INSEIKAI Tohoku Summer Camp 2023

Mathematics II: Dynamic Optimization, Financial Mathematics and Economic Modeling

TRAN Quang-Thanh (Tedd) †

ZHANG Ye (Tony) ‡ §

Graduate School of Economics and Management
Tohoku University

August 12, 2023

†tran.quang.thanh.p1@dc.tohoku.ac.jp
‡zhang.ye.q8@dc.tohoku.ac.jp
§To access the materials, visit:
Spring: https://github.com/thanhqtran/tohoku_bootcamp/tree/main/spring2023/math
Summer: https://github.com/thanhqtran/tohoku_bootcamp/tree/main/summer2023/math
Camp homepage: https://thanhqtran.github.io/tohoku_bootcamp/
Github: https://github.com/thanhqtran/tohoku_bootcamp/tree/main

Contents

Preface

Continuing from the Spring camp, this Summer Camp provides a more advanced background on the mathematical tools used in many economic models. We first briefly study difference equations and stability analysis, then review and expand the static optimization exercises. After that, we learn a different branch of literature dealing with dynamic optimization. Here, both discrete and continuous dynamic optimization methods are introduced. Finally, we study the mathematics often used in finance, such as Brownian motions in financial markets, derivative securities, and asset pricing.

Please read these notes WITH A GRAIN OF SALT as it potentially contains many errors and typos.

Syllabus

(1) Difference Equations	(2) Unconstraint Optimization
(3) Constraint Optimization	(4) Inequality Optimization
(5) Economic Applications	(6) Differential Equations
(7) Dynamic Programming	(8) Optimal Control
(9) RCK Model	(10) RBC Model
(11) Continuous OLG or DSGE	(12) Brownian motion & Martingales
(13) Stochastic calculus	(14) Derivative securities
(15) Binomial model	(16) Black-Scholes pricing model

Textbooks: These are all excellent textbooks and materials. They are self-contained and extremely valuable to your research life.

1. Sydsæter and Hammond (2008); Sydsæter et al. (2008) [accessible math]
2. Chiang (1984, 1992) [old but gold]
3. Simon and Blume (1994) [the bible in math and economic theory]
4. Shone (2002) [concise, lucid in dynamical economics]
5. de la Croix and Michel (2002) [foundations in OLG and growth]
6. Chu (2021) [a lot of good exercises on optimal control]
7. McCandless (2008) [foundations in business cycle]
8. Heer and Maussner (2009) [foundations in DSGE]
9. Campante et al. (2021) [easy to read on macroeconomic theories]
10. Acemoglu (2008) [rigor]
11. Michaillat (2023) [concise, practical with a lot of exercises]

Chapter 1

First-Order Difference Equations

In economic growth theory, in studies of the extraction of natural resources, in many models in environmental economics, and in several other areas of economics that have one key variable moves based on its past values, you will have to deal with dynamics. If we talk about dynamics, we talk about difference (in discrete time), or differential (in continuous time) equations. For the scope of this course, we are only concerned about the first-order difference equation, that is, tomorrow's value only depends on today, not including yesterday.

In preparation of the materials presented here, we reference Sydsæter and Hammond (2008); Sydsæter et al. (2008) and Chiang (1984, p.616). For differential equations, read Sydsæter et al. (2008, Chapter 5, 6), Simon and Blume (1994, Chapter 23,24,25)

1.1 One Variable Difference Equations

1.1.1 Linear case

A simple example of a linear first-order difference equation is

$$x_{t+1} = ax_t$$

for $t = 0.1, \ldots$ and a is a constant. Suppose x_0 is given, if we repeatedly applying the function for different t, we get

$$\begin{aligned}
x_1 &= ax_0, \\
x_2 &= ax_1 = a^2 x_0, \\
x_3 &= ax_2 = a^3 x_0, \ldots
\end{aligned}$$

which we can generalize it as

$$x_t = x_0 a^t$$

So that at any time t, given a_0 is known, we can calculate the current value of x_t. We can expand it by adding a constant b[(1)]. so the difference equation becomes

$$x_{t+1} = ax_t + b \tag{1.1}$$

[(1)] If $b = g(t)$, that is, the difference equation also depends on time t, then it is called non-autonomous. If the difference equation does not depend on t, then it is called autonomous which gives us the SOLUTION as follows.

$$x_t = a^t \left(x_0 - \frac{b}{1-a} \right) + \frac{b}{1-a} \qquad (a \neq 1)$$

We now discuss the notion of **stationary point**. Consider the solution of (1.1). If

$$x_0 = \frac{b}{1-a}$$

then

$$x_t = \frac{b}{1-a} \; \forall t$$

This solution $\bar{x} = b/(1-a)$ is called an equilibrium, or stationary, or steady state of (1.1). To find such an equilibrium state $\bar{x}$, we need to find $\bar{x}$ such that

$$\bar{x} = a\bar{x} + b$$

Suppose that $|a| < 1$ then

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

implying that

$$\lim_{t \to \infty} x_t = \frac{b}{1-a}$$

Hence, so long as $-1 < a < 1$, the solution converges to the equilibrium state as $t \to \infty$ and we say that the difference equation is globally asymptotically stable. But would happen otherwise?

Let us now discuss **stability** by the following illustration.

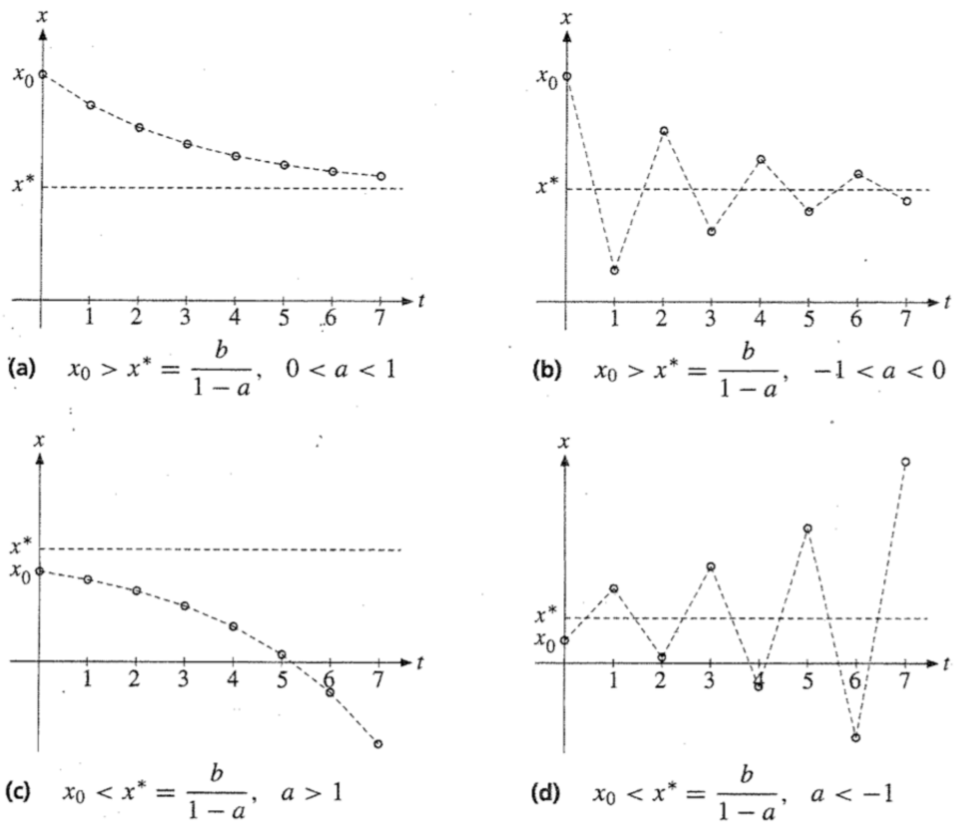


Fig. 1.1. Dynamics of stable and unstable equations (Sydsæter et al., 2008, p.393)

We summarize the cases in Fig.1.1 as follows.

(a) x_t decreases monotonically and converges to the equilibrium state x^*.

(b) x_t exhibits decreasing fluctuations, or **damped oscillations** around x^*.

(c) x_t tends to $-\infty$ monotonically (it never reaches x^*).

(d) x_t exhibits increasing fluctuations, or **explosive oscillations** around x^*. It does cross x^* at some point but never converges to it.

Thus, the condition that $|a| < 1$ is necessary to guarantee that the dynamics of x_t converge to the steady state x^*.

1.1.2 Nonlinear case

So far, we only consider linear difference equations. Although this case is easy, we almost never encounter them in economics because most dynamics in economics are nonlinear. Let us consider an autonomous first-order difference equation of one variable

$$x_t = f(x_{t-1}) \tag{1.2}$$

The procedure to find the stationary point is still the same. We still need to solve for x^* such that

$$x^* = f(x^*)$$

The solution to this is called a fixed point. How can we determine the stability of this fixed point?

It turns out, the idea behind the stability conditions stems from linear approximation. The Taylor expansion shows how the function behaves about one specific point. The behavior of x_t about this point is

$$f(x_{t-1}) = f(x^*) + f'(x^*)(x_{t-1} - x^*) + R_2(x_{t-1}, x^*)$$

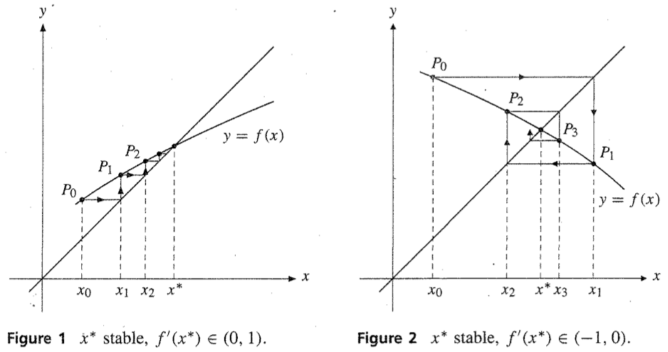
R is the remainder. If at the initial point, x_0 is sufficiently close to x^*, then $R \approx 0$. Since we are estimating the point close to x^*, we can ignore R. The point x_t about x^* can be now expressed as

$$x_t = f(x^*) + f'(x^*)(x_{t-1} - x^*).$$

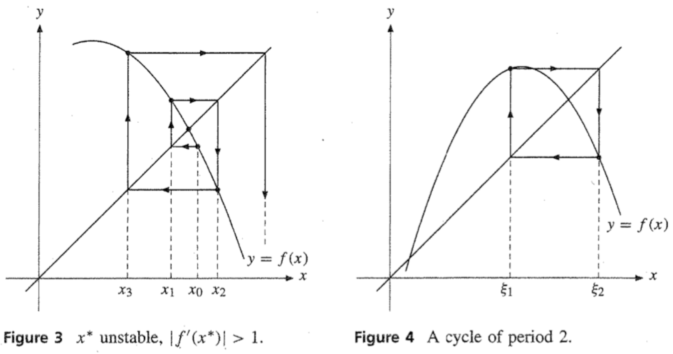
This is just similar to (1.1), so we just apply their conditions to ours, and we arrive at the following conclusion.

Theorem 1.1.1 (Stability condition)**.** *Let $\bar{x}$ be a stationary state for the difference equation $x_{t+1} = f(x_t)$, and suppose that f is differentiable in an open interval around x^*.*

1. *If $|f'(x^*)| < 1$, then x^* is **locally asymptotically stable***
2. *If $|f'(x^*)| > 1$, then x^* is **unstable***
3. *If $|f'(x^*)| = 1$, the situation is inconclusive (you will need to analyze the higher orders) – see Shone (2002, p. 89)*



(a) Stable $f'(x_t) < 1$



(b) Unstable $f'(x_t) > 1$

Fig. 1.2. Dynamics of some nonlinear equations (Sydsæter et al., 2008, p.420)

For formal proof, see Sydsæter et al. (2008, p.419)

Again, let us examine the stability illustrated in Fig.1.2.

1. Here, $0 < f'(x^*) < 1$ (derivative is positive) so the sequence monotonically converges to x^*.
2. In this case, $-1 < f'(x^*) < 0$ (derivative is negative) so x_t alternates between values above and below the equilibrium state x^*. Eventually, it converges to x^*.
3. This is the first case of unstable dynamics. Since $f'(x^*) > 1$, x is getting farther away from the equilibrium point.
4. This is a special case where solutions exhibit cyclic behavior (in this case, a cycle of period 2).

Let us explore the last case in more detail. A cycle of period 2 has the following property

$$\begin{aligned} \text{first cycle} \quad & x_0 = x_2 = x_4 = \dots, \\ \text{second cycle} \quad & x_1 = x_3 = x_5 = \dots, \\ \text{the 2 cylces are different} \quad & x_0 \neq x_1 \end{aligned}$$

This case happens if and only if there are 2 solutions for (1.2). Let us call them ξ_1, ξ_2. (pronounced /ksai/ or /zai/). Hence

$$\xi_1 = f(\xi_1),$$
$$\xi_2 = f(\xi_2).$$

If we let $F = f \circ f$, it is clear that ξ_1 and ξ_2 must be fixed points of F and hence the equilibria of the difference equation

$$y_{t+1} = F(y_t) = f(f(y_t)),$$

where $y_t = x_t x_{t+1}$. Now, we change the focus only to the stability of y_t. Applying Theorem 1.1.1, the dynamic is stable if and only if $F'(y_t) < 1$. By Chain rule

$$F'(x) = f'(f(x)) \cdot f'(x)$$

so that

$$F'(\xi_1) = f'(f(\xi_2))f'(\xi_1) = f'(\xi_2)f'(\xi_1) = F'(\xi_2).$$

Therefore, we can state

Theorem 1.1.2. *If (1.2) admits a cycle of period 2, alternating between values ξ_1 and ξ_2, then:*

1. *If $|f'(\xi_1)f'(\xi_2)| < 1$, the cycle is locally asymptotically stable.*
2. *If $|f'(\xi_1)f'(\xi_2)| > 1$, the cycle is unstable.*

1.1.3 Economic Applications

Ex. 1.1. Find the fixed point and determine their stability (Shone, 2002, p.97)

1. $y_{t+1} = 2y_t - y_t^2$,
2. $y_{t+1} = 3.2y_t - 0.8y_t^2$

Ex. 1.2 (Harrod-Domar growth)**.** Consider the following model

$$\begin{aligned} S_t &= sY_t, \\ I_t &= \nu(Y_t - Y_{t-1}), \\ S_t &= I_t \end{aligned}$$

1. Write the fundamental difference equation relating $Y_t = F(Y_{t-1})$.
2. When does the economy grow without bounds?

Ex. 1.3 (Solow Growth)**.** Consider the economy where

$$\begin{aligned} S_t &= sY_t, \\ I_t &= K_t - (1-\delta)K_{t-1}, \\ S_t &= I_t, \\ Y_t &= AK_t^{\alpha}L_t^{1-\alpha}, \\ L_t &= (1+n)L_{t-1} \end{aligned}$$

1. Define $k_t = K_t/L_t$, write the fundamental equation.
2. Derive the steady state k^*.
3. Let $A = 5, \alpha = 0.25, s = 0.1, n = 0.02, \delta = 0.05, k_0 = 20$, derive numerically.
4. Use linear approximation to investigate its stability.

1.2 System of 2 Difference Equations

Things get more complicated when there are 2 variables revolving around each other. Consider the following case system

$$\begin{aligned} x_{t+1} &= f(x_t, y_t) \\ y_{t+1} &= g(x_t, y_t) \end{aligned} \tag{1.3}$$

Functions ϕ, ψ can be linear or nonlinear, just like the case with one variable. Before diving into things, let us take a detour and do some exercises on finding the eigenvalues and eigenvectors of a matrix.

1.2.1 Eigenvalues and Eigenvectors

Definition: Given a square matrix $\mathbf{A}$, an eigenvalue of $\mathbf{A}$ is a scalar λ for which there exists a non-zero vector v such that the following equation holds:

$$\mathbf{A}v = \lambda v$$

Here, v is called an eigenvector corresponding to the eigenvalue λ. In other words, when the matrix $\mathbf{A}$ is applied to the vector v, the resulting vector is a scalar multiple of v (scaled by λ). [2]

Steps in finding the Eigenvectors and Eigenvalues

$$\begin{aligned} \mathbf{A}\mathbf{x} &= \lambda\mathbf{x} \\ \Rightarrow (\mathbf{A} - \lambda)\mathbf{x} &= 0 \\ \Rightarrow (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} &= 0 \\ \iff \det(\mathbf{A} - \lambda\mathbf{I}) &= 0 \end{aligned}$$

Solving this expression gives us the Eigenvalues λ and the eigenvector $\mathbf{x}$.

Checks when you found the eigenvalues

1. (trace) the sum of all the eigenvalues will be the sum of the diagonal of $\mathbf{A}$
2. (determinants) the product of all the eigenvalues is the determinant

The eigenvectors tell you the directions that do not change during some linear transformation, while the eigenvalues tell you the scaling vector of these eigenvectors.

Proof. Suppose $\mathbf{A}$ is a square 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Then finding the eigenvalues is to solve

$$\left| \begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \right| = \begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = 0$$

And the eigenvalues λ are the solutions of the following **Characteristic Equation**:

$$\begin{aligned} (a - \lambda)(d - \lambda) - bc &= 0 \\ \Leftrightarrow \lambda^2 - \underbrace{(a + d)}_{\beta} \lambda + \underbrace{(ad - bc)}_{\alpha} &= 0 \end{aligned} \tag{1.4}$$

[2] To understand more intuitively, visit: https://www.youtube.com/watch?v=PFDu9oVAE-g&list=PL0-GT3co4r2y2YErbmuJw2L5tW4Ew2O5B&index=14

This is called the characteristic polynomial and the eigenvalues are the roots. You can find the solutions by using quadratic formula.

Let

$$\triangle = \beta^2 - 4\alpha$$

1. if $\triangle = 0$, there is one real root $\lambda = -\dfrac{\beta}{2}$
2. if $\triangle > 0$, there are 2 real roots $\lambda_{1,2} = \dfrac{-\beta \pm \sqrt{\triangle}}{2}$.
3. if $\triangle < 0$, there are 2 complex roots $\lambda_{1,2} = \dfrac{-\beta \pm i\sqrt{|\triangle|}}{2}$ where $i = \sqrt{-1}$.

The verification process actually is a corollary of the Vieta's formulas.

$$\lambda_1 + \lambda_2 = -\beta \equiv a + d = tr(\mathbf{A}),$$
$$\lambda_1\lambda_2 = \alpha \equiv (ad - bc) = \det(\mathbf{A})$$

For each eigenvalue λ, find the corresponding eigenvector by solving the system of equations $(A - \lambda I)\mathbf{v} = \mathbf{0}$:

$$(A - \lambda \mathbf{I})\mathbf{v} = \begin{pmatrix} a-\lambda & b \\ c & d-\lambda \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

For λ_1:

$$\begin{pmatrix} a-\lambda_1 & b \\ c & d-\lambda_1 \end{pmatrix} \begin{pmatrix} v_{1,1} \\ v_{1,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

For λ_1:

$$\begin{pmatrix} a-\lambda_2 & b \\ c & d-\lambda_2 \end{pmatrix} \begin{pmatrix} v_{2,1} \\ v_{2,2} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

■

Example 1.1. Find the eigenvalues and eigenvectors of

$$\begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix}$$

Solution:

1. (eigenvalues) We need to solve

$$\begin{vmatrix} 2-\lambda & 3 \\ 5 & 4-\lambda \end{vmatrix} = 0$$

   which yields $\lambda_1 = 7$ and $\lambda_2 = -1$. These are the eigenvalues. We can do a quick verification.

   (a) (trace) $\lambda_1 + \lambda_2 = 6 = 2 + 4$ (diagonal of $\mathbf{A}$)

   (b) (determinants) $\lambda_1 \times \lambda_2 = -7 = \det \mathbf{A} (= 2 \times 4 - 5 \times 3)$.

2. (eigenvector) Let the eigenvectors be $\mathbf{v} = (v_1, v_2), \mathbf{u} = (u_1, u_2)$

   Now, for $\lambda_1 = 7$, one need to solve

$$\begin{pmatrix} 2-7 & 3 \\ 5 & 4-7 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   v_1 and v_2 are the solutions of

$$\begin{aligned} -5v_1 + 3v_2 &= 0 \\ 5v_1 - 3v_2 &= 0 \end{aligned}$$

   which gives the eigenvector $\mathbf{v} = \left(\frac{3}{5}, 1\right)$ for theeigenvalue $\lambda_1 = 7$.

   Similarly, we can find the other eigenvector $\mathbf{u} = (-1, 1)$ for the eigenvalue $\lambda_2 = -1$.

Ex. 1.4. Find the eigenvalues and verify them for the following matrices

$$\mathbf{A} = \begin{pmatrix} 3 & 8 \\ 4 & -1 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 2 & 6 \\ -1 & 3 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}, \mathbf{D} = \begin{pmatrix} 5 & 1 \\ 2 & 3 \end{pmatrix},$$

$$\mathbf{E} = \begin{pmatrix} -1 & 2 \\ 4 & -5 \end{pmatrix}, \mathbf{F} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 1 \\ 0 & 1 & 4 \end{pmatrix}, \mathbf{G} = \begin{pmatrix} 4 & -2 & 1 \\ 1 & 5 & 2 \\ 0 & 1 & 6 \end{pmatrix}$$

1.2.2 System of 2 Linear Difference Equations

We consider the linear dynamics in $\mathbb{R}^2$ as follows

$$\begin{aligned} x_{t+1} &= ax_t + by_t \\ y_{t+1} &= cx_t + dy_t \end{aligned} \tag{1.5}$$

Since we want a system involving both x and y, we assume that $b, d \neq 0$. For example

$$\begin{aligned} x_{t+1} &= 0.9x_t - 0.2y_t \\ y_{t+1} &= 0.1x_t + 0.7y_t \end{aligned} \tag{1.6}$$

We can exploit the matrix notation and write it as

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} a & b \\ c & d \end{pmatrix}}_{\mathbf{J}} \begin{pmatrix} x_t \\ y_t \end{pmatrix} \tag{1.7}$$

Yes, $\mathbf{J}$ is the Jacobian matrix. To characterize the properties of the Jacobian, we need to involve the notion of eigenvalues. But first, let's find the steady state by solving

$$\begin{aligned} \bar{x} &= a\bar{x} + b\bar{y} \\ \bar{y} &= c\bar{x} + d\bar{y} \end{aligned}$$

which can be written as

$$\begin{aligned} \bar{x} &= \frac{b}{1-a}\bar{y}, \\ \bar{y} &= \frac{d}{1-c}\bar{x} \end{aligned}$$

implying that

$$\bar{x} = \frac{bd}{(1-a)(1-c)}\bar{x}$$

Now, since $b, d \neq 0$, $\frac{bd}{(1-a)(1-c)} \neq 0$. Assume further that $a, c \neq 1$, then the only solution to the above equation is

$$\bar{x} = 0.$$

Thus, the only equilibrium point is $(\bar{x}, \bar{y}) = (0, 0)$.

In studying the stability around the steady state, the eigenvalues of a Jacobian matrix provide valuable insights into the behavior of a dynamical system near its equilibrium points. They indicate how nearby trajectories behave over time. If the real parts of the eigenvalues are negative (i.e., the absolute value of the eigenvalues are less than 1), trajectories that start near the equilibrium point will converge towards it, indicating **stability**. If the real parts are positive, trajectories will diverge, leading to instability.

Theorem 1.2.1. *Let the eigenvalues of the Jacobian matrix be $\lambda_1 \neq \lambda_2 \in \mathbb{R}$, then:*

1. *If $|\lambda_1| \leq |\lambda_2| < 1$, then*

$$\text{for all } \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}, \lim_{t\to\infty} \begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   and the steady state $(0, 0)$ is ***stable*** *in $\mathbb{R}^2$. Any value for (x_0, y_0) will lead the dynamics to the steady state. The steady state $(0, 0)$ is said to be a* ***sink****. Furthermore. assume that $|\lambda_2| > |\lambda_1|$.*

   - *if $\lambda_2 > 0$, the long run dynamics are monotonic*
   - *if $\lambda_2 > 0$, the long run dynamics are oscillating.*

2. *If $|\lambda_2| \geq |\lambda_1| > 1$, all trajectories starting from*

$$\begin{pmatrix} x_0 \\ y_0 \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

   explode*. The steady state $(0, 0)$ is* ***unstable****. It is said to be a* ***source****.*

3. *If $|\lambda_1| < 1 < |\lambda_2|$, there exists a unique direction along which the dynamics converge to $(0, 0)$. This implies that for a given x_0, there is only one value of y_0 such that the trajectory converges to the steady state. The steady state $(0, 0)$ is said to be a saddle point.*

So far, we have assume that the 2 eigenvalues λ_1, λ_2 are real and not equal each other. The following analyzes such cases.

Repeated Eigenvalues

Write them as $\lambda_1 = \lambda_2 = \lambda$, then if

1. If $|\lambda < 1|$, all trajectories converge to $(0, 0)$, which is globally stable in $\mathbb{R}^2$.
2. If $|\lambda > 1|$, all trajectories starting from $(x_0, y_0) \neq (0, 0)$ explode and $(0, 0)$ is unstable.

Complex Eigenvalues
We can write them as

$$\lambda_1 = \alpha + i\beta,$$
$$\lambda_2 = \alpha - i\beta$$

There are then two possibilities:

1. If $\alpha^2 + \beta^2 = |\lambda_1|^2 = |\lambda_2|^2 < 1$, all trajectories converge to $(0, 0)$, which is globally stable in $\mathbb{R}^2$.

2. If $\alpha^2 + \beta^2 > 1$, all trajectories starting from $(x_0, y_0) \neq (0, 0)$ explode and $(0, 0)$ is unstable.

Example 1.2. Let's work through the stability analysis for the system (1.6).
Step 1: Equilibrium Points
To find the equilibrium points, we need to solve the equations:

$$x_{t+1} = x_t$$
$$y_{t+1} = y_t$$

For the given system, the equilibrium points are found by setting each equation to its corresponding variable:

$$0.9x - 0.2y = x$$
$$0.1x + 0.7y = y$$

Solving these equations simultaneously, the equilibrium point is $(x^*, y^*) = (0, 0)$.
Step 2: Jacobian Matrix
The Jacobian matrix is given by:

$$J = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$$

Where f and g are the functions defining the system.
For the given system, we have:

$$f(x, y) = 0.9x - 0.2y$$
$$g(x, y) = 0.1x + 0.7y$$

Calculating the partial derivatives, we get:

$$J = \begin{bmatrix} 0.9 & -0.2 \\ 0.1 & 0.7 \end{bmatrix}$$

Step 3: Eigenvalues and Stability
Evaluate the Jacobian matrix at the equilibrium point $(0, 0)$:

$$J^* = \begin{bmatrix} 0.9 & -0.2 \\ 0.1 & 0.7 \end{bmatrix}$$

Calculate the eigenvalues of J^*. The characteristic equation is given by:

$$\det(J^* - \lambda I) = 0$$

Solving this equation, we find that the eigenvalues are approximately 0.8 and 0.8.
Since both eigenvalues have negative real parts, the equilibrium point $(0, 0)$ is stable.
In conclusion, the equilibrium point $(0, 0)$ for the given system is stable due to the negative real parts of the eigenvalues of the Jacobian matrix.

1.2.3 System of 2 Nonlinear Difference Equations

We focus on the nonlinear dynamics, as they are the most common in economics.

Consider the linear dynamics in $\mathbb{R}^2 \to \mathbb{R}^2$ following (1.3). Given the initial state (x_0, y_0). Assume that

$$\begin{aligned} \bar{x} &= f(\bar{x}, \bar{y}), \\ \bar{y} &= g(\bar{x}, \bar{y}) \end{aligned} \tag{1.8}$$

be the steady state $(\bar{x}, \bar{y})$ of the system (1.3). It is is locally stable if for any initial value (x_0, y_0) near enough to $(\bar{x}, \bar{y})$, the dynamics starting from (x_0, y_0) converge to $(\bar{x}, \bar{y})$.

Let us take a first-order Taylor expansion of $f(\cdot)$ around a steady state:

$$f(x, y) - f(\bar{x}, \bar{y}) \approx f'_x(\bar{x}, \bar{y})(x - \bar{x}) + f'_y(\bar{x}, \bar{y})(y - \bar{y})$$

Similarly for $g(\cdot)$:

$$g(x, y) - g(\bar{x}, \bar{y}) \approx g'_x(\bar{x}, \bar{y})(x - \bar{x}) + g'_y(\bar{x}, \bar{y})(y - \bar{y})$$

From (1.3),(1.8), we can write them in matrix form [(3)] as

$$\begin{pmatrix} x_{t+1} - \bar{x} \\ y_{t+1} - \bar{y} \end{pmatrix} = \underbrace{\begin{pmatrix} f'_x(\bar{x}, \bar{y}) & f'_y(\bar{x}, \bar{y}) \\ g'_x(\bar{x}, \bar{y}) & g'_y(\bar{x}, \bar{y}) \end{pmatrix}}_{\mathbf{J}} \begin{pmatrix} x_t - \bar{x} \\ y_t - \bar{y} \end{pmatrix} \tag{1.9}$$

where, as well all know by now, $\mathbf{J}$ is the Jacobian matrix. The system has been "linearized" and can be analyzed similarly to the linear case.

Theorem 1.2.2. *Let λ_1, λ_2 be the eigenvalues of the Jacobian matrix $\mathbf{J}$ evaluated at the steady state $(\bar{x}, \bar{y})$. Then*

1. *If $|\lambda_1| \leq |\lambda_2| < 1$, the steady state is locally stable. Any initial condition will lead the dynamics to the steady state. The steady state $(\bar{x}, \bar{y})$ is said to be a* ***sink****.*

2. *If $|\lambda_1| \geq |\lambda_2| > 1$, the steady state is unstable: for any initial condition different from the steady state, the trajectories are locally exploding. The steady state is said to be a* ***source****.*

3. *If $|\lambda_1| < 1 < |\lambda_2|$, the steady state is a saddle point. For a given initial condition on one variable, there is only one initial value of the other variable such that the trajectory converges to the steady state. Any other value for this variable would lead the trajectory to locally explode.*

When the eigenvalues are real and their moduli (absolute value) lie on the same side of 1, the steady state is also called a (stable or unstable) node.

From a practical point of view, it is often easier to use the **trace** ($\mathbf{T}$) and the **determinant**($\mathbf{D}$) of the Jacobian matrix. The results are summarized as follows.

[(3)] This is called to "linearize" around the steady state

Theorem 1.2.3. *We have*

$$\mathbf{T} = tr(\mathbf{J}) = f'_x + g'_y$$

and

$$\mathbf{D} = \det(\mathbf{J}) = f'_x g'_y - f'_y g'_x$$

Then:

1. *If* $|1 + \mathbf{D}| < |\mathbf{T}|$, *the steady state is a saddle.*
2. *If* $|1 + \mathbf{D}| > |\mathbf{T}|$ *and* $|\mathbf{D}| < 1$, *the steady state is a sink.*
3. *If* $|1 + \mathbf{D}| > |\mathbf{T}|$ *and* $\mathbf{D}| > 1$, *the steady state is a source.*

This is an important and neat result. After the linearization around the steady state, the LHS of (1.9) becomes the "amount of change" between 2 periods. Naturally, if the change converges toward zero and stays there as stability is guaranteed, it means that the steady-state sustains forever.

1.3 Economic Applications

The following exercises are from Sydsæter et al. (2008, p.395–399).

Ex. 1.5. Find the solution and equilibrium point. The last is from Shone (2002, p.139)

$$\begin{aligned}
&(a)\ x_{t+1} = 2x_t + 4,\ \ x_0 = 1,\\
&(b)\ 3x_{t+1} = x_t + 2,\ \ x_0 = 2,\\
&(c)\ 2x_{t+1} + 3x_t + 2 = 0,\ \ x_0 = -1,\\
&(d)\ x_{t+1} - x_t + 3 = 0,\ \ x_0 = 3,\\
&(e)\ x_{t+1} = 3.84x_t(1 - x_t),\ \ x_0 = 0.1 \text{ (3 cycles, use Python or Excel)}
\end{aligned}$$

Ex. 1.6 (Cobweb Model (Kaldor, 1934) [(4)]). Assume the total cost of raising q pigs is

$$C(q) = \alpha q + \beta q^2$$

Suppose there are N identical pig farms. Let the demand function for pigs be

$$D(p) = \gamma - \delta p$$

where p is the price, $\alpha, \beta, \gamma, \delta$ are positive constants. Each farmer behaves competitively and takes price as given to maximize their profit according to

$$\pi(q) = pq - C(q)$$

1. Find the quantity q^* that maximizes profit.
2. Find the Aggregate Supply $S(p)$.

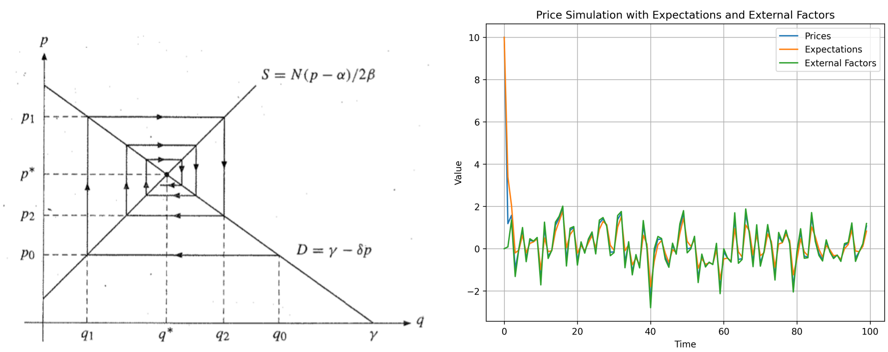
[(4)] There is no guy named "cobweb". Kaldor was the first to analyze the model and coined such a term because it looks like a web.

3. Now, suppose it takes 1 period to raise each pig. When choosing the number of pigs to raise for sale at time $t+1$, each farmer remembers the price p_t and expects p_{t+1} to be the same as p_t. Thus, aggregate supply at time $t+1$ is $S(p_t)$. Find the equilibrium price satisfying

$$S(p_t) = D(p_{t+1})$$

4. Write solution of p_t in terms of p_0 and a time path.

5. Find the equilibrium. Analyze its stability. When is it stable? When is it not?



(a) Stable cobweb (Sydsæter et al., 2008) **(b)** Inflation expectation

Fig. 1.3. The cobweb dynamics

Below are some variations of the cobweb model (Evans and Honkapohja, 2001). For numerical methods and simulation, you can visit `https://python.quantecon.org/re_with_feedback.html`.

Ex. 1.7 (Lucas (1973)'s model). Aggregate output

$$q_t = \bar{q} + \pi(p_t - p_t^e) + \zeta_t$$

where $\pi > 0$, p_t^e is the price expectation. Aggregate demand is

$$m_t + v_t = p_t + q_t$$

Money supply is random around a constant mean

$$m_t = \bar{m} + u_t$$

where all variables are in log form and u_t, v_t, ζ_t are white noise shocks. Can you achieve a reduced form of $p_t = F(p_t^e)$ [(5)].

Ex. 1.8 (Cagan (1956)'s model of hyperinflation). Demand for money depends linearly on expected inflation (the change in prices on the RHS).

$$m_t - p_t = -\psi(p_{t+1}^e - p_t), \quad \psi > 0$$

m_t, p_t, p_{t+1}^e are logs of money supply, price level and expectation of next-period price formed at time t. m_t, again, is i.i.d. around a constant mean $\bar{m}$.

Solve for p_t as a function of price expectation $F(p_{t+1}^e)$. [(6)]

[(5)] Ans: $p_t = (1+\pi)^{-1}(\bar{m} - \bar{q}) + \pi(1+\pi)^{-1}p_t^e + (1-\pi)^{-1}(u_t + v_t - \zeta_t)$
[(6)] Ans: $p_t = \alpha p_{t+1}^e + \beta m_t$ where $\alpha = \psi(1+\psi)^{-1}, \beta = (1+\psi)^{-1}$.

Chapter 2

Static Optimization and Economic Modeling

2.1 Unconstraint Optimization

Say we want to find the solutions of n choice variables ($\mathbf{x} = (x_1, \ldots, x_n)$)

$$\max_{\mathbf{x}} F(\mathbf{x})$$

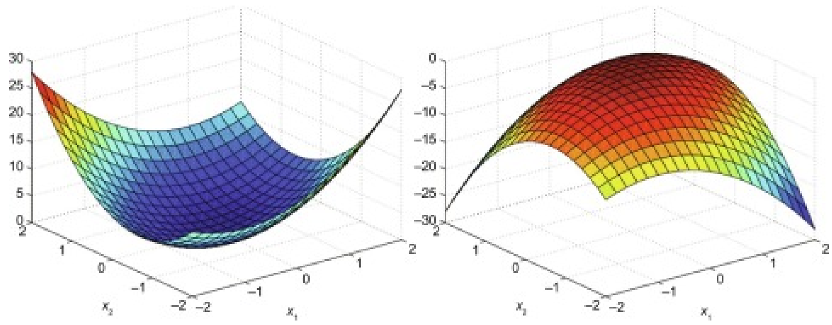


Fig. 2.1. Find the min (left) or max (right).

2.1.1 Necessary Conditions

This condition requires that the solution $\mathbf{x}^*$ must be a critical point of f, that is $f'(\mathbf{x}^*) = 0$. $\mathbf{x}^*$ will not be the endpoint of the interval under consideration, which means it lies in the INTERIOR of the domain of f.

Theorem 2.1.1. *Let $F : U \mapsto \mathbb{R}^1$ be a C^1 function defined on a subset U of $\mathbb{R}^n$. If $\mathbf{x}^*$ is a local max or min of F in U, and if $\mathbf{x}^*$ is an interior point of U, then*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, \ldots, n$$

Basically, the FOCs to every choice variable must be 0. We can write the condition in the form of Jacobian

$$DF(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial F}{\partial x_1}(\mathbf{x}^*) & \ldots & \frac{\partial F}{\partial x_n}(\mathbf{x}^*) \end{pmatrix} = \mathbf{0}$$

The solution $\mathbf{x}^* \neq 0$ for this problem is called a "non-trivial solution". Otherwise, if $\mathbf{x}^* = 0$, then it is called a "trivial" or corner solution, which is usually uninteresting in economics.

2.1.2 Sufficient Conditions

We need to use a condition on the second derivatives of F to determine whether the critical point is a max or a min. A C^2 function of n variables has n^2 second-order partial derivatives at each point in its domain. We combine them into a $n \times n$ matrix called the **Hessian** of F

$$D^2F(\mathbf{x}^*) = \begin{pmatrix} \frac{\partial^2 F}{\partial x_1^2}(\mathbf{x}^*) & \cdots & \frac{\partial^2 F}{\partial x_1 \partial x_n}(\mathbf{x}^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x_n \partial x_1}(\mathbf{x}^*) & \cdots & \frac{\partial^2 F}{\partial x_n^2}(\mathbf{x}^*) \end{pmatrix}$$

The Hessian is always a symmetric matrix. Whether the critical point is a min or max or neither depends on the definiteness of the Hessian matrix at that point.

Theorem 2.1.2. *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^n$. Suppose that $\mathbf{x}^*$ is a critical point of F, then*

1. *If the Hessian $D^2F(\mathbf{x}^*)$ is a NEGATIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MAX of F,*
2. *If the Hessian $D^2F(\mathbf{x}^*)$ is a POSITIVE DEFINITE symmetric matrix, then $\mathbf{x}^*$ is a strict LOCAL MIN of F,*
3. *If the Hessian $D^2F(\mathbf{x}^*)$ is INDEFINITE, then $\mathbf{x}^*$ is neither a local max nor a local min of F.*

In general, there are 2 methods to test for definiteness.

(1) The Signs of the Leading Minors

Theorem 2.1.3 (Sufficient Conditions for a MAX)**.** *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \text{ for } i = 1, \ldots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ alternate in sign

$$|F''_{x_1\ x_1}| < 0, \begin{vmatrix} F''_{x_1\ x_1} & F''_{x_2\ x_1} \\ F''_{x_1\ x_2} & F''_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_1\ x_3} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} < 0, \ldots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local max of F.

Theorem 2.1.4 (Sufficient Conditions for a MIN)**.** *Let $F : U \mapsto \mathbb{R}^1$ be a C^2 function whose domain is an open set U in $\mathbb{R}^1$. Suppose that*

$$\frac{\partial F}{\partial x_i}(\mathbf{x}^*) = 0 \textit{ for } i = 1, \ldots, n$$

and that the n leading principal minors of $D^2F(\mathbf{x}^*)$ <u>all positive</u>

$$|F_{x_1\ x_1}| > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} \end{vmatrix} > 0, \begin{vmatrix} F_{x_1\ x_1} & F_{x_2\ x_1} & F_{x_3\ x_1} \\ F_{x_1\ x_2} & F_{x_2\ x_2} & F_{x_3\ x_2} \\ F_{x_1\ x_3} & F_{x_2\ x_3} & F_{x_3\ x_3} \end{vmatrix} > 0, \ldots$$

at $\mathbf{x}^*$. Then $\mathbf{x}^*$ is a strict local min of F.

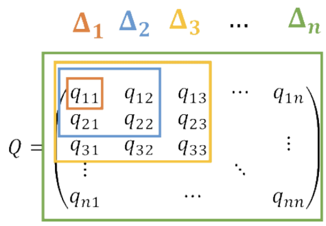


Fig. 2.2. Principal Minors

(2) The Signs of the Eigenvalues

Another way is to evaluate the Eigenvalues of the Hessian at critical points.

Theorem 2.1.5 (Eigenvalues Test for Sufficient Conditions)**.** *If the Hessian at a given point has **all positive eigenvalues**, it is said to be positive-definite, meaning the function is **concave up (convex)** at that point. If all the **eigenvalues are negative**, it is said to be a negative-definite, equivalent to **concave down**.*

To find the eigenvalues λ of a matrix $\mathbf{A}$, solve the following

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0$$

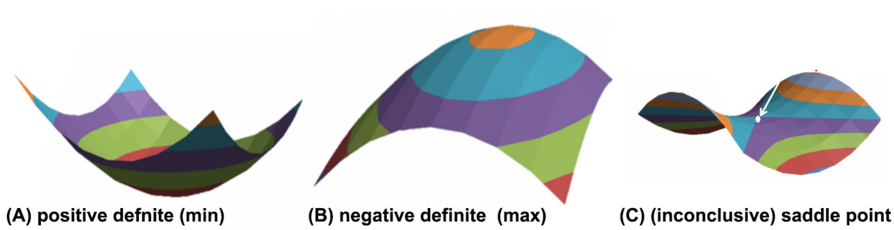


Fig. 2.3. The definiteness of the Hessian matrix

Intuitively, a negative definite Hessian matrix at the optimal point suggests that the objective function is concave in the vicinity of that point. This means that the function curves downward and resembles a bowl-like shape around the optimal point. In other words, if you move a little bit away from the optimal point in any direction, the function value will decrease. This curvature indicates that you are on the highest point in that particular region, and there is no other point nearby that can provide a higher value for the function. Mathematically, a matrix is negative definite if all its eigenvalues are negative. In the context of the Hessian matrix, the negative eigenvalues indicate that the curvature of the function in the corresponding directions is downward, which aligns with the idea of concavity.

2.1.3 Examples

Example 2.1 (Optimization)**.** Suppose

$$f(x, y) = x^4 + y^2 - xy$$

The critical point is found by

$$\begin{aligned}(x) : \frac{\partial f}{\partial x} &= 4x^3 - y = 0 \iff y = 4x^3, \\ (y) : \frac{\partial f}{\partial y} &= 2y - x = 0 \iff y = x/2.\end{aligned}$$

Solving for x yields the following critical points

$$(x^*, y^*) = (0, 0),\ (\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}}),\ (-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$$

To verify the extremum, we evaluate the Hessian matrix at the critical points

$$H = \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix}$$

The first principal minor is $12x^2$. The second principal minor is H itself where

$$|H| = 24x^2 - 1$$

Thus, we conclude that $(0, 0)$ is a saddle point.

The other 2 critical points $(\frac{1}{2\sqrt{2}}, \frac{1}{4\sqrt{2}}),\ (-\frac{1}{2\sqrt{2}}, -\frac{1}{4\sqrt{2}})$ are minima.

Example 2.2. Consider the maximization problem

$$\max\ f(x) = -x^2 + 2ax + 4a^2$$

What is the effect of an increase in a on the maximal value of $f(x, a)$?

First, Find the critical point. Taking FOC yields $x^* = a$

1. (Direct Solution) Inserting to the objective function $f(a) = 5a^2$ and so the effect of a on f is $df/da = 10a$.

2. (Envelop Theorem)

$$f'(a) = \frac{\partial f}{\partial a} = 2x + 8a$$

evaluate at $x = a$ also yields $10a$.

2.1.4 Economic Applications

Ex. 2.1. Find the optimal solution for

$$
\begin{aligned}
&(a)\ \min_x x^2 - 4x + 7,\\
&(b)\ \max_x \ln(x+1) \text{ for } x \geq 0\\
&(c)\ \max_x x^3 - 6x^2 + 9x \text{ for } x \in [-1, 4]\\
&(d)\ \max_x [800x - 2x^2] - (100 + 150x)
\end{aligned}
$$

Ex. 2.2. Find the eigenvalues and evaluate them at given points, and determine whether the matrix is negative-definite, positive-definite, or indefinite.

$$
\begin{aligned}
&(a) \begin{pmatrix} 12x^2 & -1 \\ -1 & 2 \end{pmatrix} \text{ at } (3,1),\\
&(b) \begin{pmatrix} 6x & 0 \\ 0 & 6y \end{pmatrix} \text{ at } (-1,2),\\
&(c) \begin{pmatrix} -2y^2 & -4xy \\ -4xy & -2x^2 \end{pmatrix} \text{ at } (1,-1) \text{ and } (1,0)
\end{aligned}
$$

Ex. 2.3. Find the optimal solutions for

$$
\begin{aligned}
&(a)\ \max_{x,y} f(x,y) = -2x^2 - 3y^2 + 4xy\\
&(b)\ \max_{x,y} f(x,y) = -x^3 - 2y^3 + 3xy\\
&(c)\ \min_{x,y,z} f(x,y,z) = x^2 + 6xy + y^2 - 3yz + 4z^2 - 10x - 5y - 21z.
\end{aligned}
$$

Verify with SOC condition.

Ex. 2.4. Endogenous Fertility (de la Croix, 2012) p.12.

$$\max_{s_t, n_t} \ln(w(1 - \phi n_t) - s_t) + \beta \ln((1+r)s_t) + \gamma \ln(n_t)$$

where n is fertility decision, ϕ is the time-cost of raising children, s, w, r are savings, wage rate, and interest rate.

1. Find the optimal fertility decision.
2. Add a good cost of childcare per child (say $\theta > 0$), will it change the fertility decision?

Ex. 2.5. Endogenous Fertility with Bequest (de la Croix, 2012) p.189.

$$\max_{l_t, n_t, b_t} \ln((1 - l_t - \phi N_t^\sigma n_t)k_t - n_t b_t) + \varphi \ln(l_t) + \gamma \ln(n_t k_{t+1})$$

where l_t, n_t, b_t are leisure, fertility, and educational bequests. The idea is that population size asserts a negative externality on having children.

Productive assets accumulate according to

$$k_{t+1} = \mu b_t^\eta k_t^\tau$$

1. Write the first-order conditions.
2. Find the optimal values l^*, n^*, b^*.

Ex. 2.6. Principal-Agent problem (Varian, 1992), p.453
The Agent's problem is

$$\max_a \delta + \gamma a - \frac{\gamma^2 r}{2}\sigma^2 - c(a)$$

where a is the agent's effort. The Principal's problem is

$$\max_{\delta,\gamma,a}(1-\gamma)a - \delta$$

where $\delta + \gamma a - \frac{\gamma^2 r}{2}\sigma^2 - c(a) = 0$. Let $c(a)$ be a convex function.

1. Solve the Agent's problem to obtain a
2. For the Principal's problem, first, extract δ as a function of γ, a. Then, replace it back to the objective function, use the results from 1. and solve for a.
3. Assume $c(a) = 0.5a^2$. Derive the explicit solution.

Ex. 2.7 (Maximum Likelihood Estimation (MLE))**.** Application in Statistical Inference.

1. (1 parameter) Suppose a sample $x_1, \ldots, x_n$ is modeled by a **Poisson** distribution with parameter denoted by λ, so that

$$f_X(x|\lambda) = \frac{\lambda^x}{x!}e^{-\lambda}$$

for some $\lambda > 0$. Estimate λ by MLE.

2. The **Gaussian** probability density function of a normally distributed i.i.d $N(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)^2\right]$$

In this problem, the probability density function is characterized by 2 parameters σ and μ. Use MLE to find them.

Hint: steps to estimate parameters θ using MLE.

1. Write the likelihood function

$$L_n(\theta) = \prod_{i=1}^{n} f(X_i|\theta)$$

2. Take logs to obtain log-likelihood function

$$\ell_n(\theta) = \log L_n(\theta) = \log(f(X_1|\theta)) + \cdots + \log(f(X_n|\theta)) = \sum_{i=1}^{n} \log f(X_i|\theta)$$

3. Find the estimator that maximizes this function

$$\hat{\theta} = \arg\max \ell_n(\theta)$$

2.2 Constraint Optimization

Let us consider an optimization problem for n variables with k constraints s.t.

$$\max_{\mathbf{x}}(\min) f\underbrace{(x_1,\ldots,x_n)}_{\mathbf{x}}$$
$$s.t.\ h_i(\mathbf{x}) = c_i \quad \text{for} \quad i = 1,\ldots,k.$$

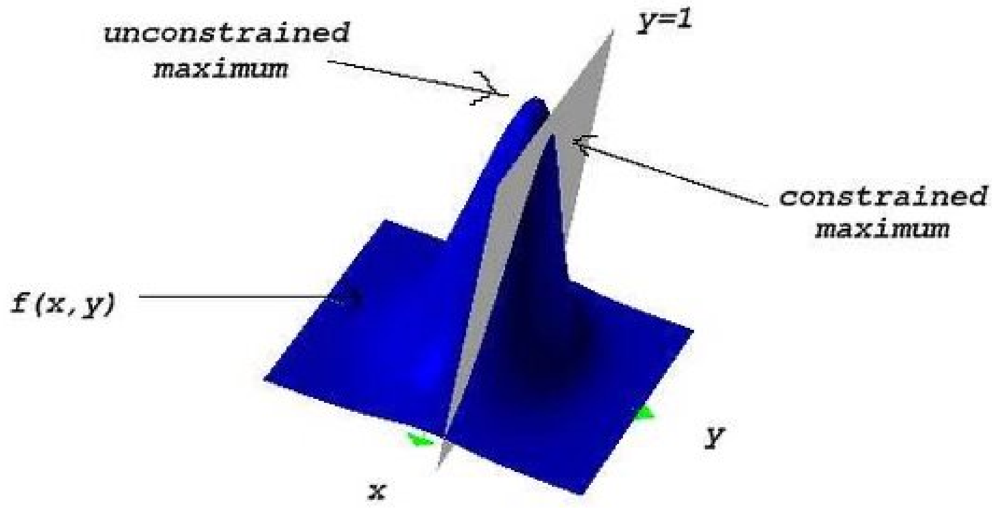


Fig. 2.4. An example of Constraint Optimization

2.2.1 Necessary First-order Conditions

Assume that NDQC is satisfied. There will be k Lagrangian multipliers λ_i for $i = 1,\ldots,k$. The Lagrangian is

$$\mathcal{L}(\mathbf{x},\lambda_i) = f(\mathbf{x}) - \sum_{i=1}^{k} \lambda_i(h_i(\mathbf{x}) - c_i)$$

As usual, the FOC is just

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0 \quad \text{for } i = 1,\ldots,n$$

Let us show what it means in a problem of 1 function of 2 variables.

Theorem 2.2.1. *Let f and h be C^1 functions of 2 variables. Suppose that $\mathbf{x}^* = (x_1^*, x_2^*)$ is a solution of the problem*

$$\max f(x_1,x_2)$$
$$s.t.\ h(x_1,x_2) = c$$

Suppose further that (x_1^*, x_2^*) is not a critical point of h. Then, there is a real number λ^* such that $(x_1^*, x_2^*, \lambda^*)$ is a critical point of the Lagrangian function

$$\mathcal{L}(x_1,x_2,\lambda) \equiv f(x_1,x_2) - \lambda[h(x_1,x_2) - c].$$

In other words, at $(x_1^*, x_2^*, \lambda^*)$ we can obtain the First-order conditions

$$\frac{\partial \mathcal{L}}{\partial x_1} = \frac{\partial \mathcal{L}}{\partial x_2} = \frac{\partial \mathcal{L}}{\partial \lambda} = 0.$$

or

$$\nabla \mathcal{L}(x_1, x_2, \lambda) = 0$$

Example 2.3. The problem is

$$\max_{x,y} \ f(x,y) = xy$$
$$s.t. \ 2x + y = 100$$

Method 1: Lagrangian Method
The Lagrangian is

$$\mathcal{L} = xy - \lambda(2x + y - 100)$$

The FOCs:

$$\begin{aligned} \mathcal{L}'_x &= y - 2\lambda = 0, \\ \mathcal{L}'_y &= x - \lambda = 0, \\ \mathcal{L}'_\lambda &= 2x + y - 100 = 0, \end{aligned}$$

which yields the solution $(x^*, y^*) = (25, 50)$.
Method 2: Substitution Method or "Naive" Method
We can turn the constrained optimization problem into an unconstrained problem. From the constraint, we have $y = 100 - x$, the problem becomes

$$\max_x x(100 - 2x)$$

The FOC is

$$100 - 4x = 0,$$

which also gives $x^* = 25, y^* = 50$.

2.2.2 Sufficient Conditions

For sufficient conditions, we need to use the notion of Bordered Hessian Matrix. The construction of such a matrix is

$$H = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & B_{11} & \dots & B_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & B_{m1} & \dots & B_{mn} \\ \hline B_{11} & \dots & B_{m1} & a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{1n} & \dots & B_{mn} & a_{1n} & \dots & a_{nn} \end{array}\right)$$

in short, it looks like this

$$H = \begin{pmatrix} 0 & \mathbf{B} \\ \mathbf{B}^T & \text{Ł} \end{pmatrix}$$

where $\mathbf{B}$ is the matrix of derivatives of constraints h_i wrt to $\mathbf{x}$, and $\mathbb{L}$ is the matrix of second-order derivatives of the Lagrangian $\mathcal{L}$ wrt to $\mathbf{x}$.

In our case, it looks like this

$$H = \left(\begin{array}{ccc|ccc} 0 & \dots & 0 & \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial h_k}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_n} \\ \hline \frac{\partial h_1}{\partial x_1} & \dots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_1^2} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_1}{\partial x_n} & \dots & \frac{\partial h_k}{\partial x_1} & \frac{\partial^2 \mathcal{L}}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 \mathcal{L}}{\partial x_n^2} \end{array}\right)$$

This $(k+n) \times (k+n)$ matrix has $k+n$ leading principal minors (the biggest one is H itself). The first m matrices $H_1, \dots, H_k$ are zero matrices. The next $k-1$ matrices $H_{k+1}, \dots, H_{2k-1}$ have zero determinant.

The determinant of the next minor H_{2k} is $\pm \det(H')^2$ where H' is the upper $k \times k$ minor of H after block of zeros, so $\det H_{2k}$ does not contain information about f.

And only the determinants of the last $n-k$ leading principal minors

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

carry information about both, the objective function f and the constraints h_i. Exactly these minors are essential for the following sufficient condition for constraint optimization.

Theorem 2.2.2 (Constraint SOC)**.** *Suppose $\mathbf{x}^*$ satisfies the FOCs.*

1. *For the bordered Hessian matrix H, the last $n-k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point ALTERNATE in signs where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^n$, then $\mathbf{x}^*$ is a LOCAL MAX.

2. *For the bordered Hessian matrix H, the last $n-k$ leading principal minors*

$$H_{2k+1}, H_{2k+2}, \dots, H_{2k+(n-k)=k+n} \equiv H$$

evaluated at the critical point have the SAME sign where the last minor $H_{n+k} \equiv H$ has the sign as $(-1)^k$, then $\mathbf{x}^*$ is a LOCAL MIN.

which can be summarized as

	H_{2k+1}	H_{2k+2}	$\dots$	H_{k+n-1}	$H_{k+n} \equiv H$
max	$(-1)^{k+1}$	$(-1)^{k+2}$	$\dots$	$(-1)^{n-1}$	$(-1)^n$
min	$(-1)^k$	$(-1)^k$	$\dots$	$(-1)^k$	$(-1)^k$

We provide here only the sufficient conditions for a problem of **2 variables and 1 constraint**, which is the most common.

Theorem 2.2.3. *Let f, h be C^2 functions on $\mathbb{R}^2$. Consider the problem*

$$\max_{x,y} f(x,y) \quad \textbf{\textit{s.t.}} \;\; h(x,y) = c \;\; \textit{for } c \in C_h \textit{(constraint set)}$$

The Lagrangian is

$$\mathcal{L}(x,y,\lambda) = f(x,y) - \lambda(h(x,y) - c)$$

Suppose that (x^*, y^*, λ^*) satisfies the following FOCs

$$\mathcal{L}'_x = \mathcal{L}'_y = \mathcal{L}'_\lambda = 0 \quad \textit{at } (x^*, y^*, \lambda^*)$$

and the bordered Hessian matrix is

$$H = \begin{pmatrix} 0 & \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} \\ \frac{\partial h}{\partial x} & \frac{\partial^2 \mathcal{L}}{\partial x^2} & \frac{\partial^2 \mathcal{L}}{\partial x \partial y} \\ \frac{\partial h}{\partial y} & \frac{\partial^2 \mathcal{L}}{\partial y \partial x} & \frac{\partial^2 \mathcal{L}}{\partial y^2} \end{pmatrix}$$

1. *if $\det(H) > 0$ at (x^*, y^*), then (x^*, y^*) is the local MAX of f on C_h.*
2. *if $\det(H) < 0$ at (x^*, y^*), then (x^*, y^*) is the local MIN of f on C_h.*

2.2.3 Examples

Example 2.4 (1 objective function of 2 variables and 1 constraint)**.** Find the extremum of

$$\begin{aligned} F(x,y) &= xy \\ s.t. \; h(x,y) &= x + y = 6. \end{aligned}$$

The Lagrangian is

$$L(x,y) = xy - \lambda(x + y - 6)$$

The FOCs are

$$\begin{aligned} (x) &: \frac{\partial L}{\partial x} = y - \lambda = 0, \\ (y) &: \frac{\partial L}{\partial y} = x - \lambda = 0, \\ (\lambda) &: \frac{\partial L}{\partial \lambda} = x + y - 6 = 0, \end{aligned}$$

which gives

$$x^* = y^* = 3, \;\; \lambda = 3$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$H = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

We have $n = 2, k = 1$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 2 > 0$ has the same sign with $(-1)^n = (-1)^2 > 0$ so our critical point is a MAX.

Example 2.5 (1 objective function of 2 variables and 2 constraint)**.** Find the extremum of

$$\begin{aligned} F(x, y, z) &= x^2 + y^2 + z^2 \\ s.t.\ h_1(x, y, z) &= 3x + y + z = 5, \\ h_2(x, y, z) &= x + y + z = 1 \end{aligned}$$

The Lagrangian is

$$L(x, y) = x^2 + y^2 + z^2 - \lambda_1(3x + y + z - 5) - \lambda_2(x + y + z - 1)$$

The FOCs are

$$\begin{aligned} (x) &: \frac{\partial L}{\partial x} = 2x - 3\lambda_1 - \lambda_2 = 0, \\ (y) &: \frac{\partial L}{\partial y} = 2y - \lambda_1 - \lambda_2 = 0, \\ (z) &: \frac{\partial L}{\partial z} = 2z - \lambda_1 - \lambda_2 = 0, \\ (\lambda_1) &: \frac{\partial L}{\partial \lambda_1} = 3x + y + z - 5 = 0, \\ (\lambda_2) &: \frac{\partial L}{\partial \lambda_2} = x + y + z - 1 = 0 \end{aligned}$$

which gives

$$x^* = 2, \quad y^* = -1/2, \quad z^* = -1/2, \quad \lambda_1 = 5/2, \quad \lambda_2 = -7/2$$

To tell the nature of its extremum, we test the second-order conditions. The bordered Hessian is

$$H = \left(\begin{array}{cc|ccc} 0 & 0 & 3 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ \hline 3 & 1 & 2 & 0 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 1 & 1 & 0 & 0 & 2 \end{array}\right)$$

We have $n = 3, k = 2$ so we have to check the $n - k = 1$ last leading principal minors, a.k.a, H itself. Calculation shows that $\det H = 16 > 0$ has the same sign with $(-1)^k = (-1)^2 > 0$ so our critical point is a MIN.

2.2.4 Economic Applications

Ex. 2.8. Find the extremum, then verify it is either max or min.

$(a)\ f(x, y) = 7 - y - x^2$	s.t	$h(x, y) = x + y = 0;$
$(b)\ f(x, y) = x(y + 4)$	s.t	$h(x, y) = x + y = 8;$
$(c)\ x_1^2 + x_2^2 + x_3^2$	s.t	$x_1 + x_2 + x_3 = 1;$
$(d)\ yz + xz$	s.t	$y^2 + z^2 = 1,\ xz = 3.$

Ex. 2.9 (Beckerian trade-off)**.** A parent's problem is

$$\begin{aligned} \max_{c_t, d_{t+1}, n_t, e_t} \quad & \ln(c_t) + \beta \ln(d_{t+1}) + \gamma \ln(\Pi_t n_t) \\ \text{s.t. } & c_t + s_t + e_t n_t = (1 - \phi n_t) w_t, \\ & d_{t+1} = (1 + r_{t+1}) s_t, \\ & \Pi_t = (\theta + e_t)^\eta \end{aligned}$$

where c, d, s, n, e are consumption when young, consumption when old, saving, number of children, and children's education. The parameters $\beta, \gamma, \phi, \theta, \eta \in (0, 1)$. This problem is a simplified version of the model in de la Croix (2012), p.22.

1. Find the optimal solutions.
2. (if time allows) verify by constructing a Hessian (you should use the Naive method)
3. Is there a trade-off between children's quantity and quality?

Ex. 2.10 (Renewable Resources)**.** Section 9.2 of Farmer and Bednar-Friedl (2010) (p.120). A country has a stock of renewable resource R_t such that

$$R_{t+1} = R_t + g(R_t) - X_t$$

where $g(R_t)$ is the rate of regeneration, while X_t is harvested stock (think of fish). We can assume a simple regenerate form

$$g(R_t) = \delta R_t - \gamma R_t^2 \text{ where } \delta > 1, \gamma < 1$$

Household's budget constraint when young is

$$c_t + k_{t+1} + p_t R_t = q_t X_t + w_t$$

LHS: expenses, including hoarding renewable resources. LHS: harvest then sell + wage. When old, his constraint is

$$d_{t+1} = (1 + r) k_{t+1} + p_{t+1} R_{t+1}$$

Utility function is $\ln(c_t) + \beta \ln(d_{t+1})$. Household's choice variables are c_t, d_{t+1}, X_t, R_t. Find the optimal solutions by forming the Lagrangian and take the FOC wrt all the choice variables.

Ex. 2.11 (New Technology)**.** Without technology, a country solves

$$\begin{aligned} & \max \log(c_0) + \beta \log(c_1), \\ & s.t. c_0 + k_1 = f(k_0), \\ & c_1 = f(k_1) \end{aligned}$$

If she invests in new technology, she solves

$$\begin{aligned} & \max \log(c_0) + \beta \log(c_1), \\ & s.t. c_0 + s_0 = f(k_0), \\ & s_0 = k_1 + \lambda k_e \\ & c_1 = h(k_e) f(k_1). \end{aligned}$$

Equation (2) means capital is used to save and make New Tech.

Let us assume $f(k) = \gamma k$, $h(x) = ax + 1$.

1. Under which condition does the country invest in new technology?
2. Under what condition, investing in the New Technology is better?

2.3 Constraint Inequality Optimization

2.3.1 KKT First-order Conditions for MAX

In this branch of problems, the constraint has inequality signs.

$$\max f(x, y) \;\; s.t. \;\; g(x, y) \leq c.$$

We solve this problem by employing the cookbook method called KKT conditions (Karush-Kuhn-Tucker).

Theorem 2.3.1 (The KKT Conditions for MAX). *Suppose we have 2 choice variables and 1 inequality constraint.*

$$\max f(x, y) \; s.t. \; g(x, y) \leq c$$

1. *Construct the Lagrangian*

$$\mathcal{L}(x, y) = f(x, y) - \lambda(g(x, y) - c)$$

2. *FOCs*

$$\begin{aligned}
&\mathcal{L}'_x = f'_x - \lambda g'_x = 0, \\
&\mathcal{L}'_y = f'_y - \lambda g'_y = 0, \\
&\lambda \cdot (g(x, y) - c) = 0, \\
&\lambda \geq 0, \\
&g(x, y) \leq c
\end{aligned}$$

3. *Complimentary slackness condition*

$$\begin{aligned}
&\lambda > 0, \;\; \textit{the constraint binds so that } g(x, y) = c \\
&\lambda = 0, \;\; \textit{the constraint does not bind so that } g(x, y) < c
\end{aligned}$$

4. *For a minimum problem, the FOCs are the same, except that* $\lambda \leq 0$.

The two inequalities $\lambda \geq 0$ and $g(x, y) \leq c$ are complementary in the sense that at most one can be "slack" – that is, at most one can hold with inequality. Equivalently, at least one must be an equality. Failure to observe that it is possible to have both $\lambda = 0$ and $g(x, y) = c$ in the complementary slackness condition is the most common error when solving nonlinear programming problems.

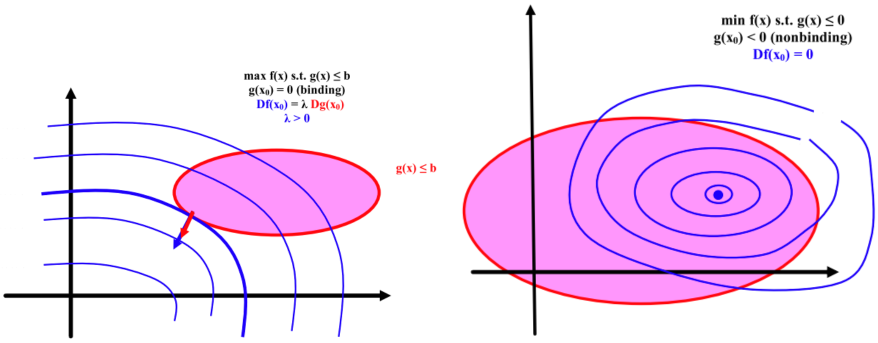
2.3.2 KKT First-order Conditions for MIN

For a minimization problem, you have 3 options

1. **Flip the sign of the objective function**, then we will turn a Minimization problem into a Maximization problem, and its FOCs follow suit.

2. Keep the constraints as is (where all constraints are $\leq$), and the FOCs are the same as the MAXIMIZATION problem **except that** $\lambda \leq 0$.

3. Flip the **signs of the constraints so that they have the form** $\geq$, then the FOCs are the same as the MAXIMIZATION problem where $\lambda \geq 0$.



(a) Binding constraint **(b)** Non-binding constraint

Fig. 2.5. Constraint binding and non-binding cases

It is easier to follow the last option.

Example 2.6.

$$\begin{aligned} \min f(x,y) &= 2y - x^2 \\ s.t.\ x^2 + y^2 &\leq 1 \end{aligned}$$

Rewrite the problem as

$$\begin{aligned} \min f(x,y) &= 2y - x^2 \\ s.t.\ \ -x^2 - y^2 &\geq -1 \end{aligned}$$

The Lagrangian is

$$L(x,y,\lambda) = 2y - x^2 - \lambda(-x^2 - y^2 + 1)$$

FOCs:

$$\begin{aligned} &(i)\ \frac{\partial L}{\partial x} = 0 \iff -2x + 2\lambda x = 0 \\ &(ii)\ \frac{\partial L}{\partial y} = 0 \iff 2 + 2\lambda y = 0 \\ &(iii)\ \lambda \cdot (-x^2 - y^2 + 1) = 0, \\ &(iv)\ \lambda \geq 0 \text{ (if } \lambda > 0\text{, constraint binds).} \end{aligned}$$

From $(i),(ii)$, we can derive $\lambda = 1, y = -1$. Since $\lambda > 0$, the constraint binds and we have $x^2 + y^2 = 1$. Since $y = -1$, we have $x = 0$, as the optimum.

2.3.3 Multiple Inequality Constraints

Consider an optimization problem of n choice variables and m inequality constraints

$$\begin{aligned} &\max f\underbrace{(x_1,\ldots,x_n)}_{\mathbf{x}} \\ &s.t. g_1(\mathbf{x}) \leq c_1, \\ &\quad \ldots, \\ &\quad g_m(\mathbf{x}) \leq c_m \end{aligned} \tag{2.1}$$

Theorem 2.3.2 (KKT Formulation)**.** *Steps in solving the problem* (2.1)

1. *Write down the Lagrangian*

$$\mathcal{L}(\mathbf{x}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - c_j)$$

2. *FOCs:*

$$\frac{\partial \mathcal{L}}{\partial x_i} = 0,$$

for each $i = 1, \ldots, n$*.*

3. *Complementary slackness*

$$\lambda_j \geq 0, g_j(\mathbf{x}) = c_j \ \textbf{\textit{or}}\ \lambda_j = 0, g_j(\mathbf{x}) < c_j$$

for $j = 1, \ldots, m$*. Can also be summarized as*

$$\lambda_j \cdot g_j(\mathbf{x}) = 0.$$

4. *Find all* $\mathbf{x} = (x_1, \ldots, x_n)$ *associated with their* $\lambda_1, \ldots, \lambda_m$ *that satisfy FOCs and the complementary slackness. These are the solution candidates, and at least 1 of them solves the problem if it has a solution.*

Example 2.7. The problem is

$$\max\ x + 3y - 4e^{-x-y}$$
$$s.t. \begin{cases} 2 - x \geq 2y \\ x - 1 \leq -y \end{cases}$$

Write the problem as

$$\max\ x + 3y - 4e^{-x-y}$$
$$s.t. \begin{cases} x + 2y \leq 2 \\ x - 1 \leq -y \end{cases}$$

The Lagrangian is

$$\mathcal{L}(x, y) = x + 3y - 4e^{-x-y} - \lambda_1(x + 2y - 2) - \lambda_2(x + y - 1)$$

KKT conditions

$$\begin{aligned}
&(i)\ \mathcal{L}'_x = 1 + 4e^{-x-y} - \lambda_1 - \lambda_2 = 0, \\
&(ii)\ \mathcal{L}'_y = 3 + 4e^{-x-y} - 2\lambda_1 - \lambda_2 = 0, \\
&\lambda_1 \cdot (x + 2y - 2) = 0, \\
&\lambda_2 \cdot (x - 1 + y) = 0, \\
&(iii)\ \lambda_1 \geq 0\ (\lambda_1 = 0 \iff x + 2y < 2), \\
&(iv)\ \lambda_2 \geq 0\ (\lambda_2 = 0 \iff x + y < 1)
\end{aligned}$$

Since $\mathcal{L}$ is concave, the KKT conditions are both necessary and sufficient for optimality.

From (ii), (i), we get $\lambda_1 = 2 > 0$, thus making (iii) binds such that $x + 2y = 2$. Suppose $\lambda_2 = 0$, there is a contradiction. Suppose $\lambda_2 > 0$, from (iv) we deduce $x+y = 1$. Using (i) and (ii), we can find $\lambda_2 = e^{-1}(4 - e) > 0$. Thus, the solution is

$$(x^*, y^*, \lambda_1, \lambda_2) = (0,\ 1, 2\ ,\ e^{-1}.(4 - e))$$

2.3.4 Nonnegativity Constraints

Most oftentimes, in economics, we want to restrict the choice variables to take nonnegative values.

Theorem 2.3.3 (Reduced KKT conditions for Nonnegativity)**.** *Consider the problem*

$$\begin{aligned} &\max\ f(x,y) \\ &s.t.\ g(x,y) \le c, \\ &\quad x \ge 0, \\ &\quad y \ge 0 \end{aligned}$$

Rewrite the problem to

$$\begin{aligned} &\max f(x,y) \\ &s.t. g(x,y) \le c, \\ &- x \le 0, \\ &- y \le 0 \end{aligned}$$

The Lagrangian is [a]

$$\mathcal{L}(x,y) = f(x,y) - \lambda[g(x,y) - c] - \mu_1(-x) - \mu_1(-y)$$

The KKT conditions

$$\begin{aligned} &(i)\ \mathcal{L}'_x = f'_x - \lambda g'_x + \mu_1 = 0 \\ &(ii)\ \mathcal{L}'_y = f'_y - \lambda g'_y + \mu_2 = 0 \\ &(iii)\ \lambda \ge 0,\ \textit{with}\ \lambda = 0\ \ \textit{if}\ \ g(x,y) < c \\ &(iv)\ \mu_1 \ge 0,\ \textit{with}\ \mu_1 = 0\ \ \textit{if}\ \ x > 0 \\ &(v)\ \mu_2 \ge 0,\ \textit{with}\ \mu_2 = 0\ \ \textit{if}\ \ y > 0 \end{aligned}$$

Combining (i) and (iv) yields

$$f'_x - \lambda g'_x \le 0,\ \ \boldsymbol{with\ equality\ if}\ x > 0$$

Combining (ii) and (v) yields

$$f'_y - \lambda g'_y \le 0,\ \ \boldsymbol{with\ equality\ if}\ y > 0$$

So the KKT conditions are reduced to just

$$f'_x - \lambda g'_x \leq 0, \ \textbf{\textit{with equality if}}\ x > 0$$
$$f'_y - \lambda g'_y \leq 0, \ \textbf{\textit{with equality if}}\ y > 0$$
$$\lambda \geq 0, \ \textit{with}\ \lambda = 0 \quad \textit{if} \quad g(x, y) < c$$

[a]tips: You should denote the Lagrangian multipliers for the main constraint by λ and nonnegativity constraints by μ for easier handling.

2.3.5 Examples

Example 2.8. Consider the utility maximization problem where there are 2 goods x, y, price of good x is p and price of good y is normalized to 1, the budget is m. Find the optimal x, y.

$$\begin{aligned} &\max x + \ln(1 + y) \\ &s.t.\ px + y \leq m, \\ &x \geq 0 \\ &y \geq 0. \end{aligned}$$

Solutions: The Lagrangian is

$$L = x + \ln(1 + y) - \lambda(px + y - m)$$

Assume the solution (x^*, y^*) exists, it must satisfy the following KKT conditions

$$\begin{aligned} &(i) && L'_x = 1 - p\lambda \leq 0, \ \textbf{with}\ 1 - p\lambda = 0 \iff x^* > 0, \\ &(ii) && L'_y = \frac{1}{1 + y^*} - \lambda \leq 0 \ \textbf{with}\ \frac{1}{1 + y^*} = 0 \iff y^* > 0 \\ &(iii) && \lambda \cdot (px^* + y^* - m) = 0, \\ &(iv) && \lambda \geq 0, px^* + y^* \leq m \end{aligned}$$

The objective function is concave in (x, y), the constraint is linear, thus the Lagrangian is concave, so the FOC is also sufficient.

Observe from condition (i) that λ cannot be zero, then condition (iii) implies that $\lambda > 0$ and the constraint binds such that

$$(iv)\ px^* + y^* = m$$

Regarding which constraints $x \geq 0, y \geq 0$ bind, we need to consider 4 cases

1. $x^* = 0, y^* = 0$: Since $m > 0$, this is impossible.
2. $x^* > 0, y^* = 0$: From (ii), we get $\lambda \geq 1$, then (i) implies that

   $$p = \frac{1}{\lambda} \leq 1$$

   Then from (iv), we have

   $$\begin{aligned} x^* &= m/p, \\ \lambda &= 1/p \end{aligned}$$

   if $0 < p < 1$.

3. $x^* = 0, y^* > 0$: By (iv), we have

$$y^* = m$$

Then (ii) yields

$$\lambda = \frac{1}{1+y^*} = \frac{1}{1+m}$$

Then from (i) we get the condition for this is that

$$p \geq m+1$$

4. $x^* > 0, y^* > 0$: With equality in both (i) and (ii), we have

$$\lambda = 1/p = 1/(1+y^*)$$

It follows that

$$\begin{aligned} &y^* = p-1, \\ &p > 1 \text{ (because } y^* > 0) \end{aligned}$$

Equation (iv) yields

$$\begin{aligned} &x^* = \frac{m+1-p}{p}, \\ &p < m+1 \text{ (because } x^* > 0) \end{aligned}$$

In summary

1. If $0 < p \leq 1$, then $(x^*, y^*) = (m/p, 0)$ with $\lambda = 1/p$
2. if $1 < p < m+1$, then $(x^*, y^*) = (\frac{m+1-p}{p}, p-1)$ with $\lambda = 1/p$
3. if $p \geq m+1$, then $(x^*, y^*) = (0, m)$ with $\lambda = 1/(1+m)$

In the 2 extreme cases (1) and (3), it is optimal to spend everything on only the cheaper good – x in case (1) and y in case (3).

2.3.6 Economic Applications

Ex. 2.12. Solve the problem

$$\begin{aligned} &\max \ f(x,y) = x^2 + y^2 + y - 1, \\ &s.t. \ g(x,y) = x^2 + y^2 \leq 1 \end{aligned}$$

Ex. 2.13 (Cost Function)**.** From Varian (1992), p.54–58.

1. Minimizing the cost function for the Cobb-Douglas technology

$$\begin{aligned} &\min_{x_1,x_2} c(\mathbf{w}, y) := w_1 x_1 + w_2 x_2, \\ &\text{s.t. } A x_1^\alpha x_2^\beta = y. \end{aligned}$$

Derive the optimal demand for x_1, x_2. Let $A = 1, \alpha + \beta = 1$, find the cost function.

2. Minimizing the cost function for CES technology

$$\min_{x_1,x_2} w_1x_1 + w_2x_2,$$
$$\text{s.t. } x_1^\rho x_2^\rho = y^\rho.$$

3. Write the cost function for Leontief technology

$$\min_{x_1,x_2} w_1x_1 + w_2x_2,$$
$$\text{s.t. } f(x_1,x_2) = \min(ax_1, bx_2) = y.$$

4. Minimizing the cost function for linear technology (using KKT).

$$\begin{aligned}&\min_{x_1,x_2} w_1x_1 + w_2x_2,\\ &\text{s.t. } ax_1 + bx_2 = y,\\ &\quad\; x_1 \geq 0,\\ &\quad\; x_2 \geq 0.\end{aligned}$$

Ex. 2.14 (Quasilinear Utility)**.** From Varian (1992), p. 164–165. Solve the problem

$$\begin{aligned}\max_{x_0,x_1} \log(x_1) + x_0,\\ p_1x_1 + x_0 = m,\\ x_0 \geq 0.\end{aligned}$$

Ex. 2.15 (Religion)**.** Based on Fan (2008); Farmer and Schelnast (2021) (p.164). Parents derive utility from consumption, child's future earnings, and religious activities ρ_t. Their problem is

$$\max \ln(c_t) + \beta \ln(w_{t+1}) + \zeta\rho_t$$

Parents do not work when old and accumulate no physical capital, the constraint is

$$c_t = (1-\rho_t)w_t$$

Human capital is

$$h_{t+1} = h_t^\alpha \rho_t^{1-\alpha}.$$

and wage is determined by human capital

$$w_{t+1} = h_{t+1}.$$

Form the problem such that $\rho_t \geq 0$.

1. What are the optimal allocations of c_t, d_{t+1}, ρ_t?

2. When will parents not spend time on religious activities?

Ex. 2.16 (Endogenous Retirement)**.** . In this case, workers are allowed to work in the second period of life (Tran, 2022). Find the solutions to

$$\max_{c_t, d_{t+1}, l_{t+1}} \ln(c_t) + \pi[\beta \ln(d_{t+1}) + \gamma \ln(l_{t+1})]$$
$$s.t.\ c_t + s_t = (1-\tau)w_t,$$
$$d_{t+1} = \frac{R_{t+1}}{\pi} s_t + (1 - l_{t+1})(1-\tau)\varepsilon w_{t+1} + p_{t+1} l_{t+1}$$
$$l_{t+1} \leq 1$$

where $\pi, l_{t+1} \in [0,1]$ are the old-age survival rate, and retirement time portion, τ, p are social security tax and pension rate. Finally, $\varepsilon \in (0,1)$ is the old-age productivity compared to young-age. The government runs a balanced budget every period such that

$$\tau w_t = p_{t+1}$$

1. Write the KKT first-order conditions
2. Derive the Euler equation
3. Assume that in the equilibrium, $l_{t+1} = 1$ is equivalent to full retirement (individuals do not work in the second period) and $l_{t+1} < 1$ is equivalent to partial retirement (individuals work for a portion of time in the second period), derive a threshold level $\hat{\varepsilon}$ that separates these 2 equilibria.

Solution: See Appendix.C.2

2.4 Economic Modeling

In previous sections, you have learned how to solve a partial equilibrium problem. That is, we have only considered one part of the equation (household OR firm). In this part, we bring them together in a equilibrium analysis (household AND firm).

2.4.1 Endogenous fertility OLG

Ex. 2.17. We first specify the model, solve the problem for each actor inside the model, then bring them together in equilibrium, then study the dynamics of a variable of interest.

Household

A representative household chooses the optimal savings and number of children to maximize their utilities as

$$\max_{s_t,n_t} U_t = \ln(c_t) + \beta \ln(R_{t+1}.s_t) + \gamma \ln(n_t),$$

subject to

$$c_t = w_t(1 - \phi n_t) - s_t,$$

given $\beta, \gamma, \phi \in (0, 1)$, $w > 0$.

1. Derive the FOCs that solve the problem for s_t, n_t.
2. Derive s^*, n^* as a function of w and other parameters.
3. Verify the SOC by checking the signs of the leading minors of the Hessian.
4. Verify the SOC by evaluating the Eigenvalues of the Hessian.

Firm

Suppose there is a representative firm of Cobb-Douglas production function

$$Y_t = K_t^{\alpha} L_t^{1-\alpha}$$

so that its profit is

$$\pi_t = p_t^Y Y_t - R_t K_t - w_t L_t$$

where input prices R_t, w_t are perfectly flexible.

1. Normalize p_t^Y to 1. Derive the optimal R_t, w_t to maximize π_t.
2. Let the capital-labor ratio as $k_t = K_t/L_t$, write w_t, R_t in terms of k_t.

Dynamics of the Intertemporal Equilibrium

The dynamics of capital and labor markets are characterized by

$$K_{t+1} = s_t L_t,$$
$$L_{t+1} = L_t(1 + n_t).$$

1. Write the law of motion of capital $k_{t+1} \equiv K_{t+1}/L_{t+1}$ in terms of k_t. Denote this equation as $k_{t+1} = \phi(k_t)$.

2. **[Existence of a Steady State]** Write the difference equation $\triangle k_t \equiv k_{t+1} - g(k_t)$. Show that $\lim_{k_t \to \infty} \triangle k_t$ and $\lim_{k_t \to 0} \triangle k_t$ have opposite signs. In that case, since $\triangle k_t$ is continuous, it must take a value of 0 somewhere, by the Intermediate Value Theorem (de la Croix and Michel, 2002) (Proposition 1.2, p.20)

3. **[The Steady State is Interior]** Also say "Existence of a nontrivial steady state" (trivial means zero). To show this, prove that

$$\text{(saving is large enough)} \lim_{k \to 0} \frac{k_{t+1}}{k_t} > 1 \text{ or } \lim_{k \to 0} \frac{s_t}{k_t} > 1 + n,$$
$$\text{(capital is bounded)} \lim_{k \to \infty} f'(k_t) = 0$$

Also known as "Absence of Catching Point (Poverty Trap)" (de la Croix and Michel, 2002) (Proposition 1.7, p.36), or Inada conditions.

4. **[Uniqueness of the Steady State]** Show that for all $k > 0$

$$\phi'(k_t) > 0, \qquad \phi''(k_t) < 0$$

de la Croix and Michel (2002) (Proposition 1.3, p.24) requires only the first condition (necessary) while Galor and Ryder (1989) demands the second condition for sufficiency. Why unique? [Hint: monotone]

5. **[Find the Steady State]** Solve for k^* by setting $k_{t+1} = k_t = k^*$

6. **[Local Stability]** Related to section 1.1.2. The steady state is locally stable if

$$\left|\phi'(k^*)\right| < 1$$

It is unstable if $|\phi'(k^*)| > 1$. If it equals 1, then the stability cannot be stated from the first-order derivative basis (de la Croix and Michel, 2002) (p.42).

7. Write a code in Python or MATLAB to simulate the model. Use the following parameters $\beta = .99^{120}, \gamma = .271, \phi = .15, \alpha = 0.3, K_0 = L_0 = 1$. What are the fertility rate, saving rate, and steady-state capital? What will happen to k^* when ϕ increases?

Remark 1. 1. The SOC check in the Household problem would be unnecessary if the constraints are linear in the choice variables (de la Croix, 2012) (p.26).

2. Normally, if you assume a Cobb-Douglas production function and a log utility function, the existence and uniqueness, and stability of a non-trivial steady state are (mostly) guaranteed.

3. In the "Existence of a nontrivial Steady State", these conditions are known as strengthened Inada condition (Galor and Ryder, 1989). The original Inada conditions are stated in the Solow model where

$$\lim_{k \to 0} f'(k_t) = \infty, \qquad \lim_{k \to \infty} f'(k_t) = 0.$$

Solutions: See Appendix.C.1.

2.4.2 Endogenous human capital OLG

Ex. 2.18. We modify Exercise 2.9 and use some assumptions from Hirazawa and Yakita (2017). In particular, we bring the model into a general equilibrium framework.

The Household sector is

$$\begin{aligned}\max_{c_t,d_{t+1},n_t,e_t} \quad & \ln(c_t) + \beta \ln(d_{t+1}) + \gamma \ln(n_t) + \gamma \ln(\theta + e_t)\\ \text{s.t. } & c_t + s_t + e_t n_t w_t = (1 - \phi n_t) w_t h_t,\\ & d_{t+1} = (1 + r_{t+1}) s_t,\\ & h_{t+1} = 1 + \mu e_t^{\eta}\end{aligned}$$

where $\alpha, \beta, \gamma, \phi, \eta \in (0, 1)$. Assume that the production sector has the form

$$Y_t = K_t^{\alpha} L_t^{1-\alpha}$$

where $L_t = N_t h_t$ is the labor in efficiency units.

1. Find the optimal solutions for Households and the representative firm.
2. Find the changes in n_t, e_t when w_t changes.
3. Can $e_t = 0$? Why?
4. Assume an interior solution for e_t, work out the dynamics of k_t and h_t.
5. Find the steady states of k and h (if possible).
6. Change the human capital accumulation function to $h_{t+1} = \mu e_t^{\eta} h_t^{1-\eta}$ and see if the problems can be solved more easily.

Remark 2. At this point, an usual procedure is to analyze the stability of this steady state. However, since there are 2 dynamic variables k_{t+1} and h_{t+1}, stability analysis can be difficult. The idea is still the same, we want to show that the "derivatives" of the future variable wrt the current variable be smaller than 1. Check section 1.2 and Appendix B for an introduction to dynamic stability analysis. The idea is implemented in the following exercise 2.19.

Ex. 2.19 (Inter-generational Taste Externalities)**.** This exercise is from de la Croix and Michel (2002) p.248, which is a simplified version of de la Croix (1996). Consider the following utility function

$$U_t = \ln(c_t - \theta a_t) + \beta \ln(d_{t+1})$$

where $\theta \in (0, 1)$ measure the intensity of inter-generational spillover (aspiration). And

$$\begin{aligned}\text{(habit formation) } & a_t = c_{t-1},\\ \text{(budget constraint) } & c_t + s_t = w_t.\end{aligned}$$

1. Find the solution for the saving function.
2. Assume that the production function is Cobb-Douglas where $f(k_t) = Ak_t^{\alpha}$. Find the competitive wage and interest rate.
3. Assume $k_{t+1} = s_t$. Write the system of 2 dynamic equations $k_{t+1} = \phi(k_t, a_t)$ and $a_{t+1} = \psi(k_t, a_t)$. Find their steady states $\bar{k}, \bar{a}$.
4. Under what conditions are the steady states locally stable?

Chapter 3

First-Order Differential Equations

3.1 Definition

Take China's GDP growth since 2000 for example. The trend is almost linear, so we

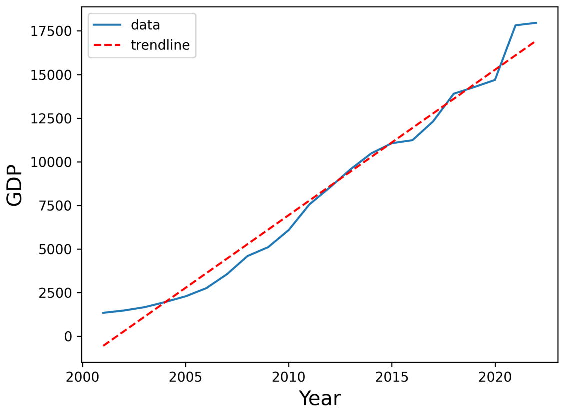


Fig. 3.1. China's GDP in billions USD (Data source: World Bank).

can write the change in each year as

$$y(t+1) = (1+r)y(t)$$

The growth rate can be calculated as

$$\frac{y(t+1) - y(t)}{y(t)} = r$$

If the time interval is smaller, say a smaller fraction $\triangle t$ of t, the total change is divided into smaller portions of time. For example, if the annual growth rate is r, then the quarterly growth rate is approximately $r/4$. The equation is then slightly changed to

$$\frac{y(t+\triangle t) - y(t)}{y(t)\triangle t} = r$$

Cross multiplying yields

$$\frac{y(t+\triangle t) - y(t)}{\triangle t} = r \cdot y(t)$$

Note that $\triangle t$ can get extremely small. Daily growth rate is $r/365$, hourly is $r/(365 \times 24)$. By letting $\triangle t \to 0$, by definition of derivatives, the LHS becomes

$$\lim_{t \to 0} \frac{y(t+\triangle t) - y(t)}{\triangle t} = \frac{dy}{dt}(t).$$

we denote its derivative with respect to time by a "dot":

$$\dot{y}(t) = \frac{dy}{dt}$$

to arrive at the differential equation

$$\dot{y}(t) = ry(t).$$

In general, this equation can be

$$\dot{y}(t) = f(t)$$

Then, its solution is **THE FUNCTION** $y(t)$ such that it satisfies the above equation. It can be worked out by integrating $\dot{x}(t)$ for the whole time line

$$y(t) = \int f(t)dt + k$$

given $y(0)$. This chapter lays out the framework on how to find it. For more exposure to differential equations and solution methods, check Tran and Zhang (2023, Ch 8.6).

3.2 One-variable: Linear case

3.2.1 Autonomous

Constant Growth

This is the simplest case. Like the above example, consider the function

$$\dot{x}(t) = \lambda x(t) \tag{3.1}$$

Theorem 3.2.1. *The solution to $\dot{x}(t) = \lambda x(t)$ is $x(t) = x(0)e^{\lambda t} \equiv x(0)\exp(\lambda t)$.*

Proof. We use the technique called separation. If the differential equation has the form

$$\dot{x}(t) = \frac{dx}{dt} = f(t)g(x)$$

1. Separate the variables (x to the left, not x to the right)

$$\frac{1}{g(x)}dx = f(t)dt$$

2. Integrate both sides

$$\int \frac{1}{g(x)}dx = \int f(t)dt$$

3. Isolate $x(t)$ for solution.

Let us apply this to our case.

1. Separate

$$\frac{\dot{x}(t)}{x(t)} = \frac{1}{x(t)}\frac{dx}{dt} = \lambda$$

2. Integrate both sides (wrt to t)

$$\int \frac{1}{x(t)}\frac{dx}{dt}dt = \int \lambda dt$$

3. Isolate $x(t)$ for solution.

$$\text{LHS} = \int \frac{1}{x(t)}dx = \ln(x(t)) + C_1,$$
$$\text{RHS} = \lambda t + C_2$$

4. Grouping constants yields

$$\ln(x(t)) = \lambda t + \xi$$

Then take exponential of both sides (remmeber: $e^{\ln(x)} = x$) so that

$$x(t) = e^{\lambda t+\xi} = e^{\lambda t}e^{\xi} = ke^{\lambda t} \quad \text{where } k \equiv e^{\xi}$$

At time $t = 0$, we have

$$x(0) = e^0 k = k$$

Plugging back, we obtain

$$x(t) = x(0)e^{\lambda t} \equiv x(0)\exp(\lambda t).$$

■

What is the steady state in this case? Alternatively, that steady state should the the solution to $\dot{x}(t) = 0$? Based on this form, so long as $\lambda \neq 0$, the only solution is $x^* = 0$.

Constant Growth plus a Constant

In a more general framework, we also interested in finding the solution for

$$\dot{x}(t) = \lambda x(t) + b \tag{3.2}$$

where b is some nonzero constant.

Theorem 3.2.2. *The solution to* $\dot{x}(t) = \lambda x(t) + b$ *is* $x(t) = \dfrac{-b}{\lambda} + ke^{\lambda t}$ *where* k *contains* $x(0)$.

Proof. Just plug the solution candidate into the equation and see if it works.

$$(LHS) : \dot{x}(t) = \frac{d}{dt}\left(\frac{-b}{\lambda} + ke^{\lambda t}\right) = \lambda ke^{\lambda t},$$
$$(RHS) : \lambda x(t) = \lambda\left[\frac{-b}{\lambda} + ke^{\lambda t}\right] + b = \lambda ke^{\lambda t} = \dot{x}(t)$$

But how did one come up with such an ingenius solution? Well, they multiply both sides by an integrating factor. In this case, it is $e^{-\lambda t}$. By doing so (3.2) becomes

$$\dot{x}(t)e^{-\lambda t} - \lambda x(t)e^{-\lambda t} = be^{-\lambda t}$$

Notice that the LHS is actually the derivative of $x(t)e^{-\lambda t}$ wrt t, thus

$$\frac{d}{dt}\left(x(t)e^{-\lambda t}\right) = be^{-\lambda t}$$

Take integrals of both sides

$$x(t)e^{-\lambda t} = \int be^{-\lambda t}dt = -\frac{b}{\lambda}e^{-\lambda t} + k. \text{ (k is a constant)}$$

Multiplying both sides with $e^{\lambda t}$ yields

$$x(t) = \frac{-b}{\lambda} + ke^{\lambda t}$$

Let $t = 0$, then $x(0) = -b/\lambda + k$ so that $k = x(0) + b/\lambda$. ■

Note that when $k = 0$, the differential equation becomes

$$x(t) = -\frac{b}{\lambda} = x^*$$

implying that this x^* is the steady state. Another method is to let $\dot{x} = 0$ in (3.2):

$$0 = \lambda x^* + b$$

3.2.2 Stability

Differential equations of these form (3.1) and (3.2) are called *autonomous* because they do not depend on time, but depend on the previous value of the variable. Their steady states are the solutions that $\dot{x}(t) = 0$

$$\begin{aligned} \dot{x}(t) &= \lambda x(t) && \Rightarrow x^* = 0, \\ \dot{x}(t) &= \lambda x(t) + b && \Rightarrow x^* = -\frac{b}{\lambda} \end{aligned}$$

The graphical solution of the steady state can be expressed as follows.

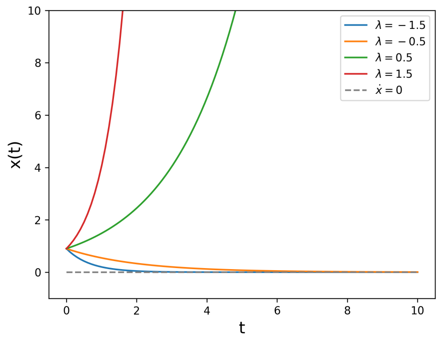
We conclude the condition for stability as follows

Theorem 3.2.3. *For a first-order linear ordinary differential equation of the form*

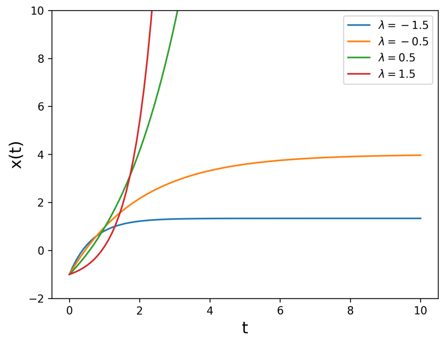
$$\dot{x}(t) = \lambda x(t) + b$$

The stability condition for a one-variable linear differential equation is as follows:

1. *If λ is negative, then the solution $x(t)$ will decay to 0 as $t \to \infty$. This is called asymptotic stability.*
2. *If λ is positive, then the solution $x(t)$ will grow without bound as $t \to \infty$. This is called instability.*



(a) $\dot{x}(t) = \lambda x(t)$



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

Fig. 3.2. Stability of the steady state

3.2.3 Nonautonomous

We move on to the nonautonomous cases where the RHS is also variable with time.

Constant Growth plus a variable Coefficient

Now, b can also depend on time, such as

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

Theorem 3.2.4. *The solution to* $\dot{x}(t) = \lambda x(t) + b(t)$ *is* $x(t) = e^{\lambda t}\left(k + \int e^{-\lambda t} b(t)dt\right)$.

Proof. Multiplying by the integrating factor $e^{-\lambda t}$, we obtain

$$\dot{x}(t)e^{-\lambda t} - \lambda x(t)e^{-\lambda t} = b(t)e^{-\lambda t},$$

which is equivalent to

$$\frac{d}{dt}\left(x(t)e^{-\lambda t}\right) = e^{-\lambda t}b(t)$$

so that

$$x(t)e^{-\lambda t} = \int e^{-\lambda t}b(t)dt + k, \;\; (k \text{ is a constant})$$

Multiplying both sides by $e^{\lambda t}$ yields

$$x(t) = e^{\lambda t}\left(k + \int e^{-\lambda t}b(t)dt\right)$$

∎

General case

Finally, we move to the most general case

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

so that even the growth rate varies with time. We still employ the same trick, multiplying both sides by a suitably chosen integrating factor $e^{-A(t)}$ to obtain

$$\dot{x}(t)e^{-A(t)} - \lambda(t)x(t)e^{-A(t)} = b(t)e^{-A(t)} \tag{3.3}$$

Now, we need to find $A(t)$ such that the LHS equals the derivative of $x(t)e^{-A(t)}$. Notice that the derivative of $x(t)e^{-A(t)}$ is

$$\frac{d}{dt}(x(t)e^{-A(t)}) = \dot{x}(t)e^{-A(t)} - \dot{A}(t)x(t)e^{-A(t)}$$

We, therefore, need to make sure that

$$\dot{A}(t) = \lambda(t)$$

This is done by choosing

$$A(t) = \int \lambda(s)ds. \tag{3.4}$$

The time of $A(t)$ itself can be different from the time path t of $x(t)$. Imagine they have 2 different clocks of their own so it's better to use a different notation. Now, (3.3) becomes

$$\frac{d}{dt}x(t)e^{-A(t)} = b(t)e^{-A(t)}.$$

Integrating both sides

$$x(t)e^{-A(t)} = \int b(t)e^{-A(t)}dt + k, \quad \text{where } k \text{ is a constant}$$

Multiplying both sides by $e^{A(t)}$ yields

$$x(t) = e^{A(t)}\left[k + \int b(t)e^{-A(t)}dt\right]$$

or if we want to expand it using (3.4)

$$x(t) = e^{\int \lambda(s)ds}\left[k + \int e^{-\int \lambda(s)ds}b(t)dt\right]$$

See Sydsæter et al. (2008, p. 204) for the initial value solution.

Ex. 3.1. Solve the following differential equations, and determine the stability of the equilibrium points.

1. $\dot{x}(t) = -2x(t) + 4$

2. $\dot{x}(t) = x(t)(3 - x(t))$

3. $\dot{x}(t) = x(t) \cdot (2 - x(t)) \cdot (x(t) - 1)$

4. $\dot{N}(t) = rN(t)(1 - N(t)/K)$ where r is the growth rate, K is the carrying capacity.

5. $\dot{x}(t) = e^{2t}/x^2$

3.3 System of 2 Differential Equations

Similar to the difference equations, differential equations can also involve a linear or nonlinear system of 2 or more variables. For example, the solution to the consumption-saving problem with CRRA utility is characterized by two first-order linear differential equations:

$$\begin{aligned}\text{(asset accumulation): } \dot{a}(t) &= ra(t) - c(t),\\ \text{(optimal consumption): } \dot{c}(t) &= \frac{r-\rho}{\gamma}c(t).\end{aligned}$$

The general form would be

$$\begin{aligned}\dot{x}(t) &= f(x, y, t)\\ \dot{y}(t) &= g(x, y, t)\end{aligned}$$

This section presents a method to solve such linear systems of differential equations.

3.3.1 Linear Homogeneous System

Consider the following *homogeneous* system, that is, a system of the form

$$\begin{aligned}\dot{x}_1(t) &= f(x_1(t), x_2(t))\\ \dot{x}_2(t) &= g(x_1(t), x_2(t))\end{aligned}$$

Solution:

For brevity, we omit the (t) notation. Since the growth rates are constants, we can write the system as

$$\begin{aligned}\dot{x}_1 &= ax_1 + bx_2,\\ \dot{x}_2 &= cx_1 + dx_2\end{aligned}$$

We can write it in matrix form as

$$\underbrace{\begin{pmatrix}\dot{x}_1\\ \dot{x}_2\end{pmatrix}}_{\dot{\mathbf{x}}} = \underbrace{\begin{pmatrix}a & b\\ c & d\end{pmatrix}}_{\mathbf{A}}\underbrace{\begin{pmatrix}x_1\\ x_2\end{pmatrix}}_{\mathbf{x}}$$

Assume that $|\mathbf{A}| \neq 0$. We have transformed the system to

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}. \tag{3.5}$$

If $\mathbf{A}$ is diagonalizable, there exists a matrix $\mathbf{V}^{2\times 2}$ such that

$$\mathbf{A} = \mathbf{V}\Lambda\mathbf{V}^{-1}$$

where

$$\mathbf{V} = \begin{pmatrix}\mathbf{v}_1 & \mathbf{v}_2\end{pmatrix}$$

be the 1×2 matrix whose columns are the **eigenvector** of $\mathbf{A}$ and

$$\Lambda = \begin{pmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{pmatrix}$$

be a square matrix 2×2 with λ_1, λ_2 be the **eigenvalues** of $\mathbf{A}$. We transform (3.5) to

$$\dot{\mathbf{x}} = \mathbf{V}\Lambda\mathbf{V}^{-1}\mathbf{x}$$

Multiply both sides with $\mathbf{V}^{-1}$ yields

$$\mathbf{V}^{-1}\dot{\mathbf{x}} = \Lambda\mathbf{V}^{-1}\mathbf{x}.$$

Define

$$\mathbf{V}^{-1}\mathbf{x} = \mathbf{y}$$

then the system becomes

$$\dot{\mathbf{y}} = \Lambda\mathbf{y} \tag{3.6}$$

We have transformed it to something similar to the very simpliest case at (3.1)[(1)]

The system (3.6) implies 2 independent differential equations

$$\begin{pmatrix} \dot{y}_1(t) \\ \dot{y}_2(t) \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 y_1(t) \\ \lambda_2 y_2(t) \end{pmatrix}$$

Each $y_i(t)$ is growing at a constant rate λ_i. Previous analyses give us the solution to each differential equation as

$$y_i(t) = k_i e^{\lambda_i t} \quad \text{for } i = 1, 2$$

Finally, solutions for $\mathbf{x}$ is given by

$$\mathbf{x} = \mathbf{V}\mathbf{y}$$

Writing explicitly

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \\ &= \mathbf{v}_1 y_1 + \mathbf{v}_2 y_2 \\ &= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} y_1 + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} y_2 \\ &= k_1 e^{\lambda_1 t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + k_2 e^{\lambda_2 t} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \end{aligned}$$

where the 2 eigenvectors are

$$\mathbf{v}_1 = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}, \qquad \mathbf{v}_2 = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

Stability

The stability conditions are very similar to the linear case, we just need to generalize the Jacobian matrix. For the proof, you can see Simon and Blume (1994, Ch 22) and Sydsæter et al. (2008, Ch 6.8, p.251) We summarize the stability analysis as follows

---

[(1)] $\dot{x}(t) = \lambda x(t)$

Theorem 3.3.1 (Stability Conditions)**.** *For a two-variable homogeneous system of differential equations, the steady state is* $(0,0)$*. Let* λ_1, λ_2 *be the eigenvalues of the Jacobian matrix, then*

1. *If* $\lambda_1 < 0$ *AND* $\lambda_2 < 0$, $|\mathbf{A}| > 0$*, the system is a sink, and the steady state is stable.*
2. *If* $\lambda_1 > 0$ *AND* $\lambda_2 > 0$, $|\mathbf{A}| > 0$*, the system is a source, and the steady state is unstable.*
3. *If* λ_1 *AND* λ_2 *have opposite signs,* $|\mathbf{A}| < 0$*, one part of the solution is stable (it converges to* 0 *as* $t \to 0$*), and the other is unstable (it converges to* ∞ *as* $t \to 0$*), the system is a saddle.*

Similar to the case of the difference equation, it is often more practical to use the trace and determinants of the Jacobian.

1. (stable) $tr(\mathbf{A}) < 0$ and $\det(\mathbf{A}) > 0$.
2. (unstable) $tr(\mathbf{A}) > 0$ and $\det(\mathbf{A}) > 0$.
3. (saddle) $\det(\mathbf{A}) < 0$.

3.3.2 Nonlinear Homogeneous Systems

The system is generalized as

$$\dot{x} = f(x, y),$$
$$\dot{y} = g(x, y)$$

Both equations must be **differentiable**. Denote the steady state as $(\bar{x}, \bar{y})$. We write the system in matrix notation as follows

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$$

where, of course, the Jacobian matrix is

$$\mathbf{A} = \begin{pmatrix} f'_x & f'_y \\ g'_x & g'_y \end{pmatrix}$$

then its stability can be stated as follows.

Theorem 3.3.2 (Lyapunov)**.** *Suppose that* f *and* g *are* C^1 *functions.*

Let $(\bar{x}, \bar{y})$ *be the steady (equilibrium) state for the system*

$$\dot{x} = f(x, y),$$
$$\dot{y} = g(x, y)$$

Let $\mathbf{A}$ *be the Jacobian matrix*

$$\mathbf{A} = \begin{pmatrix} f'_x & f'_y \\ g'_x & g'_y \end{pmatrix}$$

with the trace as

$$tr(\mathbf{A}) = f'_x(\bar{x}, \bar{y}) + f'_y(\bar{x}, \bar{y}),$$

and the determinant as

$$|\mathbf{A}| = f'_x(\bar{x}, \bar{y})g'_y(\bar{x}, \bar{y}) - f'_y(\bar{x}, \bar{y})g'_x(\bar{x}, \bar{y}) > 0$$

Then:

1. *If $tr(\mathbf{A}) < 0$ and $|\mathbf{A}| > 0$, both eigenvalues of $\mathbf{A}$ have **negative** real parts, then $(\bar{x}, \bar{y})$ is locally asymptotically stable.*
2. *If $tr(\mathbf{A}) > 0$ and $|\mathbf{A}| > 0$, both eigenvalues of $\mathbf{A}$ have **positive** real parts, then $(\bar{x}, \bar{y})$ is unstable.*
3. *If $|\mathbf{A}| < 0$, the eigenvalues are nonzero real numbers of OPPOSITE signs, so $(\bar{x}, \bar{y})$ is a saddle.*

Last but not least, since the solutions to the systems (either 1 variable or multivariate) sometimes cannot be attained explicitly, we need to show that such an equilibrium point exists. Fortunately, the theorem is pretty straightforward.

Theorem 3.3.3 (Existence and Uniqueness)**.** *Consider the first-order differential equations*

$$\begin{aligned} \dot{x} &= f(x, y, t), \\ \dot{y} &= g(x, y, t) \end{aligned} \tag{3.7}$$

1. *If f and g are continuous functions in a neighborhood of (x_0, y_0, t_0), then there exist functions $x^*(t)$ and $y^*(t)$ about that point such that they satisfy the system (3.7) and $x(t_0) = x_0$, $y(t_0) = y_0$.*
2. *If f and g are differentiable (C^1) at the point (x_0, y_0, t_0), then the solution $x(t), y(t)$ are unique.*

Ex. 3.2. Determine the local stability of the following (Sydsæter et al., 2008, p.254)

(a) $\dot{x} = -x + 0.5y^2, \; \dot{y} = 2x - 2y$ at $(0, 0)$

(b) $\dot{x} = x - 3y + 2x^2 + y^2 - xy, \; \dot{y} = 2x - y - e^{x-y}$ at $(1, 1)$

(c) $\dot{x} = -x^3 - y, \; \dot{y} = x - y^3$ at $(0, 0)$

(d) $\dot{x} = y - x, \; \dot{y} = -x^2 + 8x - 2y$ at $(6, 6)$

(e) $\dot{x} = -3x - 2y + 8x^2 + y^3, \; \dot{y} = 3x + y - 3x^2y^2 + y^4$ at $(0, 0)$

Ex. 3.3 (Pollution)**.** Suppose the economy is described by the system

$$\text{(capital) } \dot{K} = K(sK^{\alpha-1} - \delta)$$
$$\text{(pollution) } \dot{P} = K^{\beta} - \gamma P$$

where $s \in (0, 1), \alpha \in (0, 1), \delta > 0, \gamma > 0, \beta > 1$.

1. Find the equilibrium point (K^*, P^*). Check its stability.
2. Find an explicit expression for $K(t)$ given $K(0) = K_0 \geq 0$. Examine its limit as $t \to \infty$.

3.4 Phase Diagrams

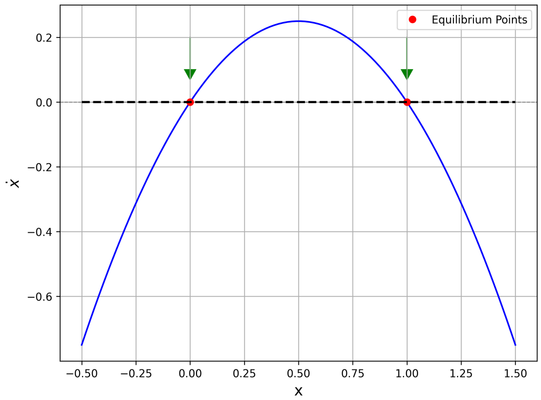
Phase diagrams are particularly useful to visualize the solutions and predict their stability without explicitly solving them. We are only interested in autonomous functions.

3.4.1 Single Variable Differential Equations

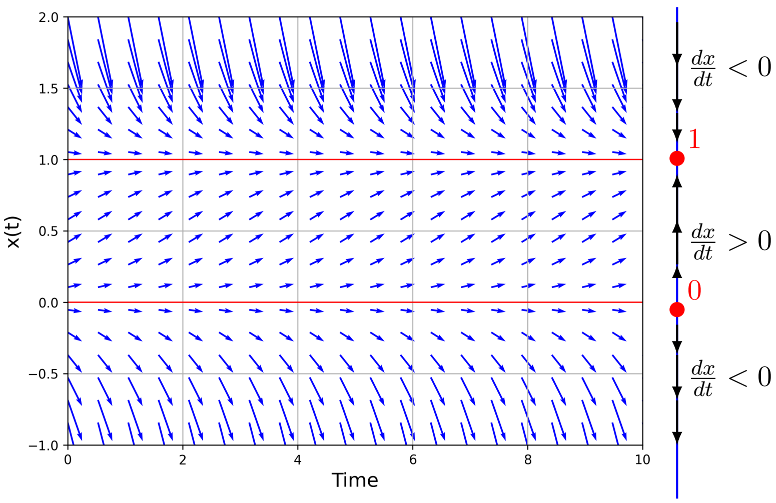
Consider the following [2].

$$\dot{x} = x(1-x)$$

It is easy to see that there are 2 fixed points 0 and 1. How is the stability at each point?



Without using the eigenvalues, we can use phase diagrams to investigate their stability.



How to construct it?

1. (Nullclines) First, we need to draw the red lines, which represent the fixed points that solve $\dot{x} = 0$.
2. (Directional Arrows). To construct the directional arrows (that is, to determine how the points below and above $x(t)$ behave around x^*. This process involves differentiation. Just try a number and see the sign of $\dot{x}$.

[2] You should watch this https://youtu.be/swt-let4pCI and this https://youtu.be/cC2w2z_i2DA

3.4.2 System of 2 Linear Differential Equations

Next, we move on to a system of 2 linear equations. Construction-wise, it is similar to the case of 1 variable, but we are now concerned with 2 variables instead of 1.

Let us consider the system

$$\begin{aligned} \dot{x} &= ax + by + \kappa_1, \\ \dot{y} &= cx + dy + \kappa_2. \end{aligned} \tag{3.8}$$

where $a < 0, b < 0, c < 0, d > 0, \kappa_1 > 0, \kappa_2 > 0$. You can see that

$$ad - bc < 0$$

so the eigenvalues of the Jacobian of this system have OPPOSITE signs. The dynamical system is a "saddle". Check Fig. 3.4 while reading the following.

1. (A. Nullclines) We need to plot the nullclines, which are the loci $\dot{x} = 0$ and $\dot{y} = 0$. The first locus $\dot{x} = 0$ is given by

   $$y = -\frac{a}{b}x - \frac{\kappa_1}{b}$$

   since $a < 0, b < 0$, the locus is a straight line with a negative slope in the (x, y) plane. The second locus $\dot{y} = 0$ is given by

   $$y = -\frac{c}{d}x - \frac{\kappa_2}{d}$$

   since $c < 0, d > 0$, the locus is a straight line with a positive slope.

2. (B. Steady State) The steady state is given by the intersection of the two nullclines. Let's call it (x^*, y^*). Together with the nullclines, it divides the (x, y) plane into 4 areas.

3. (C. Directional Arrows) These arrows determine the direction of the system's trajectories over time anywhere on the phase. Look at the system (3.8).

   - From the $\dot{x}$ function, we see that

     $$\frac{d\dot{x}}{dx} = a < 0 \qquad\qquad \frac{d\dot{x}}{dy} = b < 0$$

     so $\dot{x}$ is decreasing in x (and decreasing in y). Any point above the $\dot{x} = 0$ line must have $\dot{x} < 0$ and any point below $\dot{x} = 0$ must have $\dot{x} > 0$.

   - From the $\dot{y}$ function, we see that

     $$\frac{d\dot{y}}{dy} = d > 0 \qquad\qquad \frac{d\dot{y}}{dx} = c < 0$$

     so $\dot{y}$ is increasing in y (and decreasing in x). Any point above the $\dot{y} = 0$ line must have $\dot{x} > 0$ and any point below $\dot{x} = 0$ must have $\dot{x} < 0$.

   - (D. Trajectories) Using the information above, we can draw trajectories that satisfy the system. These are solutions to the system. To select a specific solution among all possible solutions, we will need to specify either an initial condition or a final condition. Among all the trajectories, we highlight the saddle path for the system. We know that such a saddle path exist because the eigenvalues of the system have opposite sign. The saddle path is the straight line that goes through the steady state.

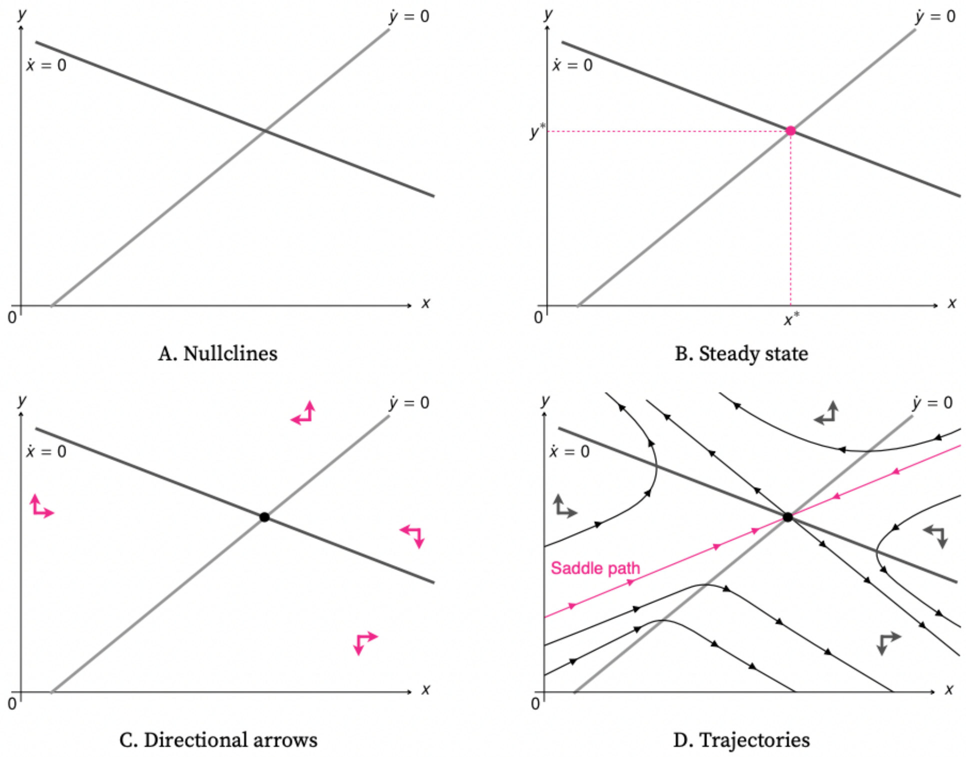


Fig. 3.4. Phase diagram of the dynamical system (3.8) (Michaillat, 2023) .

Ex. 3.4. The exercises here are from Michaillat (2023); Shone (2002)

(a) (Problem 3) Consider the linear system of differential equations given by

$$\begin{pmatrix}\dot{x}\\ \dot{y}\end{pmatrix} = \begin{pmatrix}1 & 1\\ 4 & 1\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix}$$

1. Find the general solution of the system using either eigenvalue method, given $(x_0, y_0) = (\kappa_1, \kappa_2)$.
2. Draw the phase diagram that plots the trajectories of the system

(b) The IS-LM continuous time model (Shone, 2002, p.10.3).

$$\text{(IS)} : c(t) = a + b(1-\tau)y(t) - hr(t),$$
$$\text{(LM)} : m(t) = ky(t) - ur(t)$$

where $a > 0, b \in (0,1), \tau \in (0,1), h > 0$ and c, a, b, τ, y, h, r are consumption, fixed consumption, marginal propensity to consume, tax, real income, coeff. of investment, and interest rate. On the other hand, $k, u > 0$ control the demand for real money. Assume that the dynamics of the 2 markets are

$$\dot{y} = \alpha(c(t) - y(t)) \quad (\alpha > 0),$$
$$\dot{r} = \beta(m(t) - m_0) \quad (\beta > 0)$$

where income adjusts according to the excess demand in that market and interest rates adjust according to the excess demand in the money market.

1. Express the system explicitly in $\dot{y}, \dot{r}$.
2. Find the fixed point y^*, r^*.
3. Draw the phase diagram with the 2 loci $\dot{y} = 0$ and $\dot{r} = 0$.

3.4.3 System of 2 Nonlinear Differential Equations

In macroeconomics, most of the advanced dynamics fall into this category. These systems, in general, are difficult to solve explicitly. Nevertheless, without solving them, we can still characterize their properties by constructing the phase diagrams.

Consider the following nonlinear system that describes a typical growth model

$$\begin{aligned} \dot{k} &= f(k) - c - \delta k, \\ \dot{c} &= [f'(k) - (\delta + \rho)]c \end{aligned} \tag{3.9}$$

where $\rho > 0, \delta \in (0,1)$ are parameters, the capital stock $k(t)$ is a state variable with the initial value k_0 given and $c(t)$ is the control variable. The production function satisfies the Inada conditions as follows

$$\begin{aligned} &f(0) = 0, \\ &f' > 0, f'' < 0, \\ &\lim_{k \to \infty} f'(k) = 0, \lim_{k \to 0} f'(k) = \infty. \end{aligned}$$

1. (A. Nullclines) As always, we draw the nullclines. The $\dot{k} = 0$ curve is defined by

   $$c = f(k) - \delta k$$

   The $\dot{c} = 0$ curve is defined by

   $$f'(k) = \delta + \rho.$$

   In the (k, c) plane, the $\dot{k} = 0$ curve is concave in k (because $f(k)$ is concave) while the $\dot{c} = 0$ curve is just a vertical line.

2. (B. Steady State) The intersection of these two loci is the steady state (k^*, c^*).

3. (C. Directional Arrows) To do this, we investigate the partial derivatives

   $$\begin{aligned} \frac{\partial \dot{k}}{\partial c} &= -1 < 0, \\ \frac{\partial \dot{c}}{\partial k} &= c \cdot f''(k) < 0 \end{aligned}$$

   What do they mean?

   - The first equation says $\dot{k}$ is a decreasing function in c. So when c increases, $\dot{k}$ turns more negative, implying that k decreases (points above $\dot{k} = 0$ moves leftward); and when c decreases, the reverse happens, k increases and the points move rightward.
   - The second equation says $\dot{c}$ is decreasing in k. So when k increases, $\dot{c}$ turns negative, implying that c decreases (the points to the right of $\dot{c} = 0$ moves downward); and when k decreases, c increases and the points to the left of $\dot{c} = 0$ moves upward.

4. (D. Tracjectories) The directional arrows drawn describe a saddle around the steady state. The only way for the economy to converge to the steady state is on the saddle path leading to it. This means that given any initial capital k_0, initial consumption c_0 is such that the pair (k_0, c_0) lies on the saddle path.

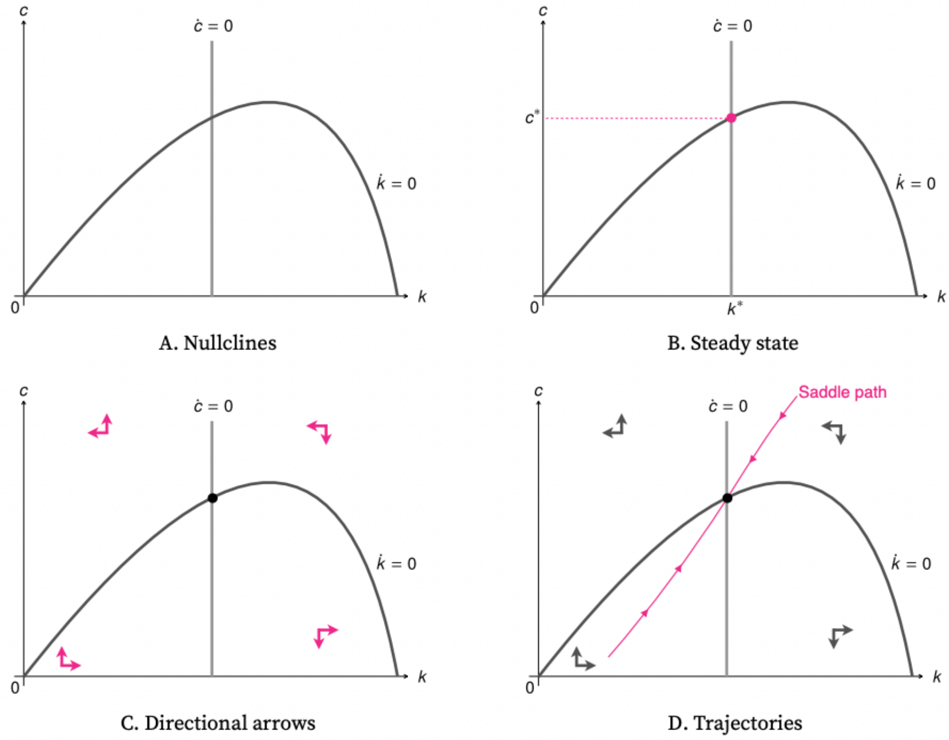
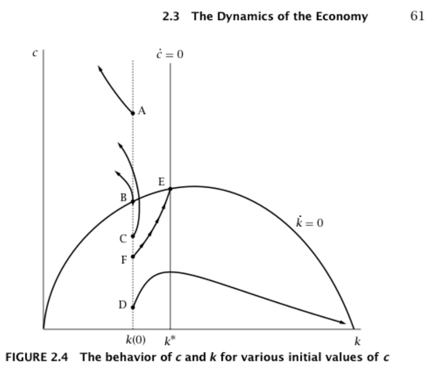
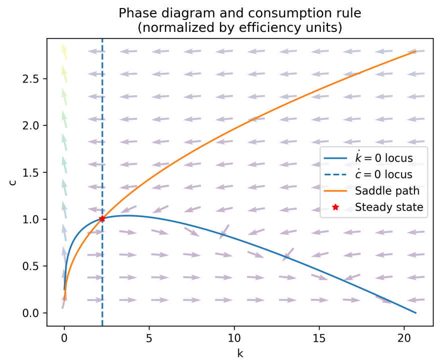


Fig. 3.5. Phase Diagram of system (3.9) (Michaillat, 2023).



(a) Trajectories (Romer, 2011)



(b) A closer look.
Source code: https://github.com/Mv77/Ramsey_growth/blob/master/Ramsey.py

Fig. 3.6. More on the trajectories of this model.

Lineariziation around the Steady State

The phase diagram indicates that the system is a saddle around the steady state in Fig. 3.5. We can also obtain this result by linearizing the nonlinear system (3.9) using a first-order Taylor expansion about the steady state:

$$\dot{k} = \dot{k^*} + (k - k^*)\frac{\partial \dot{k}}{\partial k} + (c - c^*)\frac{\partial \dot{k}}{\partial c},$$
$$\dot{c} = \dot{c^*} + (k - k^*)\frac{\partial \dot{c}}{\partial k} + (c - c^*)\frac{\partial \dot{c}}{\partial c}.$$

Since at the steady state $\dot{k^*} = \dot{c^*} = 0$, the system reduced to

$$\begin{pmatrix} \dot{k} \\ \dot{c} \end{pmatrix} = \mathbf{J}^* \begin{pmatrix} k - k^* \\ c - c^* \end{pmatrix}$$

where $\mathbf{J}^*$ is the Jacobian matrix evaluated at the steady state

$$\mathbf{J}^* = \begin{pmatrix} \frac{\partial \dot{k}}{\partial k}|_{(k^*,c^*)} & \frac{\partial \dot{k}}{\partial c}|_{(k^*,c^*)} \\ \frac{\partial \dot{c}}{\partial k}|_{(k^*,c^*)} & \frac{\partial \dot{c}}{\partial c}|_{(k^*,c^*)} \end{pmatrix}.$$

More explicitly

$$\mathbf{J}^* = \begin{pmatrix} f'(k^*) - \delta & -1 \\ cf''(k^*) & f'(k^*) - (\delta + \rho) \end{pmatrix}.$$

The result of evaluation is

$$\mathbf{J}^* = \begin{pmatrix} \rho & -1 \\ cf''(k) & 0 \end{pmatrix}.$$

Calculating its determinant shows that

$$\det(\mathbf{J}^*) = cf''(k) < 0$$

Thus, the 2 eigenvalues have opposite sign, so the system around the steady state is a saddle.

Ex. 3.5. Given the system

$$\begin{aligned} \dot{k}(t) &= \frac{q(t) - 1}{\chi} k(t), \\ \dot{q}(t) &= rq(t) - f'(k(t)) - 0.5\chi^{-1}(q(t) - 1)^2 \end{aligned}$$

where $\chi, r > 0$ and

$$f(k(t)) = Ak(t)^\alpha$$

Draw the phase diagram and show that the steady state is a saddle point locally.

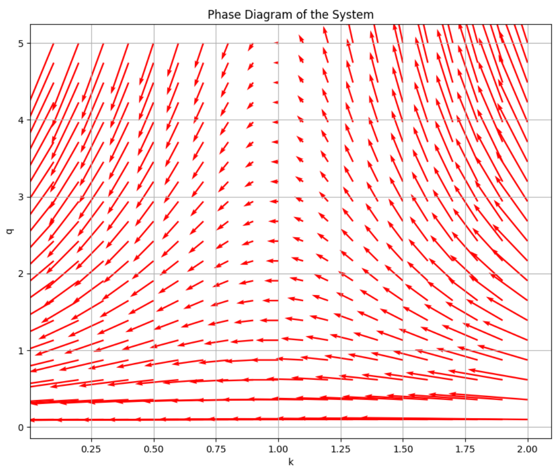


Fig. 3.7. Illustration with $r = 0.1, \chi = 0.9, A = 2, \alpha = 0.3$.

Chapter 4

Dynamic Programming

There are 2 versions of dynamic optimization methods. The first deals with continuous time, where a variable is differentiable with time. The second deals with discrete time, where a variable is not differentiable with respect to time.

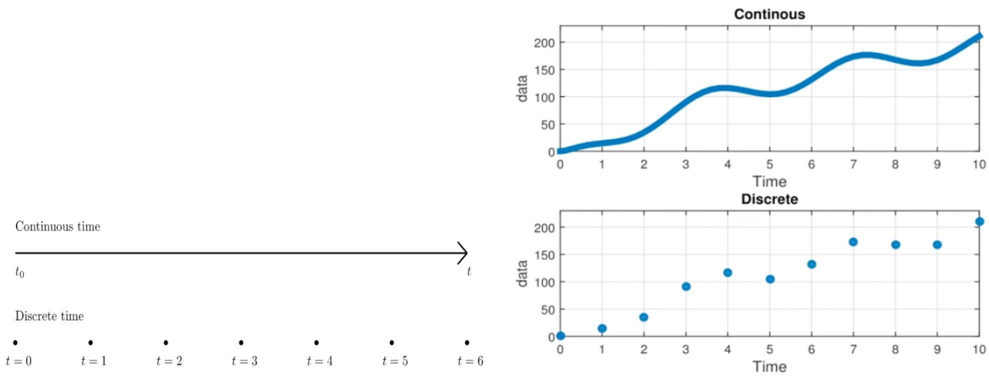


Fig. 4.1. Conceptualization of Time. Source: Klaus Prettner

In discrete time, we often use the Bellman method, while in continuous time, we often use the Hamiltonian method. Both of them are useful for analysis, and then we often use numerical simulations or a phase diagram to work out the dynamics of the model. The notes are based on Chiang (1992); Sydsæter et al. (2008); Heer and Maussner (2009); Campante et al. (2021), with references to a lot of online materials such as

1. http://www.chrisedmond.net/phd2019.html
2. http://www.chrisedmond.net/hons2019.html
3. https://pascalmichaillat.org/c3/

4.1 Motivation

Also known as Discrete-Time Optimization. Household solves

$$\max_{c_t} \sum_{t=0}^{\infty} \beta^t u(c_t),$$
$$s.t.\ c_t = f(k_t) - k_{t+1} + (1-\delta)k_t.$$

Assume a nonzero initial condition, the constraint gives us the equation of motion of the state variable

$$\begin{array}{ll} (\textbf{transition equation}) & k_{t+1} = f(k_t) + (1-\delta)k_t - c_t, \\ (\textbf{initial condition}) & k_0 > 0 \end{array}$$

Find the solution for the optimal choice of c_t.

4.2 Lagrangian Method

The Lagrangian:

$$\mathcal{L} = \sum_{t=0}^{\infty} \beta^t u(c_t) + \sum_{t=0}^{\infty} \lambda_t \left[f(k_t) + (1-\delta)k_t - k_{t+1} - c_t \right]$$

Grouping all the summation and rewriting the Lagrangian:

$$\mathcal{L} = \sum_{t=0}^{\infty} \left[\beta^t u(c_t) + \lambda_t \left(f(k_t) + (1-\delta)k_t - k_{t+1} - c_t \right) \right]$$

The term inside the sum should be optimized at each point in time. The modified Lagrangian:

$$L = \beta^t u(c_t) + \lambda_t \left(f(k_t) + (1-\delta)k_t - k_{t+1} - c_t \right)$$

FOC: (Note that k_{t+1} appears twice at time t and $t+1$)

$$(c_t) : \frac{\partial L}{\partial c_t} = \beta^t u'(c_t) - \lambda_t = 0 \tag{4.1}$$

$$(k_{t+1}) : \frac{\partial L}{\partial k_{t+1}} = -\lambda_t + \lambda_{t+1} \left[f'(k_{t+1}) + (1-\delta) \right] = 0 \tag{4.2}$$

By virtue of Eq. (4.1), we see that:

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

Plugging back to Eq. (4.2) and rearranging give us the Euler equation:

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

Consider the utility function $u(c_t) = \ln(c_t)$, we have:

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

In the steady-state, $c_{t+1} = c_t = \bar{c}$ and $k_{t+1} = k_t = \bar{k}$.

Provided with a functional form of the production function $f(k)$, one can find the steady-state values of $\bar{c}, \bar{k}$ and then use backward induction to figure out the dynamics from a given k_0.

4.3 Bellman Method

The following steps describe Bellman's method.

1. Set up the Bellman equation

$$V(k_t) = \max\left[u(c_t) + \beta V(k_{t+1})\right]$$

   where one can replace

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

   so the problem changes from optimizing c_t to optimizing k_{t+1}. Let's rewrite the Value Function as

$$V(k_t) = \max_{k_{t+1}}\left[u(f(k_t) - k_{t+1} + (1-\delta)k_t) + \beta V(k_{t+1})\right]$$

2. Maximizing the Value function wrt. the control variable k_{t+1}:

$$\frac{\partial V(k_t)}{\partial k_{t+1}} = 0 \Leftrightarrow \frac{\partial u(k_{t+1})}{\partial k_{t+1}} + \beta\frac{\partial V(k_{t+1})}{\partial k_{t+1}} = 0$$

3. Do we know $\frac{\partial V(k_{t+1})}{\partial k_{t+1}}$? Yes, it can be obtained by differentiating the Value Function to the state variable k_t

$$\frac{\partial V(k_t)}{\partial k_t} = (f'(k_t) + 1 - \delta)u'(f(k_t) - k_{t+1} + (1-\delta)k_t)$$

   implying that

$$\begin{aligned}\frac{\partial V(k_{t+1})}{\partial k_{t+1}} &= (f'(k_{t+1}) + 1 - \delta)u'(f(k_{t+1}) - k_{t+2} + (1-\delta)k_{t+1}) \\ &= (f'(k_{t+1}) + 1 - \delta)u'(c_{t+1})\end{aligned}$$

4. Derive the Euler equation relating the dynamics of the choice variable.

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

   where the transversality condition holds

$$\lim_{t\to 0} \beta^t u'(c_t)k_{t+1} = 0.$$

4.4 Examples

4.5 Exercises

Chapter 5

Optimal Control

Also known as Continuous-Time Optimization. Let us consider the problem

$$\max_{\{c_t\}_{t=0}^{T}} e^{-(\rho-n)t}u(c_t)dt, \tag{5.1}$$

$$s.t.\ c_t + \dot{k}_t = f(k_t) - \delta k_t. \tag{5.2}$$

The control variable is c_t, and the state variable is a_t. Assume a nonzero initial condition, the constraint gives us the equation of motion for the state variable

(**transition equation**)	$\dot{k}_t = f(k_t) - \delta k_t - c_t,$
(**initial condition**)	$k_0 > 0$

5.1 Lagrangian Method

Everything begins with the Lagrangian function. In fact, it's more intuitive to start here than going straight to Hamiltonian:

$$\mathcal{L} = \int_0^\infty e^{-\rho t}u(c(t))dt + \int_0^\infty \lambda(t)\left[f(k(t)) - \delta k(t) - c(t) - \dot{k}(t)\right] dt$$

It's the sum of 2 integrals with the same boundary and the same variables (time), so we can sum them up and rewrite the Lagrangian:

$$\mathcal{L} = \int_0^\infty \left\{e^{-\rho t}u(c(t)) + \lambda(t)\left[f(k(t)) - \delta k(t) - c(t) - \dot{k}(t)\right]\right\} dt$$

We want to maximize $c(t)$ with respect to $k(t)$, but $\dot{k}(t)$ is not independent with $k(t)$. So $\dot{k}(t)$ needs to be got rid of, and can somehow be expressed in terms of $k(t)$.

First, separate the $\dot{k}(t)$ term from the integral.

$$\mathcal{L} = \int_0^\infty \left\{e^{-\rho t}u(c(t)) + \lambda(t)\left[F(k(t)) - \delta k(t) - c(t)\right]\right\} dt - \int_0^\infty \lambda(t)\dot{k}(t)dt$$

Using integration by parts for the second integral term:

$$\int_0^\infty \lambda(t)\dot{k}(t)dt = \lambda(t)k(t)\Big|_0^\infty - \int_0^\infty \dot{\lambda}(t)k(t)dt = \lambda(\infty)k(\infty) - \lambda(0)k(0) - \int_0^\infty \dot{\lambda}(t)k(t)dt$$

Notice that if k at time $t = \infty$ is indeed ∞, and $\lambda(\infty)$ (the penalty term) is also non-zero, there is no reason to consume now. The household will wait until the very end of time close to infinity since the utility is essentially infinity. On the other hand, if k at $t = \infty$ is $-\infty$, there is no reason to save, it's better to consume everything and leave nothing left. Such extreme cases would appear if $\lambda(\infty)k(\infty)$ is not bounded, and it will dominate everything.

To counter that problem, we introduce the transversality (no Ponzi scheme) condition.

$$\lim_{t\to\infty} \lambda(t)k(t) = 0$$

We also assume $k(0)$ is given and is a constant (the economy needs something to start with), meaning it will disappear when we derive the FOC. As it is given, there should be no penalty term ($\lambda(0) = 0$). Thus, we have:

$$\int_0^{\infty} \lambda(t)\dot{k}(t)dt = -\int_0^{\infty} \dot{\lambda}(t)k(t)dt$$

Plugging back to the Lagrangian $\mathcal{L}$, we obtain:

$$\mathcal{L} = \int_0^{\infty} \left\{e^{-\rho t}u(c(t)) + \lambda(t)\left[F(k(t)) - \delta k(t) - c(t)\right]\right\} dt + \int_0^{\infty} \dot{\lambda}(t)k(t)dt$$

Sum them back to under 1 integral again:

$$\mathcal{L} = \int_0^{\infty} \left\{e^{-\rho t}u(c(t)) + \lambda(t)\left[F(k(t)) - \delta k(t) - c(t)\right] + \dot{\lambda}(t)k(t)\right\} dt$$

The Lagrangian enables us to maximize the function at each point in time, so we only care about maximizing the functional form inside the integral (we don't care about time anymore). Rewrite the term inside the integral as L, and call this the *modified Lagrangian*. The following two expressions are equivalent.

$$L = e^{-\rho t}u(c(t)) + \lambda(t)\left[f(k(t)) - \delta k(t) - c(t)\right] - \lambda(t)\dot{k}(t) \tag{5.3}$$

$$L = \underbrace{e^{-\rho t}u(c(t)) + \lambda(t)\left[f(k(t)) - \delta k(t) - c(t)\right]}_{\text{Hamiltonian}} + \dot{\lambda}(t)k(t) \tag{5.4}$$

Terminologically speaking: $c(t)$ is the control variable since it can jump (taking any value at any point in time). In contrast, $k(t)$ is the state variable because it accumulates over time. $\lambda(t)$ is the co-state variable (also known as *shadow price* in economics because it shows the marginal cost of violating the constraints at each point in time).

Sticking with the Lagrangian method, we derive the FOC:

$$(c(t)) : \frac{dL}{dc(t)} = e^{-\rho t}u'(c(t)) - \lambda(t) = 0 (\text{using } (5.3)) \tag{5.5}$$

$$(k(t)) : \frac{dL}{dk(t)} = \lambda(t)[f'(k(t)) - \delta] + \dot{\lambda}(t) = 0 \text{ (using Eq.(5.4))} \tag{5.6}$$

$$(\lambda(t)) : \frac{dL}{d\lambda(t)} = f(k(t)) - \delta k(t) - c(t) = 0 \text{ (using (5.3))} \tag{5.7}$$

We need to get rid of λ. From $(c(t))$ condition, we derive:

$$\lambda(t) = e^{-\rho t}u'(c(t))$$

Thus, $\dot{\lambda}(t)$ can be derived as:

$$\dot{\lambda}(t) = \frac{d\lambda(t)}{dt} = -\rho e^{-\rho t} u'(c(t)) + e^{-\rho t} u''(c(t))\dot{c}(t)$$

And we obtain:

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -\rho \frac{u'(c(t))}{u'(c(t))} + \frac{u''(c(t))}{u'(c(t))}\dot{c}(t) = -\rho + \frac{u''(c(t))}{u'(c(t))}\dot{c}(t)$$

We can explicitly derive the condition for a functional form of u, especially with CRRA (constant relative risk aversion). For simplicity, let $u(c) = \ln(c)$, then $u'(c(t)) = \frac{1}{c(t)}$ and $u''(c(t)) = \frac{-1}{c^2(t)}$. Plugging back into the above equation, we obtain:

$$\frac{\dot{\lambda}(t)}{\lambda(t)} = -\rho - \frac{\dot{c}(t)}{c(t)} \quad (5)$$

From the FOC of $k(t)$, we know that:

$$-\frac{\dot{\lambda}(t)}{\lambda(t)} = f'(k(t)) - \delta$$

Combining with Eq. (5) yields the Euler equation:

$$\frac{\dot{c}(t)}{c(t)} = f'(k(t)) - \delta - \rho$$

This condition constitutes the optimal path of consumption chosen by the household.

5.2 Hamiltonian

A cookbook method (Pontryagin's maximum principle) for this problem. The maximum principle can be summarized as a series of steps with the economic intuition explained as follows (Campante et al., 2021).

1. Set at the present-value Hamiltonian

$$H_t = u(c_t)e^{-\rho t} + \lambda_t(\underbrace{f(k_t) - c_t - \delta k_t}_{\dot{k}_t}) \tag{5.8}$$

   where λ_t is called the co-state variable. The Hamiltonian is similar to the way you obtain the Lagrangian (but without integral). The interpretation of the co-state variable λ_t is the same as the Lagrangian multiplier. It tells you the marginal benefit of a marginal addition to the stock of the state variable k_t (in economic jargon, it is called "the shadow value/ price of the state variable"), which is also the value of the state variable at $t = 0$.

2. Take the FOC of Hamiltonian wrt the control variable(s)

$$\frac{\partial H_t}{\partial c_t} = 0 \Leftrightarrow e^{-\rho t} u'(c_t) = \lambda_t. \tag{5.9}$$

   This is legitimate since the Hamiltonian is static at each t, so the optimal value of the choice variable should be done in a static optimization fashion. (5.9) states that the marginal utility gain from increasing consumption has to be equal to the marginal cost (λ) of not adding such amount to the stock of your assets (the state variable).

3. Derive the optimal path of the state and co-state variable(s)

$$\dot{k}_t = \frac{\partial H_t}{\partial \lambda_t} = f(k_t) - c_t - \delta k_t, \tag{5.10}$$

$$\dot{\lambda}_t = -\frac{\partial H_t}{\partial k_t} = -\lambda_t(f'(k_t) - \delta) \tag{5.11}$$

The Hamilton is static, but our model is dynamic. This means that, at any instant, we must figure out that whatever we leave for the next instant is consistent with optimization. This is the key insight into the problem. Furthermore, what we care about in an infinite time span can be broken down into a sequence of choices between the current instant and the next in an infinitesimally small time frame.

To figure out that path, the maximization principle tells you that you need to satisfy the co-state equations. The first equation (5.10) links one instant of the state variable to the next and must be satisfied at any time. The second equation (5.11) is an "asset pricing" condition. Basically, the term $\dot{\lambda}_t$ is the appreciation in the marginal value of capital, so $-\dot{\lambda}_t$ is its depreciation. If you carry more capital to the next, its value should depreciate more as the volume increase. The term $\partial H/\partial k$, on the other hand, shows the marginal return of capital at this instant, contributing to the utility and production (which are encompassed in the Hamiltonian). Obviously, the equilibrium is brought by equalizing the two terms. (Read more)

4. Set the transversality condition

$$\lim_{t \to \infty} \lambda_t k_t = 0. \tag{5.12}$$

While the transition (5.10), and Euler equations ((5.9), (5.11)) jointly govern the temporal behavior of consumption and capital (state variable), the initial and transversality conditions specify the state of the economy at the boundaries (i.e., $t = 0$ and $t = \infty$). Initially, the economy must start with something it can work with – $k_0 > 0$. The transversality condition, in particular, prevents the over-accumulation of capital by requiring zero present value of the capital in an infinitely distant future.

In a finite time horizon (i.e., you must die at some point T), it makes sense that $\lambda(T) = 0$, so you have no incentive to save for the next period, you must consume everything. However, things get more complicated as we progress to the infinite horizon. Without this equation, households will keep accumulating capital indefinitely as capital has value λ and households still have k_t in expense, which implies that a steady state cannot be secured. To prevent this case from happening, condition (5.12) is necessary. The requirement is that at the limit $t \to \infty$, if the terminal capital has a positive present value, then you must consume it all to make $k(t) = 0$. Otherwise, it must have zero valuation λ_t so that you are indifferent about leaving it unexploited. This transversality condition that imposes a non-negativity constraint on capital is also known as the "no-Ponzi scheme" condition

5. Derive the Euler equation relating the dynamics of the choice variable

$$\rho - \frac{u''(c_t)\dot{c}_t}{u'(c_t)} = f'(k_t) - \delta. \tag{5.13}$$

Ultimately, we are interested in the dynamic behavior of the control and state variable over time. This is done by taking (5.9) and differentiating both sides with respect to time (product and chain rule)

$$\dot{\lambda}_t = -\rho e^{-\rho t} u'(c_t) + e^{-\rho t} u''(c_t)\dot{c}_t \tag{5.14}$$

which is the co-state variable. Substituting it to (5.11), we have

$$-\rho e^{-\rho t} u'(c_t) + e^{-\rho t} u''(c_t)\dot{c}_t = -\lambda_t (f'(k_t) - \delta)$$

and using (5.9) for λ_t, we obtain

$$-\rho e^{-\rho t} u'(c_t) + e^{-\rho t} u''(c_t)\dot{c}_t = -e^{-\rho t} u'(c_t)(f'(k_t) - \delta)$$

Dividing both sides by $e^{-\rho t} u'(c_t)$, we obtain (5.13). If we assume a log utility function, then it is easy to obtain the same result with Lagrangian

$$\frac{\dot{c}_t}{c_t} = f'(k_t) - \delta - \rho.$$

5.3 Examples

5.4 Exercises

Chapter 6

Some Dynamic Economic Models

6.1 Ramsey-Cass-Koopsman (RCK)

This section deals with the neoclassical growth model, also known as the Ramsey-Cass-Koopsman, in continuous time. Ramsey (1928) first solved this problem before he died at the age of 26. He was so ahead of his time that nearly 4 decades later, by the works of Cass (1965) and Koopmans (1963), the model became as well-known as today.

6.1.1 Endogenous Labor

The economy solves

$$
\begin{aligned}
& \max \int_0^\infty \beta^t \left[\ln(c_t) + \gamma \ln(1 - l_t)\right] dt, \\
\text{s.t. } & c_t + i_t = y_t, \\
& y_t = f(k_t) = A k_t^\alpha, \\
& \dot{k}_t = i_t - \delta k_t
\end{aligned}
$$

6.2 Real Business Cycle (RBC)

6.3 Programing

Sources:

1. https://python-advanced.quantecon.org/discrete_dp.html
2. https://macroeconomics.github.io/Dynamic%20Programming.html
3. https://github.com/lnsongxf/NumEcon/tree/master/numecon/macro

6.3.1 Backward Induction

6.3.2 Value Function Iteration

6.3.3 Policy Function Iteration

Bibliography

Acemoglu, D. (2008). Introduction to modern economic growth. Princeton university press.

Cagan, P. (1956). The monetary dynamics of hyperinflation. Studies in the Quantity Theory if Money.

Campante, F., Sturzenegger, F., and Velasco, A. (2021). Advanced macroeconomics: an easy guide. LSE Press.

Cass, D. (1965). Optimum growth in an aggregative model of capital accumulation. The Review of economic studies, 32(3):233–240.

Chiang, A. (1984). Fundamental methods of mathematical economics (3rd. ed). MacGraw-Hill.

Chiang, A. (1992). Elements of Dynamic Optimization. MacGraw-Hill.

Chu, A. C. (2021). Advanced Macroeconomics: An introduction for undergraduates. World Scientific.

Dannan, F. M., Elaydi, S. N., and Ponomarenko, V. (2003). Stability of hyperbolic and nonhyperbolic fixed points of one-dimensional maps. Journal of difference equations and Applications, 9(5):449–457.

de la Croix, D. (1996). The dynamics of bequeathed tastes. Economics Letters, 53(1):89–96.

de la Croix, D. (2012). Fertility, Education, Growth, and Sustainability. Cambridge University Press.

de la Croix, D. and Michel, P. (2002). A theory of economic growth: dynamics and policy in overlapping generations. Cambridge University Press.

Evans, G. W. and Honkapohja, S. (2001). Learning and expectations in macroeconomics. Princeton University Press.

Fan, C. S. (2008). Religious participation and children's education: A social capital approach. Journal of Economic Behavior & Organization, 65(2):303–317.

Farmer, K. and Bednar-Friedl, B. (2010). Intertemporal resource economics: An introduction to the overlapping generations approach. Springer.

Farmer, K. and Schelnast, M. (2021). Growth and International Trade: An Introduction to the Overlapping Generations Approach. Springer.

Galor, O. and Ryder, H. E. (1989). Existence, uniqueness, and stability of equilibrium in an overlapping-generations model with productive capital. Journal of Economic Theory, 49(2):360–375.

Heer, B. and Maussner, A. (2009). Dynamic general equilibrium modeling: computational methods and applications. Springer Science & Business Media.

Hirazawa, M. and Yakita, A. (2017). Labor supply of elderly people, fertility, and economic development. Journal of Macroeconomics, 51:75–96.

Kaldor, N. (1934). A classificatory note on the determinateness of equilibrium. The review of economic studies, 1(2):122–136.

Koopmans, T. C. (1963). On the concept of optimal economic growth. Cowles Foundation Discussion Papers, (163).

Lucas, R. E. (1973). Some international evidence on output-inflation tradeoffs. The American economic review, pages 326–334.

McCandless, G. (2008). The ABCs of RBCs: An introduction to dynamic macroeconomic models. Harvard University Press.

Michaillat, P. (2023). Lecture notes for math camp.

Ramsey, F. P. (1928). A mathematical theory of saving. The economic journal, 38(152):543–559.

Romer, D. (2011). Advanced macroeconomics fourth edition. McGraw-Hill.

Shone, R. (2002). Economic Dynamics: Phase diagrams and their economic application. Cambridge University Press.

Simon, C. P. and Blume, L. (1994). Mathematics for economists. Norton New York.

Sydsæter, K., Hammond, P., Seierstad, A., and Strom, A. (2008). Further mathematics for economic analysis. Pearson education.

Sydsæter, K. and Hammond, P. J. (2008). Essential mathematics for economic analysis. Pearson Education.

Tran, Q.-T. (2022). The aging tax on potential growth in asia. Journal of Asian Economics, 81:101495.

Tran, Q.-T. and Zhang, Y. (2023). Inseikai tohoku spring camp 2023 — Mathematics I: Algebra, calculus and static optimization.

Varian, H. R. (1992). Microeconomic analysis. Norton New York, 3 edition.

Appendix A

Cheatsheets

A.1 Differentiation

See Tran and Zhang (2023, Ch. 5).

Basic Rules

$$(cf)' = cf' \qquad (f+g)' = f' + g'$$

$$(fg)' = f'g + fg' \qquad (f/g)' = \frac{f'g - fg'}{g^2} \quad (g \neq 0)$$

$$(f \circ g)' = f'(g(x)) \cdot g'(x)$$

Common Derivatives

$$\frac{d}{dx}(x^n) = nx^{n-1} \qquad \frac{d}{dx}(e^x) = e^x \qquad \frac{d}{dx}(\ln x) = \frac{1}{x} \qquad \frac{d}{dx}\left(\frac{1}{x}\right) = -\frac{1}{x^2}$$

Chain Rule

$$\frac{d}{dx}(f(g(x))) = f'(g(x)) \cdot g'(x)$$

Product Rule

$$\frac{d}{dx}(uv) = u'v + uv'$$

Quotient Rule

$$\frac{d}{dx}\left(\frac{u}{v}\right) = \frac{u'v - uv'}{v^2} \quad (v \neq 0)$$

Power Rule

$$\frac{d}{dx}(x^n) = nx^{n-1}$$

Exponential Rule

$$\frac{d}{dx}(e^{u(x)}) = u'(x)e^{u(x)}$$

A.2 Integration

See Tran and Zhang (2023, Ch. 8)

Basic Integrals

$$\int k\,dx = kx + C \quad \text{(where } k \text{ is a constant)}$$
$$\int x^n\,dx = \frac{1}{n+1}x^{n+1} + C \quad \text{(where } n \neq -1\text{)}$$
$$\int e^x\,dx = e^x + C$$
$$\int \frac{1}{x}\,dx = \ln|x| + C$$

Common Integrals

$$\int e^{ax}\,dx = \frac{1}{a}e^{ax} + C$$
$$\int \ln x\,dx = x\ln x - x + C$$

Integration by Parts

$$\int u\,dv = uv - \int v\,du$$

A.3 Taylor Series Formula

See Tran and Zhang (2023, Ch. 6.6). The Taylor series expansion of a function $f(x)$ at a point a is given by:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \cdots$$

Or, in sigma notation:

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

Order 1

The Taylor series formula of order 1 for a function $f(x)$ centered at a is given by:

$$f(x) = f(a) + f'(a)(x-a)$$

Order 2

The Taylor series formula of order 2 for a function $f(x)$ centered at a is given by:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2$$

A.4 Implicit Function Theorem

See Tran and Zhang (2023, Ch 10.6). Let $F(x, y)$ be a continuously differentiable function defined in a neighborhood of a point (a, b). If $F(a, b) = 0$ and $\frac{\partial F}{\partial y} \neq 0$ at (a, b), then there exists an open interval I containing a and an open interval J containing b, and a unique continuously differentiable function $f : I \to J$, such that for all x in I, the equation $F(x, f(x)) = 0$ holds.

Furthermore, the derivative $f'(x)$ where $y = f(x)$ is given by:

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}$$

Sometimes, when explicit solutions are not attainable, you can perform comparative statics with this theorem.

Example A.1. Find $y'(x)$ when $xy = 5$.

Let $F(x, y) = xy$. Then $F'_x = y, F'_y = x$. For $x \neq 0$, the IFT says that

$$y' = -\frac{F'_x}{F'_y} = -\frac{y}{x}$$

A.5 Intermediate Value Theorem

See Tran and Zhang (2023, Ch 6.10). Let f be a continuous function on the closed interval $[a, b]$. If y is any number between $f(a)$ and $f(b)$, then there exists at least one number c in the open interval (a, b) such that $f(c) = y$.

Application: Mostly to prove the existence of a solution.

We can use the Intermediate Value Theorem (IVT) to show that certain equations have solutions, or that certain polynomials have roots. For instance, the polynomial $f(x) = x^4 + x - 3$ is complicated, and finding its roots is very complicated. However, it's easy to check that $f(-1) = -3$ and $f(2) = 15$. Since $-3 < 0 < 15$, there has to be a point c between -1 and 2 with $f(c) = 0$. In other words, $f(x)$ has a root somewhere between -1 and 2. We don't know where, but we know it exists.

In a more general concept, if you need to solve

$$f(x) = g(x)$$

Sometimes, solving it is difficult. Instead, you can use numerical methods (that is, let the computer do the hard part). However, you may still want to prove the existence of such an x^*. Then you can show that $f(x)$ is increasing from $[-\infty, \infty]$, while $g(x)$ is decreasing from $[-\infty, \infty]$. Thus, they must cross somewhere, and that somewhere is x^*. And if f, g are monotone, then this x^* is unique.

Example A.2. Prove that the equation

$$2x - 5e^{-x}(1 + x^2) = 0$$

has a unique solution, which lies in the interval (0, 2).

Solution:

Define $g(x) = 2x - 5e^{-x}(1 + x^2)$. Then $g(0) = -5$ and $g(2) = 4 - 25/e^2$. In fact $g(2) > 0$ because $e > 5/2$. According to the intermediate value theorem, therefore, the continuous function g must have at least one zero in $(0, 2)$. Moreover, note that $g'(x) = 2 + 5e^{-x}(1 + x^2) - 10xe^{-x} = 2 + 5e^{-x}(1 - 2x + x^2) = 2 + 5e^{-x}(x - 1)^2$.But then $g'(x) > 0$ for all x, so g is strictly increasing. It follows that g can have only one zero.

A.6 Matrix Algebra

See Tran and Zhang (2023, Ch. 9)

A.6.1 Matrix Operations

Addition and Subtraction

$$A + B = B + A \quad \text{(Commutative)}$$

$$(A + B) + C = A + (B + C) \quad \text{(Associative)}$$

Scalar Multiplication

$$c(A + B) = cA + cB$$

$$(c + d)A = cA + dA$$

Matrix Multiplication

$$A(BC) = (AB)C \quad \text{(Associative)}$$

$$A(B + C) = AB + AC \quad \text{(Distributive)}$$

$$(cA)B = A(cB) = c(AB)$$

A.6.2 Matrix Multiplication

Given matrices $A_{m\times p}$ and $B_{p\times n}$, their matrix product $C = AB$ is defined by the formula:

$$C_{ij} = \sum_{k=1}^{p} A_{ik} \cdot B_{kj}$$

where $1 \leq i \leq m$ and $1 \leq j \leq n$.

In other words, the entry in the i-th row and j-th column of C is obtained by multiplying the elements in the i-th row of A with the corresponding elements in the j-th column of B, and then summing up these products.

A.6.3 Transpose

A transpose $\mathbf{A}^T$ of a $k \times n$ matrix is the $n \times k$ matrix obtained by interchanging rows and columns of $\mathbf{A}$. Notation: $\mathbf{A}^T$ or $\mathbf{A}'$.

Definition A.1 (Rules for Transposition)**.**

$$\begin{aligned}
(\mathbf{A} + \mathbf{B})^T &= \mathbf{A}^T + \mathbf{B}^T, \\
(\mathbf{A} - \mathbf{B})^T &= \mathbf{A}^T - \mathbf{B}^T, \\
(\mathbf{A}^T)^T &= \mathbf{A}, \\
(r\mathbf{A})^T &= r\mathbf{A}^T, \\
(\mathbf{AB})^T &= \mathbf{B}^T\mathbf{A}^T.
\end{aligned}$$

A matrix is **symmetric** iff $\mathbf{A} = \mathbf{A}'$

A.6.4 Determinants

Order 2

The determinant of a 2×2 matrix is given by

$$|\mathbf{A}| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - cb$$

Order 3

The determinant of a 3×3 matrix is given by:

$$\begin{aligned} |A| &= \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \\ &= a_{11}[a_{22}a_{33} - a_{32}a_{23}] - a_{12}[a_{21}a_{33} - a_{31}a_{23}] + a_{13}[a_{21}a_{32} - a_{31}a_{22}] \end{aligned}$$

For higher order, see Tran and Zhang (2023, Ch. 9.4.2).

A.7 Jacobians, Gradients, and Hessians

The Gradient ∇ is a vector that points in the direction of the steepest increase of a scalar-valued (1D) function of multiple variables. Its magnitude $|\nabla|$ is the rate of change in that direction. The gradient is a vector and has the same dimension as the number of variables in the function. For a function of n variables, the gradient is an n-dimensional vector.

The Jacobian is a matrix that represents the collection of all first-order partial derivatives of a vector-valued function with respect to multiple variables. It provides information about how each component of the vector function changes as the variables change. The Jacobian is a matrix whose size is determined by the number of components in the vector-valued function and the number of variables. For a function with m components (functions) and n variables, the Jacobian is an $m \times n$ matrix.

In summary, the gradient is a vector that describes the rate of change of a scalar-valued (1D) function, while the Jacobian is a matrix that describes the rate of change of a vector-valued (n-D) function.

The Hessian is just a matrix of second-order mixed partials of a scalar field (that is, gradients). The Hessian matrix is symmetric. This means that the element in row i and column j is the same as the element in row j and column i. The mixed partial derivatives in the Hessian matrix satisfy the equality of mixed partials $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$. The eigenvalues of the Hessian matrix are indicators of the curvature of the function at a critical point. Positive eigenvalues suggest a local minimum, negative eigenvalues suggest a local maximum, and a mix of positive and negative eigenvalues suggest a saddle point. For more detail, see Tran and Zhang (2023, Ch 10).

1D (1 dimension)

Let $f(x)$ be a one-dimensional function. The Gradient ∇ is a vector of the first order derivative of f with respect to x mapping $\mathbb{R}^1 \to \mathbb{R}^1$ (a scalar field)

$$\nabla = \frac{df}{dx}$$

The Hessian is a matrix of second-order mixed partials of a scalar field.

$$\mathbf{H} = \frac{d^2 f}{dx^2}$$

The Jacobian is a matrix of gradients for components of a vector field (in this case, only 1 component), thus

$$\mathbf{J} = \frac{dF}{dx} = \frac{df}{dx}$$

So in 1D, Jacobian and Gradient are the same.

2D (2 dimensions)

Let $f(x, y)$ be a two-dimensional function. The Gradient is defined as:

$$\nabla = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \end{bmatrix}$$

The Hessian is

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix}$$

Consider $f(x, y)$ and $g(x, y)$ be a system of two-dimensional functions. The Jacobian matrix $\mathbf{J}$ for this system is defined as:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{bmatrix}$$

3D (3 dimensions)

Let $f(x, y, z)$ be a three-dimensional function. The Gradient is:

$$\nabla = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \end{bmatrix}$$

The Hessian is

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{bmatrix}$$

Consider $f(x, y, z), g(x, y, z), h(x, y, z)$ a system of three-dimensional functions. The Jacobian matrix $\mathbf{J}$ for this system is defined as:

$$\mathbf{J} = \begin{bmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} & \frac{\partial g}{\partial z} \\ \frac{\partial h}{\partial x} & \frac{\partial h}{\partial y} & \frac{\partial h}{\partial z} \end{bmatrix}$$

In a sense, The Hessian is the Jacobian of the gradient of a function that maps from n-dimension to 1-Dimension.

Appendix B

Dynamical Stability Analysis

Borrow from de la Croix and Michel (2002), Technical Appendix A.2. (p.311) and Dannan et al. (2003).

B.1 Dimension One

Let $f(x)$ be a function defined on some interval I of $\mathbb{R}$ with values in I. The time path, given an initial state $x_0 \in I$ and the equation $x_{t+1} = f(x_t)$ is uniquely defined.

A steady state solution $\bar{x}$ to $\bar{x} = f(\bar{x})$ which is interior to I is locally stable if for any initial value x_0 near enough to $\bar{x}$, the dynamics starting from x_0 converge to $\bar{x}$. Formally, there exists $\varepsilon > 0$ such that $(\bar{x} - \varepsilon, \bar{x} + \varepsilon) \in I$ and for any $x_0 \in (\bar{x} - \varepsilon, \bar{x} + \varepsilon)$ the corresponding dynamics satisfy

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

At a corner steady state like 0, when f is defined on $\mathbb{R}_{++}$, the corner local stability of 0 is defined similarly but for $x_0 \in (0, \varepsilon)$.

Definition B.1 (Hyperbolicity)**.** Assume f is continuously differentiable in I, $\bar{x}$ is a steady state $\in I$. If $f'(\bar{x}) = 1$, then $\bar{x}$ is non-hyperbolic. Its stability type cannot be determined on the basic of its first-order derivative, but only by analyzing the second-order derivatives.

Otherwise, if $f'(\bar{x}) \neq 1$, then $\bar{x}$ is hyperbolic.

Appendix C

Solutions to Some Problems

C.1 OLG Model

The problem is

$$\max_{s_t, n_t} \ln(w_t(1 - \phi n_t) - s_t) + \beta \ln(R_{t+1} s_t) + \gamma \ln(n_t)$$

FOCs

$$\begin{aligned}
\frac{\partial U_t}{\partial s_t} &= -\frac{1}{w_t(1 - \phi n_t) - s_t} + \frac{\beta}{s_t} = 0, \\
\frac{\partial U_t}{\partial n_t} &= -\frac{\phi w_t}{w_t(1 - \phi n_t) - s_t} + \frac{\gamma}{n_t} = 0,
\end{aligned}$$

which implies that

$$\begin{aligned}
(s_t) : s_t &= \beta c_t, \\
(n_t) : n_t &= \frac{\gamma c_t}{\phi w_t}
\end{aligned}$$

Using the budget constraint, we can derive

$$\begin{aligned}
s_t^* &= \frac{\beta}{1 + \beta + \gamma} w_t, \\
n_t^* &= \frac{\gamma}{\phi(1 + \beta + \gamma)}.
\end{aligned}$$

The Hessian

$$\mathbf{H} = \begin{pmatrix} -\frac{1}{\Gamma^2} - \frac{\beta}{s_t^2} & -\frac{\phi w_t}{\Gamma^2} \\ -\frac{\phi w_t}{\Gamma^2} & -\frac{\phi^2 w_t^2}{\Gamma^2} - \frac{\gamma}{n_t^2} \end{pmatrix}$$

where $\Gamma = w_t(1 - \phi n_t) - s_t$. The leading principal minors are

$$\begin{aligned}
|\mathbf{H}_1| &= -\frac{1}{\Gamma^2} - \frac{\beta}{s_t^2} < 0, \\
|\mathbf{H}| &= \begin{vmatrix} -\frac{1}{\Gamma^2} - \frac{\beta}{s_t^2} & -\frac{\phi w_t}{\Gamma^2} \\ -\frac{\phi w_t}{\Gamma^2} & -\frac{\phi^2 w_t^2}{\Gamma^2} - \frac{\gamma}{n_t^2} \end{vmatrix} = \frac{\gamma}{\Gamma^2 n_t^2} + \frac{\beta}{s_t^2} \left(\frac{\phi^2 w_t^2}{\Gamma^2} + \frac{\gamma}{n_t^2} \right) > 0
\end{aligned}$$

so the solutions obtained at the FOCs are sufficient.

We move on to the firm's problem

$$\pi_t = K_t^\alpha L_t^{1-\alpha} - R_t K_t - w_t L_t$$

The FOCs are

$$\frac{\partial \pi_t}{\partial K_t} = 0 \Leftrightarrow R_t = \alpha K_t^{\alpha-1} L_t^{1-\alpha} \equiv \alpha k_t^{\alpha-1} = f'(k_t),$$
$$\frac{\partial \pi_t}{\partial L_t} = 0 \Leftrightarrow w_t = (1-\alpha) K_t^{\alpha} L_t^{-\alpha} \equiv (1-\alpha) k_t^{\alpha} = f(k_t) - f'(k_t)k_t.$$

where

$$k_t = K_t/L_t,$$
$$y_t = Y_t/K_t = k_t^\alpha = f(k_t)$$

1. The law of motion of capital

$$k_{t+1} = \frac{s_t L_t}{L_{t+1}} = \frac{s_t(w_t(k_t), R_{t+1})}{1+n} \tag{C.1}$$

2. Existence: We need to show that there is a solution to (C.1). Let us rewrite it to

$$\triangle(k_t, w_t) \equiv (1+n)k_{t+1} - s_t(w_t(k_t), R_{t+1}) = 0$$

Since saving is always smaller than w_t by a factor $\beta/(1+\beta+\gamma)$, one has

$$0 < s_t(w_t(k_t), R_{t+1}) < w_t$$
$$\Leftrightarrow 0 < \frac{s_t(w_t(k_t), R_{t+1})}{k_t} < \frac{w_t}{k_t}$$

C.2 Endogenous Retirement

A representative agent's problem:

$$\max_{c_t, d_{t+1}, l_{t+1}} U_t = \ln(c_t) + \beta\pi \ln(d_{t+1}) + \gamma\pi \ln(l_{t+1})$$

s.t.

$$(1-\tau)w_t + \frac{\pi}{R_{t+1}}\left[(1-l_{t+1})(1-\tau)\varepsilon w_{t+1} + l_{t+1}p_{t+1}\right] - c_t - \frac{\pi}{R_{t+1}}d_{t+1} = 0$$
$$1 - l_{t+1} \geq 0$$

Lagrangian: (with $\lambda_i, \quad i = \{1,2\}$ as Lagrangian multipliers):

$$\mathcal{L} = \ln(c_t) + \beta\pi \ln(d_{t+1}) + \gamma\pi \ln(l_{t+1})$$
$$+ \lambda_{1,t}\left\{(1-\tau)w_t + \frac{\pi}{R_{t+1}}\left[(1-l_{t+1})(1-\tau)\varepsilon w_{t+1} + l_{t+1}p_{t+1}\right] - c_t - \frac{\pi}{R_{t+1}}d_{t+1}\right\}$$
$$+ \lambda_{2,t+1}(1-l_{t+1})$$

Karesh-Kuhn-Tucker (KKT) First-order conditions:

- $\mathcal{L}'(c_t) = \dfrac{1}{c_t} - \lambda_{1,t} = 0 \Leftrightarrow c_t = \dfrac{1}{\lambda_{1,t}}$
- $\mathcal{L}'(d_{t+1}) = \dfrac{\beta\pi}{d_{t+1}} - \dfrac{\lambda_{1,t}\pi}{R_{t+1}} = 0 \Leftrightarrow d_{t+1} = \dfrac{\beta R_{t+1}}{\lambda_{1,t}} = \beta R_{t+1} c_t$
- $\mathcal{L}'(l_{t+1}) = \dfrac{\gamma\pi}{l_{t+1}} - \lambda_{1,t}\pi\left[\dfrac{(1-\tau)\varepsilon w_{t+1} - p_{t+1}}{R_{t+1}}\right] - \lambda_{2,t+1} = 0$

$\Leftrightarrow \dfrac{\pi\gamma}{l_{t+1}} = \dfrac{\pi}{R_{t+1}c_t}\left[(1-\tau)\varepsilon w_{t+1} - p_{t+1}\right] + \lambda_{2,t+1}$

$\Leftrightarrow \dfrac{\gamma}{l_{t+1}} = \dfrac{(1-\tau)\varepsilon w_{t+1} - p_{t+1}}{R_{t+1}c_t} + \dfrac{\lambda_{2,t+1}}{\pi}$

- $\lambda_{1,t} > 0$
- Complimentary slackness:

$$\begin{cases} \lambda_{2,t+1} \geq 0 \\ (1 - l_{t+1}) \geq 0 \\ \lambda_{2,t+1}(1 - l_{t+1}) = 0 \end{cases} \Rightarrow \begin{cases} l_{t+1} = 1, \lambda_{2,t+1} > 0 \\ l_{t+1} < 1, \lambda_{2,t+1} = 0 \end{cases}$$

The Euler equations:

$$d_{t+1} = \beta R_{t+1} c_t \tag{C.2}$$

$$\frac{\gamma}{l_{t+1}} \geq \frac{(1-\tau)\varepsilon w_{t+1} - p_{t+1}}{R_{t+1}c_t} \tag{C.3}$$

with equality holds if $l_{t+1} < 1$

The threshold level of $\hat{\varepsilon}$ therefore can be derive by letting $l_{t+1} = 1$ where

$$\hat{\varepsilon} = \frac{\gamma R_{t+1} c_t + p_{t+1}}{(1-\tau)w_{t+1}}$$

so that if $\varepsilon \leq \hat{\varepsilon}$, agents choose $l_{t+1} = 1$ to retire fully (not working in the second half of life). Otherwise, they choose $l_{t+1} < 1$ and work a portion of time in the second period of life.