1} - k_t = 0$). This will help us to understand how consumption and capital change over time whenever they are not on their stability loci. The intuition behind the phase diagram is the following: consider the Euler equation and suppose

$$\begin{aligned} c_{t+1} > c_t \iff & \frac{U'(c_t)}{U'(c_{t+1})} > 1 \\ \iff & \beta[1 - \delta + f'(k_{t+1})] > 1 \\ \iff & f'(k_{t+1}) > \frac{1}{\beta} - 1 + \delta = \rho + \delta = f'(\bar{k}) \\ \iff & k_{t+1} < \bar{k} \end{aligned}$$

Whenever the capital stock will be less than its steady state value, the **real interest rate ($f'(k_{t+1}) - \delta$) will be high relative to the time discount rate** so the representative consumer will find it optimal to defer consumption so as to invest in capital accumulation thereby enjoying higher consumption tomorrow relative to today. Similarly, from the resource constraint

$$\begin{aligned} k_{t+1} > k_t &\iff k_{t+1} - k_t = f(k_t) - \delta k_t - c_t > 0 \\ &\iff f(k_t) - \delta k_t > c_t \end{aligned}$$

The capital stock grows whenever there is any output left over once consumption and depreciation have been taken out.

4.5 The Linear Approximation of Ramsey Model

The Ramsey model is described by the coupled system of non-linear difference equations in 4.8. To solve it we have to log-linearize it around its steady state. For sake of clarity, I derive the log-linear system step by step (not very efficient). You can verify that by applying directly the general result of log-linearization (given a function $z_t = f(x_t, y_t)$, its log-linear approximation is $\bar{z}_t \hat{z}_t \simeq f_x(\bar{x}, \bar{y})\bar{x}\hat{x}_t + f_y(\bar{x}, \bar{y})\bar{y}\hat{y}_t$) you get the same result.

4.5.1 Approximation of the Euler Equation

First consider the Euler equation 4.7

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

Let's start with the LHS: first I substitute the consumption with its log deviation form and then I apply a the Taylor expansion around the steady state.[1]

$$\begin{aligned} U'(\bar{c}e^{\hat{c}_t}) &\simeq U'(\bar{c}e^0) + U''(\bar{c}e^0)\bar{c}e^0(\hat{c}_t - 0) = \\ &= U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_t \end{aligned}$$

For the RHS we get

$$\begin{aligned} &\beta U'(\bar{c}e^{\hat{c}_{t+1}})[1 - \delta + f'(\bar{k}e^{\hat{k}_{t+1}})] \\ &\beta U'(\bar{c}e^0)[1 - \delta + f'(\bar{k}e^0)] + \beta U''(\bar{c}e^0)\bar{c}e^0[1 - \delta + f'(\bar{k}e^0)](\hat{c}_{t+1} - 0) + \beta U'(\bar{c}e^0)f''(\bar{k}e^0)\bar{k}e^0(\hat{k}_{t+1} - 0) \\ &\beta U'(\bar{c})\underbrace{[1 - \delta + f'(\bar{k})]}_{1/\beta \text{ from SS}} + \beta U''(\bar{c})\bar{c}\underbrace{[1 - \delta + f'(\bar{k})]}_{1/\beta \text{ from SS}}\hat{c}_{t+1} + \beta U'(\bar{c})f''(\bar{k})\bar{k}\hat{k}_{t+1} \\ &U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_{t+1} + \beta U'(\bar{c})f''(\bar{k})\bar{k}\hat{k}_{t+1} \end{aligned}$$

[1]Remember that that the Taylor expansion for a function in two variables $f(x, y)$ around a given point (x_0, y_0) is given by: $f(x, y) \simeq f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$

Finally putting together the LHS and the RHS we obtain

$$\begin{aligned}
U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_t &= U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_{t+1} + \beta U'(\bar{c})f''(\bar{k})\bar{k}\hat{k}_{t+1} \\
U''(\bar{c})\bar{c}\hat{c}_{t+1} + \beta U'(\bar{c})f''(\bar{k})\bar{k}\hat{k}_{t+1} &= U''(\bar{c})\bar{c}\hat{c}_t \\
\hat{c}_{t+1} + \frac{U'(\bar{c})}{U''(\bar{c})\bar{c}}\beta f''(\bar{k})\bar{k}\hat{k}_{t+1} &= \hat{c}_t \\
\hat{c}_{t+1} - \frac{\beta f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})}\hat{k}_{t+1} &= \hat{c}_t
\end{aligned}$$

where the number

$$\mathcal{R}(\bar{c}) = -\frac{U''(\bar{c})\bar{c}}{U'(\bar{c})}$$

is the so-called Arrow/Pratt measure of relative risk aversion (i.e., a measure of the local concavity of the utility function).

4.5.2 Approximation of the Resource Constraint

Consider the resource constraint

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

The LHS approximation is

$$\bar{k}e^{\hat{k}_{t+1}} = \bar{k}e^0 + \bar{k}e^0(\hat{k}_{t+1} - 0) = \bar{k} + \bar{k}\hat{k}_{t+1}$$

and the RHS

$$\begin{aligned}
&f(\bar{k}e^{\hat{k}_t}) + (1-\delta)\bar{k}e^{\hat{k}_t} - \bar{c}e^{\hat{c}_t} \\
&f(\bar{k}e^0) + (1-\delta)\bar{k}e^0 - \bar{c}e^0 + [f'(\bar{k}e^0)\bar{k}e^0 + (1-\delta)\bar{k}e^0](\hat{k}_t - 0) - \bar{c}(\hat{c}_t - 0) \\
&f(\bar{k}) + (1-\delta)\bar{k} - \bar{c} + [f'(\bar{k})\bar{k} + (1-\delta)\bar{k}]\hat{k}_t - \bar{c}\hat{c}_t.
\end{aligned}$$

Putting together the LHS and the RHS we get

$$\begin{aligned}
\bar{k} + \bar{k}\hat{k}_{t+1} &= \underbrace{f(\bar{k}) + (1-\delta)\bar{k} - \bar{c}}_{\bar{k}\text{ from resource constr.}} + [f'(\bar{k})\bar{k} + (1-\delta)\bar{k}]\hat{k}_t - \bar{c}\hat{c}_t \\
\bar{k}\hat{k}_{t+1} &= \underbrace{[f'(\bar{k}) + 1 - \delta]}_{1/\beta\text{ from SS}}\bar{k}\hat{k}_t - \bar{c}\hat{c}_t \\
\hat{k}_{t+1} &= \frac{1}{\beta}\hat{k}_t - \frac{\bar{c}}{\bar{k}}\hat{c}_t
\end{aligned}$$

Finally we got the log-linear approximation of the coupled system of non-linear difference equations which describes the path to equilibrium given an initial condition k_0

$$\begin{array}{c} \hat{c}_{t+1} - \frac{\beta f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})}\hat{k}_{t+1} = \hat{c}_t \\ \hat{k}_{t+1} = \frac{1}{\beta}\hat{k}_t - \frac{\bar{c}}{\bar{k}}\hat{c}_t \end{array} \tag{4.10}$$

4.6 Dynamic Properties of the Model

The path to the equilibrium of the model is then described by the system of equation 4.10. To analyze it we have to re-write in matrix form, that is

$$\begin{pmatrix} 1 & -\frac{\beta f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -\frac{\bar{c}}{\bar{k}} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \hat{c}_t \\ \hat{k}_t \end{pmatrix}$$

$$\begin{aligned} \begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} &= \begin{pmatrix} 1 & \frac{\beta f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -\frac{\bar{c}}{\bar{k}} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \hat{c}_t \\ \hat{k}_t \end{pmatrix} = \\ &= \begin{pmatrix} 1 - \frac{\beta f''(\bar{k})\bar{c}}{\mathcal{R}(\bar{c})} & \frac{f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})} \\ -\frac{\bar{c}}{\bar{k}} & \frac{1}{\beta} \end{pmatrix} \begin{pmatrix} \hat{c}_t \\ \hat{k}_t \end{pmatrix} \end{aligned}$$

To determine the stability properties of the model, we need to know the eigenvalues (λ_1, λ_2) of the 2-by-2 coefficient matrix A, where

$$A = \begin{pmatrix} 1 - \frac{\beta f''(\bar{k})\bar{c}}{\mathcal{R}(\bar{c})} & \frac{f''(\bar{k})\bar{k}}{\mathcal{R}(\bar{c})} \\ -\frac{\bar{c}}{\bar{k}} & \frac{1}{\beta} \end{pmatrix}$$

To compute its eigenvalues we need to solve

$$\begin{aligned} \det(A - \lambda I) &= 0 \\ \det \begin{pmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{pmatrix} &= 0 \end{aligned}$$

and solving for the determinant yields

$$\begin{aligned} (a_{11} - \lambda)(a_{22} - \lambda) - a_{21}a_{12} &= 0 \\ \lambda^2 - (a_{11}a_{22})\lambda + a_{11}a_{22} - a_{21}a_{12} &= 0 \end{aligned}$$

which is known as the **characteristic polynomial** $p(\lambda)$. For a 2-by-2 system, this is just a quadratic equation (higher dimensional A lead to higher dimensional polynomials, but that need not bother us here). notice also that the characteristic polynomial can be written as follows

$$p(\lambda) = \lambda^2 - tr(A)\lambda + \det(A)$$

That is, finding the eigenvalues of A just means finding the roots of $p(\lambda) = \lambda^2 - (a_{11}a_{22})\lambda + a_{11}a_{22} - a_{21}a_{12} = 0$. The solutions can be obtained with the quadratic formula.

$$\lambda_1, \lambda_2 = \frac{(a_{11}a_{22}) \pm \sqrt{(a_{11}a_{22})^2 - 4(a_{11}a_{22} - a_{21}a_{12})}}{2}$$

The complete computation is clearly painful. A shortcut to get insight of the nature, sign and magnitude of the eigenvalues is by considering an important result in linear algebra, the fact that the trace and determinant are respectively the sum and product of the eigenvalues (λ_1, λ_2). So

$$\begin{aligned} tr(A) &= \lambda_1 + \lambda_2 = 1 - \frac{\beta f''(\bar{k})\bar{c}}{\mathcal{R}(\bar{c})} + \frac{1}{\beta} > 0 \\ \det(A) &= \lambda_1 \cdot \lambda_2 = 1 \end{aligned}$$

Of course the roots may be real or complex. Generally, a square matrix has as many eigenvalues as its dimension, but some of these values may be repeated. As we know from the notes on systems of linear difference equations, this might represent a problem for the solution techniques. To know the natures of the eigenvalues we compute the discriminant of the characteristic polynomial[2]

$$\Delta = b^2 - 4ac = tr(A)^2 - 4\det(A).$$

It can be shown that $\Delta > 0$ (see Chris Edmond notes), so the eigenvalues are both real. Also, because the product of eigenvalues $\det(A)$ is positive, they must both have the same sign. Since the sum of the eigenvalues $tr(A)$ is also positive, and they are both of the same sign, both eigenvalues must individually be positive. Finally since they are not equal and their product is equal to one, it must be that $\lambda_1 > 1$ and $0 < \lambda_2 < 1$. We have formally established what we already guessed from the phase diagram: the linear system is saddle-path unstable.

4.7 Solving for the transitional dynamics

To summarize, we have a dynamic system that can be written in log-deviations as

$$\begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} = A \begin{pmatrix} \hat{c}_t \\ \hat{k}_t \end{pmatrix}$$

where A is a 2-by-2 matrix of coefficients with one unstable eigenvalue $\lambda_1 > 1$ and one stable eigenvalue $0 < \lambda_2 < 1$. As we have seen before, the square matrix A is not diagonal, meaning that there is feedback between the two equations and we can no longer solve them indepen-

[2]In algebra, the discriminant of a polynomial is an expression which gives information about the nature of the polynomial's roots. For example, the discriminant of the quadratic polynomial $ax^2 + bx + c$ is $\Delta = b^2 - 4ac$. Here, if $\Delta > 0$, the polynomial has two real roots, if $\Delta = 0$, the polynomial has one real root, and if $\Delta < 0$, the polynomial has no real roots.

dently (as we do for the univariate difference equations). In this case, the equations are said to be **coupled**.

Since solving diagonal system is easy, we "uncouple" or "diagonalize" the system through the following transformation

$$\begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} = Q \begin{pmatrix} \lambda_1^t & 0 \\ 0 & \lambda_2^t \end{pmatrix} Q^{-1} \begin{pmatrix} \hat{c}_t \\ \hat{k}_t \end{pmatrix} \tag{4.11}$$

where Q is the matrix of eigenvectors with columns corresponding to the eigenvalues and the initial condition k_0 is given. Equation 4.11 can be re-written as

$$\begin{aligned}
\begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} &= \begin{pmatrix} q_{11} & q_{12} \\ q_{21} & q_{22} \end{pmatrix} \begin{pmatrix} \lambda_1^t & 0 \\ 0 & \lambda_2^t \end{pmatrix} \frac{1}{\det(Q)} \begin{pmatrix} q_{22} & -q_{12} \\ -q_{21} & q_{11} \end{pmatrix} \begin{pmatrix} \hat{c}_0 \\ \hat{k}_0 \end{pmatrix} = \\
&= \begin{pmatrix} q_{11}\lambda_1^t & q_{12}\lambda_2^t \\ q_{21}\lambda_1^t & q_{22}\lambda_2^t \end{pmatrix} \begin{pmatrix} q_{22} & -q_{12} \\ -q_{21} & q_{11} \end{pmatrix} \frac{1}{\det(Q)} \begin{pmatrix} \hat{c}_0 \\ \hat{k}_0 \end{pmatrix} = \\
&= \begin{pmatrix} q_{11}q_{22}\lambda_1^t - q_{12}q_{21}\lambda_2^t & -q_{11}q_{12}\lambda_1^t + q_{12}q_{11}\lambda_2^t \\ q_{21}q_{22}\lambda_1^t - q_{22}q_{21}\lambda_2^t & -q_{21}q_{12}\lambda_1^t + q_{22}q_{11}\lambda_2^t \end{pmatrix} \frac{1}{\det(Q)} \begin{pmatrix} \hat{c}_0 \\ \hat{k}_0 \end{pmatrix} = \\
&= \begin{pmatrix} \hat{c}_0[q_{11}q_{22}\lambda_1^t - q_{12}q_{21}\lambda_2^t] + \hat{k}_0[-q_{11}q_{12}\lambda_1^t + q_{12}q_{11}\lambda_2^t] \\ \hat{c}_0[q_{21}q_{22}\lambda_1^t - q_{22}q_{21}\lambda_2^t] + \hat{k}_0[-q_{21}q_{12}\lambda_1^t + q_{22}q_{11}\lambda_2^t] \end{pmatrix} \frac{1}{\det(Q)} = \\
&= \begin{pmatrix} q_{11}\lambda_1^t[q_{22}\hat{c}_0 - q_{12}\hat{k}_0] + q_{12}\lambda_2^t[-q_{21}\hat{c}_0 + q_{11}\hat{k}_0] \\ q_{21}\lambda_1^t[q_{22}\hat{c}_0 - q_{12}\hat{k}_0] + q_{22}\lambda_2^t[-q_{21}\hat{c}_0 + q_{11}\hat{k}_0] \end{pmatrix} \frac{1}{\det(Q)}
\end{aligned}$$

Since $\lambda_1^t \to \infty$ as $t \to \infty$, we have an explosive behavior. Nonetheless, notice that with a system of difference equations, we need as many "initial conditions" as there are endogenous variables in order to pin down the solution completely. So far we have only specified k_0 as known parameter, therefore we still have one degree of freedom to use, that is we still have to specify c_0. Essentially, we can pick consumption so that we do not have an explosive path that violates either implicit non-negativity constraints on consumption and capital or the transversality condition. Apparently, setting

$$\hat{c}_0 = \frac{q_{12}}{q_{22}}\hat{k}_0$$

will ensure $q_{22}\hat{c}_0 - q_{12}\hat{k}_0 = 0$ and wipe out the unstable dynamics.

Hence our complete solution is

$$\begin{pmatrix} \hat{c}_{t+1} \\ \hat{k}_{t+1} \end{pmatrix} = \frac{1}{\det(Q)} \begin{pmatrix} q_{12}[q_{11} - q_{21}\frac{q_{12}}{q_{22}}] \\ q_{22}[q_{11} - q_{21}\frac{q_{12}}{q_{22}}] \end{pmatrix} \lambda_2^t \hat{k}_0 \tag{4.12}$$

In practice this means we can first compute the matrix A, then compute its eigenvalues and eigenvectors to get the matrix Q and pick the stable eigenvalue so that we have a bounded solution.

Chapter 5

STOCHASTIC DIFFERENCE EQUATIONS

We will consider two ways to model uncertainty. We will adopt the following notational conventions: random variables will be denoted by capital letters, like X_t and Z_t, realizations of random variables will be denoted by corresponding little letters, say x_t and z_t, a stochastic process will be a sequence of random variables, say $\{X_t\}$ and $\{Z_t\}$, and a sample path will be a sequence of realizations, say $\{x_t\}$ and $\{z_t\}$. If a random variable X_t has realizations that take values in some continuous set, we will model the process $\{X_t\}$ in terms of stochastic difference equations. If X_t has realizations that take values in some discrete set, we will model the process $\{X_t\}$ in terms of a Markov chain.

A stochastic process (SP) is a family of random variables $\{X_t \mid t \in T\}$ defined on a given probability space, indexed by the time variable t, where t varies over an index set T. Notice the following paralleslism: a random variable assigns a number to each outcome s in the sample space S, while a stochastic process assigns a sample function $x(t, s)$ to each outcome s. A sample function $x(t, s)$ is the time function associated with outcome of an experiment. The ensemble of a stochastic process (SP) is the set of all possible time functions that can result from an experiment.

5.1 Stochastic difference equations

Consider the following linear, stochastic difference equation of the form

$$X_{t+1} - aX_t = b + \sigma Z_{t+1}, \quad t = 0, 1, 2, ... \tag{5.1}$$

given scalars a, b, and $\sigma \geq 0$, an initial realization $X_0 = x_0$ and an exogenous stochastic process $\{Z_t\}$. Notice that the only stochastic element of equation 5.1 is the term Z_{t+1}. That is, in this model the primitive stochastic process is $\{Z_t\}$ induces a new process $\{X_t\}$: in fact if $\sigma \geq 0$ we have a difference equation that is buffeted by shocks. In econometrics, a linear stochastic difference equation is sometimes also known as an autoregression and the particular equation in (5.1) is known as an AR(1). Notice finally that if $\sigma = 0$, then we have as a special case the deterministic difference equation $X_{t+1} - aX_t = b$.

We will typically assume that each random variable Z_t is an independent standard normal so that

$$\begin{aligned} \mathsf{E}\{Z_t\} &= 0 \quad \text{for } \forall t \\ \mathsf{V}\{Z_t\} &= 1 \quad \text{for } \forall t \\ \mathsf{E}\{Z_t Z_{t-1}\} &= 0 \quad \text{for } \forall t \neq k \end{aligned}$$

The main reason for assuming normal shocks is that linear combinations of normal random variables are also normal random variables, so if the Z_t are normal, the induced random variables X_t will also be normal. To see how useful this is, recall that if Z is standard normal, then $X = \mu + \sigma Z$ is also normal with mean μ and standard deviation σ.

So, **conditional on** $X_0 = x_0$, the random variable X_1 is given by

$$X_1 = b + ax_0 + \sigma Z_1$$

with conditional mean and conditional variance given by

$$\begin{aligned} \mathsf{E}_0\{X_1\} &\equiv \mathsf{E}\{X_1 | X_0 = x_0\} = b + ax_0 \\ \mathsf{V}_0\{X1\} &\equiv \mathsf{V}\{X_1 | X_0 = x_0\} = \sigma^2 \end{aligned}$$

We can get the same result by noticing that $Z_1 \sim \mathcal{N}(0,1)$, i.e., is a standard normal. Therefore the distribution of X_1 conditional on $X_0 = x_0$ (its conditional distribution) is also normal and equal to

$$X_1 | X_0 \sim \mathcal{N}(b + ax_0, \sigma^2)$$

More generally

$$X_{t+1} | X_t \sim \mathcal{N}(b + ax_t, \sigma^2)$$

As in the deterministic case, if $|a| < 1$, we can solve the stochastic difference equation by iteration (notice the similarity with the AR processes)

$$\begin{aligned} X_t &= b + aX_{t-1} + \sigma Z_t = \\ &= b + a(b + aX_{t-2} + \sigma Z_{t-1}) + \sigma Z_t = b + ab + a^2 X_{t-2} + a\sigma Z_{t-1} + \sigma Z_t = \\ &= \ldots = \\ &= a^t x_0 + \sum_{i=0}^{t-1} a^i b + \sum_{i=0}^{t-1} a^i \sigma Z_{t-i} \end{aligned}$$

and using a well knoewn result of algebraic series

$$X_t = a^t x_0 + \frac{1-a^t}{1-a} b + \sigma \sum_{i=0}^{t-1} a^i Z_{t-i}$$

Since X_t is a linear combination of normal random variables, it is also normal with mean

$$\begin{aligned}
\mathsf{E}_0\{X_t\} &= \mathsf{E}_0 \left\{ a^t x_0 + \frac{1-a^t}{1-a} b + \sigma \sum_{i=0}^{t-1} a^i Z_{t-i} \right\} = \\
&= a^t x_0 + \frac{1-a^t}{1-a} b + \sigma \mathsf{E}_0 \left\{ \sum_{i=0}^{t-1} a^i Z_{t-i} \right\} = \\
&= a^t x_0 + \frac{1-a^t}{1-a} b + \sigma \sum_{i=0}^{t-1} a^i \mathsf{E}_0 \left[Z_{t-i} \right] = \\
&= a^t x_0 + \frac{1-a^t}{1-a} b
\end{aligned}$$

and variance

$$\begin{aligned}
\mathsf{V}_0\{X_t\} &= \mathsf{V}_0 \left\{ a^t x_0 + \frac{1-a^t}{1-a} b + \sigma \sum_{i=0}^{t-1} a^i Z_{t-i} \right\} = \\
&= \sigma^2 \mathsf{V}_0 \left\{ \sum_{i=0}^{t-1} a^i Z_{t-i} \right\} = \\
&= \sigma^2 \sum_{i=0}^{t-1} a^{2i} \mathsf{V}_0\{Z_{t-i}\} = \\
&= \sigma^2 \frac{1-a^{2t}}{1-a^2}
\end{aligned}$$

This last result uses the fact that the shocks Z_t are independent and so have zero covariances. Because of this, the variance of the sum is just the sum of the variances.[1] In short, the distribution of X_t conditional only the trivial initial realization x_0 is normal with

$$X_t | x_0 \sim \mathcal{N} \left(a^t x_0 + \frac{1-a^t}{1-a} b, \sigma^2 \frac{1-a^{2t}}{1-a^2} \right).$$

5.2 Stationary distributions

Now recall the usual stability criteria for deterministic difference equations: if $|a| < 1$, then a linear deterministic difference equation is globally stable and converges to a unique steady state, a number $\bar{x}$. With a stochastic difference equation we have no hope of finding a single steady state: in fact every state depends on the realization of the random variable, $Z_t = z_t$. Instead, we look for a **steady state distribution** or **stationary distribution** of $X_t \mid x_0$.

[1] In fact you have that $\mathsf{V}[XY] = \mathsf{V}[X] + \mathsf{V}[Y] + 2\mathsf{COV}[X, Y]$

Taking the limit as $t \to \infty$, we see that

$$\begin{aligned} \lim_{t\to\infty} \mathsf{E}_0\{X_t\} &= \frac{b}{1-a} \\ \lim_{t\to\infty} \mathsf{V}_0\{X_t\} &= \frac{\sigma^2}{1-a^2} \end{aligned}$$

So as we iterate, the dependence of the distribution on the initial realization x_0 vanishes and the distribution of X_t settles down to a normal distribution with a mean equal to the steady state of the deterministic difference equation (the "**non-stochastic steady state**") and a variance that depends on both the variance of the shocks σ^2, and the persistence of the difference equation as measured by a.

Two special cases are worthy of further comment.

5.2.1 The I.I.D. Case

If $a = 0$, then the difference equation reduces to

$$X_{t+1} = b + \sigma Z_{t+1}, \qquad t = 0, 1, 2, ...$$

In this case, we say that $\{X_t\}$ is **independent and identically distributed** (IID). The distribution of $X_t \mid x_0$ for $t > 0$ does not depend on x_0 and is just $X_t \mid x_0 \sim \mathcal{N}(b, \sigma^2)$. If $\{X_t\}$ is IID, then there is no persistence.

5.2.2 Random walk case

If $a = 1$, then the difference equation reduces to

$$X_{t+1} - X_t = b + \sigma Z_{t+1}, \qquad t = 0, 1, 2, ...$$

In this case, we say that $\{X_t\}$ is a **random walk**. If $b \neq 0$ then it is a **random walk with drift**.

If $\{X_t\}$ is a random walk with drift, its differences have a well-behaved distribution, because the difference

$$\Delta X_{t+1} = X_{t+1} - X_t = b + \sigma Z_{t+1} \sim IID$$

Therefore

$$\begin{aligned} \mathsf{E}_\mathsf{t}\{\Delta X_{t+1}\} &= b \\ \mathsf{V}_\mathsf{t}\{\Delta X_{t+1}\} &= \sigma^2 \end{aligned}$$

Then $\{X_t\}$ trends up or down over time with a constant expected change of $\mathsf{E}_\mathsf{t}\{\Delta X_{t+1}\} = b$.

Let's see however, that if $\{X_t\}$ is a random walk its **levels** do not have a well-behaved stationary distribution. Iterating in the usual manner gives

$$\begin{aligned} X_t &= b + X_{t-1} + \sigma Z_t = \\ &= b + (b + X_{t-2} + \sigma Z_{t-1}) + \sigma Z_t = b + b + X_{t-2} + \sigma Z_{t-1} + \sigma Z_t = \\ &= ... = \\ &= bt + X_0 + \sigma \sum_{i=0}^{t-1} Z_{t-i} \end{aligned}$$

So that the mean is

$$\mathsf{E}_\mathsf{t}\{X_t\} = \mathsf{E}_\mathsf{t}\left\{bt + x_0 + \sigma \sum_{i=0}^{t-1} Z_{t-i}\right\} = bt + x_0$$

and variance

$$\mathsf{V}_\mathsf{t}\{X_t\} = \mathsf{V}_\mathsf{t}\left\{bt + X_0 + \sigma \sum_{i=0}^{t-1} Z_{t-i}\right\} = \sigma \sum_{i=0}^{t-1} \mathsf{V}_\mathsf{t}\{Z_{t-i}\} = t\sigma^2$$

As you can clearly see

$$\begin{aligned} \lim_{t\to\infty} \mathsf{E}_\mathsf{t}\{X_t\} &= \pm\infty \\ \lim_{t\to\infty} \mathsf{V}_\mathsf{t}\{X_t\} &= \infty \end{aligned}$$

Conclusion. *There is a close relationship between stability for a deterministic difference equation and the existence of a stationary distribution for the associated stochastic difference equation.*

Chapter 6

STOCHASTIC GROWTH MODEL

6.1 The Model

Consider the followin maximization problem:

$$\max_{\{c_t,k_{t+1}\}_{t=0}^{\infty}} \mathsf{E}_0 \left\{ \sum_{t=0}^{\infty} \beta^t U(c_t) \right\} \tag{6.1}$$

with $0 < \beta < 1$ and subject to for $t = 1, 2, ...$

$$\begin{aligned} & c_t + k_{t+1} \leq z_t f(k_t) + (1-\delta)k_t \\ & c_t + k_{t+1} \geq 0 \\ & k_0, z_0 \text{ given} \end{aligned}$$

and an exogenous stochastic process for the level of technology $\{Z_t\}$.

The "present value" Lagrangian is

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t U(c_t) + \sum_{t=0}^{\infty} \mu_t [f(k_t) + (1-\delta)k_t - c_t - k_{t+1}] \tag{6.2}$$

and the FOCs with respect to c_t, k_{t+1}, μ_t (for $\forall t$) are

$$\begin{aligned} &\text{FOC}(c_t): && \beta^t U'(c_t) - \mu_t = 0 \\ &\text{FOC}(k_{t+1}): && -\mu_t + \mathsf{E}_0 \left\{ \mu_{t+1}[1-\delta + z_{t+1} f'(k_{t+1})] \right\} = 0 \\ &\text{FOC}(\mu_t): && f(k_t) + (1-\delta)k_t - c_t - k_{t+1} = 0 \end{aligned} \tag{6.3}$$

Combining the FOCs yields

$$U'(c_t) = \beta \mathsf{E}_0 \left\{ U'(c_{t+1})[1-\delta + z_{t+1} f'(k_{t+1})] \right\} \tag{6.4}$$

This is sometimes known as a **stochastic Euler equation**. It equates the marginal cost of consumption foregone to the expected marginal benefit, which depends on the marginal utility of consumption in the future (which is random) and on the gross return to capital (which is also random).

In addition to the stochastic Euler equation, optimal plans (i.e., the optimal path of the economy) also depend on the resource constraint. Moreover, as with the deterministic model, a transversality condition rules out capital accumulation paths that grow too quickly. To close the model we only need a specification of the law of motion for the exogenous shocks.

Once we have these equations (Euler equation, resource constraint, tvc, exogenous stochastic process) a simple way to solve this model is to log-linearize the stochastic Euler equation and the resource constraint around some point and to then solve the resulting system of stochastic difference equations. The point around which the log-linearization is often performed is the non-stochastic steady state.

6.2 Non-stochastic Steady State

Suppose that we shut off the shocks by setting $z_t = \bar{z}$ for all t and we then look for a steady state characterized by numbers $\bar{c} = c_t = c_{t+1}$ and $\bar{k} = k_t = k_{t+1}$.

$$\begin{aligned} 1 &= \beta[1-\delta+\bar{z}f'(\bar{k})] \quad (6.5) \\ \bar{c} &= \bar{z}f(\bar{k}) - \delta\bar{k} \end{aligned}$$

Given the level $\bar{z}$, the steady state capital stock is found from the first equation, while steady state consumption is backed out the resource constraint.

6.3 Log-Linearization

The **resource constraint** $c_t + k_{t+1} \leq z_t f(k_t) + (1-\delta)k_t$ can be written as $k_{t+1} = f(k_t, c_t, z_t)$. Remembering that given a function $z_t = f(x_t, y_t)$, its log-linear approximation is $\bar{z}_t\hat{z}_t \simeq f_x(\bar{x},\bar{y})\bar{x}\hat{x}_t + f_y(\bar{x},\bar{y})\bar{y}\hat{y}_t$, we can write

$$\bar{k}\hat{k}_{t+1} + \bar{c}\hat{c}_t = [z_t f'(\bar{k}) + (1-\delta)]\bar{k}\hat{k}_t + f(\bar{k})\bar{z}\hat{z}_t.$$

Let's re-write this difference equation in terms coefficients $(A, B, C, ...)$ that are known functions of the parameters of the model $(\beta, \delta, ...)$ to get

$$0 = A\hat{k}_{t+1} + B\hat{k}_t + C\hat{c}_t + D\hat{z}_t.$$

Concerning the **stochastic Euler equation** instead we proceed the normal way. We start with the LHS and, applying the transformation $c_t = \bar{c}e^{\hat{c}_t}$ we get

$$U'(\bar{c}e^{\hat{c}_t}) \simeq U'(\bar{c}e^0) + U''(\bar{c}e^0)\bar{c}e^0(\hat{c}_t - 0) = U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_t$$

The RHS is given instead by

$$\begin{aligned} & \beta \mathsf{E}_0 \left\{ U'(\bar{c}e^{\hat{c}_{t+1}})[1-\delta+\bar{z}e^{\hat{z}_{t+1}}f'(\bar{k}e^{\hat{k}_{t+1}})] \right\} \\ \simeq \;& \beta \mathsf{E}_0 \left\{ U'(\bar{c})[1-\delta+\bar{z}f'(\bar{k})] + U''(\bar{c})\bar{c}\hat{c}_{t+1}[1-\delta+\bar{z}f'(\bar{k})] + U'(\bar{c})f'(\bar{k})\bar{z}\hat{z}_{t+1} + U'(\bar{c})\bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \end{aligned}$$

and by factoring out $U'(\bar{c})$

$$\begin{aligned} & \beta_0 U'(\bar{c})\mathsf{E} \left\{ \underbrace{1-\delta+\bar{z}f'(\bar{k})}_{1/\beta} + \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1}[\underbrace{1-\delta+\bar{z}f'(\bar{k})}_{1/\beta}] + f'(\bar{k})\bar{z}\hat{z}_{t+1} + \bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \\ & \beta_0 U'(\bar{c})\mathsf{E}_0 \left\{ \frac{1}{\beta} + \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1}\frac{1}{\beta} + f'(\bar{k})\bar{z}\hat{z}_{t+1} + \bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \\ & U'(\bar{c})\mathsf{E}_0 \left\{ 1 + \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1} + \beta f'(\bar{k})\bar{z}\hat{z}_{t+1} + \beta\bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \\ & U'(\bar{c}) + U'(\bar{c})\mathsf{E}_0 \left\{ \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1} + \beta f'(\bar{k})\bar{z}\hat{z}_{t+1} + \beta\bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \end{aligned}$$

Now equate with the LHS

$$\begin{aligned} U'(\bar{c}) + U''(\bar{c})\bar{c}\hat{c}_t &= U'(\bar{c}) + U'(\bar{c})\mathsf{E}_0 \left\{ \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1} + \beta f'(\bar{k})\bar{z}\hat{z}_{t+1} + \beta\bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \\ \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_t &= \mathsf{E}_0 \left\{ \frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}\hat{c}_{t+1} + \beta f'(\bar{k})\bar{z}\hat{z}_{t+1} + \beta\bar{z}f''(\bar{k})\bar{k}\hat{k}_{t+1} \right\} \end{aligned}$$

and remembering that $\mathcal{R}(c) = -\frac{U''(\bar{c})}{U'(\bar{c})}\bar{c}$ we can write the last equation in terms of coefficient $(A, B, C, ...)$ that are known functions of the parameters of the model $(\beta, \delta, ...)$

$$0 = E_0 \left\{ F\hat{k}_{t+1} + J\hat{c}_{t+1} + K\hat{c}_t + L\hat{z}_{t+1} \right\}.$$

Finally assume that we can write the log-linear exogenous law of motion for productivity shocks as

$$\hat{z}_{t+1} = N\hat{z}_t + \varepsilon_{t+1}$$

6.4 Method of Undetermined Coefficients

From the previous section, we have the log-linear model

$$\begin{aligned} 0 &= A\hat{k}_{t+1} + B\hat{k}_t + C\hat{c}_t + D\hat{z}_t \\ 0 &= E_0 \left\{ F\hat{k}_{t+1} + J\hat{c}_{t+1} + K\hat{c}_t + L\hat{z}_{t+1} \right\} \\ \hat{z}_{t+1} &= N\hat{z}_t + \varepsilon_{t+1} \end{aligned} \tag{6.6}$$

The first equation contains the static resource constraint, the second equation contains the "forward-looking" Euler equation, while the last is the exogenous law of motion for the productivity shocks (which is independent of the others and, therefore, has its own solution). We can again try to solve this model with the method of undetermined coefficients. To do so, guess that the solutions are laws of motion

$$\begin{aligned}\hat{k}_{t+1} &= \mathbf{P}\hat{k}_t + \mathbf{Q}\hat{z}_t \\ \hat{c}_t &= \mathbf{R}\hat{k}_t + \mathbf{S}\hat{z}_t,\end{aligned}$$

for unknown coefficients $\mathbf{P}, \mathbf{Q}, \mathbf{R}, \mathbf{S}$. If we could solve for these coefficients, then together with the initial conditions $\hat{k}_0, \hat{z}_0$ and the exogenous law of motion for the shocks, $\hat{z}_{t+1} = N\hat{z}_t + \varepsilon_{t+1}$, we would have a complete solution to the linear model. We would then be in a position to do things like simulate the model, compute impulse response functions and the like. Then I plug the guesses in the system described by 6.6 and I get

$$\begin{aligned}0 &= A(\mathbf{P}\hat{k}_t + \mathbf{Q}\hat{z}_t) + B\hat{k}_t + C(\mathbf{R}\hat{k}_t + \mathbf{S}\hat{z}_t) + D\hat{z}_t \\ 0 &= E_0\left\{F(\mathbf{P}\hat{k}_t + \mathbf{Q}\hat{z}_t) + J(\mathbf{R}\hat{k}_{t+1} + \mathbf{S}\hat{z}_{t+1}) + K(\mathbf{R}\hat{k}_t + \mathbf{S}\hat{z}_t) + L\hat{z}_{t+1}\right\}\end{aligned}$$

and I re-order the terms

$$\begin{aligned}0 &= (A\mathbf{P} + B + C\mathbf{R})\hat{k}_t + (A\mathbf{Q} + C\mathbf{S} + D)\hat{z}_t \\ 0 &= E_0\left\{(F\mathbf{P} + K\mathbf{R})\hat{k}_t + J\mathbf{R}\hat{k}_{t+1} + (K\mathbf{S} + F\mathbf{Q})\hat{z}_t + (L + J\mathbf{S})\hat{z}_{t+1}\right\}\end{aligned}$$

and again, I plug in the guesses to get rid of the terms in $t+1$

$$\begin{aligned}0 &= (A\mathbf{P} + B + C\mathbf{R})\hat{k}_t + (A\mathbf{Q} + C\mathbf{S} + D)\hat{z}_t \\ 0 &= E_0\left\{(F\mathbf{P} + K\mathbf{R})\hat{k}_t + J\mathbf{R}(\mathbf{P}\hat{k}_t + \mathbf{Q}\hat{z}_t) + (K\mathbf{S} + F\mathbf{Q})\hat{z}_t + (L + J\mathbf{S})(N\hat{z}_t + \varepsilon_{t+1})\right\}\end{aligned}$$

and taking conditional expectations and grouping terms we get the complete system

$$\begin{aligned}0 &= (A\mathbf{P} + B + C\mathbf{R})\hat{k}_t + (A\mathbf{Q} + C\mathbf{S} + D)\hat{z}_t \qquad (6.7) \\ 0 &= (F\mathbf{P} + K\mathbf{R} + J\mathbf{R}\mathbf{P})\hat{k}_t + (K\mathbf{S} + F\mathbf{Q} + J\mathbf{R}\mathbf{Q} + NL + NJ\mathbf{S})\hat{z}_t\end{aligned}$$

As we noted before these solutions have to hold for any given $\hat{k}_t, \hat{z}_t$ which means that we have to impose the following set of restrictions (I denote the unknowns in bold to avoid confusion)

$$0 = A\mathbf{P} + B + C\mathbf{R} \qquad (6.8)$$

$$0 = A\mathbf{Q} + C\mathbf{S} + D \qquad (6.9)$$

from the first equation in (6.7) and

$$0 = F\mathbf{P} + K\mathbf{R} + J\mathbf{R}\mathbf{P} \qquad (6.10)$$

$$0 = K\mathbf{S} + F\mathbf{Q} + J\mathbf{R}\mathbf{Q} + NL + NJ\mathbf{S} \qquad (6.11)$$

from the second equation in (6.7). The set of restrictions defines a system of 4 equations in 4 unknowns (equations 6.8, 6.9, 6.10, and 6.11) that can be solved as follows. From equation 6.8

$$\begin{aligned}\mathbf{R} &= -C^{-1}(A\mathbf{P}+B)\\ \mathbf{S} &= -C^{-1}(A\mathbf{Q}+D)\end{aligned}$$

and plugging these into 6.10

$$\begin{aligned}0 &= F\mathbf{P}+K(-C^{-1}(A\mathbf{P}+B))+J(-C^{-1}(A\mathbf{P}+B))\mathbf{P}\\ 0 &= F\mathbf{P}-KC^{-1}A\mathbf{P}-KC^{-1}B-JC^{-1}A\mathbf{P}^2-JC^{-1}B\mathbf{P}\end{aligned}$$

which re-organized yields a quadratic equation

$$-JC^{-1}A\mathbf{P}^2+(F-KC^{-1}A-JC^{-1}B)\mathbf{P}-KC^{-1}B=0$$

Because this model has a saddle-path, this quadratic equation will have one stable and one unstable root. Since only the stable root will give a solution that satisfies the transversality condition, we pick that one. Then with **P** solved for, we know $\mathbf{R}=-C^{-1}(AP+B)$. With **P** and **R** solved for, and the expression for $\mathbf{S}=f(\mathbf{Q})$, we can plug all of them in equation 6.11 to get an equation in **Q**

$$0=FQ+JRQ-JC^{-1}(AQ+D)N-KC^{-1}(AQ+D)+LN$$

or

$$\mathbf{Q}=\frac{JC^{-1}DN+KC^{-1}D-LN}{F+JR-JC^{-1}AN-KC^{-1}A}$$

which we can compute because we have already solved for **P** and **R** and all the other coefficients were known to begin with. Finally, we can recover the remaining coefficient via $\mathbf{S}=-C^{-1}(A\mathbf{Q}+D)$. So, after all that, we have a method for solving for the unknown coefficients.

6.5 To recap

1. We solve for the non-stochastic steady state and construct the known matrices of coefficients
2. We then solve a quadratic equation for the critical **P**
3. We then back out the associated **R** so that we have all the coefficients on the endogenous state variables
4. Using **P** and **R** we solve for the coefficients on the shocks, **Q** and **S**
5. And then we're done and ready to do interesting things like simulating the model by iterating on the linear laws of motion.