Lecture Notes: Mathematics for Economics

Cuong Le Van, Ngoc-Sang Pham

Lecture Notes: Mathematics for Economics

Cuong Le Van* and Ngoc-Sang Pham†

September 3, 2021
(Link to get the latest version)

Abstract

We present some mathematical tools widely used in courses taught in (under)graduate programs in economics. We hope that readers can learn how to apply mathematical results in economics and how to prove them. We focus on two topics: finite-dimensional convex optimization and discrete-time dynamical systems.[1] We also present several applications in economics.

*CNRS, PSE, CASED. Email: Cuong.Le-Van@univ-paris1.fr

†EM Normandie Business School, Métis Lab. *Emails*: npham@em-normandie.fr, pns.pham@gmail.com. Webpage: https://phamns.wordpress.com/

[1]We do not cover linear algebra in this lecture. For more complete treatments of mathematics for economics, see Simon and Blume (1994), Hoy et al. (2001).

Contents

1 Some basics of mathematical analysis

An excellent introduction of mathematical analysis can be found in Rudin (1976). In the following, we only cover essential notions and results that are extremely useful in economics and will be used for the next sections.

1.1 Bounds, inferior, superior

Definition 1. *A greatest element of a subset S of a partially ordered set $(P, \leq)$[2] is an $a \in S$ satisfying $b \leq a \ \forall b \in S$.*

A least element of a subset S of a partially ordered set $(P, \leq)$ is an $a \in S$ satisfying $a \leq b \ \forall b \in S$.

Definition 2. *A lower bound of a subset S of a partially ordered set $(P, \leq)$ is an element a of P such that: $a \leq x$ for all $x \in S$.*

A lower bound a of S is called an infimum (or greatest lower bound, or meet; abbreviated inf) of S if, for all lower bounds y of S in P, $y \leq a$ (a is larger than or equal to any other lower bound).

Similarly, an upper bound of a subset S of a partially ordered set $(P, \leq)$ is an element b of P such that $b \geq x \forall x \in S$.

An upper bound b of S is called a supremum (or least upper bound, or join; abbreviated sup) of S if, for all upper bounds z of S in P, $z \geq b$ (b is less than any other upper bound).

Exercise 1. *Let $S = (0, 1)$ in $(\mathbb{R}, \leq)$.*

(i) Prove that $\inf S = 0$ and $\sup S = 1$.

(ii) Prove that there is neither any greatest element nor least element of S.

Proposition 1. *Let $A \subset \mathbb{R}$ be a nonempty set. The $\sup(A)$ is the unique element (eventually $+\infty$) such that :*

(i) If $x > \sup(A)$ then $x \notin A$

(ii) if $x < \sup(A)$ then there exists $a \in A$ such that $x < a$

(iii) There exists a sequence in A which converges to $\sup(A)$.

Proof. (i) If $x \in A$ then $x \leq \sup(A)$ since this one is an upper bound of A. The result is then obvious.

(ii) If not, for any $a \in A, a \leq x$ and x is an upper bound which is smaller than the smallest upper bound which is $\sup(A)$.

(iii) First suppose that $\sup(A)$ is finite. Then for any $k \in \mathbb{N}$, there exists $a^k \in A$ which satisfies $\sup(A) - \frac{1}{k} < a^k \leq \sup(A)$. The sequence $\{a^k\}_k$ converges to $\sup(A)$. Now suppose $\sup(A) = +\infty$.The set A is then unbounded from above. Hence, for any $k \in \mathbb{N}$, there exists $a^k \in A$ with $a^k \geq k$. Obviously, the sequence $\{a^k\}_k$ converges to $+\infty = \sup(A)$. □

Proposition 2. *Let $A \subset \mathbb{R}$ be a nonempty set. The $\inf(A)$ is the unique element (eventually $-\infty$) such that :*

(i) If $x < \inf(A)$ then $x \notin A$

(ii) if $x > \inf(A)$ then there exists $a \in A$ such that $x > a$

(iii) There exists a sequence in A which converges to $\inf(A)$.

[2] The set P and the binary relation $\leq$ constitute a partially ordered set if the binary $\leq$ is reflexive ($x \leq x$), transitive ($x \leq y$ and $y \leq z$ imply that $x \leq z$) and anti-symmetric ($x \leq y$ and $y \leq x$ imply that $x = y$).

1.2 Sequence and limit

A sequence is a function whose domain is the positive numbers. We write $(x_i)_i = (x_1, x_2, \dots,)$. The value x_i may be in $\mathbb{R}, \mathbb{R}^N$ or in other spaces.

A subsequence $(x_{k_n})_n$ of $(x_k)_k$ is an infinite sequence $x_{k_1}, x_{k_2}, \dots$ where $k_1, k_2, \dots$ is an infinite increasing sequence of integers.

We now focus on the case $x \in \mathbb{R}^N \equiv \{(a_1, a_2, \dots, a_N) : a_i \in \mathbb{R}, \forall i = 1, \dots, N\}$. We say that a sequence $(x_i)_i$ has the limit x if, for any $\epsilon > 0$, there is a positive integer n such that $\|x_i - x\| \leq \epsilon \ \forall i \geq n$, where $\|a - b\| \equiv \sqrt{\sum_{i=1}^N (a_i - b_i)^2}$ is the distance between vectors a and b.

Definition 3. *A sequence having the limit is said to be convergent (in this case, we write $\lim_{n\to\infty} x_n = x$). A sequence is divergent if it is not convergent.*

Definition 4. *A sequence of real numbers $\{x^k\}_k$ converges to $+\infty$ (notation $x^k \to +\infty$) if: for any $A \in \mathbb{R}$, there exists K such that, if $k \geq K$ then $x^k \geq A$.*

Exercise 2. *Prove that $\|a \cdot b\| \leq \|a\|\|b\|$ where $a \cdot b \equiv (a_1 b_1, \dots, a_N b_N)$.*

Exercise 3. *The above distance $\|\cdot\|$ between two vectors satisfies the Triangle Inequality:*

$$\|x + y\| \leq \|x\| + \|y\|$$

and

$$\|x - z\| \leq \|x - y\| + \|y - z\|$$

Theorem 1. *A sequence in $\mathbb{R}^n$ has at most one limit.*

Proof. Assume $\{x_k\}_k$ in $\mathbb{R}^n$ has two limits a, b. Then, given $\varepsilon > 0$, one can find N_a, N_b such that, for any $n \geq N_a$, one has $\|x_n - a\| \leq \varepsilon$ and any $n \geq N_b$, one has $\|x_n - b\| \leq \varepsilon$. In this case, if we take $N \geq \max\{N_a, N_b\}$, we have

$$\|a - b\| = \|a - x_N + x_N - b\| \leq \|a - x^N\| + \|x^N - b\| \leq 2\varepsilon \quad \forall \epsilon$$

It means that $\|a - b\| \leq 2\epsilon \ \forall \epsilon > 0$. So, we have $\|a - b\| = 0$ and hence $a = b$. □

Theorem 2. *A sequence of points $\{x^k\}_k$ in $\mathbb{R}^n$ converges to $x = (x_1, x_2, \dots, x_n)$, where $x^k = (x_1^k, x_2^k, \dots, x_n^k)$, if, and only if, $x_i^k \to x_i$ for any $i = 1, \dots, n$.*

Proof. Exercise □

Remark 1. *Prove that If the real sequences (a_n) and (b_n) converge to $x \in \mathbb{R}$, then any sequence (c_n) satisfying $a_n \leq c_n \leq b_n \ \forall n$ also converges to x.*

Properties 1. *Suppose that two real-value sequences x_n and y_n are convergent with limits x and y respectively. We have*

1. $\lim_{n\to\infty} c x_n = cx \ \forall c \in \mathbb{R}$.
2. $\lim_{n\to\infty} (x_n + y_n) = x + y$.
3. $\lim_{n\to\infty} x_n y_n = xy$.

4. $\lim_{n\to\infty}(x_n/y_n) = x/y$ *if* $y \neq 0$.

Exercise 4 (Theorem 3.20 (Rudin, 1976))**.** *Prove that*

1. If $p > 0$*, then* $\lim_{n\to\infty} 1/n^p = 0$.

2. If $p > 0$*, then* $\lim_{n\to\infty} p^{1/n} = 1$.[3]

3. $\lim_{n\to\infty} n^{1/n} = 1$.

4. If $p > 0$*, and* $\alpha \in \mathbb{R}$*, then* $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$.[4]

5. If $|x| < 1$*, then* $\lim_{n\to\infty} x^n = 0$.

Proof. Point 4. Recall the binomial theorem: $(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k}y^k = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$ where $\begin{pmatrix} n \\ k \end{pmatrix} \equiv \frac{n!}{k!\,(n-k)!}$. By applying this theorem, we have

$$\begin{aligned}(1+p)^n &= \sum_{k=0}^n \begin{pmatrix} n \\ k \end{pmatrix} 1^{n-k}p^k = \sum_{k=0}^n \frac{n(n-1)\cdots(n-k+1)}{k!}p^k \\ &\geq \frac{n(n-1)\cdots(n-k+1)}{k!}p^k \quad \forall k \leq n\end{aligned}$$

Given p and α, choose $k > \alpha, k > 0$. Choose $n > 2k$, we have

$$(1+p)^n \geq \frac{n(n-1)\cdots(n-k+1)}{k!}p^k = n(n-1)\cdots(n-k+1)\frac{p^k}{k!} > \frac{n^k}{2^k}\frac{p^k}{k!}$$

because $n-k+1 > n/2$. Therefore,

$$\frac{n^\alpha}{(1+p)^n} < \frac{2^k k!}{p^k} n^{\alpha-k}$$

Since $\alpha - k < 0$, the right hand side converges to zero when n goes to infinity. So, $\lim_{n\to\infty} \frac{n^\alpha}{(1+p)^n} = 0$. □

Exercise 5. *Find the limit (if it exists) of the following real-value sequences:*

1. $x_n = (-1)^n/n$.

2. $x_n = (-1)^n$.

Definition 5. *A real-value sequence is increasing (respectively, strictly increasing) if* $x_{n+1} \geq x_n\ \forall n$ *(respectively,* $x_{n+1} > x_n\ \forall n$*). It is decreasing (respectively, strictly decreasing) if* $x_{n+1} \leq x_n\ \forall n$ *(respectively,* $x_{n+1} < x_n\ \forall n$*).*

A sequence is bounded if it has a lower bound and an upper bound.

Theorem 3 (monotone convergence theorem)**.** *A monotonic sequence (with real values) is convergent if and only if it is bounded.*

[3]Hint: Define $x \equiv p^{1/n} - 1$. Observe that $(1+x)^n > 1 + nx\ \forall x > 0$ and n is integer.

[4]Hint: Let k be an integer and higher than α. Prove that $(1+p)^n > \frac{n^k p^k}{2^k k!}\ \forall n > 2k$.

Proof. Consider a monotonically increasing sequence (x_n).

Suppose that (x_n) is convergent to x. It is easy to see that x_n is bounded by x.

Suppose now that (x_n) is bounded. In this case, we have $\sup_n x_n < \infty$. We will prove that it is convergent and $\lim_{n\to\infty} x_n = \sup_n x_n$. We do so by using the definition. Denote $c \equiv \sup_n x_n$. Let ϵ be strictly positive. By definition of supremum, there exists an element x_m such that $c - x_m < \epsilon$ (indeed, otherwise $c - \epsilon \geq x_m \ \forall m$, which means that $c - \epsilon$ is an upper bound of (x_n) and smaller than c, a contradiction). We have $|x_n - c| = c - x_n \ \forall n$ because $x_n \leq c$.

For $n \geq m$, we have $|x_n - c| = c - x_n \leq c - x_m \leq \epsilon$. So, $\lim_{n\to\infty} x_n = \sup_n$.

□

Example 1. *(Growth, Interest rates). At date 0 (initial date), the GDP per capita of countries a and b are y_a and y_b respectively with $y_a < y_b$.*

Denote $y_{i,t}$ the GDP per capita of country i at period t.

Assume that the rate of growth of country i is r_i which is constant over time, for $i = a, b$.

Assume that $r_a > r_b > 0$.

Prove that

1. *$y_{a,t} - y_{b,t}$ converges to infinity*
2. *and there is a date t_0 such that $y_{a,t} - y_{b,t} \leq 0 \ \forall t \leq t_0$ and $y_{a,t} - y_{b,t} > 0 \ \forall t > t_0$.*

Example 2 (Compound interest)**.** *(1) The value of P usd invested at an annual rate of interest r compounded n times per year is*

$$V_n = P\Big(1 + \frac{r}{n}\Big)^n$$

When $n \to \infty$, we define

$$V = \lim_{n\to\infty} V_n = \lim_{n\to\infty} P\Big(1 + \frac{r}{n}\Big)^n.$$

This is the value of P usd invested for one year at an interest rate of r with continuous compounding.

Prove that

$$V = Pe^r.$$

where e is the natural number. It approximately equals 2.71828. Recall the definitions of e.

$$e = \lim_{n\to\infty} P\Big(1 + \frac{1}{n}\Big)^n$$
$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1} + \frac{1}{1\cdot 2} + \frac{1}{1\cdot 2\cdot 3} + \cdots$$

(2) The value of P usd invest for t years becomes

$$V_{n,t} = P\Big[\Big(1 + \frac{r}{n}\Big)^n\Big]^t$$

Prove that

$$V_t \equiv \lim_{n\to\infty} V_{n,t} = Pe^{rt}$$

Definition 6. *The limit inferior of a sequence (x_n) is defined by*

$$\liminf_{n\to\infty} x_n \equiv \lim_{n\to\infty}\Big(\inf_{m\geq n} x_m\Big) \text{ or } \liminf_{n\to\infty} x_n \equiv \sup_{n\geq 0}\Big(\inf_{m\geq n} x_m\Big)$$

The limit superior of a sequence (x_n) is defined by

$$\limsup_{n\to\infty} x_n \equiv \lim_{n\to\infty}\Big(\sup_{m\geq n} x_m\Big) \text{ or } \limsup_{n\to\infty} x_n \equiv \inf_{n\geq 0}\Big(\sup_{m\geq n} x_m\Big)$$

Note that the sequence $a_n \equiv \inf_{m\geq n} x_m$ is increasing while the sequence $b_n \equiv \sup_{m\geq n} x_m$ is decreasing.

Exercise 6. *Prove that* $\limsup_{n\to\infty}(-1)^n = 1$ *and* $\liminf_{n\to\infty}(-1)^n = -1$.

Exercise 7. *Suppose that* $\liminf_{n\to\infty} x_n > a$. *Prove that there exists a positive integer* $n_0 > 0$ *such that* $x_n > a\ \forall n > n_0$.

1.3 Series

Definition 7. *Let (a_t) be a sequence.* $s_n \equiv \sum_{t=}^{n} a_t$ *is called a series.*

Example: Present value of a stream of payments. Present value of incomes.

Remark 2. *1. If the series converges, then* $\lim_{n\to\infty} a_n = 0$.

2. The converse may not hold.

We introduce "comparison test" which is very useful.

Proposition 3. *1. If there exists N_0 such that $|a_n| \leq c_n\ \forall n \geq N_0$, and if $\sum c_n$ converges, then $\sum_n a_n$ converges*

2. If there exists N_0 such that $a_n \geq d_n\ \forall n \geq N_0$, and if $\sum d_n$ diverges, then $\sum_n a_n$ diverges

Proposition 4. *1. The series converges if* $\limsup_{n\to\infty}|\frac{a_{n+1}}{a_n}| < 1$.

2. The series diverges if there exists n_0 such that $|\frac{a_{n+1}}{a_n}| \geq 1\ \forall n \geq n_0$.

3. The series diverges if $\liminf_{n\to\infty}|\frac{a_{n+1}}{a_n}| > 1$.

Exercise 8. *Prove that*

1. $\sum_{t=1}^{n} k = \frac{n(n+1)}{2}$
2. $\sum_{t=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.
3. $\sum_{t=1}^{n} k^3 = \frac{n^2(n+1)^2}{4}$.
4. $\sum_{t=1}^{n} k^4 = \frac{n(6n^4+15n^3+10n^2-1)}{30}$.

Exercise 9. *Prove that*

1. $\sum_{k=0}^{n-1} a^k = \frac{a^n-1}{a-1}$ *if* $a \neq 1$, *and equals* n *if* $a = 1$.

2. $\sum_{k=1}^{n-1} a^k = \frac{a^n-a}{a-1}$ *if* $a \neq 1$, *and equals* $n-1$ *if* $a = 1$.

3. $\sum_{k=1}^{n} ka^k = \frac{(a-1)(n+1)a^{n+1}-a^{n+2}+a}{(a-1)^2}$ *if* $a \neq 1$.

4. $\sum_{k=0}^{\infty} a^k = \frac{1}{1-a}$ *if* $a \in (0,1)$, *and equals* ∞ *if* $a \geq 1$.

5. *(Present Value of a Stream of Payments) Given that the interest rate is strictly positive* $(r > 0)$, *prove that*
$$\sum_{k=1}^{\infty} \frac{V}{(1+r)^k} = \frac{V}{r}.$$

Exercise 10. *Prove that*

1. $\sum_{n=1}^{\infty} \frac{1}{n^a} < \infty$ *if* $a > 1$.

2. $\sum_{n=1}^{\infty} \frac{1}{n^a} = \infty$ *if* $a \leq 1$. *In particular,* $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$.

Exercise 11. *Prove that*
$$e \equiv \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + \frac{1}{1} + \frac{1}{1 \cdot 2} + \frac{1}{1 \cdot 2 \cdot 3} + \cdots < \infty$$

Exercise 12. *Consider a positive sequence* (a_n). *Prove that*
$$\lim_{n\to\infty} \prod_{i=1}^{n} (1 + a_i) < \infty \text{ if and only if } \sum_{i=1}^{n} a_i < \infty.$$

1.4 Open, closed and compact sets

For $x \in \mathbb{R}^n$, the *open ball* $B(x,r)$ with *centre* x *and radius* r is the set $B(x,r) \equiv \{y \in \mathbb{R}^n : \|y - x\| < r\}$.

- A set $S \in \mathbb{R}^n$ is *open* if for all $x \in S$ there exists $r > 0$ such that $B(x,r) \subseteq S$. In particular the ball $B(x,r)$ is an open set.
- A point x is in the *interior* of a set $S \subseteq \mathbb{R}^n$ if there exists $r > 0$ such that the open ball $B(x,r)$ is contained in S. The set of all interior points of S is denoted by int$S$.
- A set $S \in \mathbb{R}^n$ is *closed* if its complement $S^c = \{x \in \mathbb{R}^n : x \notin S\}$ is open

Properties 2.
- *The union of an arbitrary collection of open sets is again open*
- *The intersection of an arbitrary collection of open sets is not always open. Example:* $S_n = (1 - 1/n, 2 + 1/n)$. *Then* $\cap_{n\geq 1} S_n = [1,2]$ *is a closed set.*
- *The intersection of a finite collection of open sets is open*

- *The sum of two open sets is open, where we define the sum of two sets S_1, S_2 of $\mathbb{R}^n$ by*
$$S_1 + S_2 = \{x \in \mathbb{R}^n : x = x_1 + x_2, x_1 \in S_1, x_2 \in S_2\}$$
- *The union of a finite collection of closed sets is closed*
- *The union of an arbitrary collection of closed sets is not always closed. Example: $S_n = [0, 1 - 1/n]$. Then $\cup_{n\geq 1} S_n = [0, 1)$ is not closed.*
- *The intersection of an arbitrary collection of closed sets is closed*
- *The sum of two closed sets is not always closed.*

  Example: $S_1 = \{(x, y) \in \mathbb{R}^2 : x > 0, y = \frac{1}{x}\}$, $S_2 = \{(x, y) \in \mathbb{R}^2 : x > 0, y = -\frac{1}{x}\}$. Note that $(0, 0) \notin S_1 + S_2$ because, if $(x, y) \in S_1 + S_2$, we must have $x > 0$.

Definition 8. *Let $S \subset \mathbb{R}^N$. We say that S is bounded if there exists M such that $S \subseteq B(0, M)$. In other words, S is bounded if there exists $M > 0$ such that $\|x\| \leq M$ for all $x \in S$.*

Theorem 4 (the Bolzano-Weierstrass theorem)**.** *If the sequence (x_n) in $\mathbb{R}^N$ is bounded, then it has a convergent subsequence.*

Proof. Let us firstly prove the result in the space $\mathbb{R}$. Let (x_n) be a bounded sequence.

We will construct a subsequence of (x_n), which is monotonic. Then, by applying Theorem 3, this subsequence is convergent.

It remains to prove the following result.

Lemma 1. *Every infinite sequence (x_n) in $\mathbb{R}$ has a monotone subsequence.*

Proof. Define

$$A = \{n \in \mathbb{N} : \forall p \geq n, x_p \leq x_n\}$$
$$B = \{n \in \mathbb{N} : \exists p \geq n, x_p > x_n\}$$

There are two cases.

1. $card(A) = \infty$. Let $(n_k)_k$ be an increasing infinite sequence in A. Then the subsequence (x_{n_k}) of (x_n) is decreasing.
2. $card(A) < \infty$. Let N_a be the highest element of A. Hence, for any $n > N_a$, we have $n \in B$. Take $m_1 = N_a + 1$, then $m_1 \in B$ and hence there exists $m_2 > m_1$ such that $x_{m_2} > x_{m_1}$. By induction argument, there exists an increasing infinite sequence $(m_k)_k$ such that $x_{m_{k+1}} > x_{m_k}$ $\forall k$. So, the subsequence $(x_{m_k})_k$ of (x_n) is increasing.

□

We now consider the case of $\mathbb{R}^N$. Let $(x_k) \subset \mathbb{R}^N$. $x_k = (x_{1,k}, \dots, x_{N,k})$. Since (x_k) is bounded, the sequence $(x_{i,k})$ is bounded for any $i = 1, \dots, N$. So, there is a subsequence $(x_{1,k_{1,m}})_m$ which is convergent.

We consider the subsequence $x_{k_{1,m}} = (x_{1,k_{1,m}}, x_{2,k_{1,m}}, \dots, x_{N,k_{1,m}})$ of (x_k).

Since $(x_{2,k_{1,m}})$ is bounded, it has a convergent subsequence $(x_{2,k_{2,m}})$ with $(k_{2,m}) \subset (k_{1,m})$. By induction argument, we con construct subsequences $(k_{N,m}) \subset (k_{N-1,m}) \subset \cdots \subset (k_{1,m})$ and the subsequence $x_{k_{N,m}} = (x_{1,k_{N,m}}, x_{2,k_{N,m}}, \dots, x_{N,k_{N,m}})$ of (x_k) is convergent. □

Definition 9. *Let $S \subset \mathbb{R}^N$. We say that S is compact if it is* ***bounded and closed***

Theorem 5. *Let $S \subset \mathbb{R}^N$. The following statements are equivalent*

1. *S is compact.*
2. *S is bounded and closed. (The Heine–Borel theorem.)*
3. *For all sequences $\{x^k\}$ in S, there exists a subsequence $\{x^{k_m}\}_m$ which converges to a point $x \in S$.*[5]

Proof. We prove the equivalence between (2) and (3).

(3) implies (2): Assume S has the property that for all sequences $\{x^k\}$ in S, there exists a subsequence $\{x^{k_m}\}_m$ which converges to a point $x \in S$. Let us prove it is compact. First, S is bounded. If not there exists a sequence $\{x^k\}_k \subset S$ with $\lim_{k\to+\infty} \|x^k\| = +\infty$. There exists a subsequence $\{x^{k_n}\}_n$ of $\{x^k\}_k$ which converges to some $x \in S$, which is impossible since $\lim_n \|x^{k_n}\| = +\infty$.
We now prove S is closed. Suppose $\{x^k\}_k \subset S$ converges to x. There exists a subsequence which converges in S. This limit must be x. Hence, $x \in S$.

(2) implies (3): Assume that S is compact. We will assume $S \subset \mathbb{R}^2$. One can easily see that the proof can be carried on when the dimension of the space is larger than 2. Let $\{x^k\}$ be a sequence of S. Write $x^k = (x_1^k, x_2^k), \forall k$. Since S is bounded, there exists a subsequence $\{x_1^{k_n}\}_n$ which converges to some $x_1 \in \mathbb{R}$. But there exists also a subsequence $\{x_2^{k_{n_l}}\}_l$ of $\{x_2^{k_n}\}_n$ which converges to some $x_2 \in \mathbb{R}$. Since S is closed, $(x_1, x_2) \in S$ since it is the limit of $\{(x_1^{k_{n_l}}, x_2^{k_{n_l}})\}_l$. To summarize, we have found a subsequence of $\{x^k\}$ which converges to a point in S. □

Properties 3.

- *The union of a finite collection of compact sets is compact.*
- *The union of an arbitrary collection of compact sets is NOT ALWAYS compact.*
- *The intersection of an arbitrary collection of compact sets is compact.*
- *The sum of two compact sets is compact.*

1.5 Continuous functions

Mappings Let $f : S \to T$, where $S \subseteq \mathbb{R}^n$, $T \subseteq \mathbb{R}^m$. Then S is called the *domain* of f, and T is the *range* of f. For $R \subseteq S$, the *image* of R under f is

$$f(R) = \{y \in T : y = f(x), \text{ for some } x \in R\}$$

For $U \subseteq T$, the *inverse image* of U, $f^{-1}(U)$ is

$$f^{-1}(U) = \{x \in S : f(x) = y, y \in U\}$$

The set $f^{-1}(U)$ may be empty.

[5] We can also prove that S is compact if and only if every open covering of S has a finite sub-covering. Here, we define that: An open *covering* of a set S is a collection of open sets $\{U_\alpha\}_\alpha$ such that $S \subseteq \cup_\alpha U_\alpha$. This definition is in general used when $\mathbb{R}^N$ is replaced by a topological space.

- A mapping $f : S \to \mathbb{R}^m$, where $S \subseteq \mathbb{R}^n$ is continuous at $x \in S$ if for all sequences $\{x_k\}$ in S converging to x, we have that $f(x_k) \to f(x)$. Equivalently, f is continuous at $x \in S$, if for every $\varepsilon > 0$ there exists $\delta > 0$ such that if $y \in S$ and $\|y - x\| < \delta$ then $\|f(y) - f(x)\| < \varepsilon$.

- Let $f(x) = (f_1(x), \ldots, f_m(x))$ where f_i is a mapping from S to $\mathbb{R}$. Then f is continuous at $x \in S$ if, and only if, each f_i is continuous at x.

- A mapping f is continuous on S if it is continuous at any $x \in S$.

- Let $f : \mathbb{R}^n \to \mathbb{R}^m, g : \mathbb{R}^n \to \mathbb{R}^m$. Assume f, g are continuous. Then $f + g$ is continuous. If λ is a real number, then λf is continuous.

- Let $f : \mathbb{R}^n \to \mathbb{R}^m, g : \mathbb{R}^m \to \mathbb{R}^p$. Assume f, g are continuous. The mapping gof defined by $gof(x) = g(f(x))$ for any x is continuous.

- Let $f : \mathbb{R}^n \to \mathbb{R}^m$. Then f is continuous if, and only if, for any open set $A \subseteq \mathbb{R}^m$, $f^{-1}(A)$ is open in $\mathbb{R}^n$. Equivalently, f is continuous if, and only if, for any closed set $A \subseteq \mathbb{R}^m$, $f^{-1}(A)$ is closed in $\mathbb{R}^n$.

  Proof. Assume f is continuous. Let $A \in \mathbb{R}^m$ be open. We will show that $f^{-1}(A)$ is open. For that, let $x \in f^{-1}(A)$. Since A is open, there exists an open ball $B(f(x), \varepsilon) \subset A$. Since f is continuous, there exists $\delta > 0$ such that if $y \in B(x, \delta)$ then $f(y) \in B(f(x), \varepsilon) \subset A$. This implies $y \in f^{-1}(A)$. Equivalently, $B(x, \delta) \subset f^{-1}(A)$ and $f^{-1}(A)$ is open.
  Conversely, assume that for any open set $A \subset \mathbb{R}^m$, the set $f^{-1}(A)$ is open. We will prove that f is continuous. Let $A = \{z \in \mathbb{R}^m : \|z - f(x)\| < \varepsilon\}$. The set A is open. Observe that $x \in f^{-1}(A)$. Since $f^{-1}(A)$ is open, there exists an open ball $B(x, \delta) \subset f^{-1}(A)$. That means, if $\|y - x\| < \delta$ then $f(y) \in A$ or equivalently $\|f(y) - f(x)\| < \varepsilon$. We have proved that f is continuous.
  To prove that f is continuous if, and onl! y if, for any closed set $A \subseteq \mathbb{R}^m$, $f^{-1}(A)$ is closed in $\mathbb{R}^n$, one can observe that $f^{-1}(A^c) = (f^{-1}(A))^c$. $\square$

Theorem 6 (Intermediate value theorem)**.** *Let f be a continuous function on the interval $[a, b]$. If $f(a) < f(b)$ and $c \in (f(a), f(b))$, then there exists a point $x \in (a, b)$ such that $f(x) = c$.*[6]

Note that such a x is not necessarily unique.

Proof. Define $S \equiv \{d \in [a, b] : f(d) \leq c\}$. S is non-empty and bounded. So, there exists the supremum of S. Let $x \equiv sup(S)$. We can prove that $f(x) = c$.

$\square$

Example 3 (The demand, the supply and equilibrium price)**.** *Let $D(p)$ and $S(P)$ be the demand and supply functions which depend on the price $p \geq 0$. Assume that D and S are continuous.*

A price $p > 0$ is said to be an equilibrium price if $D(p) = S(P)$.

According to Theorem 6, there is an equilibrium price p is $D(0) - S(0) > 0$ and $\lim_{x \to \infty} D(x) - S(x) < 0$.

[6] A generalization of this theorem can be found in Theorem 4.22 in Rudin (1976).

Proposition 5. *Let $f : S \to \mathbb{R}^m$, where $S \subseteq \mathbb{R}^n$. Assume S is compact and f is continuous on S. Then $f(S)$ is compact.*

Proof. First, $f(S)$ is bounded. If not there exists a sequence $\{y^n\}_n \subset f(S)$ with $\lim_{n\to+\infty} \|y^n\| = +\infty$. We can write $y^n = f(x^n)$ with $x^n \in S$ for every n. Since S is compact, there exists a subsequence $\{x^{n_k}\}_k$ which converges to some $x \in S$ and $f(x^{n_k})$ converges to $f(x)$. That is a contradiction since $\|f(x^{n_k})\|$ converges to infinity too.
We now show that $f(S)$ is closed. For that, let $\{y^n\}_n \subset f(S)\to y$. We claim that $y \in f(S)$. Write $y^n = f(x^n)$ with $x^n \in S$ for every n. Since S is compact, there exists a subsequence $\{x^{n_k}\}_k$ which converges to some x in S and $f(x^{n_k})$ converges to $f(x)$. We must have $y = f(x)$. Hence, $y \in f(S)$. $\square$

Theorem 7 (Weirstrass Theorem)**.** *Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. Assume S is compact, nonempty and f is continuous on S. Then f has both a maximum and a minimum.*

Proof. Let $M = \sup(f(S))$. There exists a sequence $\{f(x^n)\}_n$ converging to M. Since S is compact, there exists a subsequence $\{x^{n_k}\}$ which converges to some $x \in S$ and $f(x^{n_k})$ converges to $f(x) \in f(S)$. We have $M = f(x)$ and hence, $M = \max(f(S))$. The proof is similar for $\min(f(S))$. $\square$

1.6 Derivatives

A motivation: The Total- and Marginal-Cost Functions. Suppose that the total cost function of a firm is $C = C(y)$ where y is its output. Let us look at

$$\frac{\Delta C}{\Delta y} = \frac{C(y+\Delta y) - C(y)}{\Delta y}$$

This represents the (average) rate of change in cost per added unit of output produced. If we know $\frac{\Delta C}{\Delta y}$, then we can compute the change in cost $C(y + \Delta y) - C(y)$ as a function of Δy by using the formula $C(y + \Delta y) - C(y) = \frac{\Delta C}{\Delta y} \times \Delta y$.

The instantaneous rate of change (the marginal-cost of production) is

$$\lim_{\Delta y\to 0} \frac{\Delta C}{\Delta y} = \lim_{\Delta y\to 0} \frac{C(y+\Delta y) - C(y)}{\Delta y}$$

This motivates us to study derivatives of functions.

Differentiability of real functions

- Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}$. We say that f is *differentiable at* $x_0 \in S$, where x_0 must be in the interior of S, if there exists the limit

  $$\lim_{x\to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

  Denote this limit by $f'(x_0)$. This is called the derivative of function f at x_0.

  Note that we have

  $$\lim_{h\to 0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0$$

- Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. We say that f is *differentiable at* $x_0 \in S$, where x_0 must be in the interior of S, if there exists a vector $a \in \mathbb{R}^n$ so that

$$\frac{f(x) - f(x_0) - a \cdot (x - x_0)}{\|x - x_0\|} \to 0, \text{ as } x \to x_0$$

  where $\|x - x_0\| \equiv \sqrt{(x_1 - x_{0,1})^2 + \cdots + (x_n - x_{0,n})^2}$.

  The vector a is called the *derivative* of f at x_0 and is denoted by $Df(x_0)$. Moreover, we say that f is *differentiable on* S if it is differentiable at every point of S. We can regard Df as a mapping from S to $\mathbb{R}^n$. If Df is continuous, we say that f is *continuously differentiable* or f is C^1.

- Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. We study the derivative via *partial derivatives*. Let $(e_1, \ldots, e_n)$ be the canonical basis of $\mathbb{R}^n$. If e_i is a vector of this basis, then the coordinates of e_i equal zero excepted the $i-$th coordinate which equals 1. The $i-$th partial derivative of f at a point x is the number $\frac{\partial f(x)}{\partial x_i}$ defined by

$$\frac{\partial f(x)}{\partial x_i} = \lim_{t \to 0} \left\{ \frac{f(x + te_i) - f(x)}{t} \right\}$$

Differentiability of mappings

- A mapping $f : S \to \mathbb{R}^m$, where $S \subseteq \mathbb{R}^n$, is *differentiable at* $x_0 \in S$, where x_0 must be in the interior of S, if there exists a $m \times n$ matrix A so that

$$\frac{f(x) - f(x_0) - A(x - x_0)}{\|x - x_0\|} \to 0, \text{ as } x \to x_0$$

  where $\|x - x_0\| \equiv \sqrt{(x_1 - x_{0,1})^2 + \cdots + (x_n - x_{0,n})^2}$ The matrix A is called the *derivative* of f at x_0 and is denoted by $Df(x_0)$. Moreover, we say that f is *differentiable on* S if it is differentiable at every point of S.

- If f is differentiable on a set S, its derivative Df can be seen as a mapping $Df : S \to \mathbb{R}^{m \times n}$. If this mapping is continuous, we say that f is *continuously differentiable* or f is C^1.

Properties 4. *Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is an open set. If f is differentiable on S, then it is continuous on s.*

Proof. Using definition. □

Theorem 8. *Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is an open set. The function f is C^1 on S if, and only if, all partial derivatives of f exist and are continuous on S. In that case we also have*

$$Df(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n} \right)$$

Second derivatives

- Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. Then the derivative $Df(x) = \left(\frac{\partial f(x)}{\partial x_1}, \frac{\partial f(x)}{\partial x_2}, \ldots, \frac{\partial f(x)}{\partial x_n}\right)$ is also a mapping from S to $\mathbb{R}^n$. If Df is differentiable then f is called *twice differentiable* with second derivative $D^2 f(x)$. The partial derivatives of the partial derivatives of f are denoted by $\frac{\partial^2 f(x)}{\partial x_i \partial x_j}$ if $i \neq j$, and by $\frac{\partial^2 f(x)}{\partial x_i^2}$ if $i = j$. In these cases we have

$$D^2 f(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} \cdots \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \cdots\cdots \\ \cdots\cdots \\ \cdots\cdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} \cdots \frac{\partial^2 f(x)}{\partial x_n^2} \end{bmatrix}$$

  This matrix is called the *Hessian of* f at x.

- When f is twice differentiable on S and each second partial derivative is a continuous function, then f is called *twice continuously differentiable* or C^2.

Theorem 9. *If f is C^2 on $S \subseteq \mathbb{R}^n$, then $D^2 f(x)$ is a symmetric matrix, i.e. $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ for all i, j and all $x \in S$.*

Notice that it can happen that $\frac{\partial^2 f(x)}{\partial x_i \partial x_j} \neq \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$ if our assumptions are not satisfied.

Exercise 13. *(Rudin, 1976) Define*

$$\begin{cases} f(0,0) & = 0 \\ f(x,y) & = \frac{xy(x^2-y^2)}{x^2+y^2} \text{ if } (x,y) \neq (0,0) \end{cases}$$

Prove that

1. *f, $D_1 f$, $D_2 f$ are continuous in $\mathbb{R}^2$.*
2. *$D_{12} f$ and $D_{21} f$ exist at every point of $\mathbb{R}^2$, and are continuous except at $(0,0)$.*
3. *$D_{12} f(0,0) = 1$ and $D_{21} f(0,0) = -1$.*

Theorem 10 (Rolle's theorem)**.** *Let f be a function from $[a,b]$ into $\mathbb{R}$. Assume that f is continuous on $[a,b]$, differentiable on (a,b), and $f(a) = f(b)$. Then there exists $c \in (a,b)$ such that $f'(c) = 0$.*

Proof. Since f is continuous, there exist $m \equiv min\{f(x) : x \in [a,b]\}$ and $M \equiv max\{f(x) : x \in [a,b]\}$.

If $m = M$, then f is constant which implies that its derivative equals zero.

If $m < M$, we have either $f(a) > m$ or $f(a) < M$. Without loss of generality, we can assume that $f(a) > m$. Then there exists c such that $f(c) = m$. We observe that $c \neq a, c \neq b$. Define $\phi(h) \equiv \frac{f(c+h)-f(h)}{h}$ with $h \not\emptyset$ and h is small enough so that $c+h \in [a,b]$. By definition, we have $\lim_{h\to 0} \phi(c) = f'(c)$.

We will prove that $f'(c) = 0$. By definition of c, we have $f(c+h) - f(h) \geq 0$, $\forall h$ satisfying $c + h \in [a,b]$. Let $h > 0$ and tend to zero, we have $f'(c) \geq 0$. Let $h < 0$ and tend to zero, we have $f'(c) \leq 0$. Therefore, $f'(c) = 0$.

□

Theorem 11 (Mean value theorem)**.** *Let f be a function from $[a, b]$ into $\mathbb{R}$. Assume that f is continuous on $[a, b]$, differentiable on (a, b). Then there exists $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.*

Proof. Without loss of generality, we assume that $a < b$.

Consider the function

$$g(x) \equiv f(x) - \frac{f(b) - f(a)}{b - a}(x - a)$$

We have $g(a) = g(b)$. Applying Rolle's theorem, we get the result. □

The following theorem is very important for Optimization and Dynamical System. We state the result without proof.

Theorem 12 (**Taylor expansion**)**.** *Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is an open set. Pick $x_0 \in S$*

1. *If f is C^1 on S, then for any $x \in S$ we can write*

$$f(x) = f(x_0) + Df(x_0) \cdot (x - x_0) + R_1(x, x_0)\|x - x_0\|$$

where $R_1(x_0, x_0) = 0$ and $R_1(x, x_0) \to 0$ as $x \to x_0$.

We explicitly write

$$f(x_1, \ldots, x_n) = f(x_{0,1}, \ldots, x_{0,n}) + \sum_{i=1}^{n} \frac{\partial f(x_0)}{\partial x_i}(x_i - x_{0,i}) + R_1(x, x_0)\|x - x_0\|$$

2. *If f is C^2 on S, then for any $x \in S$, one can write*

$$f(x) = f(x_0) + Df(x_0) \cdot (x - x_0) + \frac{1}{2}(x - x_0)' D^2 f(x_0)(x - x_0) + R_2(x, x_0)\|x - x_0\|^2$$

where $R_2(x_0, x_0) = 0$ and $R_2(x, x_0) \to 0$ as $x \to x_0$.

We explicitly write

$$\begin{aligned} f(x_1, \ldots, x_n) =& f(x_{0,1}, \ldots, x_{0,n}) + \sum_{i=1}^{n} \frac{\partial f(x_0)}{\partial x_i}(x_i - x_{0,i}) \\ &+ \frac{1}{2}\sum_{i,j} \frac{\partial^2 f(x_0)}{\partial x_i \partial x_j}(x_i - x_{0,i})(x_j - x_{0,j}) + R_2(x, x_0)\|x - x_0\| \end{aligned}$$

2 (Quasi)Convexity and (quasi)concavity

The notion of convexity is very important in economics. Indeed, in economic models, the budget set of agents are usually assumed to be convex. In many setups, we assume that preferences of agents are convex or the utility functions are (quasi)concave. These observations motivate us to study the (quasi)convex and (quasi)concave functions.

2.1 Convexity and concavity

Definition 10 (Convex set)**.** *A set A in $\mathbb{R}^n$ is convex if $\lambda a+(1-\lambda)b \in A$ $\forall a \in A$, $\forall b \in A$, $\forall \lambda \in [0,1]$ (the line segment is a subset of A).*

Definition 11 (Concavity)**.** *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty set of $\mathbb{R}^n$, into $\mathbb{R}$. We require that A is convex.*

1. *f is convex on A if*

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b) \forall a \in A, \forall b \in A, \forall \lambda \in [0,1].$$

2. *f is strictly convex on A if*

$$f(\lambda a + (1-\lambda)b) < \lambda f(a) + (1-\lambda) f(b) \forall a \in A, \forall b \in A, a \neq b, \forall \lambda \in (0,1).$$

3. *The function $f : A \to \mathbb{R}$ is concave on A if $-f$ is convex.*
   Explicitly, f is concave on A if

$$\forall a_1 \in A, \forall a_2 \in A, \forall \lambda \in [0,1], f(\lambda a_1 + (1-\lambda)a_2) \geq \lambda f(a_1) + (1-\lambda) f(a_2).$$

   It is strictly concave on A if:

$$\forall a_1 \in A, \forall a_2 \in A, a_1 \neq a_2, \forall \lambda \in]0,1[, f(\lambda a_1 + (1-\lambda)a_2) > \lambda f(a_1) + (1-\lambda) f(a_2).$$

The function $f(x) = x^2$ is strictly convex and $f(x) = x^{1/2}$ is strictly concave. The function $f(x) = Ax$ is convex and concave but neither strictly convex nor strictly concave.

Proposition 6 (Jensen inequality)**.** *Let U be a nonempty convex set of $\mathbb{R}^n$ and let $f : U \to \mathbb{R}$ be a convex function on U. Then f is convex if, and only if, for any integer $p \geq 2$, for any p elements of U, $x_1, \ldots, x_p$, for any p positive numbers $\lambda_1, \ldots, \lambda_p$, the sum of which equals 1, then*

$$f(\lambda_1 x_1 + \ldots + \lambda_p x_p) \leq \lambda_1 f(x_1) + \ldots + \lambda_p f(x_p).$$

Proof. Let $f : U \to \mathbb{R}$ where U is a convex, nonempty subset of $\mathbb{R}^n$. Assume f convex. Let $x_1, \ldots, x_p$ be p elements of U.

If $p = 2$, then obviously $f(\lambda_1 x_1 + \ldots + \lambda_p x_p) \leq \lambda_1 f(x_1) + \ldots + \lambda_p f(x_p)$.
Assume that Jensen Inequality holds for $p-1$. We will show that it holds also for p. If $\lambda_p = 1$, then the result is trivially true. So, we assume $\lambda_p \neq 1$. Let $s_{p-1} = \lambda_1 + \ldots + \lambda_{p-1}$. We have:

$$\lambda_1 x_1 + \ldots + \lambda_p x_p = s_{p-1}\big(\frac{\lambda_1}{s_{p-1}} x_1 + \ldots + \frac{\lambda_{p-1}}{s_{p-1}} x_{p-1}\big) + (1 - s_{p-1}) x_p.$$

The function f being convex, we have:

$$f(\lambda_1 x_1 + \ldots + \lambda_p x_p) \leq s_{p-1} f\big(\frac{\lambda_1}{s_{p-1}} x_1 + \ldots + \frac{\lambda_{p-1}}{s_{p-1}} x_{p-1}\big) + (1 - s_{p-1}) f(x_p).$$

We have assumed that Jensen Inequality holds for $p-1$. Hence

$$f(\frac{\lambda_1}{s_{p-1}}x_1+\ldots+\frac{\lambda_{p-1}}{s_{p-1}}x_{p-1}) \le \frac{\lambda_1}{s_{p-1}}f(x_1)+\ldots+\frac{\lambda_{p-1}}{s_{p-1}}f(x_{p-1}).$$

Thus:

$$f(\lambda_1 x_1+\ldots+\lambda_p x_p) \le \lambda_1 f(x_1)+\ldots+\lambda_p f(x_p).$$

The converse is true by taking $p=2$. □

The following theorem is very important and useful.

Theorem 13 (first-order convexity condition)**.** *Let f be a differentiable function from its domain ($dom(f)$) into $\mathbb{R}$. Assume that $dom(f)$ is convex.*

1. *f is convex if and only if, for any $x, y \in dom(f)$, we have $f(y)-f(x) \ge Df(x)\cdot(y-x)$.*

   f is strictly convex if and only if, for any $x, y \in dom(f)$, $x \ne y$ we have $f(y)-f(x) > Df(x)\cdot(y-x)$.

2. *f is concave if and only if, for any $x, y \in dom(f)$, we have $f(y)-f(x) \le Df(x)\cdot(y-x)$.*

   f is strictly concave if and only if, for any $x, y \in dom(f)$, $x \ne y$ we have $f(y)-f(x) < Df(x)\cdot(y-x)$.

Proof. Notice that we cannot directly use Taylor's expansion. Let us prove the first point.

Suppose that f is convex. We have to prove that $f(y)-f(x) \ge Df(x)\cdot(y-x)$ for any $x, y \in dom(f)$. Let $x, y \in dom(f)$. Consider $z(\lambda) \equiv (1-\lambda)x+\lambda y$ with $\lambda \in [0,1]$. Since $dom(f)$ is convex, we have $z(\lambda) \in dom(f)$, $\forall \lambda \in [0,1]$. By the convexity of f, we have

$$f((1-\lambda)x+\lambda y) \le (1-\lambda)f(x)+\lambda f(y) \Leftrightarrow f(x+\lambda(y-x)) \le f(x)+\lambda(f(y)-f(x))$$

or equivalently

$$\frac{f(x+\lambda(y-x))-f(x)}{\lambda} \le f(y)-f(x).$$

Let λ go to zero, we get that $Df(x)\cdot(y-x) \le f(y)-f(x)$.

Suppose now that

$$f(y)-f(x) \ge Df(x)\cdot(y-x), \forall x, y \in dom(f) \tag{1}$$

Consider $z(\lambda) \equiv (1-\lambda)x+\lambda y$ with $\lambda \in [0,1]$. Applying (1), we have

$$f(x)-f(z) \ge Df(z)\cdot(x-z) \Rightarrow (1-\lambda)\big(f(x)-f(z)\big) \ge (1-\lambda)Df(z)\cdot(x-z) \tag{2}$$

$$f(y)-f(z) \ge Df(z)\cdot(y-z) \Rightarrow \lambda\big(f(y)-f(z)\big) \ge \lambda Df(z)\cdot(x-z) \tag{3}$$

Taking the sum of both sides, we get that

$$\begin{aligned}(1-\lambda)\big(f(x)-f(z)\big)+\lambda\big(f(y)-f(z)\big) &\ge (1-\lambda)Df(z)\cdot(x-z)+\lambda Df(z)\cdot(x-z)\\ &\Leftrightarrow (1-\lambda)f(x)+\lambda f(y) \ge f(z)\end{aligned}$$

So, the function f is convex. □

Corollary 1. *Consider $f : dom(f) \to \mathbb{R}$. Assume that f is convex. If $x \in dom(f)$ such that $Df(x) = 0$, then $f(x) = \min_{a \in dom(f)}(f(a))$.*

The following properties are very useful. We can prove them by using definition.

Properties 5. *We have the following properties.*

1. *If f, g are convex from a convex set U into $\mathbb{R}$ and if λ is a nonnegative real number, then $f + g$ and (λf) are convex.*

2. *Let f, g be two convex functions from a convex set U, into $\mathbb{R}$. Then $\max(f, g)$ is convex. More generally, consider a collection of convex functions $\{f_i\}_{i=1,\dots,I}$, from U into $\mathbb{R}$. Then $\max\{f_i \mid i = 1, \dots, I\}$ is convex from U into $\mathbb{R}$.*

3. *If functions $f : U \to \mathbb{R}$, $g : V \to \mathbb{R}$, with $f(U) \subset V$, are convex, if g is nondecreasing, then $g \circ f$ is convex.*

4. *If $(f_i)_{i\in\mathbb{N}}$ is a sequence of convex functions from $\mathbb{R}^n$ into $\mathbb{R}$ which converges pointwise, i.e. $\forall x \in \mathbb{R}^n$, the sequence $(f_i(x))_{i\in\mathbb{N}}$ converges in $\mathbb{R}$, then the function defined for all $x \in \mathbb{R}^n$ by $f(x) = \lim_{i\to+\infty} f_i(x)$ is convex from $\mathbb{R}^n$ into $\mathbb{R}$.*

Proof. Let us prove point 2. Denote $f = max_i f_i$. Let $a, b \in U$ and $\lambda \in [0, 1]$. We have

$$f_i(\lambda a + (1-\lambda)b) \leq \lambda f_i(a) + (1-\lambda)f_i(b) \leq \lambda f(a) + (1-\lambda)f(b).$$

Taking the maximum over i, we have $f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b)$. So, the function f is convex.

□

Example 4 (Simple composition results)**.**

1. *If g is convex then $e^{g(x)}$ is convex.*

2. *If g is concave and positive, then log $g(x)$ is concave.*

3. *If g is concave and positive, then $1/g(x)$ is convex.*

4. *If g is convex and nonnegative and $p \geq 1$, then $(g(x))^p$ is convex.*

5. *If g is convex then $-ln(-g(x))$ is convex on $\{x : g(x) < 0\}$.*

2.1.1 Convexity and continuity

Theorem 14. *If f is convex from an open set $U \subset \mathbb{R}^n$ into $\mathbb{R}$, then it is continuous in U.*[7]

Proof. It is not easy to prove this result. We need an intermediate step.

Lemma 2. *Let f be a convex function from an open set U of $\mathbb{R}^n$ into $\mathbb{R}$. Assume $0 \in U$. Then, there exists a closed ball $\bar{B}(0, r) \subset U$ such that f is bounded above on $\bar{B}(0, r)$.*

[7] See Theorem 5.2.1 in Florenzano and Le Van (2001) for a stronger result.

Proof. Since $0 \in U$, one can choose $\alpha > 0$ sufficiently small such that the convex hull $V = co\{\alpha e^1, \dots, \alpha e^n, -\alpha e^1, \dots, -\alpha e^n\}$,[8] where the e^i are the vectors of the canonical basis of $\mathbb{R}^n$, is contained in U.

V has a nonempty interior, 0 is in the interior of V. Thus, V contains a closed ball $\bar{B}(0,r)$.[9] Any x in V may be expressed as: $x = \alpha \sum_{i=1}^n \lambda_i e^i - \alpha \sum_{i=1}^n \lambda_i' e^i$ with $\lambda_i \geq 0, \lambda_i' \geq 0, \forall i$, $\sum_{i=1}^n (\lambda_i + \lambda_i') = 1$. Since f is convex,

$$f(x) \leq \sum_{i=1}^n \left(\lambda_i f(\alpha e^i) + \lambda_i' f(-\alpha e^i)\right) \leq \max\left\{f(\alpha e^1), \dots, f(\alpha e^n), f(-\alpha e^1), \dots, f(-\alpha e^n)\right\}.$$

Therefore f is bounded above on $\bar{B}(0,r)$. □

We now prove Theorem 14. Let $x_0 \in U$. Define $V \equiv U - \{x_0\}$. Observe that V is convex, open and $0 \in V$. Consider the function $h : V \to \mathbb{R}$ defined by $h(x) = f(x + x_0) - f(x_0)$. Obviously h is convex and $h(0) = 0$. Moreover, f is continuous at x_0 if, and only if, h is continuous at 0.

We will prove that h is continuous at 0. Let $\{x^n\}_n \subset V$ converge to 0. We have to prove that $h(x^n)$ converges to $h(0) = 0$.

Since V is open and $0 \in V$, there exists a closed ball $\bar{B}(0,r) \subset V$. Since $\{x^n\}_n$ converge to 0, there exists n_0 such that $\{x^n\}_{n \geq n_0} \subset \bar{B}(0,r)$. Let S denote the sphere of radius r: $S \equiv \{x \in \mathbb{R}^n : \|x\| = r\}$. Define $y^n = \frac{rx^n}{\|x^n\|}$, $z^n = -\frac{rx^n}{\|x^n\|}$. Then $y^n \in S$, $z^n \in S, \forall n$. One can see that

$$\begin{aligned} x^n &= \frac{\|x^n\| y^n}{r} + \left(1 - \frac{\|x^n\|}{r}\right) 0 \\ 0 &= \frac{r}{r + \|x^n\|} x^n + \left(\frac{\|x^n\|}{r + \|x^n\|}\right) z^n \end{aligned}$$

Since h is convex, we then get

$$\begin{aligned} h(x^n) &\leq \frac{\|x^n\|}{r} h(y^n) + \left(1 - \frac{\|x^n\|}{r}\right) h(0) = \frac{\|x^n\|}{r} h(y^n) \\ &\leq \frac{\|x^n\|}{r} \sup_{y \in S} h(y) \\ 0 = h(0) &\leq \frac{r}{r + \|x^n\|} h(x^n) + \left(\frac{\|x^n\|}{r + \|x^n\|}\right) h(z^n) \\ \Rightarrow 0 &\leq h(x^n) + \frac{\|x^n\|}{r} h(z^n) \leq h(x^n) + \frac{\|x^n\|}{r} \sup_{y \in S} h(y). \end{aligned}$$

Let $n \to +\infty$. Then $\|x^n\| \to 0$ and

$$\begin{aligned} \limsup_n h(x_n) &\leq 0 = h(0) \\ 0 &\leq \liminf_n h(x^n). \end{aligned}$$

[8] Let $T = \{a_1, a_2, \dots, a_m\} \subset \mathbb{R}^n$. We define: the convex hull of T denoted by coT is the set

$$coT = \left\{x : x = \sum_{i=1}^m \lambda_i a_i\right\}, \lambda_i \geq 0, \forall i, \sum_{i=1}^m \lambda_i = 1$$

[9] See Lemma 1.2.1 in Florenzano and Le Van (2001).

Summing up

$$0 = h(0) \leq \liminf_n h(x^n) \leq \limsup_n h(x_n) \leq 0 = h(0)$$

Hence $0 = h(0) = \lim_n h(x^n)$. $\square$

2.1.2 Testing concavity and convexity

Definition 12 (Definite and semidefinite matrices). *Let A be an $n \times n$ matrix. Then A is said to be*

- *positive definite if $x'Ax > 0$ for all $x \in \mathbb{R}^n, x \neq 0$*
- *positive semidefinite if $x'Ax \geq 0$ for all $x \in \mathbb{R}^n$*
- *negative definite if $x'Ax < 0$ for all $x \in \mathbb{R}^n, x \neq 0$*
- *negative semidefinite if $x'Ax \leq 0$ for all $x \in \mathbb{R}^n$*

Example 5. *The matrix $I = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where $a, b > 0$, is positive-definite.*

In particular, the identity matrix $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is positive-definite.

The matrix $I = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, where $a, b < 0$, is negative-definite.

Proposition 7 (second-order convexity condition). *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty, convex set of $\mathbb{R}^n$, into $\mathbb{R}$. Assume that f is twice continuously differentiable.*

1. *f is concave if and only if $D^2 f(x)$ is negative semidefinite for every $x \in A$.*

   If $D^2 f(x)$ is negative definite for every $x \in A$, then the function is strictly concave.[10]

2. *f is convex if and only if $D^2 f(x)$ is positive semidefinite for every $x \in A$.*

   If $D^2 f(x)$ is positive definite for every $x \in A$, then the function is strictly convex.

Proof. To present a proof. Let us prove point 1.

$\square$

Corollary 2. *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty, convex set of $\mathbb{R}$, into $\mathbb{R}$. Assume that f is twice continuously differentiable.*

1. *f is concave if and only if $f''(x) \leq 0$ for every $x \in A$.*

   If $f''(x) < 0$ for every $x \in A$, then the function is strictly concave.

2. *f is convex if and only if $f''(x) \geq 0$ for every $x \in A$.*

   If $f''(x) > 0$ for every $x \in A$, then the function is strictly convex.

[10] We do not have "if and only if" for the strictly concave function. Indeed, the function $f(x) = -x^4$ is strictly concave, but its second derivative equals 0 at $x = 0$.

Exercise 14 (Function of one variable). *1. Exponential.* e^{ax} *is convex on* $\mathbb{R}$*, for any* $a \in \mathbb{R}$.

2. *Power.* x^a *is convex on* $\mathbb{R}_{++}$ *when* $a \geq 1$ *or* $a \leq 0$*, and concave for* $a \in [0, 1]$.

3. *Power of absolute value:* $|x|^p$*, for* $p \geq 1$*, is convex on* $\mathbb{R}$.

4. *Logarithm:* $log(x)$ *is concave on* $\mathbb{R}_{++}$.

5. *Negative entropy:* $xlog(x)$ *(either on* $\mathbb{R}_{++}$*, or on* $\mathbb{R}_{+}$*, defined as 0 for* $x = 0$*) is convex. is convex.*

Proposition 7 leads to the following result.

Corollary 3. *Let* $f : A \to \mathbb{R}$ *be a function defined from* A*, a nonempty, convex set of* $\mathbb{R}^2$*, into* $\mathbb{R}$*. Assume that* f *is twice continuously differentiable.*

1. f *is concave if and only if*

$$f_{11}(x_1, x_2) \leq 0 \tag{4a}$$

$$f_{22}(x_1, x_2) \leq 0 \tag{4b}$$

$$f_{11}(x_1, x_2)f_{22}(x_1, x_2) - \big(f_{12}(x_1, x_2)\big)^2 \geq 0 \tag{4c}$$

2. f *is strictly concave if and only if*

$$f_{11}(x_1, x_2) < 0 \tag{5a}$$

$$f_{11}(x_1, x_2)f_{22}(x_1, x_2) - \big(f_{12}(x_1, x_2)\big)^2 > 0. \tag{5b}$$

This result is very useful when we need to verify the concavity (convexity) of a function of 2 variables. When the function is of several variables, we need conditions to check whether the Hessian matrix $D^2 f$ is negative or positive (semi)definite. Linear Algebra answers this question.

Positive (semi)definite matrix: a test

Matrices are commonly written in box brackets or parentheses:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij}) \in \mathbb{R}^{m \times n}.$$

Given a matrix $\mathbf{A}$, we denote by A_k the $k \times k$ submatrix of A formed by taking just the first k rows and columns of A. It means that

$$A_k = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \end{bmatrix}$$

We give without proof the following tests in Linear Algebra.

Lemma 3 (Sylvester's criterion). *Let A be a* **symmetric** $n \times n$ **matrix**.

(i) A is positive definite if, and only if, $detA_k > 0$ for all $k = 1, \ldots, n$

(ii) A is negative definite if, and only if, $detA_k > 0$ for all even $k \in \{1, \ldots, n\}$ and $detA_k < 0$ for all odd $k \in \{1, \ldots, n\}$. (i.e., $(-1)^i det(A_i) > 0 \ \forall i = 1, \ldots, n$).

There is NO equivalence of the above result for positive or negative SEMIDEFINITE matrices.

Lemma 4. *Let A be a* **symmetric** $n \times n$ **matrix**. *A is negative semidefinite if and only if $(-1)^i det(A_i^\pi) > 0 \ \forall i = 1, \ldots, n$ and for every permutation π of the indices $\{1, \ldots, n\}$.*[11]

These two lemmas and Proposition 7 allow us to verify whether a function f which is twice continuously differentiable is (strictly)concave or (strictly)convex via the Hessian of f.

Exercise 15 (Function of several variables). *1. Norms. $f(x) = \|x\|$ is convex on $\mathbb{R}^n$.*

2. *Max function. $f(x) = max_i(x_i)$ is convex on $\mathbb{R}^n$.*
3. *Quadratic-over-linear function. The function $f(x, y) = \frac{x^2}{y}$ defined on the domain*
$$dom(f) = \mathbb{R} \times \mathbb{R}_{++}$$
is convex.
4. *Log-sum-exp function. $f(x) = ln(e^{x_1} + \cdots + e^{x_n})$ is convex on $\mathbb{R}^n$.*
5. *Geometric mean: The function $f(x) = \Big(x_1 x_2 \cdots x_n\Big)^{\frac{1}{n}}$ is convex on $dom(f) \equiv \mathbb{R}^n_{++}$.*
6. *Quadratic functions: $f(x) = \frac{1}{2} x^T Q x + c^T x$ on $\mathbb{R}^n$, where $Q \in \mathbb{R}^{n \times n}$ is a positive definite matrix and $c \in \mathbb{R}^n$.*

2.2 Quasiconvexity and quasiconcavity

Definition 13 (Quasi-concavity). *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty set of $\mathbb{R}^n$, into $\mathbb{R}$. We require that A is convex.*

1. *We say that f is quasi-concave if its upper contour sets $\{x \in A : f(x) \geq t\}$ are convex sets for any t, that is*
$$f(\lambda x + (1 - \lambda)x') \geq t \ if \ min(f(x), f(x')) \geq t$$
for any $t \in \mathbb{R}$, $x, x' \in A$, $\lambda \in [0, 1]$.
2. *We say that f is strictly quasi-concave if it satisfies:*
$$f(\lambda x + (1 - \lambda)x') > t \ if \ min(f(x), f(x')) \geq t$$
for any $t \in \mathbb{R}$, $x, x' \in A$, $x \neq x'$, $\lambda \in (0, 1)$.

[11] A permutation of the set $S \equiv \{1, \ldots, n\}$ is defined as a bijection from S to itself.

3. *f is call (strictly) quasiconvex if $-f$ is (strictly) quasiconcave. Precisely, f is call quasiconvex if $\{x \in A : f(x) \leq t\}$ are convex sets for any t.*

4. *A function that is both quasiconvex and quasiconcave is quasilinear.*

Example 6. 1. *Prove that $ln(x)$ on $\mathbb{R}_{++}$ is quasiconvex and quasiconcave.*

2. *Prove that $sin(x)$ on $\mathbb{R}$ is neither quasiconvex nor quasiconcave.*

3. *Consider the function $f : \mathbb{R}^2 \to \mathbb{R}$ with $domain(f) \equiv \mathbb{R}^2_+$ and $f(x_1, x_2) = x_1 x_2$.*

   (a) *Prove that f is neither convex nor concave.*

   (b) *Prove that f is quasiconcave on $\mathbb{R}^2_+$ but not quasiconcave on $\mathbb{R}^2$.*

Properties 6. *(i) f is quasiconcave if and only if*

$$f(\lambda x + (1-\lambda)x') \geq min(f(x), f(x'))$$

for any $x, x' \in A$, $\lambda \in [0,1]$.
f is strictly quasiconcave if and only if

$$f(\lambda x + (1-\lambda)x') > min(f(x), f(x'))$$

for any $t \in \mathbb{R}$, $x, x' \in A$, $\lambda \in [0,1]$.
(ii) Any increasing function of one variable is quasiconcave.
(iii) A concave function is quasiconcave. However, the converse is not true.

Properties 7. 1. *If f_i is quasiconvex and $\lambda > 0$, then λf_i is quasiconvex.*

2. *If f_i is quasiconvex for all $i \in I$, then $\sup_i(f_i)$ is quasiconvex.*

3. *If f is quasiconvex and g is non-decreasing from $\mathbb{R}$ into $\mathbb{R}$, then $g \circ f$ is quasiconvex.*

Notice that the sum of two quasiconvex functions need not be quasiconvex.

Exercise 16. *Suppose that $f(x,y)$ is quasiconvex and C is a convex set. Prove that*

$$g(x) \equiv \inf_{y \in C} f(x,y)$$

is quasiconvex

2.2.1 Testing quasiconcavity and quasiconvexity

Theorem 15 (first-order quasiconvexity condition)**.** *Let f be a differentiable function from its domain ($dom(f)$) into $\mathbb{R}$. Assume that $dom(f)$ is convex.*

1. *f is quasiconvex if and only if, for any $x, y \in dom(f)$, we have that:*

$$f(y) \leq f(x) \Rightarrow Df(x) \cdot (y - x) \leq 0. \tag{6}$$

2. *f is quasiconcave if and only if, for any $x, y \in dom(f)$, we have that.*

$$f(y) \geq f(x) \Rightarrow Df(x) \cdot (y - x) \geq 0. \tag{7}$$

Proof. Let us prove point 2.

Necessary condition. Suppose that f is quasiconcave.

Let $x, y \in dom(f)$. Assume that $f(y) \geq f(x)$. We have to prove that $Df(x)\cdot(y-x) \geq 0$.

Since f is quasiconcave, we have

$$f((1-\lambda)x + \lambda y) \geq min(f(x), f(y)) = f(x)$$

for any $\lambda \in [0, 1]$. Observe that $(1-\lambda)x + \lambda y = x + \lambda(y-x)$. Hence, we have

$$\frac{f(x + \lambda(y-x)) - f(x)}{\lambda} \geq 0 \ \forall \lambda \in (0, 1].$$

Let λ go to zero, we get that $Df(x) \cdot (y-x) \geq 0$.

Sufficient condition. Suppose that f satisfies condition (7). We will prove that f is quasiconcave, i.e.,

$$f((1-\lambda)x + \lambda y) \geq min(f(x), f(y))$$

for any $x, y \in A$, $\lambda \in [0, 1]$.

Suppose that this is not true. That is there exists $x, y \in A$, $\lambda \in [0, 1]$ such that $f((1-\lambda)x + \lambda y) < min(f(x), f(y))$. We have to find a contradiction.

Without loss of generality, we assume that $f(y) \geq f(x)$. So, we have

$$f((1-\lambda)x + \lambda y) < f(x) \leq f(y).$$

For each $\theta \in [0, 1]$, we define $z_\theta \equiv (1-\theta)x + \theta y = x + \theta(y-x)$ and $g(\theta) \equiv f((1-\theta)x + \theta y)$. Since $dom(f)$ is convex, we have $z_\theta \in dom(f)$.

We present two ways to get a contradiction

1. **Proof 1**. Since $g(\lambda) < g(0)$, there exists $\lambda_1 \in (0, \lambda)$ such that $g'(\lambda_1) = \frac{g(\lambda) - g(0)}{\lambda} < 0$.

   Since $g(\lambda) < g(1)$, there exists $\lambda_2 \in (\lambda, 1)$ such that $g'(\lambda_2) = \frac{g(\lambda) - g(1)}{\lambda - 1} > 0$. To sum up, we have $0 < \lambda_1 < \lambda < \lambda_2 < 1$ and $g'(\lambda_2) > 0 > g'(\lambda_1)$. This implies that $g'(\lambda_1)(\lambda_2 - \lambda_1) < 0$ and $g'(\lambda_2)(\lambda_1 - \lambda_2) < 0$.

   If $g(\lambda_1) \geq g(\lambda_2)$, i.e., $f((1-\lambda_1)x + \lambda_1 y) \geq f((1-\lambda_2)x + \lambda_2 y)$, we apply condition (7) to get get $g'(\lambda_2)(\lambda_1 - \lambda_2) \geq 0$, a contradiction.

   If $g(\lambda_1) \leq g(\lambda_2)$, i.e., $f((1-\lambda_1)x + \lambda_1 y) \leq f((1-\lambda_2)x + \lambda_2 y)$, we apply condition (7) to get get $g'(\lambda_1)(\lambda_2 - \lambda_1) \geq 0$, a contradiction.

2. **Proof 2**. Define $m \equiv \min_{\theta \in [0,1]}\{g(\theta)\}$ and $\theta^* \equiv \inf\{\theta \in [0, 1] : g(\theta) = m\}$. Notice that $m < f(x)$.

   By continuity of the function g and $m < f(x)$, there exists $\epsilon > 0$ such that $g(\theta) < f(x)$ $\forall \theta \in (\theta^* - \epsilon, \theta^*)$.

   By definition of θ^*, we have $g(\theta) > m \ \forall \theta < \theta^*$.

   So, we have $g(\theta) > m = g(\theta^*) \ \forall \theta \in (\theta^* - \epsilon, \theta^*)$. Take $\theta \in (\theta^* - \epsilon, \theta^*)$. Applying the mean value theorem, there exists $\tau \in (\theta^* - \epsilon, \theta^*)$ such that

$$g'(\tau) = \frac{g(\theta) - g(\theta^*)}{\theta - \theta^*} < 0$$

This means that

$$Df(z_\tau) \cdot (y - x) > 0. \tag{8}$$

Now, since $\tau \in (\theta^* - \epsilon, \theta^*)$, we have $g(\tau) < f(x)$, i.e., $f(x) > f((1-\tau)x + \tau y)$. Recall that $f(y) \geq f(x)$, we have $f(y) > f((1-\tau)x + \tau y)$. Using condition (7), we get that $Df((1-\tau)x + \tau y) \cdot ((1-\tau)(y-x)) \geq 0$ which is equivalent to

$$Df(z_\tau) \cdot (y - x) \geq 0.$$

which is a contradiction with (8). We have finished our proof.

□

Theorem 16 (second-order quasiconvexity condition)**.** *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty, convex set of $\mathbb{R}^n$, into $\mathbb{R}$. Assume that f is twice continuously differentiable.*

1. *If f is quasiconvex, we have that:*

$$x \in dom(f), y \in \mathbb{R}^n, y^T Df(x) = 0 \Rightarrow y^T D^2 f(x) y \geq 0.$$

2. *f is quasiconvex if f satisfies*

$$y^T Df(x) = 0 \Rightarrow y^T D^2 f(x) y > 0$$

for any $x \in dom(f)$ and $y \in \mathbb{R}^n$, $y \neq 0$.

Proof. See Florenzano and Le Van (2001) or Boyd and Vandenberghe (2004) among others. □

For other second-order conditions for quasiconvexity, see Exercise 3.44 in Boyd and Vandenberghe (2004).

We present here a simple test for functions of 2 variables.

Proposition 8. *Let $f : A \to \mathbb{R}$ be a function defined from A, a nonempty, convex set of $\mathbb{R}^2$, into $\mathbb{R}$. Suppose that (1) f is twice continuously differentiable, (2) f is strictly increasing in each component ($f_1 \equiv \frac{\partial f}{\partial x_1} > 0, f_2 \equiv \frac{\partial f}{\partial x_2}$ on A).*

1. *If f is quasiconcave, then the determinant of the bordered Hessian is non-negative*

$$\text{the bordered Hessian: } B_2 \equiv \begin{bmatrix} 0 & f_1 & f_2 \\ f_1 & f_{11} & f_{12} \\ f_2 & f_{21} & f_{22} \end{bmatrix}$$

that is

$$2f_1(x_1,x_2)f_2(x_1,x_2)f_{12}(x_1,x_2) - \big[f_1(x_1,x_2)\big]^2 f_{22}(x_1,x_2) - \big[f_2(x_1,x_2)\big]^2 f_{11}(x_1,x_2) \geq 0$$

$\forall (x_1, x_2) \in A.$

2. *If*

$$2f_1(x_1,x_2)f_2(x_1,x_2)f_{12}(x_1,x_2) - \big[f_1(x_1,x_2)\big]^2 f_{22}(x_1,x_2) - \big[f_2(x_1,x_2)\big]^2 f_{11}(x_1,x_2) > 0$$

$\forall (x_1, x_2) \in A$, *then f is strictly quasiconcave.*

Proof. See Simon and Blume (1994), page 530, among others. □

2.3 Other exercises

Exercise 17. *For each of the following functions, determine if the function is convex, concave, quasiconvex, or quasiconcave.*

1. $f(x) = e^x - 1$ *on* $\mathbb{R}$.
2. $f(x) = \frac{x^{1-\sigma}}{1-\sigma} - 2$ *on* $\mathbb{R}_+$.
3. $f(x_1, x_2) = x_1 x_2$ *on* $\mathbb{R}^2_{++}$.
4. $f(x_1, x_2) = \frac{1}{x_1 x_2}$ *on* $\mathbb{R}^2_{++}$.
5. $f(x_1, x_2) = \frac{x_1}{x_2}$ *on* $\mathbb{R}^2_{++}$.
6. $f(x_1, x_2) = \frac{x_1^2}{x_2}$ *on* $\mathbb{R} \times \mathbb{R}_{++}$.

More exercises can be found in Florenzano and Le Van (2001) and Boyd and Vandenberghe (2004).

The following exercises are important in Economics.

Exercise 18 (Probability and Statistics). *Let X be a real-valued random variable defined by $P(X = a_i) = Probability(X = a_i) = p_i$ where $a_1 < a_2 < \cdots < a_n$. For each of the following functions of p on the simplex Δ_n (defined by $\Delta_n \equiv \{p \in \mathbb{R}^n_+ : p_1 + \cdots + p_n = 1\}$), determine if the function is convex, concave, quasiconvex, or quasiconcave.*

1. $\mathbb{E}(X) \equiv \sum_i p_i a_i$.
2. $P(X \geq \alpha)$
3. $P(\alpha \leq X \leq \beta)$.
4. $\sum_i p_i ln(p_i)$ *(the negative entropy of the distribution).*
5. $Var(X) \equiv \mathbb{E}(X - \mathbb{E}(X))^2 = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
6. $Quartile(X) \equiv \inf\{\beta : P(X \leq \beta) \geq 0.25\}$

Exercise 19 (Functions in economics). *For each of the following functions, determine if the function is convex, concave, quasiconvex, or quasiconcave.*

1. *Cobb-Douglas function:* $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$ *on* $\mathbb{R}^2_{++}$*, where* $\alpha \in [0, 1]$.
2. *Cobb-Douglas function:* $f(x_1, x_2) = x_1^\alpha x_2^\beta$ *on* $\mathbb{R}^2_{++}$*, where* $\alpha > 0, \beta > 0$.
3. *Leontief function:* $f(x_1, x_2) = min(\frac{x_1}{a}, \frac{x_2}{b})$ *on* $\mathbb{R}^2_{++}$*, where* $a > 0, b > 0$.
4. *Constant elasticity of substitution (CES) function:* $f(x_1, x_2) = (ax_1^r + bx_2^r)^{\frac{1}{r}}$ *on* $\mathbb{R}^2_{++}$*, where* $a > 0, b > 0$.[12]

[12]Hint: Look at two cases $r < 1$ and $r > 1$.

5. *The sum of squared residuals (in the multiple linear regression):*

$$f(\beta_0, \beta_1, \cdots, \beta_k) \equiv \sum_{i=1}^{n} \Big(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{i,j} \Big)^2$$

on $\mathbb{R}^{k+1}$. *(Readers may like to consider the case* $k = 1$ *to simplify the calculation.)*

3 Finite-dimensional optimization

3.1 Motivating examples

1. Utility maximization problem
2. Profit maximization problem
3. Cost minimization problem

3.2 Convex optimization in finite-dimensional spaces

This section uses Florenzano and Le Van (2001).

Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$ is a nonempty set.

- A point $x \in S$ is a *global maximum* (respectively, global minimum) of f on S if $f(y) \leq f(x)$ (respectively, $f(y) \geq f(x)$) for all $y \in S$.
- A point $x \in S$ is a *local maximum* (respectively, local minimum) of f on S if there exists $r > 0$ such that $f(y) \leq f(x)$ (respectively, $f(y) \geq f(x)$) for all $y \in S \cap B(x, r)$.
- A point $x \in S$ is a *strict local maximum* (respectively, strict local minimum) of f on S if there exists $r > 0$ such that $f(y) < f(x)$ (respectively, $f(y) > f(x)$) for all $y \in S \cap B(x, r), y \neq x$.
- A point $x \in S$ is an *unconstrained local maximum* (respectively, unconstrained local minimum) of f on S if there exists $r > 0$ such that $B(x, r) \subseteq S$ and $f(y) \leq f(x)$ (respectively, $f(y) \geq f(x)$) for all $y \in B(x, r)$.

Remark 3. *Assume that the set S is convex. Any local minimum of a convex function is also a global minimum. A strictly convex function will have at most one global minimum.*

Proof. Let x be a local minimum. It means that there exists $r > 0$ such that $f(y) \leq f(x)$ for all $y \in S \cap B(x, r)$.

We have to prove that $f(y) \geq f(x)$ for all $y \in S$. Suppose that there exists $y \in S$ such that $f(y) < f(x)$ Define $z_\lambda \equiv (1 - \lambda)x + \lambda y$. Since S is convex, we have that $z_\lambda \in S$. Observe that

$$z_\lambda = x + \lambda(y - x)$$

So, we can choose $\lambda > 0$ small enough so that $z_\lambda \in B(x, r)$.

We now look at $f(z_\lambda)$. By the convexity of the function f, we have

$$f(z_\lambda) = f((1 - \lambda)x + \lambda y) \leq (1 - \lambda)f(x) + \lambda f(y)$$

Since $f(y) < f(x)$, we get that $f(z_\lambda) < f(x)$. This is a contradiction to the fact that x is a local minimum. We have finished our proof. $\square$

We aim to give the necessary and sufficient conditions for a point to be an optimal solution to Problem (P):

$$(P) \qquad \text{Minimize } f_0(x) \text{ under the constraints } \begin{cases} f_i(x) \leq 0, \ \forall i \in I \\ g_i(x) \leq 0, \ \forall i \in J \\ g_i(x) = 0, \ \forall i \in K. \end{cases}$$

where $f_0 : \mathbb{R}^n \to \mathbb{R}$ is a convex function, I, J and K are finite and possibly empty sets, for all $i \in I$, f_i is convex, non-affine function from $\mathbb{R}^n$ into $\mathbb{R}$, for all $i \in J \cup K$, g_i is a non-null affine function.

The function f_0 is called the *objective function*. A *feasible point* is a point $x \in \mathbb{R}^n$ that satisfies all the constraints. An *optimal solution* to (P) or simply a *solution* to (P) is a feasible point $\overline{x}$, such that for all feasible point x, $f_0(x) \geq f_0(\overline{x})$.

Let I be a finite set, $card(I)$ denotes the number of elements of I.

3.2.1 Separation theorems

H is a hyperplane in $\mathbb{R}^n$ if there exists $p \in \mathbb{R}^n, p \neq 0$ and $\alpha \in \mathbb{R}$ such that $H = \{x \in \mathbb{R}^n : p \cdot x = \alpha\}$.

Proposition 9. *[First separation theorem] Let A and B be two nonempty disjoint convex subsets of $\mathbb{R}^n$. There there exist α, β, $\alpha \leq \beta$ and $p \in \mathbb{R}^n$, $p \neq 0$, such that $p \cdot a \leq \alpha \leq \beta \leq p \cdot b$, for all $a \in A$, all $b \in B$ (i.e., we can separate A and B by a hyperplane).*

Proof. See Florenzano and Le Van (2001). $\square$

Proposition 10. *[Second separation theorem] Let A and B be two nonempty disjoint closed convex subsets of $\mathbb{R}^n$. If one of them is compact, then there there exist α, β, $\alpha < \beta$ and $p \in \mathbb{R}^n$, $p \neq 0$, such that $p \cdot a \leq \alpha < \beta \leq p \cdot b$, for all $a \in A$, all $b \in B$ (i.e., we can strictly separate A and B by a hyperplane).*

Proof. See Florenzano and Le Van (2001). $\square$

Notice that the compactness in Proposition 10 is an important assumption.

Exercise 20. *Let $A = \{(x, y) : xy \geq 1, x \geq 0\}$ and $B = \{(x, y) : x \leq 0\}$. Prove that A and B are non-empty, convex, closed and disjoint. Prove that X and Y cannot be strictly separated, i.e., there do not exist α, β, $\alpha < \beta$ and $p \in \mathbb{R}^n$, $p \neq 0$, such that $p \cdot a \leq \alpha < \beta \leq p \cdot b$, for all $a \in A$, all $b \in B$.*

Note that A and B are not bounded in this example. Therefore, they are not compact.

The following result is very important.

Proposition 11. *Let C be nonempty convex sets in $\mathbb{R}^n$ and $P = (\mathbb{R}^m_- \times \{0_{\mathbb{R}^r}\})$ where, $m \geq 0, r \geq 0$, $m + r = n$. Suppose that $C \cap (\mathbb{R}^m_- \times \{0_{\mathbb{R}^r}\}) = \emptyset$. Then there exists $p \in \mathbb{R}^n$, $p \neq 0$ such that*

$$p \cdot x \leq 0 \leq p \cdot y, \forall x \in P, \ \forall y \in C$$

and there exists $z \in C$ such that $p \cdot z > 0$.

Proof. Applying Proposition 3.2.6 in Florenzano and Le Van (2001). □

Let us illustrate this result by considering a particular case: $n = 2$, $m = 2$, $r = 0$. In this case, $P = \{(x, y) : x \leq 0, y \leq 0\}$. A line $p = (p_1, p_2)$ is represented by an equation

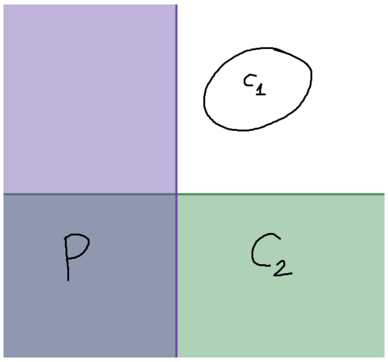


Figure 1: Monotonic convergence versus oscillatory convergence

$p_1 x + p_2 y = 0$. Our result says that there is a line that separates P and C and there is a point $z = (z_1, z_2) \in C$ such that z is above the line p, i.e., $p_1 z_1 + p_2 z_2 > 0$.

When the set C is C_1 in the graphic, it is easy to geometrically check our result.

When the set C is $C_2 \equiv \{(x, y) : x > 0, y \leq 0\}$, we can choose the line "$x = 0$", i.e., $p = (p_1, p_2) = (1, 0)$. It is easy to see that this line separates P and C_2. Moreover, for any $z \in C_2$, we have $p_1 z_1 + p_2 z_2 = z_1 > 0$.

3.2.2 Necessary and sufficient condition for optimality

Lemma 5. *(i) Let f be a linear function on $\mathbb{R}^n$ (in the sense that $f(x) = Ax = A_1 x_1 + \cdots + A_n x_n$). If $f(x) \geq 0$ for any x then $f = 0$.*

(ii) Hence, if g is an affine function (i.e., $g(x) = Ax + B$) which satisfies $g(x) \geq 0$ for any x, then g equals a nonnegative constant.

Proof. (i) Suppose $f \neq 0$. There exists x such that $f(x) > 0$. But we have a contradiction $0 > -f(x) = f(-x) \geq 0$. Hence $f = 0$.

(ii) We can write $\forall x,\ g(x) = f(x) + b$ where f is linear and b is a real constant. Let $x \in \mathbb{R}^n$. We have $f(x) + b \geq 0$. Let $\lambda > 0$. We also have $\lambda f(x) + b = f(\lambda x) + b \geq 0,\ \forall \lambda > 0$. This is equivalent to $f(x) + \frac{b}{\lambda} \geq 0$ for all $\lambda > 0$. Let $\lambda \to +\infty$. We get $f(x) \geq 0$. But x has been arbitrarily chosen. That means $f(x) \geq 0,\ \forall x$. Thus, $f = 0$ and $g(x) = b,\ \forall x$. And $b \geq 0$. □

Lemma 6. *Let I, J and K be finite possibly empty sets in $\mathbb{N}$, and for all $i \in I$, f_i is a convex, non-affine function from $\mathbb{R}^n$ into $\mathbb{R}$, and for all $i \in J \cup K$, g_i is non-null affine function. Assume there exists x_0 such that*

$$\begin{cases} g_i(x_0) \leq 0, & \forall i \in J \\ g_i(x_0) = 0, & \forall i \in K. \end{cases}$$

If the system:

$$\begin{cases} f_i(x) < 0, \ \forall i \in I \\ g_i(x) \leq 0, \ \forall i \in J \\ g_i(x) = 0, \ \forall i \in K. \end{cases}$$

has no solution, then there exist nonnegative real scalars $(\lambda_i)_{i\in I}$, $(\mu_i)_{i\in J}$, and real numbers $(\mu_i)_{i\in K}$, at least one of the $(\lambda_i)_{i\in I}$ is not zero, which verify:

$$\sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J} \mu_i g_i(x) + \sum_{i\in K} \mu_i g_i(x) \geq 0, \ \forall x.$$

Proof. Let $p = card(I)$, $q = card(J)$, $r = card(K)$, and

$$Z = \left\{ (z_i)_{i\in I\cup J\cup K} \in \mathbb{R}^{p+q+r} \;\middle|\; \exists x, \ \begin{array}{l} \forall i \in I, \ f_i(x) < z_i \\ \forall i \in J \cup K, \ g_i(x) = z_i \end{array} \right\}$$

The set Z is convex, nonempty and $Z \cap (\mathbb{R}^{p+q}_{-} \times \{0_{\mathbb{R}^r}\}) \ = \emptyset$. From Proposition 11 there exist real scalars $(\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$, all of them not being equal to zero, which verify:

$$\sum_{i\in I} \lambda_i z_i + \sum_{i\in J\cup K} \mu_i z_i \geq \sum_{i\in I} \lambda_i \zeta_i + \sum_{i\in J} \mu_i \zeta_i, \ \forall z \in Z, \ \forall \zeta \in \mathbb{R}^{p+q}_{-}$$

and there exists $z \in Z$ such that $\sum_{i\in I} \lambda_i z_i + \sum_{i\in J\cup K} \mu_i z_i > 0$.

If for some i, $\lambda_i < 0$, then letting z_i tend to $+\infty$, we get a contradiction. Thus $\lambda_i \geq 0, \ \forall i \in I$.

If for some $i \in J$, one has $\mu_i < 0$, then letting ζ_i tend to $-\infty$, we have another contradiction. Hence, $\mu_i \geq 0, \forall i \in J$.

Let $\varepsilon > 0$ and $x \in \mathbb{R}^n$. Define $z \in Z$ by

$$\begin{cases} z_i = f_i(x) + \varepsilon, & \forall i \in I, \\ z_i = g_i(x), & \forall i \in J \cup K. \end{cases}$$

We have $\sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x) + \varepsilon \sum_{i\in I} \lambda_i \geq 0$. Let ε tend to zero. We obtain that $\sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x) \geq 0, \forall x \in \mathbb{R}^n$. ZZ

To end the proof, it remains to show that at least one of the $(\lambda_i)_{i\in I}$ is strictly positive. Assume the contrary. We then have $\sum_{i\in J\cup K} \mu_i g_i(x) \geq 0, \ \forall x \in \mathbb{R}^n$ and hence $\sum_{i\in J\cup K} \mu_i g_i(x_0) = 0$. The affine function $\sum_{i\in J\cup K} \mu_i g_i$ has a minimum in $\mathbb{R}^n$. From Lemma 5, it must be equal to zero. But since $\sum_{i\in I} \lambda_i z_i + \sum_{i\in J\cup K} \mu_i z_i > 0$ for some $z \in Z$, and since all the $(\lambda_i)_{i\in I}$ are equal to zero, there exists x such that $\sum_{i\in J\cup K} \mu_i g_i(x) > 0$. That contradicts that $\sum_{i\in J\cup K} \mu_i g_i$ is equal to zero. □

Definition 14. *The Problem (P) satisfies Slater Condition (S) if there exists x_0 such that:*

$$f_i(x_0) < 0, \ \forall i \in I,$$
$$g_i(x_0) \leq 0, \ \forall i \in J,$$
$$g_i(x_0) = 0, \ \forall i \in K.$$

Proposition 12. *Consider Problem (P). Assume that (P) satisfies Slater Condition (S). If (P) has an optimal solution $\overline{x}$, then there exists scalars $(\lambda_i)_{i\in I}$, $(\mu_i)_{i\in J\cup K}$ such that:*
(i) $\forall i \in I,\ \lambda_i \geq 0, \lambda_i f_i(\overline{x}) = 0,$
$\forall i \in J,\ \mu_i \geq 0, \mu_i g_i(\overline{x}) = 0.$
(ii)

$$f_0(\overline{x}) + \sum_{i\in I} \lambda_i f_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i g_i(\overline{x}) \leq f_0(x) + \sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x)$$

for any x.
(iii) If $f_0,\ (f_i)_{i\in I}, (g_i)_{i\in J\cup K}$ are differentiable, then

$$0 = Df_0(\overline{x}) + \sum_{i\in I} \lambda_i Df_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i Dg_i(\overline{x})$$

Proof. Let $\alpha = f_0(\overline{x})$. Then the following system has no solution (why? Readers should explain this point).

$$\begin{cases} f_0(x) < \alpha, \\ f_i(x) < 0,\ \forall i \in I, \\ g_i(x) \leq 0,\ \forall i \in J, \\ g_i(x) = 0,\ \forall i \in K. \end{cases}$$

The Slater condition allows us to apply Lemma 6. According to Lemma 6, there exist $\lambda_0, (\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$ such that $\lambda_0 \geq 0, \lambda_i \geq 0,\ \forall i \in I, \mu_i \geq 0,\ \forall i \in J$, at least one of the $\lambda_0, (\lambda_i)_{i\in I}$ is strictly positive, and

$$\lambda_0(f_0(x) - \alpha) + \sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x) \geq 0,\ \forall x. \tag{10}$$

We claim that $\lambda_0 > 0$. Indeed, if $\lambda_0 = 0$, then $\forall x$, $\sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x) \geq 0$. Moreover, since there exists at least one $i \in I$ such that $\lambda_i > 0$, we deduce $\sum_{i\in I} \lambda_i f_i(x_0) + \sum_{i\in J\cup K} \mu_i g_i(x_0) < 0$, a contradiction.

Thus $\lambda_0 > 0$ and one can, without loss of generality, suppose $\lambda_0 = 1$.

Define the convex function h by

$$h(x) = f_0(x) + \sum_{i\in I} (\lambda_i f_i)(x) + \sum_{i\in J\cup K} \mu_i g_i(x),\ \forall x \in \mathbb{R}^n.$$

The inequality (10) can be equivalently rewritten as $h(x) \geq \alpha\ \forall x$. But $h(\overline{x}) - \alpha = \sum_{i\in I} \lambda_i f_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i g_i(\overline{x}) \leq 0$. Hence, $\alpha = h(\overline{x})$. Thus $\sum_{i\in I} \lambda_i f_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i g_i(\overline{x}) = 0$ and since $\overline{x}$ is feasible, one has $\forall i \in I,\ \lambda_i f_i(\overline{x}) = 0$ and $\forall i \in J,\ \mu_i g_i(\overline{x}) = 0$. We have proved Assertion (i).

Now, $h(x) \geq \alpha\ \forall x$ is equivalent to

$$f_0(\overline{x}) + \sum_{i\in I} \lambda_i f_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i g_i(\overline{x}) \leq f_0(x) + \sum_{i\in I} \lambda_i f_i(x) + \sum_{i\in J\cup K} \mu_i g_i(x)$$

for any x. We have proved assertion (ii). Statement (iii) is obvious when $f_0,\ (f_i)_{i\in I}, (g_i)_{i\in J\cup K}$ are differentiable. The proof is now complete. □

Definition 15. *The real numbers $(\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$ are called Lagrange parameters, Lagrange multipliers or Kuhn-Tucker coefficients or more simply multipliers of Problem (P).*

Definition 16. *We say that $\overline{x}$, $(\lambda_i)_{i\in I}$, $(\mu_i)_{i\in J\cup K}$ satisfy Kuhn-Tucker Conditions of Problem (P) if they satisfy Conditions (i), (ii) and (iii).*
(i) $\forall i \in I$, $\lambda_i \geq 0, f_i(\overline{x}) \leq 0, \lambda_i f_i(\overline{x}) = 0$,
$\forall i \in J$, $\mu_i \geq 0, g_i(\overline{x}) \leq 0, \mu_i g_i(\overline{x}) = 0$.
(ii) $\forall i \in K$, $g_i(\overline{x}) = 0$.
(iii) $f_0(\overline{x}) + \sum_{i\in I}\lambda_i f_i(\overline{x}) + \sum_{i\in J\cup K}\mu_i g_i(\overline{x}) \leq f_0(x) + \sum_{i\in I}\lambda_i f_i(x) + \sum_{i\in J\cup K}\mu_i g_i(x)$ *for any* x.

Conditions (i), (ii) and (iii) are called Kuhn-Tucker Conditions for Problem (P).

Definition 17. *The Lagrangian of Problem (P) is the function $L : \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}^{r-q} \times \mathbb{R}^n \to \mathbb{R}$ defined by: for all $(\lambda, \mu, x) = ((\lambda_i)_{i\in I}, (\mu_i)_{i\in J}, (\mu_i)_{i\in K}, x)$,*

$$L(\lambda, \mu, x) = f_0(x) + \sum_{i\in I}\lambda_i f_i(x) + \sum_{i\in J}\mu_i g_i(x) + \sum_{i\in K}\mu_i g_i(x).$$

where $p = card(I), q = card(J)$ and $r = card(J \cup K)$.

Proposition 13. *Let $\overline{x}, (\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$ verify Kuhn-Tucker Conditions for Problem (P). Then $\overline{x}$ is a solution to (P).*

Proof. Let $\forall x \in \mathbb{R}^n$, $h(x) = f_0(x) + \sum_{i\in I}(\lambda_i f_i)(x) + \sum_{i\in J\cup K}\mu_i g_i(x)$. Condition (iii) is equivalent to $h(x) \geq h(\overline{x}), \forall x$. Combining conditions (i) and (ii), we get $h(x) \geq f(\overline{x}), \forall x$. Let x satisfy the contraints of Problem (P). We obtain $f_0(x) \geq f_0(\overline{x})$, $\forall x$. □

Theorem 17 (Kuhn-Tucker (minimization problem))**.** *Assume that Slater Condition is satisfied for Problem (P). Then $\overline{x}$ is a solution to (P) if, and only if, there exists coefficients $(\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$ which, together with $\overline{x}$, satisfy Kuhn-Tucker Conditions for Problem (P).*

Proof. The statement follows from Proposition 12 and Proposition 13. □

As particular case of Proposition 12, we get the following result when the problem is without convex constraints.

Corollary 4. *Consider Problem (P) without convex constraints, i.e.*

$$\min f_0(x) \text{ under the constraints } \begin{cases} g_i(x) \leq 0, & \forall i \in J \\ g_i(x) = 0, & \forall i \in K. \end{cases}$$

where $f_0 : \mathbb{R}^n \to \mathbb{R}$ is convex, and for all $i \in J \cup K$, g_i is affine.
For this problem, Slater Condition is: $\exists x_0$ such that $g_i(x_0) \leq 0$, $\forall i \in J$, and $g_i(x_0) = 0$, $\forall i \in K$.

The problem (P) has an optimal solution $\overline{x}$ if and only if there exist scalars $(\mu_i)_{i\in J\cup K}$ verifying:
(i) $\forall i \in J$, $\mu_i \geq 0, \mu_i g_i(\overline{x}) = 0$,
(ii) $f_0(\overline{x}) + \sum_{i\in J\cup K}\mu_i g_i(\overline{x}) \leq f_0(x) + \sum_{i\in J\cup K}\mu_i g_i(x)$ *for any* x

Proof. Obvious □

Remark 4. *(i) The Slater condition is very important even in the one-dimensional case. Consider the following example:*

$$\min\{f(x) = x \mid x^2 \leq 0\}.$$

The problem has a unique solution $\overline{x} = 0$. The Slater condition is not satisfied. There exists no $\lambda \geq 0$ such that $0 = \overline{x} + \lambda\overline{x}^2 \leq x + \lambda x$ for any $x \in \mathbb{R}$.
(ii) It is important to notice that Slater Condition is not necessary to obtain Kuhn-Tucker Conditions. In the previous example, Slater Condition is not satisfied and there is no Kuhn-Tucker coefficient. Now replace this problem by an equivalent problem which is $\min\{f(x) = x \mid g(x) = |x| \leq 0\}$. As before, Slater Condition does not hold. The unique solution is always $\overline{x} = 0$. Let $\lambda = 1$, we have successively $g(0) = 0, \lambda g(0) = 0$ and $0 = f(0) + \lambda g(0) \leq f(x) + \lambda g(x) = x + |x|$, for any $x \in \mathbb{R}$. In other words, Kuhn-Tucker Conditions hold.
(iii) In the previous example, one can check that Kuhn-Tucker Conditions are sufficient for 0 to be a solution. This result is quite general as it will be proved in the next proposition.

Maximization problem

In many economic models, we need to maximize a function subject to several constraints (physical constraint, financial constraint, legal constraint, ...). We present here the result concerning the following maximization problem (P'):

$$(P') \qquad \text{Maximize } f_0(x) \text{ under the constraints } \begin{cases} f_i(x) \leq 0, \ \forall i \in I \\ g_i(x) \leq 0, \ \forall i \in J \\ g_i(x) = 0, \ \forall i \in K. \end{cases}$$

where $f_0 : \mathbb{R}^n \to \mathbb{R}$ is a **concave function**, I, J and K are finite and possibly empty sets, for all $i \in I$, f_i is convex, non-affine function from $\mathbb{R}^n$ into $\mathbb{R}$, for all $i \in J \cup K$, g_i is a non-null affine function.

Definition 18. *The Lagrangian of Problem (P') is the function $L : \mathbb{R}^p_+ \times \mathbb{R}^q_+ \times \mathbb{R}^{r-q} \times \mathbb{R}^n \to \mathbb{R}$ defined by: for all $(\lambda, \mu, x) = ((\lambda_i)_{i\in I}, (\mu_i)_{i\in J}, (\mu_i)_{i\in K}, x)$,*

$$L(\lambda, \mu, x) = f_0(x) - \sum_{i\in I} \lambda_i f_i(x) - \sum_{i\in J} \mu_i g_i(x) - \sum_{i\in K} \mu_i g_i(x).$$

where $p = card(I), q = card(J)$ and $r = card(J \cup K)$.

Theorem 18 (Kuhn-Tucker (maximization problem))**.** *Assume that functions f_i, g_i are differentiable.*

Assume that Slater Condition is satisfied for Problem (P'). Then $\overline{x}$ is a solution to (P) if, and only if, there exists coefficients $(\lambda_i)_{i\in I}, (\mu_i)_{i\in J\cup K}$ which, together with $\overline{x}$, satisfy Kuhn-Tucker Conditions for Problem (P), i.e.,

1. $\forall i \in I,\ \lambda_i \geq 0, \lambda_i f_i(\overline{x}) = 0,$
   $\forall i \in J,\ \mu_i \geq 0, \mu_i g_i(\overline{x}) = 0.$

2. $$f_0(\overline{x}) - \sum_{i\in I} \lambda_i f_i(\overline{x}) - \sum_{i\in J\cup K} \mu_i g_i(\overline{x}) \geq f_0(x) - \sum_{i\in I} \lambda_i f_i(x) - \sum_{i\in J\cup K} \mu_i g_i(x) \qquad (11)$$

for any x.

If f_0, $(f_i)_{i\in I}, (g_i)_{i\in J\cup K}$ are differentiable, then condition (11) is replaced by

$$Df_0(\overline{x}) = \sum_{i\in I} \lambda_i Df_i(\overline{x}) + \sum_{i\in J\cup K} \mu_i Dg_i(\overline{x}) \tag{12}$$

Proof. This is a direct consequence of Theorem 17. □

3.3 Applications

Theorems 17 and 18 have many applications in economics, econometrics and finance. In the following, we present some applications.

3.3.1 Consumer maximization problem

The consumer maximizes her utility by choosing allocation $(x_1, \dots, x_n)$ of commodity

$$\max_{(x_1,\dots,x_n)} U(x_1, \dots, x_n)$$
$$\text{subject to: } p_1x_1 + \cdots p_nx_n \leq w$$
$$x_1 \geq 0, \dots, x_n \geq 0$$

where p_i is price of commodity i and w is the consumer's income. (p_i) and w are endogenously given.

Assume that the function $U : \mathbb{R}^n_+ \to \mathbb{R}$ is continuously differentiable. Assume that $p_i > 0\ \forall i$, $w > 0$. Assume that U is concave. Assume that $\frac{\partial U}{\partial x_i}(x) > 0\ \forall i, \forall x$.

We consider a simple example

Example 7. Consider $n = 2$ and $u(x) = x_1^{\alpha_1} x_2^{\alpha_2}$ where $\alpha_i > 0$. We want to solve the following problem

$$\max_{(x_1,x_2)} x_1^{\alpha_1} x_2^{\alpha_2}$$
$$\text{subject to: } p_1x_1 + p_2x_2 \leq w$$
$$x_1 \geq 0, x_2 \geq 0$$

where $p_i > 0$ is price of commodity i and $w > 0$ is the consumer's income.

Notice that our objective function is quasiconcave but not concave. We cannot directly apply Theorem 18. We proceed as follows.

If (a_1, a_2) is a solution to the above problem, then $a_1 > 0, a_2 > 0$ (why?). This implies that (a_1, a_2) is a solution of the following problem

$$(P_e) \quad \max_{(x_1,x_2)} f(x_1, x_2) = ln(x_1^{\alpha_1} x_2^{\alpha_2}) = \alpha_1 ln(x_1) + \alpha_2 ln(x_2)$$
$$\text{subject to: } p_1x_1 + p_2x_2 \leq w$$
$$x_1 \geq 0, x_2 \geq 0$$

The objective function is concave now. It is differentiable on $\mathbb{R}^2_{++}$. Applying Theorem 18, there exist non-negative multipliers $\lambda_1, \lambda_2, \lambda$ such that

$$\frac{\partial f}{\partial x_1}(a_1, a_2) = -\lambda_1 + p_1\lambda$$
$$\frac{\partial f}{\partial x_2}(a_1, a_2) = -\lambda_2 + p_2\lambda$$
$$\lambda_1 x_1 = 0, \quad \lambda_2 x_2 = 0, \lambda(p_1x_1 + p_2x_2 - w) = 0$$

Since $a_1 > 0, a_2 > 0$, we have $\lambda_1 = \lambda_2 = 0$. From this, we find that $\frac{p_1a_1}{\alpha_1} = \frac{p_2a_2}{\alpha_2}$. Observe that $\frac{\partial f}{\partial x_1}(a_1, a_2) = p_1\lambda$ which implies that $\lambda > 0$. Thus, we have $p_1a_1 + p_2a_2 = w$.[13]

Therefore, we get that

$$p_1a_1 = \frac{\alpha_1}{\alpha_1 + \alpha_2}w, \quad p_2a_2 = \frac{\alpha_2}{\alpha_1 + \alpha_2}w.$$

We have proved that if (a_1, a_2) is a solution, then $p_1a_1 = \frac{\alpha_1}{\alpha_1+\alpha_2}w, \quad p_2a_2 = \frac{\alpha_2}{\alpha_1+\alpha_2}w$. Lastly, we see that such a pair (a_1, a_2) with the above multipliers satisfy the Kuhn-Tucker conditions for the problem (P_e). We conclude that $(\frac{\alpha_1}{\alpha_1+\alpha_2}\frac{w}{p_1}, \frac{\alpha_2}{\alpha_1+\alpha_2}\frac{w}{p_2})$ is the unique solution of the original problem.

Exercise 21. *Consider the function* $u : \mathbb{R}^2_+ \to \mathbb{R}_+$ *defined by* $u(c_1, c_2) = a_1\frac{c_1^\alpha}{\alpha} + a_2\frac{c_2^\alpha}{\alpha}$ *where* $a_1 > 0, a_2 > 0, \alpha \in (0, 1)$.

Let $w > 0, p_1 > 0, p_2 > 0$.

We want to solve the following problem:

$$(P): \quad \textit{Maximize } u(c_1, c_2) \textit{ under the constraints: } \begin{cases} p_1c_1 + p_2c_2 \le w \\ c_1 \ge 0, c_2 \ge 0 \end{cases}$$

1. *Prove that the set* $B \equiv \{(c_1, c_2) : p_1c_1 + p_2c_2 \le w, c_1 \ge 0, c_2 \ge 0\}$ *is convex, compact.*
2. *Prove that the problem (P) has a solution.*
3. *Prove that the function u is strictly concave on* $\mathbb{R}^2_+$.
4. *Prove that: if* (c_1, c_2) *is a solution, then* $p_1c_1 + p_2c_2 = w$.
5. *Assume that* (c_1, c_2) *is a solution and* $c_1 > 0, c_2 > 0$.
   (a) *Is the Slater condition satisfied and why? If this is the case, write the Kuhn-Tucker conditions.*
   (b) *Find this solution* (c_1, c_2).
   (c) *When* p_1 *increases, how does the solution* (c_1, c_2) *change? Provide an economic interpretation.*

[13] In general, if the function f is strictly increasing in one component, then $p_1a_1 + p_2a_2 = w$. Indeed, if $p_1a_1 + p_2a_2 < w$, then we can increase a_1 and/or a_2 by ϵ with ϵ is low enough such that $p_1(a_1 + \epsilon) + p_2(a_2 + \epsilon) < w$ to get a strictly higher value of the objective function).

6. *We want to prove that the solution found in Question 5 is the unique solution of the problem (P).*

   (a) *Assume that $c_1 = 0$ at optimum. Find c_2 and compute $u_0 \equiv u(c_1, c_2)$*

   (b) *Prove that $(x, \frac{w-p_1x}{p_2})$ is in the set B for any x satisfying $\in [0, w/p_1]$.*

   (c) *Compute $\lim_{x\to 0, x>0} \frac{u(x,\frac{w-p_1x}{p_2})-u_0}{x}$ and prove that $\frac{u(x,\frac{w-p_1x}{p_2})-u_0}{x} > 0$ for $x > 0$ small enough.*

   (d) *Prove that c_1 cannot be zero.*

   (e) *Prove that $c_1 > 0, c2 > 0$ at optimum and hence the solution (c_1, c_2) in Question 5 is the unique solution of the problem (P).*

7. *We now use another method to prove that the solution found in Question 5 is the unique solution of the problem (P).*

   (a) *Assume that (c_1', c_2') is a solution of the problem (P) and that $(c_1', c_2') \neq (c_1, c_2)$. Prove that $u(\frac{c_1+c_1'}{2}, \frac{c_2+c_2'}{2}) > u(c_1, c_2)$.*

   (b) *How can we get a contradiction?*

Proof. 1. Convexity: let $c = (c_1, c_2)$ and $c' = (c_1', c_2')$ be in B. We have to check that $\lambda c + (1-\lambda)c' \in B$ for any $\lambda \in [0, 1]$. This is easy.

   Compactness: we see that $p_ic_i \leq w \ \forall i = 1, 2$. Since $p_i > 0$, we have $c_i \in [0, w/p_i]$. So, B is bounded. It is easy to see that B is closed. Therefore, B is compact.

2. The function u is continuous and B is compact. So, the problem (P) has a solution.

3. The function $a_i\frac{c_i^\alpha}{\alpha}$ is strictly concave for any $i = 1, 2$. So, the function u is strictly concave.

   Remark: We can also compute the Hessian matrix and then prove that u is strictly concave.

4. It is easy to see that, if (c_1, c_2) is a solution, then $p_1c_1 + p_2c_2 = w$. Otherwise, we can improved a little bit c_1 or c_2 to get a strictly higher utility.

5. Assume that (c_1, c_2) is a solution and $c_1 > 0, c_2 > 0$.

   (a) The Problem (P) satisfies *Slater Condition (S)* if there exists (c_1, c_2) such that:

$$\begin{aligned} p_1c_1 + p_2c_2 &\leq w \\ -c_1 &\leq 0 \\ -c_2 &\leq 0. \end{aligned}$$

   It is easy to see that Slater condition is satisfied. So, we can write the first-order conditions

$$\begin{aligned} a_1c_1^{\alpha-1} &= \lambda p_1 \\ a_2c_2^{\alpha-1} &= \lambda p_2 \end{aligned}$$

(b) From the first-order condition, we can compute c_1 and c_2 as functions of λ. Then, we use $p_1c_1 + p_2c_2 = w$ to get an equation of λ. Solving this equation, we find λ and then compute c_1, c_2:

$$c_1 = \frac{\left(\frac{a_1}{p_1}\right)^{\frac{1}{1-\alpha}}}{\frac{a_1^{\frac{1}{1-\alpha}}}{p_1^{\frac{\alpha}{1-\alpha}}} + \frac{a_2^{\frac{1}{1-\alpha}}}{p_2^{\frac{\alpha}{1-\alpha}}}}, \quad c_2 = \frac{\left(\frac{a_2}{p_2}\right)^{\frac{1}{1-\alpha}}}{\frac{a_1^{\frac{1}{1-\alpha}}}{p_1^{\frac{\alpha}{1-\alpha}}} + \frac{a_2^{\frac{1}{1-\alpha}}}{p_2^{\frac{\alpha}{1-\alpha}}}}$$

(c) When p_1 increases, we see that the optimal value c_1 decreases (this is a version of "Law of demand"). We also observe that c_2 increases (this is a substitution effect).

6. If $c_1 = 0$, we have $p_2c_2 = w$ and hence $c_2 = w/p_2$. We have $u_0 = u(0, w/p_2) = \frac{a_2}{\alpha}(\frac{w}{p_2})^\alpha$

   It is easy to see that $(x, \frac{w-p_1x}{p_2})$ is in the set B for any x satisfying $\in [0, w/p_1]$.

   We have

$$\begin{aligned}\frac{u(x, \frac{w-p_1x}{p_2}) - u_0}{x} &= \frac{a_1\frac{x^\alpha}{\alpha} + a_2\frac{(\frac{w-p_1x}{p_2})^\alpha}{\alpha} - \frac{a_2}{\alpha}(\frac{w}{p_2})^\alpha}{x} \\ &= a_1\frac{x^{\alpha-1}}{\alpha} + \frac{a_2\frac{(\frac{w-p_1x}{p_2})^\alpha}{\alpha} - \frac{a_2}{\alpha}(\frac{w}{p_2})^\alpha}{x}\end{aligned}$$

   The first terms converges to infinity while the second converges to $a_2\frac{p_1}{p_2}(\frac{w}{p_2})^{\alpha-1}$ when x goes to zero (because $\alpha - 1 < 0$).

   Since $\frac{u(x,\frac{w-p_1x}{p_2})-u_0}{x}$ is continuous in x, we have $\frac{u(x,\frac{w-p_1x}{p_2})-u_0}{x} > 0$ for $x > 0$ small enough.

   We have seen that $u(x, \frac{w-p_1x}{p_2}) > u_0$ for $x > 0$ small enough. So, the allocation $(0, w/p_2)$ cannot be optimal. It means that c_1 cannot be zero. By using the same argument, c_2 cannot be zero. So, we have $c_1 > 0$ and $c_2 > 0$ at optimal.

7. We now use another method to prove that the solution found in Question 5 is the unique solution of the problem (P).

   Assume that (c_1', c_2') is a solution of the problem (P) and that $(c_1', c_2') \neq (c_1, c_2)$. Since u is strictly concave, we have

$$u(\frac{c_1 + c_1'}{2}, \frac{c_2 + c_2'}{2}) > \frac{1}{2}u(c_1, c_2) + \frac{1}{2}u(c_1', c_2') \geq \frac{1}{2}u(c_1, c_2) + \frac{1}{2}u(c_1, c_2) = u(c_1, c_2).$$

   So, $(\frac{c_1+c_1'}{2}, \frac{c_2+c_2'}{2})$ is in the set B and $u(\frac{c_1+c_1'}{2}, \frac{c_2+c_2'}{2})$ is strictly higher that the maximum value $u(c_1, c_2)$. This is a contradiction. So, we have proved the uniqueness.

$\square$

Exercise 22. *(1) For the following utility functions, check whether there are continuous, monotomic, strictly monotonic, concave, quasi-concave.*

$$\begin{aligned}
\textit{Leontief preference:}\quad & u(x) = min\{\alpha_1 x_1, \alpha_2 x_2\}, \quad \alpha_h > 0, h = 1,2\\
& u(x) = c_1x_1 + c_2x_2, \quad c_h > 0\\
& u(x) = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}, \quad \alpha_h > 0\\
& u(x) = (x_1^{\alpha} + x_2^{\alpha})^{\frac{1}{\alpha}}, \quad \alpha \in (0,1)\\
& u(x) = x_1(1+\sqrt{x_2})\\
& u(x) = x_1 + \sqrt{x_2})\\
& u(x) = -e^{-\alpha_1 x_1} - e^{-\alpha_2 x_2}, \quad \alpha_h > 0.
\end{aligned}$$

(2) For each of above functions, find the solution of the following maximization problem

$$\begin{aligned}
&\max_{(x_1,x_2)} U(x_1,x_2)\\
&\quad \textit{subject to: } p_1x_1 + p_2x_2 \le w\\
&\qquad\qquad\quad x_1 \ge 0, x_2 \ge 0
\end{aligned}$$

where $p_i > 0$ is price of commodity i and $w > 0$ is the consumer's income.

Proof. Let consider the Leontief utility function. $u(x) = min\{\alpha_1 x_1, \alpha_2 x_2\}, \quad \alpha_h > 0, h = 1,2$. This function is continuous. It is increasing in each component but it may not be strictly increasing. Indeed, if $min\{\alpha_1 x_1, \alpha_2 x_2\} = \alpha_1 x_1$, then we have $u(x_1, x_2') = u(x_1, x_2)$ for all $x_2' > x_2$.

Notice that this function is not differentiable at the point (x_1, x_2) satisfying $\alpha_1 x_1 = \alpha_2 x_2$. So, we cannot use our tests based on derivatives to verify the (quasi)concavity of u. We can prove the concavity of this function by using definition. See also Proposition 5. □

Exercise 23. *In a two-good economy. The consumption set of a consumer is $X = \{x \in \mathbb{R}^2_+ : x_1 + x_2 \ge 1\}$. Let $u(x_1,x_2) = x_1 + 4x_2$. Let the price $(p_1,p_2) = (1,2)$.*

1. *Find the solution of the following maximization problem*

$$\begin{aligned}
&\max_{(x_1,\dots,x_n)} U(x_1,x_2)\\
&\quad \textit{subject to: } x \in X\\
&\qquad\qquad\quad p_1x_1 + p_2x_2 \le w
\end{aligned}$$

   where $w > 0$ is the consumer's income.

2. *The solution (x_1,x_2) depends on w. Does the demand of the commodity 1 (i.e., $x(1)$) increase with the income w?*

3.3.2 A two-period optimal growth model

An agent living for two periods (present and future, represented by 0 and 1) wants to choose consumption allocation (c_0, c_1) and physical capital k_1 to maximize her(his) utility $U(c_0, c_1)$

$$\max_{(c_0,c_1,k_1)} U(c_0, c_1) \tag{13}$$

$$\text{subject to: } c_1 \geq 0, c_1 \geq 0, k_1 \geq 0 \tag{14}$$

$$c_0 + k_1 \leq w_0 \tag{15}$$

$$c_1 \leq w_1 + F(k_1) \tag{16}$$

where w_0, w_1 are given and strictly positive. The function F is assumed to be strictly increasing, concave, continuously differentiable and $F(0) = 0$.

We would like to solve this problem to understand the optimal value of (c_0, c_1, k_1). Notice that in this setup, k_t can be also interpreted as investment.

First, it is easy to see that the Slater condition is satisfied. So, we can write the Lagrangian

$$L = U(c_0, c_1) + \mu_0 c_0 + \mu_1 c_1 + \mu_k k_1 + \lambda_0(w_0 - c_0 - k_1) + \lambda_1(w_1 + F(k_1) - c_1) \tag{17}$$

Assume that $U(c_0, c_1) = u(c_0) + \beta u(c_1)$ where the function u is strictly concave, increasing, continuously differentiable, $u'(0) = \infty$. Parameter β represents the rate of time preference.

Since $u'(0) = \infty$, we have $c_0 > 0, c_1 > 0$ at optimum (@reader: why?), which implies that $\mu_0 = \mu_1 = 0$. Therefore, the first-order conditions become

$$\frac{\partial U(c_0, c_1)}{\partial c_0} - \lambda_0 \Leftrightarrow u'(c_0) = \lambda_0$$

$$\frac{\partial U(c_0, c_1)}{\partial c_1} - \lambda_1 \Leftrightarrow \beta u'(c_1) = \lambda_1$$

$$\mu_k - \lambda_0 + \lambda_1 F'(k_1) = 0 \Leftrightarrow \lambda_0 = \lambda_1 F'(k_1) + \mu_k$$

$$\mu_k k_1 = 0, \mu_k \geq 0$$

So, we obtain that $u'(c_0) = \beta F'(k_1)u'(c_1) + \mu_k$. At optimum, we must have $c_0 + k_1 = w_0$ and $c_1 = w_1 + F(k_1)$. Thus, we get that

$$u'(w_0 - k_1) = \beta F'(k_1)u'(w_1 + F(k_1)) + \mu_k \tag{18}$$

Notice that at this stage, we do not require that $F'(0) = \infty$.

There are two cases.

1. $k_1 > 0$. In this case, we have $\mu_k = 0$ and hence k_1 is determined by

$$H(k_1) \equiv u'(w_0 - k_1) - \beta F'(k_1)u'(w_1 + F(k_1)) = 0 \tag{19}$$

The function H is strictly increasing in k_1. $H(0) = u'(w_0) - \beta F'(0)u'(w_1)$ while $H(w_0) = \infty$ because $u'(0) = \infty$. So, the existence of a strictly positive solution k_1 requires that

$$u'(w_0) - \beta F'(0)u'(w_1) < 0.$$

It means that the productivity $F'(0)$, the rate of time preference β, the endowment at initial date w_0 are high and the endowment at date 1 is low.

2. $k_1 = 0$. Condition (18) becomes

$$u'(w_0) - \beta F'(0)u'(w_1) \leq 0.$$

To sum up, we obtain the following result:

Proposition 14. *Assume that above conditions hold. Assume also that $u'(0) = \infty$. The optimal choice k_1 is strictly positive if and only if*

$$u'(w_0) - \beta F'(0)u'(w_1) < 0$$

In such a case, k_1 is the unique solution to the equation $H(k_1) = 0$.

It is useful to consider some particular cases.

1. Assume that $u(c) = ln(c)$, $F'(k) = Ak^\alpha$ where $\alpha \in (0, 1)$, and $w_1 = 0$. In this case, $F'(0) = \infty$. We have

$$\frac{1}{w_0 - k_1} = \beta\alpha A k_1^{\alpha-1}\frac{1}{w_1 + Ak_1^\alpha} = \beta\alpha A k_1^{\alpha-1}\frac{1}{Ak_1^\alpha} \tag{20}$$

$$\Leftrightarrow k_1 = \frac{\alpha\beta}{1+\alpha\beta}w_0 \tag{21}$$

   The investment k_1 is increasing in the initial endowment w_0 and the rate of time preference.

2. Assume that $u(c) = ln(c)$, $F'(k) = Ak$, $w_1 > 0$. We have that

   (a) If $w_1 < \beta A w_0$, then k_1 is strictly positive and we can compute that $k_1 = \frac{\beta A w_0 - w_1}{A(1+\beta)}$. The investment is increasing in the productivity A, the rate of time preference β, the initial endowment but decreasing in the endowment in the future.

   (b) If $w_1 \geq \beta A w_0$, then $k_1 = 0$. The intuition: when the productivity, the initial endowment, the rate of time preference are low, but the endowment in the future is high, we do not need to save/invest.

3.3.3 Cost minimization problem

Let prices be strictly positive $p_1 > 0, \dots, p_n > 0$. The firm minimizes its cost by choosing allocation $(x_1, \dots, x_n)$ of inputs such that the production $F(x_1, \dots, x_n)$ is not less than a given level y.

$$\min_{(x_1,\dots,x_n)} p_1x_1 + \dots + p_nx_n \tag{22}$$

$$\text{subject to: } x_1 \geq 0, x_n \geq 0 \tag{23}$$

$$F(x_1, \dots, x_n) \geq y \tag{24}$$

Example 8. *Assume that $n = 2$, $F(x_1, x_2) = Ax_1^{a_1}x_2^{a_2}$ with $a_1, a_2 > 0$ and $a_1 + a_2 \leq 1$. Find the solution of the cost minimization problem.*

Solution. This is a minimization problem. The objective function is linear and hence convex.

We rewrite $F(x_1, x_2) \geq y$ as $y - F(x_1, x_2) \leq 0$. Since the function F is concave, the function $-F$ is convex. We can easily check that the Slater condition is satisfied. So, we can apply Theorem 17. The Lagrangian is

$$L = p_1x_1 + p_2x_2 + \lambda\big(y - Ax_1^{a_1}x_2^{a_2}\big) + \mu_1(-x_1) + \mu_2(-x_2)$$

The first order conditions are

$$\begin{aligned} \frac{\partial L}{\partial x_1} &= p_1 - \lambda a_1 A x_1^{a_1-1}x_2^{a_2} - \mu_1 = 0 \\ \frac{\partial L}{\partial x_2} &= p_2 - \lambda a_2 A x_1^{a_1}x_2^{b_1-1} - \mu_2 = 0 \\ \mu_1 x_1 = 0&, \quad \mu_2 x_2 = 0 \end{aligned}$$

Since $F(x_1, x_2) = Ax_1^{a_1}x_2^{a_2} \geq y > 0$, we have $x_1 > 0, x_2 > 0$. Thus, $\mu_1 = \mu_2 = 0$.

According to the FOCs, we have

$$\frac{p_1}{p_2} = \frac{\lambda a_1 A x_1^{a_1-1}x_2^{a_2}}{\lambda a_2 A x_1^{a_1}x_2^{a_2-1}} = \frac{a_1x_2}{a_2x_1} \Rightarrow x_1 = \frac{a_1}{a_2}\frac{p_2}{p_1}x_2$$

At optimum, we must have $y = F(x_1, x_2)$, and hence,

$$y = Ax_1^{a_1}x_2^{a_2} = A\big(\frac{a_1}{a_2}\frac{p_2}{p_1}x_2\big)^{a_1}x_2^{a_2} = A\big(\frac{a_1}{a_2}\frac{p_2}{p_1}\big)^{a_1}x_2^{a_1+a_2}$$

Therefore, we find that

$$\begin{aligned} x_2 &= A^{\frac{-1}{a_1+a_2}}\big(\frac{a_2}{a_1}\frac{p_1}{p_2}\big)^{\frac{a_1}{a_1+a_2}}y^{\frac{1}{a_1+a_2}} \\ x_1 &= A^{\frac{-1}{a_1+a_2}}\big(\frac{a_1}{a_2}\frac{p_2}{p_1}\big)^{\frac{a_1}{a_1+a_2}}y^{\frac{1}{a_1+a_2}} \end{aligned}$$

The optimal quantity of input 1 is increasing in the output y and the price of input 2 but decreasing in the price of input 1 and the productivity A.

We can also compute the cost function

$$\begin{aligned} Cost = p_1x_1 + p_2x_2 &= \frac{a_1}{a_2}p_2x_2 + p_2x_2 = \frac{a_1+a_2}{a_1}p_2x_2 \\ &= \frac{a_1+a_2}{a_2}p_2A^{\frac{-1}{a_1+a_2}}\big(\frac{a_2}{a_1}\frac{p_1}{p_2}\big)^{\frac{a_1}{a_1+a_2}}y^{\frac{1}{a_1+a_2}} \\ &= \frac{a_1+a_2}{a_2^{\frac{a_2}{a_1+a_2}}a_1^{\frac{a_1}{a_1+a_2}}}A^{\frac{-1}{a_1+a_2}}p_1^{\frac{a_1}{a_1+a_2}}p_2^{\frac{a_2}{a_1+a_2}}y^{\frac{1}{a_1+a_2}} \end{aligned}$$

The cost is increasing in input prices p_1, p_2, the output y, but decreasing in the productivity A. □

Proposition 15 (Shephard's lemma)**.** *Consider the cost minimization problem. Assume that F is in C^2 (i.e., twice continuously differentiable). Assume that F is strictly increasing in each component and concave. Let $x^*(y,p)$ denote the solution of the problem and $C(y,p)$ the optimal value (cost function). Suppose that $p_i > 0$ and $x_i^* > 0$ $\forall i$. Prove that*

$$x_i^*(y,p) = \frac{\partial C(y,p)}{\partial p_i} \forall i \tag{25}$$

$$\frac{\partial x_i^*}{\partial p_j} = \frac{\partial x_j^*}{\partial p_i} \forall i,j. \tag{26}$$

Proof. The Lagrangian is

$$L = \sum_i p_i x_i + \lambda\big(y - F(x_1,\ldots,x_n)\big) + \sum_i \mu_i(-x_i)$$

The first order conditions imply that, for any i,

$$\frac{\partial L}{\partial x_i} = p_i - \lambda \frac{\partial F}{\partial x_i}(x_1^*,\ldots,x_n^*) - \mu_i = 0$$
$$x_i^* \mu_i = 0$$

Under assumption $x_i^* > 0$, we have $\mu_i = 0$. So, we get that

$$p_i = \lambda \frac{\partial F}{\partial x_i}(x_1^*,\ldots,x_n^*), \forall i$$

The cost function is $C(y,p) = \sum_i p_i x_i^*$. Thus, we can compute

$$\frac{\partial C(y,p)}{\partial p_i} = x_i^* + \sum_{t=1}^n p_t \frac{\partial x_t^*}{\partial p_i} \tag{27}$$

By taking the derivative with respect to p_i of both sides of the equation $F(x_1^*,\ldots,x_n^*) = y$, we have

$$\sum_{t=1}^n \frac{\partial F}{\partial x_t}(x_1^*,\ldots,x_n^*)\frac{\partial x_t^*}{\partial p_i} = 0$$

Since $\frac{\partial F}{\partial x_t}(x_1^*,\ldots,x_n^*) = p_t/\lambda$, we get that $\sum_{t=1}^n p_t \frac{\partial x_t^*}{\partial p_i} = 0$. So, by combining with (27), we obtain $\frac{\partial C(y,p)}{\partial p_i} = x_i^*$. From this, we compute

$$\frac{\partial x_i^*}{\partial p_j} = \frac{\partial^2 C(y,p)}{\partial p_j \partial p_i}$$
$$\frac{\partial x_j^*}{\partial p_i} = \frac{\partial^2 C(y,p)}{\partial p_i \partial p_j}$$

Since $\frac{\partial^2 C(y,p)}{\partial p_j \partial p_i} = \frac{\partial^2 C(y,p)}{\partial p_i \partial p_j}$, we obtain (26). □

3.3.4 Profit maximization problem

Let prices of inputs be strictly positive $p_1 > 0, \dots, p_n > 0$. Denote p be the price of output. The firm maximizes its profit by choosing allocation $(x_1, \dots, x_n)$ of inputs:

$$\max_{(x_1,\dots,x_n)} pF(x_1, \dots, x_n) - (p_1x_1 + \dots + p_nx_n) \tag{28}$$

$$\text{subject to: } x_1 \geq 0, x_n \geq 0 \tag{29}$$

Assume that the function $F : \mathbb{R}^n_+ \to \mathbb{R}$ is continuously differentiable. Assume that F is concave. Assume that F is increasing in each component. $\frac{\partial F}{\partial x_i}(x) > 0 \ \forall i, \forall x$.

In the case of two inputs, the problem becomes

$$\max_{(x_1,x_2)} pF(x_1, x_2) - (p_1x_1 + p_2x_2) \tag{30}$$

$$\text{subject to: } x_1 \geq 0, x_2 \geq 0 \tag{31}$$

Exercise 24. *Find the solution and compute the maximum profit of the profit maximization problem for the following cases:*

1. $n = 1$, $F(x) = ax$ *where* $a > 0$.
2. $n = 1$, $f(x) = Ax^\alpha$ *where* $A > 0$, $\alpha \in (0, 1)$.
3. $n = 2$, $F(x_1, x_2) = Ax_1^{\alpha_1}x_2^{\alpha_2}$ *where* $A > 0, \alpha_1 > 0, \alpha_2 > 0, \alpha_1 + \alpha_2 \leq 1$.
4. $n = 2$, $f(x_1, x_2) = ax_1 + bx_2$, *where* $a, b > 0$.
5. $n = 2$, $f(x_1, x_2) = min(ax_1, bx_2)$ *where* $a, b > 0$.
6. $f(x_1, x_2) = \left(Ax_1^r + Bx_2^r\right)^{\frac{1}{r}}$, *where* $r < 1, r \neq 0$.
7. $n = 3$, $F(x_1, x_2, x_3) = Ax_1^{\alpha_1}x_2^{\alpha_2}x_3^{\alpha_3}$ *where* $A > 0, \alpha_1 > 0, \alpha_2 > 0, \alpha_3 > 0, \alpha_1+\alpha_2+\alpha_3 \leq 1$.

   When $\alpha_1 + \alpha_2 + \alpha_3 = 1$, *compute the cost share* $\frac{p_ix_i}{\sum_i p_ix_i}$.

Linear production functions. Let us consider the case $f(x_1, x_2) = ax_1 + bx_2$, where $a, b > 0$. Then $pf(x_1, x_2) - p_1x_1 - p_2x_2 = p(ax_1 + bx_2) - p_1x_1 - p_2x_2$ The profit maximization problem becomes

$$\max_{(x_1,\, x_n)} (pa - p_1)x_1 + (pb - p_2)x_2$$

$$\text{subject to: } x_1 \geq 0, x_2 \geq 0$$

Suppose that (x_1, x_2) is a solution. It is easy to see that

- If $pa = p_1$, then x_1 takes any value.
- If $pa > p_1$, then $x_1 = \infty$.
- If $pa < p_1$, then $x_1 = 0$.

$\square$

Leontief production functions. We now consider $f(x_1, x_2) = min(ax_1, bx_2)$ where $a, b > 0$. The profit maximization problem becomes

$$\max_{(x_1, x_n)} pmin(ax_1, bx_2) - p_1x_1 - p_2x_2 \text{ subject to: } x_1 \geq 0, x_2 \geq 0$$

The objective function is not differentiable. So, we cannot apply the Kuhn-Tucker theorem. We will prove that: if (x_1, x_2) is a solution, then $ax_1 = bx_2$. Indeed, if, $ax_1 > bx_2$, then $min(ax_1, bx_2) = bx_2$. We can introduce $x_1' = x_1 - \epsilon$ where $ax_1' > bx_2$. Then, the new profit $\pi(x_1', x_2)$ is strictly higher than $\pi(x_1, x_2)$, a contradiction. Hence, we cannot have $ax_1 > bx_2$. Similarly, we cannot have $ax_1 < bx_2$. So, $ax_1 = bx_2$. The remaining is simple.

$\square$

Cobb-Douglas production function (the case of one variable). The problem becomes

$$\max_{x \geq 0} pAx^\alpha - p_1x$$

where $\alpha \in (0, 1)$ while p, p_1 are the prices of output, input respectively.

First, a solution $\bar{x}$ must be strictly positive $\bar{x} > 0$ (why?).[14] The FOC implies that we have $\alpha pA(\bar{x})^{\alpha-1} = p_1$ at optimal. So, the solution is

$$\bar{x} = (\frac{\alpha pA}{p_1})^{\frac{1}{1-\alpha}}.$$

So, the optimal choice of input of the firm is increasing in the productivity A and the output price p while decreasing in the input price p_1. $\square$

Cobb-Douglas production function (the case of two variables). The problem becomes

$$\max_{x_1 \geq 0, x_2 \geq 0} pAx_1^{\alpha_1}x_2^{\alpha_2} - p_1x_1 - p_2x_2$$

where $\alpha_1 + \alpha_2 \leq 1$.

Denote $\Pi \equiv \max_{x_1 \geq 0, x_2 \geq 0} pAx_1^{\alpha_1}x_2^{\alpha_2} - p_1x_1 - p_2x_2$. Observe that $\Pi \geq 0$. Moreover, at optimum, if $x_1 = 0$ or $x_2 = 0$, then $x_1 = x_2 = 0$.

1. Assume that $\alpha_1 + \alpha_2 = 1$. Let (x_1, x_2) be a solution.

   If $x_1 = 0$ or $x_2 = 0$, then $\Pi = 0$.

   If $x_1 > 0, x_2 > 0$. We have the FOCs

$$\alpha_1 pAx_1^{\alpha_1-1}x_2^{\alpha_2} - p_1 = 0$$
$$\alpha_2 pAx_1^{\alpha_1}x_2^{\alpha_2-1} - p_2 = 0.$$

[14] Indeed, by definition of $\bar{x}$, we have $pA(\bar{x})^\alpha - p_1\bar{x} \geq pAx^\alpha - p_1x, \forall x$. If $\bar{x} = 0$, we have $0 \geq pAx^\alpha - p_1x = x(pAx^{\alpha-1} - p_1), \forall x$. Hence, $0 \geq x(pAx^{\alpha-1} - p_1), \forall x$. Let x be closed to zero but still strictly positive so that $pAx^{\alpha-1} - p_1 > 0$, we have a contradiction. Therefore, we have $\bar{x} > 0$.

Hence,

$$\frac{p_1}{p_2} = \frac{\alpha_1 pAx_1^{\alpha_1-1}x_2^{\alpha_2}}{\alpha_2 pAx_1^{\alpha_1}x_2^{\alpha_2-1}} = \frac{\alpha_1 x_2}{\alpha_2 x_1} \Leftrightarrow x_2 = \frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1.$$

From this and the FOCs, we have

$$p_1 = \alpha_1 pAx_1^{\alpha_1-1}\big(\frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1\big)^{\alpha_2} \Leftrightarrow pA = \frac{p_1^{1-\alpha_2}p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2}\alpha_2^{\alpha_2}}$$

To sum up, we have the following result:

(a) If $pA < \frac{p_1^{1-\alpha_2}p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2}\alpha_2^{\alpha_2}}$, then the unique solution is $x_1 = x_2 = 0$. In this case, the profit $\Pi = 0$.

(b) If $pA = \frac{p_1^{1-\alpha_2}p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2}\alpha_2^{\alpha_2}}$, then any coupe (x_1, x_2) satisfying $x_2 = \frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1 \geq 0$ is a solution. In this case, the profit $\Pi = 0$.

(c) If $pA > \frac{p_1^{1-\alpha_2}p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2}\alpha_2^{\alpha_2}}$, there is no solution. Indeed, let us consider an allocation $x_2 = \frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1$, we have that

$$pAx_1^{\alpha_1}x_2^{\alpha_2} - p_1x_1 - p_2x_2 = pAx_1^{\alpha_1}(\frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1)^{\alpha_2} - p_1x_1 - p_2\frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1 \tag{32}$$

$$= pAx_1^{\alpha_1+\alpha_2}\frac{p_1^{\alpha_2}}{p_2^{\alpha_2}}\frac{\alpha_2^{\alpha_2}}{\alpha_1^{\alpha_2}} - p_1x_1(1+\frac{\alpha_2}{\alpha_1}) = pAx_1\frac{p_1^{\alpha_2}}{p_2^{\alpha_2}}\frac{\alpha_2^{\alpha_2}}{\alpha_1^{\alpha_2}} - \frac{p_1x_1}{\alpha_1} \tag{33}$$

$$= x_1\frac{p_1}{\alpha_1}\Big(pA\frac{\alpha_2^{\alpha_2}\alpha_1^{1-\alpha_2}}{p_1^{1-\alpha_2}p_2^{\alpha_2}} - 1\Big) \tag{34}$$

Since $pA\frac{\alpha_2^{\alpha_2}\alpha_1^{1-\alpha_2}}{p_1^{1-\alpha_2}p_2^{\alpha_2}} - 1 > 0$, when we let x_1 tend to infinity, the profit of this allocation (x_1, x_2) tends to infinity.

2. Assume that $\alpha_1 + \alpha_2 - 1 < 1$. A solution (x_1, x_2) must satisfy $x_1 > 0, x_2 > 0$ (why?). We have the FOCs

$$\alpha_1 pAx_1^{\alpha_1-1}x_2^{\alpha_2} - p_1 = 0$$
$$\alpha_2 pAx_1^{\alpha_1}x_2^{\alpha_2-1} - p_2 = 0.$$

Hence, we get

$$\frac{p_1}{p_2} = \frac{\alpha_1 pAx_1^{\alpha_1-1}x_2^{\alpha_2}}{\alpha_2 pAx_1^{\alpha_1}x_2^{\alpha_2-1}} = \frac{\alpha_1 x_2}{\alpha_2 x_1} \Leftrightarrow x_2 = \frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1.$$

From this and the FOCs, we have

$$p_1 = \alpha_1 pAx_1^{\alpha_1-1}\big(\frac{p_1}{p_2}\frac{\alpha_2}{\alpha_1}x_1\big)^{\alpha_2} \Leftrightarrow pAx_1^{\alpha_1+\alpha_2-1} = \frac{p_1^{1-\alpha_2}p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2}\alpha_2^{\alpha_2}}$$

Since $\alpha_1 + \alpha_2 - 1 < 1$, we can easily find that

$$x_1 = \left(\frac{pA}{\frac{p_1^{1-\alpha_2} p_2^{\alpha_2}}{\alpha_1^{1-\alpha_2} \alpha_2^{\alpha_2}}} \right)^{\frac{1}{1-\alpha_1-\alpha_2}}$$

This is increasing in the productivity A, the output price p but decreasing in the input prices p_1, p_2.

□

3.3.5 Least squares regression

The sum of squared residuals (in the multiple linear regression):

$$f(\beta_0, \beta_1, \cdots, \beta_k) \equiv \sum_{i=1}^{n} \Big(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{i,j} \Big)^2$$

on $\mathbb{R}^{k+1}$.

1. Let us start by considering the case $k = 1$.

$$f(\beta_0, \beta_1) \equiv \sum_{i=1}^{n} \Big(y_i - \beta_0 - \beta_1 x_i \Big)^2$$

We want to solve the following problem

$$\begin{aligned} &\text{Minimize}_{(\beta_0, \beta_1)} f(\beta_0, \beta_1) \equiv \sum_{i=1}^{n} \Big(y_i - \beta_0 - \beta_1 x_i \Big)^2 \\ &\text{subject to } \beta_0 \in \mathbb{R}, \beta_1 \in \mathbb{R} \end{aligned}$$

This is an unconstrained convex minimization problem.

The objective function is convex (why?).

The first-order conditions (or Kuhn-Tucker conditions) give us

$$\beta_1 \sum_{i=1}^{n} x_i (x_i - \bar{x}) = \sum_{i=1}^{n} x_i (y_i - \bar{y})$$
$$\beta_0 = \bar{y} - \beta_1 \bar{x}$$

where $\bar{x} \equiv \frac{\sum_{i=1}^{n} x_i}{n}$, $\bar{y} \equiv \frac{\sum_{i=1}^{n} y_i}{n}$ are the sample averages.

Notice that $\sum_{i=1}^{n} x_i (x_i - \bar{x}) = \sum_{i=1}^{n} (x_i - \bar{x})^2$ and $\sum_{i=1}^{n} x_i (y_i - \bar{y}) = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$.

2. We now consider the general case. We want to solve the following problem

$$\begin{aligned} &\text{Minimize}_{(\beta_0, \ldots, \beta_k)} \sum_{i=1}^{n} \Big(y_i - \beta_0 - \sum_{j=1}^{k} \beta_j x_{i,j} \Big)^2 \\ &\text{subject to } \beta_0 \in \mathbb{R}, \cdots, \beta_k \in \mathbb{R} \end{aligned}$$

This is an unconstrained convex minimization problem.

3.4 Non-convex optimization (optional)

We follow the exposition of Florenzano and Le Van (2001).

3.4.1 Unconstrained Optimization

First-order conditions for an unconstrained optimum

Theorem 19. *Suppose $x^* \in S$ is either an unconstrained local minimum or an unconstrained local maximum. Then $Df(x^*) = 0$.*

Proof. Suppose that x^* is an unconstrained local maximum. There exists $r > 0$ such that $B(x^*, r) \subseteq S$ and $f(y) \leq f(x^*)$ for all $y \in B(x^*, r)$. Assume $a = Df(x^*) \neq 0$. From the Taylor expansion, we have

$$f(x) = f(x^*) + Df(x^*) \cdot (x - x^*) + R_1(x, x^*)\|x - x^*\|$$

For any real number t, define $x_t = x^* + ta$. For t close enough to zero, we have that $x_t \in B(x^*, r) \cap S$. Then

$$\begin{aligned} f(x_t) &= f(x^*) + a \cdot (x_t - x^*) + R_1(x_t, x^*)\|x_t - x^*\| \\ &= f(x^*) + t\|a\|^2 + R_1(x^* + ta, x^*)|t|\|a\| \\ &= f(x^*) + t\|a\| \left(\|a\| + R_1(x^* + ta, x^*)\right), \text{ when } t \geq 0 \\ &> f(x^*) \end{aligned}$$

when t is small enough. We get a contradiction. Hence $Df(x^*) = 0$. The proof is similar when x^* is an unconstrained local minimum. □

A point x is an optimum for f if it is either a maximum or a minimum of f. If it is an unconstrained local optimum, we have proved that $Df(x) = 0$.

Second order conditions for an unconstrained local optimum

Lemma 7. *(i) Let M be a positive definite $n \times n$ matrix. Let $S(0, 1)$ denote the unit-sphere of $\mathbb{R}^n$. Then $\min_{x \in S(0,1)} x'Mx > 0$.*

(ii) Let M be a negative definite $n \times n$ matrix. Let $S(0, 1)$ denote the unit-sphere of $\mathbb{R}^n$. Then $\max_{x \in S(0,1)} x'Mx < 0$.

Proof. (i) The function $\psi : S(0, 1) \to \mathbb{R}_+$ defined by $\psi(x) = x'Mx$ for $x \in S(0, 1)$ is continuous and positive for any $x \in S(0, 1)$. Since $S(0, 1)$ is compact, ψ has a minimum on $S(0, 1)$ which is positive, i.e. there exists $\bar{x} \in S(0, 1)$ which satisfies $0 < \psi(\bar{x}) = \bar{x}'M\bar{x} = \min_{x \in S(0,1)} x'Mx$.

(ii) The proof is similar. □

Theorem 20. *Let $f : S \to \mathbb{R}$, where $S \subseteq \mathbb{R}^n$. Assume x_0 is an unconstrained local optimum of f. If $D^2 f(x_0)$ is negative definite then x_0 is an unconstrained local maximum. If $D^2 f(x_0)$ is positive definite, then x_0 is an unconstrained local mimimum.*

Proof. We must have $Df(x_0) = 0$. Consider the Taylor expansion

$$\begin{aligned} f(x) &= f(x_0) + Df(x_0)\cdot(x-x_0) + \frac{1}{2}(x-x_0)'D^2f(x_0)(x-x_0) + R_2(x,x_0)\|x-x_0\|^2 \\ &= f(x_0) + \frac{1}{2}(x-x_0)'D^2f(x_0)(x-x_0) + R_2(x,x_0)\|x-x_0\|^2 \end{aligned}$$

Assume $D^2f(x_0)$ positive definite. When $x \neq x_0$, let $u = \frac{x-x_0}{\|x-x_0\|} \in S(0,1)$. We know that $\min_{x\in S(0,1)} x'D^2f(x_0)x = \alpha > 0$. Therefore,

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{2}(x-x_0)'D^2f(x_0)(x-x_0) + R_2(x,x_0)\|x-x_0\|^2 \\ &= f(x_0) + \frac{1}{2}\|x-x_0\|^2u'D^2f(x_0)u + R_2(x,x_0)\|x-x_0\|^2, \text{ when } x\neq x_0 \\ &\geq f(x_0) + \frac{1}{2}\|x-x_0\|^2\left[\alpha + 2R_2(x,x_0)\right], \text{ when } x \neq x_0 \end{aligned}$$

When x is close to x_0 but different from x_0, we have $\alpha + 2R_2(x,x_0) > 0$ and thus $f(x) \geq f(x_0)$. We have that x_0 is an unconstrained local minimum.

Similarly, when $D^2f(x_0)$ is negative definite then x_0 is an unconstrained local maximum. □

3.4.2 Constrained optimization

Definition 19. *Let f be a continuously differentiable mapping from an open, nonempty convex set U of $\mathbb{R}^n$ into $\mathbb{R}^n$. Let $a \in U$. Then $f(a) = (f_1(a),\ldots,f_n(a))$.The Jacobian matrix $J_f(a)$ is*

$$J_f(a) = \begin{bmatrix} Df_1(a) \\ Df_2(a) \\ \cdots \\ Df_n(a) \end{bmatrix}$$

where the row-vector $Df_i(a)$ is the derivative of f_i at point a.

Let f be a function from an open, convex, nonempty set U of $\mathbb{R}^n$ into $\mathbb{R}$. We say that f is *locally minimal* at $\overline{x}$ under the constraints $x \in \Gamma$, if $\overline{x} \in \Gamma$ and if there exists a neighborhood V of $\overline{x}$ such that $\forall x \in V \cap \Gamma$, $f(x) \geq f(\overline{x})$.

Consider the following problem (P):

$$(\tilde{P}) \qquad \min f_0(x) \text{ under the constraints } \begin{cases} x \in U \\ f_i(x) \leq 0, \ \forall i = 1,\ldots,I \\ g_i(x) = 0, \ \forall i = 1,\ldots,K \end{cases}$$

where f_0 and f_i, for $i = 1,\ldots,I$, g_i, for i$i = 1,\ldots,K$ are continuously differentiable functions from an open, convex, nonempty set U of $\mathbb{R}^n$ into $\mathbb{R}$.

Lemma 8. *Let f be a differentiable function from an open, convex set $U \subset \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q$ into $\mathbb{R}$. We suppose $0 \in U$. Suppose that the function $f : (x,y,z) \to f(x,y,z)$ is locally minimal at 0 under the constraints $x \geq 0, y = 0$. Then:*

$$f'_x(0,0,0) \geq 0, \ f'_z(0,0,0) = 0.$$

Proof. Let $C = \{(x, y, z) \in \mathbb{R}^m \times \mathbb{R}^p \times \mathbb{R}^q \mid x \geq 0, y = 0\}$. Let $(r, s, t) \in C$. Then there exists $\lambda_1 > 0$ such that $\forall \lambda \in [0, \lambda_1[$, one has $(\lambda r, \lambda s, \lambda t) \in C \cap U$. The function F defined by $\forall \lambda \geq 0$, $F(\lambda) = f(\lambda r, \lambda s, \lambda t)$ is locally minimal at 0. Hence, $F'(0) \geq 0$, i.e.

$$f'(0, 0, 0) \cdot (r, s, t) \geq 0, \ \forall (r, s, t) \in C.$$

In particular, $f'_x(0, 0, 0) \cdot r \geq 0$, $\forall r \in \mathbb{R}^m$ such that $r \geq 0$, $f'_z(0, 0, 0) \cdot t \geq 0$, $\forall t \in \mathbb{R}^q$. Hence: $f'_x(0, 0, 0) \geq 0$ and $f'_z(0, 0, 0) = 0$. □

Let x be a *feasible point*, i.e. $f_i(x) \leq 0$, $\forall i = 1, \ldots, I$, and $g_i(x) = 0$, $\forall i = 1, \ldots, K$. Let $I(x) = \{i \mid f_i(x) = 0\}$. We say that the *constraints of Problem (P) are regular at x* if the gradients $(Df_i(x))_{i \in I(x)}, (Dg_i(x))_{i=1,\ldots,K}$ are linearly independent.

Theorem 21. *Suppose that f_0 is locally minimal at $\overline{x}$ under the constraints of Problem (P). Suppose also that the constraints are regular at $\overline{x}$. Then there exist non-negative scalars $(\lambda_i)_{i \in I}$, and scalars $(\mu_i)_{i \in K}$ such that:*
(i) $Df_0(\overline{x}) = -\sum_{i=1,\ldots,I} \lambda_i Df_i(\overline{x}) + \sum_{i=1,\ldots,K} \mu_j Dg_j(\overline{x})$,
(ii) $\lambda_i f_i(\overline{x}) = 0$, $\forall i = 1, \ldots, I$.

Proof. Observe that f_0 is also locally minimal at $\overline{x}$ under the constraints

$$\begin{cases} f_i(x) \leq 0, \ \forall i \in I(\overline{x}), \\ g_i(x) = 0, \ \forall i = 1, \ldots, K \end{cases}$$

where $I(\overline{x}) = \{i \in I \mid f_i(\overline{x}) = 0\}$.
Suppose that $I(\overline{x})$ has cardinal J. Since, by assumption, the constraints are regular at $\overline{x}$, there exists $\theta_1, \ldots, \theta_q$, vectors of $\mathbb{R}^n$, with $q = n - J - K$, such that the matrix $((Df_i(\overline{x}))_{i=1,\ldots,I}; (Dg_i(\overline{x}))_{i=1,\ldots,K}, \theta_1, \ldots, \theta_q)$ is invertible. Define $\tilde{\theta}_i(x) = \theta_i \cdot (x - \overline{x})$, $\forall i = 1, \ldots, q, \forall x \in \mathbb{R}^n$, and the map φ from $\mathbb{R}^n$ into $\mathbb{R}^n$ by

$$\varphi(x) = ((-f_i(x))_{i \in I(\overline{x})}, (g_j(x))_{i=1,\ldots,K}, (\tilde{\theta}_i(x))_{i=1,\ldots,q}).$$

One has $\varphi(\overline{x}) = 0$, and $J_\varphi(\overline{x})$ is invertible, where $J_\varphi(\overline{x})$ is the Jacobian matrix of φ at $\overline{x}$. From the Local inversion theorem (Theorem 22), φ has, in a neighborhood V of $\overline{x}$, an inverse φ^{-1} which is continuously differentiable.
Define, on $\varphi(V)$, the map $F = f_0 \circ \varphi^{-1}$. One has: $F(u, v, w) = f_0(x)$ if $\varphi(x) = (u, v, w)$. In particular, $F(0, 0, 0) = f_0(\overline{x})$. Hence $F(u, v, w) \geq F(0, 0, 0)$ for $u \geq 0, v = 0$. From Lemma 8, $F'_u(0, 0, 0) \geq 0$ and $F'_w(0, 0, 0) = 0$, that means: $F'(0, 0, 0) = (\lambda, \mu, 0)$ with $\lambda \in \mathbb{R}^J_+$ and $\mu \in \mathbb{R}^K$. But $Df_0(\overline{x}) = F'(0, 0, 0)(J_\varphi(\overline{x}))$. Since

$$J_\varphi(\overline{x}) = ((-Df_i(\overline{x}))_{i \in I(\overline{x})}, (Dg_i(\overline{x}))_{i \in K}, (\theta_i)_{i=1,\ldots,q}),$$

one gets:

$$Df_0(\overline{x}) = -\sum_{i \in I(\overline{x})} \lambda_i Df_i(\overline{x}) + \sum_{i=1,\ldots,K} \mu_i Dg_i(\overline{x})$$

with $\lambda_i \geq 0$, $\forall i \in I(\overline{x})$. Define $\lambda_i = 0$, $\forall i \notin I(\overline{x})$. Relation (i) is thus proved. We have $\lambda_i f_i(\overline{x}) = 0, \forall i = 1, \ldots, I$, that is Condition (ii). □

Exercise 25. *Solve*

$$\min\{(3\sqrt{2}x + 3y - 1)$$

under the constraints

$$\Big\{(x, y) \in \mathbb{R}^2 : x^2 + \frac{y^2}{2} = 1\}$$

4 Comparative Statics

We are interested in a fundamental question: What is the effect of a change in an exogenous variable on the solution value of the endogenous variable? The issue is referred to "comparative statics".

4.1 Mathematical tools

We present fundamental tools that help us to do comparative statics.

Theorem 22. *[The Inverse Function Theorem] Let f be a continuously differentiable mapping from an open, nonempty set E of $\mathbb{R}^n$ into $\mathbb{R}^n$. Consider a point $a \in E$. Denote $b = f(a)$.*

Assume that the Jacobian matrix $J_f(a)$ (derivatives of f)

$$\mathbf{J} = \begin{bmatrix} \dfrac{\partial \mathbf{f}}{\partial x_1} & \cdots & \dfrac{\partial \mathbf{f}}{\partial x_n} \end{bmatrix} = \begin{bmatrix} \nabla^{\mathrm{T}} f_1 \\ \vdots \\ \nabla^{\mathrm{T}} f_n \end{bmatrix} = \begin{bmatrix} \dfrac{\partial f_1}{\partial x_1} & \cdots & \dfrac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial x_1} & \cdots & \dfrac{\partial f_n}{\partial x_n} \end{bmatrix}$$

is invertible. $\nabla^{\mathrm{T}} f_i$ is the transpose (row vector) of the gradient of the i component.

Then

1. *there exists an open set U and V in $\mathbb{R}^n$ such that $a \in U$ and $b \in V$, f is one-to-one on U, and $f(U) = V$.*

2. *the inverse function of f (denoted by f^{-1}) is continuously differentiable on V.*

   Recall: The inverse function f^{-1} is defined by: for $y \in V$, take $x \in U$ such that $f(x) = y$ (this value x is uniquely determined because f is one-to-one on U). Then we defined $f^{-1}(y) = x$, i.e., $f^{-1}(f(x)) = x$.

Proof. See Rudin (1976), page 221. □

Theorem 23. *[The implicit Functions Theorem] Let f be a continuously differentiable function from an open, nonempty set $E = U \times V$ of $\mathbb{R}^{n+m}$ ($U \subset \mathbb{R}^n$, $V \in \mathbb{R}^m$) into $\mathbb{R}^n$. Suppose that $f(a, b) = 0$ for some point $(a, b) \in U \times V$. Assume that that $Df_a(a, b)$ is invertible, where $Df_a(a, b)$ is the partial derivative of f with respect to the variable (a) at point (a, b) defined by*

$$Df_a(a,b) = \begin{bmatrix} \dfrac{\partial f_1}{\partial a_1}(a,b) & \cdots & \dfrac{\partial f_1}{\partial a_n}(a,b) \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_n}{\partial a_1}(a,b) & \cdots & \dfrac{\partial f_n}{\partial a_n}(a,b) \end{bmatrix}$$

1. *Then there exists an open set $U_1 \subset U$ containing a, an open set $V_1 \in V$ containing b, an open set W containing 0 and a function $g : V_1 \times W \to U_1$ such that*

   $$f(x, y) = z \Leftrightarrow x = g(y, z) \quad \forall x \in U_1, \forall y \in V_1, \forall z \in W$$

   Moreover, g is continuously differentiable on $V_1 \times W$.

2. *In particular, we have that:*

$$f(x,y) = 0 \Leftrightarrow x = \phi(y) \quad \forall x \in U_1, \forall y \in V_1 \tag{35}$$

where ϕ is differentiable on V_1. Moreover, we can compute the derivative of ϕ by using $f(\phi(y), y) = 0$

$$Df_x(\phi(y), y)D\phi(y) + Df_y((\phi(y), y)) = 0$$

Note that $Df_x(\phi(y), y)$ is an $n \times n$ matrix while $Df_y((\phi(y), y))$ is an $n \times m$ matrix.

$$Df_x = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}, \quad D\phi = \begin{bmatrix} \frac{\partial \phi_1}{\partial y_1} & \cdots & \frac{\partial \phi_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial \phi_n}{\partial y_1} & \cdots & \frac{\partial \phi_n}{\partial y_m} \end{bmatrix}$$

$$Df_y = \begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_m} \end{bmatrix}$$

The function ϕ is implicitly defined by (35).

4.2 Applications in Economics

Let us explicitly write $f(a, b) = 0$ as follows

$$\begin{aligned} &f_1(a_1, \ldots, a_n, b_1, \ldots, b_m) = 0 \\ &\cdots \\ &f_n(a_1, \ldots a_n, b_1, \ldots, b_m) = 0. \end{aligned}$$

In economics, $a_1, \ldots, a_n$ are viewed as endogenous variables while $b_1, \ldots, b_m$ exogenous. To study the effect of exogenous variables (b_j) on endogenous variables (a_i) or to understand how (a_i) changes when (b_i) changes, we can apply the implicit function theorem and compute the derivatives of (a_i) as functions of (b_j).

Let us consider a two-period optimal growth model introduced in Section 3.3.2. Let conditions in Proposition 14 be satisfied. Then, the optimal physical capital k_1 is determined by

$$u'(w_0 - k_1) - \beta F'(k_1)u'(w_1 + F(k_1)) = 0 \tag{36}$$

Notice that w_0, w_1, β are exogenous parameters while k_1 is endogenous and depends on w_0, w_1, β.

Denote $f(k_1, w_0, w_1, \beta) \equiv u'(w_0 - k_1) - \beta F'(k_1)u'(w_1 + F(k_1))$. Observe that the function f is striclty increasing in k_1 and w_1, but strictly decreasing in w_0 and β

Assume that u' and F' are continuously differentiable. Applying the implicit functions theorem, the optimal value k_1 determined by $f(k_1, w_0, w_1, \beta) = 0$ can be expressed as a differentiable function of w_0, w_1, β. We write $k_1 = k_1(w_0, w_1, \beta)$.

We now look at the role of the initial endowment w_0. Taking the derivative with respect to w_0 of both sides of the equation $f(k_1, w_0, w_1, \beta) = 0$, we have

$$\frac{\partial f}{\partial k_1}(k_1, w_0, w_1, \beta)\frac{\partial k_1}{\partial w_0}(w_0, w_1, \beta) + \frac{\partial f}{\partial w_0}(k_1, w_0, w_1, \beta) = 0$$

Since $\frac{\partial f}{\partial k_1} > 0$ and $\frac{\partial f}{\partial w_0} < 0$, we get that $\frac{\partial k_1}{\partial w_0}(w_0, w_1, \beta) > 0$. It means that the optimal value k_1 is increasing in the initial endowment.

5 Discrete dynamical systems (difference equations)

5.1 Motivating examples

The basic idea is that, in some cases the (economic) outcomes at a period depend on the variables in the pass. For instance,

$$\begin{aligned} x_{1,t+1} &= f_1(x_{1,t}, \cdots, x_{n,t}) \\ &\cdots \\ x_{n,t+1} &= f_n(x_{1,t}, \cdots, x_{n,t}). \end{aligned}$$

So, investigating the evolution of the sequence $(x_{1,t}, \ldots, x_{n,t})$ requires us to deal with this dynamical system.[15]

5.2 One-dimensional, first-order systems

Consider the one-dimentional autonomous, first-order difference equation:

$$y_{t+1} = f(y_t) \forall t \geq 0 \tag{37}$$

y_0 is given and $f : \mathbb{R} \to \mathbb{R}$ is a real function. If we consider the system

$$y_{t+1} = f_t(y_t) \forall t \geq 0 \tag{38}$$

where the function f_t depends on time, then it is called "non-autonomous"

Definition 20. *A solution of this difference equation is a trajectory $(x_t)_{t\geq 0}$ that satisfies (37) and initial condition $x_0 = y_0$.*

Definition 21. *Let f be a function and $x \in \mathbb{R}$ (or in general, in the domain of f). We define $f^n(x)$ the nth iterate of f under f by the following relationship: $f^n(x) = f(f^{n-1}(x)) \forall n \geq 1$, $f^0(x) \equiv x$.*

Definition 22.
1. *A steady-state equilibrium (or equilibrium point, or fixed point) is a value x^* satisfying $f(x^*) = x^*$.*
2. *x is an eventually equilibrium (fixed) point if there exists a positive integer r and an equilibrium point x^* such that $f^r(x) = x^*$, $f^{r-1}(x) \neq x^*$, where $f^n(x)$ is the nth iterate of f under f.*

[15] See Bosi and Ragot (2011) for an excellent introduction of discrete dynamical systems.

5.2.1 Linear cases

Consider a linear, first-order, autonomous difference equation

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

where a, b are constants.

We can proved that

$$x_t = \begin{cases} a^t x_0 + b\frac{1-a^t}{1-a} \forall t \geq 0, \text{ if } a \neq 1 \\ x_0 + bt \text{ if } a = 1. \end{cases}$$

Consequently, if $a \in (-1, 1)$, we have $\lim_{t\to\infty} x_t = b/(1-a)\ \forall y_0$.

Notice that $x^* \equiv b/(1-a)$ is the unique steady state when $a \neq 1$. We also observe that

$$x_t = \begin{cases} a^t(y_0 - y^*) + y^* \forall t \geq 0, \text{ if } a \neq 1 \\ y_0 + bt \text{ if } a = 1. \end{cases}$$

Graphics here.

1. $a \in (0, 1)$. (Monotonic Convergence)
2. $a \in (-1, 0)$. (Oscillatory Convergence)
3. $a \geq 1, b \neq 0$. (Go to infinity. Monotonic Divergence)
4. $a = -1$. (Two-period cycles)
5. $a < -1$. (Oscillatory Divergence)

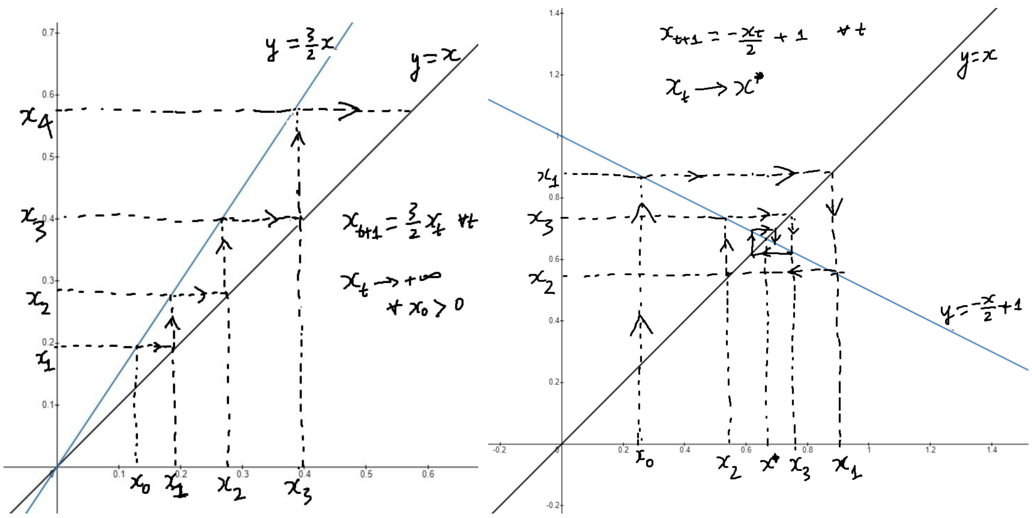


Figure 2: Monotonic convergence versus oscillatory convergence

Exercise 26. *Consider a linear, first-order, nonautonomous difference equation*

$$y_{t+1} = a_t y_t + b_t$$

where a_t, b_t are real values.

Prove that

$$y_t = \Big(\prod_{i=1}^{t-1} a_i\Big) y_0 + \sum_{k=0}^{t-1} b_k \Big(\prod_{i=k}^{t-1} \frac{a_i}{a_k}\Big)$$

Exercise 27. *Study the following system*

$$x_{t+1} = A x_t^{\alpha}$$

where $x_0 > 0$ is given, $A > 0$ and $\alpha > 0$.

Find the fixed points.

Find conditions (based on x_0, A, α) under which x_t converges (diverges).

Hint: Consider $\alpha = 1, \alpha > 1, \alpha < 1$.

Exercise 28. *Study the following system*

$$x_{t+1} = A \max(x_t - b, 0)$$

where $x_0 > 0$ is given, $A > 0$ and $b > 0$.

Find the fixed points.

Find conditions (based on x_0, A, b) under which x_t converges (diverges).

Hint: Draw the graph of the function $f(x) = Amax(x-b, 0)$. Consider $A = 1, A > 1, A < 1$.

Proof. Let $x > 0$ be a fixed point. We have $x = Amax(x - b, 0)$. Since $x > 0$, we have $max(x - b, 0) > 0$. This implies that $x - b > 0$. Hence, $max(x - b, 0) = x - b$. From this, we can find x by $x = A(x - b)$, i.e., $x(A - 1) = Ab$. □

Exercise 29. *Suppose that aggregate consumption in period t is given by*

$$C_t = A + BY_{t-1}$$

where Y_t represents the income in period t and $B \in (0, 1)$ is the marginal propensity to consume out of the previous year's income.

Assume that

$$\begin{aligned} Y_t &= C_t + I_t \\ I_t &= (1+g)^t \end{aligned}$$

where $g > 0$ is the exogenous growth rate in investment spending.

1. *Prove that*

$$\begin{aligned} Y_t &= B^t Y_0 + \sum_{k=0}^{t-1} \big(A + (1+g)^k\big) B^{t-1-k} \\ &= B^t Y_0 + A\frac{1-B^t}{1-B} + A\Big((1+g)^t - B^t\Big) \end{aligned}$$

2. Moreover, we have

$$\lim_{t\to\infty} \frac{C_t}{Y_{t-1}} = B$$
$$\lim_{t\to\infty} \frac{Y_t}{I_t} = A$$

3. When happens if $g < 0$?

5.2.2 Nonlinear, first-order, autonomous difference equation

When the function f is not linear, it is not easy to explicitly compute $f^t(x_0)$ or investigate the evolution of $f^t(x_0)$. There may exist multiple steady-state equilibria. Indeed, let consider the system $k_{t+1} = f(k_t)$ where $f(x) = Ax^{0.5}(1+x)$, $A > 0$, and $k_0 > 0$. We can check that if $A < 0.5$, then there are two steady states (see Example 9).

5.2.3 Convergence and global stability

In economics, we are particularly interested in the convergence and stability of the sequence $x_t = f^t(x_0)$. We would like to find conditions under which the sequence x_t converges.

Proposition 16. *Let $f : S \to S$ where S is a closed set of $\mathbb{R}$. Suppose that f is increasing. Suppose that there exists a unique x^* such that $f(x^*) = x^*$. Moreover, $f(x) < x \ \forall x > x^*$, and $f(x) > x \ \forall x < x^*$. Then the system (x_t) defined by $x_{t+1} = f(x_t)$ converges to x^* for any $x_0 \in S$.*

Proof. If $x_0 = x^*$, then $x_t = x^*$, $\forall t$.

If $x_0 < x^*$, then $f(x_0) > x_0$, or, equivalently, $x_1 > x_0$. By induction, we can easily prove that the sequence x_t is increasing: $x_{t+1} \geq x_t, \forall t$. Indeed, this holds for $t = 0$. Assume that it holds until date t. Since $x_t) \geq x_{t-1}$, we have $f(x_t) \leq f(x_{t-1})$, which is equivalent to $x_{t+1} \leq x_t$. So, x_t is increasing. Notice that $x_{t+1} \geq x_t$ means that $f(x_t) \geq x_t$. According to our assumption, this implies that $x_t \leq x^*$. To sum up, x_t is increasing and bounded from above. Thus, it converges to some value $\bar{x}$. We will prove that $\bar{x} = x^*$. Since $x_{t+1} = f(x_t)$, by letting t tend to infinity, we have $\bar{x} = f(\bar{x})$. Since the fixed point is unique, we obtain that $\bar{x} = x^*$.

For the case $x_0 > x^*$, we can use a similar argument to prove that x_t decreasingly converges to x^*. □

Definition 23 (contraction mapping). *A function $f : \mathbb{R} \to \mathbb{R}$ is a contraction mapping (or function) if there exists $\theta \in (0, 1)$ such that $|f(x) - f(y)| \leq \theta|x - y| \ \forall x, y$.*

If f is differentiable and $\sup_x |f'(x)| < 1$, then f is a contraction mapping (why?).

Notice that the condition $|f(x) - f(y)| \leq \theta|x - y| \ \forall x, y$ is quite restrictive. It is not satisfied for the function $f(x) = x^2$. Indeed, we have $|f(x)-f(y)| = |(x+y)(x-y)| \geq |x-y|$ if $|(x + y)| \geq 1$.

We present the following result which is a simple case of Banach Fixed Point Theorem that shows the existence and uniqueness of the fixed point of a contraction mapping in a complete metric space.

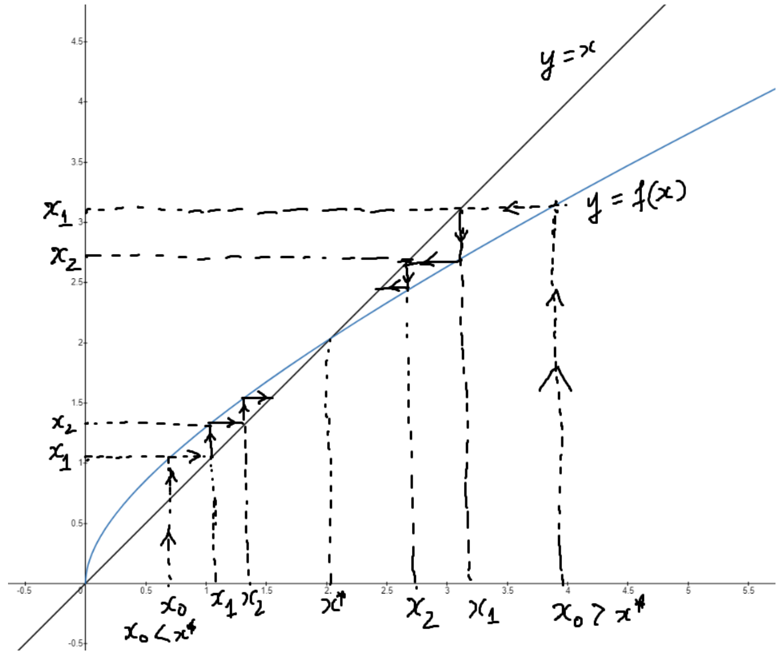


Figure 3: $\lim_{t\to\infty} x_t = x^*$, $\forall x_0 > 0$. See Proposition 16.

Proposition 17 (The Contraction Mapping Theorem). *Let $f : \mathbb{R} \to \mathbb{R}$ be contraction mapping. Then*

1. *f is a unique fixed point x^* (i.e., there is a unique x^* such that $f(x^*) = x^*$).*
2. *$|f^n(x_0) - x^*| \leq \theta^n |x_0 - x^*|$ $\forall n \geq 1$.*

Corollary 5. *Consider the difference equation $x_{t+1} = f(x_t)$ where $f : \mathbb{R} \to \mathbb{R}$.*

1. *If f is a contraction mapping, then the system has a unique steady-state x^* and $\lim_{t\to\infty} x_t = x^*$ for any x_0.*
2. *Assume that f is differentiable and $\sup_x |f'(x)| < 1$, then the system has a unique steady-state x^* and $\lim_{t\to\infty} x_t = x^*$ for any x_0.*

We also state the results in the space $\mathbb{R}^n$.

Proposition 18. *Let $f : \mathbb{R}^n \to \mathbb{R}^n$ be contraction mapping, i.e., there exists $\theta \in (0, 1)$ such that*

$$\|f(x) - f(y)\| \leq \theta \|x - y\| \quad \forall x, y.$$

1. *f has a unique fixed point x^* (i.e., there is a unique x^* such that $f(x^*) = x^*$).*
2. *For every $x_0 \in \mathbb{R}$, the sequence (x_t), determined by $x_{t+1} = f(x_t)$ $\forall t$, converges to x^*.*

Proof. Although we consider the space $\mathbb{R}^n$, the proof in a complete metric space is similar.

Uniqueness of x^*. Suppose that there is $x \neq x^*$ such that $f(x) = x$. Since f is a contraction mapping, we have

$$\|f(x) - f(x^*)\| \leq \theta \|x -^* \|.$$

However, $f(x) - f(x^*) = x - x^*$. Hence, $\|x -^* \| \leq \theta \|x -^* \| < \|x -^* \|$, a contradiction. Therefore, we have proved the uniqueness of x^*.

Existence of x^*. Let $x_0 \in \mathbb{R}^n$. Consider the sequence (x_t) defined by $x_{t+1} = f(x_t), \forall t$, i.e, $x_t = f^t(x_0)$. We will prove that x_t converges to x^*.

First, we prove that (x_t) is a Cauchy sequence in the sense that: $\forall \epsilon > 0$, there exists n_0 such that $\|x_n - x_m\| < \epsilon, \forall n > n_0, m > n_0$. Indeed, we have

$$\begin{aligned} \|x_2 - x_1\| &= \|f(x_1) - f(x_0)\| \leq \theta \|x_1 - x_0\| \\ \|x_{n+1} - x_n\| &= \|f(x_n) - f(x_{n-1})\| \leq \theta \|x_n - x_{n-1}\| \leq \cdots \leq \theta^n \|x_1 - x_0\| \end{aligned}$$

and hence, for $n > m$, we have

$$\begin{aligned} \|x_n - x_m\| &= \|x_n - x_{n-1} + \cdots + x_{m+1} - x_m\| \leq \|x_n - x_{n-1}\| + \cdots + \|x_{m+1} - x_m\| \\ &\leq \theta^{n-1}\|x_1 - x_0\| + \cdots \leq \theta^m \|x_1 - x_0\| = \|x_1 - x_0\| \theta^m (1 + \theta + \cdots + \theta^{n-1-m}) \\ &\leq \|x_1 - x_0\| \theta^m \frac{1}{1-\theta} \end{aligned}$$

For any $\epsilon > 0$, since $\theta \in (0,1)$, we can choose n_0 such that $\|x_1 - x_0\| \theta^{n_0} \frac{1}{1-\theta} < \epsilon$. Then, for any $n > n_0, m > n_0$, we have

$$\|x_n - x_m\| \leq \|x_1 - x_0\| \theta^m \frac{1}{1-\theta} \leq \|x_1 - x_0\| \theta^{n_0} \frac{1}{1-\theta} < \epsilon.$$

Second, since (x_t) is a Cauchy sequence, it converges to some value $\bar{x} \in \mathbb{R}^n$ (see Rudin (1976)). By definition, we have $x_{t+1} = f(x_t)$. Let t tend to infinity, we get that $\bar{x} = f(\bar{x})$. However, since the fixed point is unique, we obtain that $\bar{x} = x^*$. It means that x_t converges to x^*.

□

5.2.4 Local stability

Definition 24. 1. *The equilibrium point x^* is stable (or Lyapunov stable) if, for any $\epsilon > 0$, there exists $\delta > 0$ such that: $|x_0 - x^*| < \delta$ implies that $|f^n(x_0) - x^*| < \epsilon$ $\forall n > 0$.*

If x^* is not stable, then it is called unstable.

2. *x^* is said to be attracting if there exists $\eta > 0$ such that:*

$$|x_0 - x^*| < \eta \text{ implies } \lim_{t \to \infty} x_t = x^*$$

x^* is called global attracting if $\eta = \infty$.

3. *x^* is said to be asymptotically stable equilibrium point if it is stable and attracting.*

If $\eta = \infty$, then x^* is said to be globally asymptotically stable.

Definition 25. *The equilibrium point x^* is said to be hyperbolic if $|f'(x^*)| \neq 1$.*

Theorem 24. *Let x^* be an equilibrium point of the difference equation $x_{t+1} = f(x_t)$. Assume that f is continuously differentiable at x^*. We have that:*

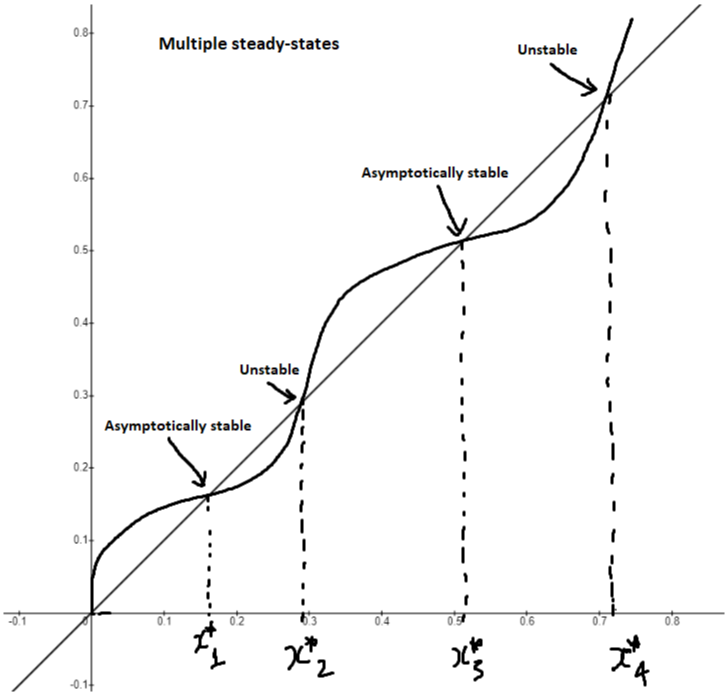


Figure 4: Monotonic convergence versus oscillatory convergence

1. *If* $|f'(x^*)| < 1$, *then* x^* *is asymptotically stable.*
2. *If* $|f'(x^*)| > 1$, *then* x^* *is unstable.*

Proof. (1) Suppose that $|f'(x^*)| < 1$. We can choose M such that $|f'(x^*)| < M < 1$. Since $f'(x)$ is continuous, there exists an interval $J = (x^* - \gamma, x^* + \gamma)$, where $\gamma > 0$, such that $|f'(x)| < M < 1 \; \forall x \in J$.

Let $x_0 \in J$. We have

$$x_1 - x^* = f(x_0) - x^* = f(x_0) - f(x^*) = f'(\xi)(x_0 - x^*)$$

where ξ is between x_0 and x^*. (Recall that x_0 may be lower or higher than x^*.) By consequence, we have

$$|x_1 - x^*| = |f'(\xi)||x_0 - x^*| \leq M|x_0 - x^*| \leq |x_0 - x^*| < \gamma$$

This implies that $x_1 \in J$. By induction we get that $x_n \in J \; \forall n$ and

$$|x_n - x^*| \leq M^n |x_0 - x^*| \forall n.$$

Let $\epsilon > 0$, choose $\delta < \min(\gamma, \epsilon)$, then we have

$$|x_n - x^*| \leq M^n |x_0 - x^*| \leq |x_0 - x^*| < \delta < \epsilon \forall n.$$

So, x^* is stable. x^* is attracting because: $|x_0 - x^*| < \delta$ implies that $\lim_{t\to\infty} x_t = x^*$. Therefore, x^* is stable.

(2) We now suppose that $|f'(x^*)| > 1$. Since $f'(x)$ is continuous and $|f'(x^*)| > 1$, there exist $\gamma > 0$ and $M > 1$ such that $|f'(x)| > M > 1 \; \forall x \in J \equiv (x^* - \gamma, x^* + \gamma)$.

Suppose that x^* is stable. Let ϵ be in the interval $(0, \gamma)$. So, there exist $\delta > 0$ such that: $|x_0 - x^*| < \delta$ implies that $|f^n(x_0) - x^*| < \epsilon \ \forall n > 0$.

Take $0 < \alpha < \min(\epsilon, \delta)$. We have: $|x_0 - x^*| < \alpha$ implies that $|f^n(x_0) - x^*| < \epsilon \ \forall n > 0$.

Take x_0 be such that $0 < |x_0 - x^*| < \alpha$. We have $|x_n - x^*| < \epsilon < \gamma \ \forall n > 0$. Hence, $x_n \in J \ \forall n$. So, we get that

$$|x_n - x^*| = |f'(\xi_n)||x_{n-1} - x^*| \geq \cdots \geq M^n |x_0 - x^*|$$

where ξ_n is between x_0 and x^* and hence belongs the interval J.

Since $M > 1$, $|x_n - x^*|$ must tend to infinity, a contradiction (because $|x_n - x^*| < \epsilon$). Therefore, x^* is not stable. $\square$

Example 9 (increasing return to scale and middle-income trap)**.** *Consider the system $k_{t+1} = f(k_t)$ where $f(x) = Ax^{0.5}(1 + x)$, $A > 0$, and $k_0 > 0$ is given. k_t can be interpreted as the physical capital stock of the economy at date t (see Section 5.2.6).*

Solving the problem. The equation determining the steady states is $x = Ax^{0.5}(1 + x)$, or equivalently $Ax - x^{0.5} + A = 0$.

Denote $\Delta = 1 - 4A^2$. There is no positive steady state if and only if $\Delta < 0$ or equivalently $A > 1/2$. In this case, we can check that $Ax^{0.5}(1+x) > x$, $\forall x > 0$. By consequence, k_t is strictly increasing in t. Since there is no positive steady state, k_t converges to infinity.

Economically, we can say that when the productivity A is high (in the sense that $A > 1/2$), the economy grows without bound (k_t converges to infinity).

There are 2 positive steady states if and only if $\Delta > 0$ or equivalently, $0 < A < 1/2$. We can easily compute these two steady state:[16]

$$x_L = \Big(\frac{1 - \sqrt{1 - 4A^2}}{2A}\Big)^2, \quad x_H = \Big(\frac{1 + \sqrt{1 - 4A^2}}{2A}\Big)^2.$$

We can see in the graph that $f(x) > x, \forall x \in (0, x_L)$ or $x \in (x_H, \infty)$, and $f(x) < x, \forall x \in (x_L, x_H)$.

1. We can see that x_L is asymptotically stable because $|f'(x_L)| < 1$. (Actually, $f'(x_L) \in (0, 1)$.)

2. We now prove k_t increasingly converges to ∞ if $k_0 > x_H$. Then, x_H is not stable.

   For $k_0 > x_H$, we can show that k_t is strictly increasing in t (use the fact that $f(x) > x, \forall x \in (x_H, \infty)$). By consequence, k_t converges. It cannot converge to a steady state because $k_{t+1} > k_t > x_H$, $\forall t$. So, it must converge to infinity.

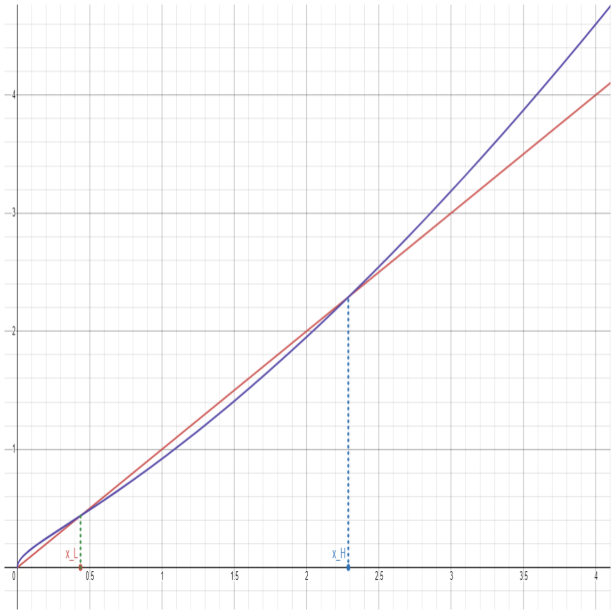
   The point x_H is not stable because k_t converges to infinity for any $k_0 > x_H$.

3. We now want to prove that: $\lim_{t\to\infty} k_t = x_L \ \forall k_0 \in (0, x_H)$.

   First, observe that $f(x) \geq x$ if $x \in (0, x_L)$ and $f(x) < x$ if $x \in (x_L, x_H)$.

   (a) We prove that k_t increasingly converges to x_L if $k_0 \in (0, x_L)$.
   Proof: Let $k_0 \in (0, x_L)$. By applying the above remark, we have $f(k_0) \leq k_0$, and hence, $k_1 \leq k_0$.

[16] However, we can eventually provide analysis without computing x_L, x_H.



Since f is increasing, we have $f(k_0) \leq f(x_L) = x_L$ or equivalently $k_1 \leq x_L$.
By using the induction argument, we can prove that $k_{t+1} \leq k_t$ and $k_t \leq x_L$ for any t. Since the sequence k_t is increasing and bounded from above, it converges to some value, say k^*, and $f(k^*) = k^*$. Since $k^* \leq x_L$, we have $k^* = x_L$.

(b) By using a similar argument, we can prove that k_t decreasingly converges to x_L if $k_0 \in (x_L, x_H)$.

4. We now look at the role of A and provide an economic interpretation.

   Observe that x_H is decreasing in A. The point x_H can be viewed as a middle income trap in the sense that k_t converges to infinity for any $k_0 > x_H$ while $k_t < x_H$, $\forall t$ (this means that we cannot overcome this threshold x_H) if $k_0 \leq x_H$.

   The middle income trap is decreasing in the productivity A. This leads to an interpretation: The higher the level of productivity A, the lower the middle income trap, the higher possibility we can get the growth (k_t converges to infinity).

□

Exercise 30. *Consider the system* $x_{t+1} = x_t\big(1 + r(1 - x_t)\big)$.

1. *Prove that* $a = 1$ *is a fixed point.*
2. *Prove that if* $r \in (0, 1)$*, then* $a = 1$ *is asymptotically stable.*
3. *If* $r = 0$*, then* $a = 1$ *is not asymptotically stable.*
4. *Study the (asymptotically) stability of this fixed point in the case* $r = 2$.

Exercise 31. *Let us consider the system* $x_{t+1} = f(x_t)$ *with* $f(x) = x^{0.5}(0.5 + 0.4x)$. *(1) Find two positive steady states. (2) Prove that among these points, one steady state is asymptotically stable while another one is unstable.*[17]

[17] Readers may like to use https://www.desmos.com/calculator?lang=en to draw the graph of the function f.

5.2.5 Nonhyperbolic fixed points

Proposition 19. *Let x^* be a fixed point and $f'(x^*) = 1$. Assume that f is in C^3. We have*

1. *If $f''(x^*) \neq 0$, then x^* is unstable.*
2. *If $f''(x^*) = 0$ and $f'''(x^*) > 0$ then x^* is unstable.*
3. *If $f''(x^*) = 0$, and $f'''(x^*) < 0$ then x^* is asymptotically stable.*

Proof. By using the Taylor theorem and $f'(x^*) = 1$, we have

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + R_2(x, x^*)\|x - x^*\|^2 \tag{39}$$

$$= f(x^*) + (x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + R_2(x, x^*)\|x - x^*\|^2 \tag{40}$$

$$f(x) = f(x^*) + f'(x^*)(x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + \frac{1}{6}f'''(x^*)(x - x^*)^3 + R_3(x, x^*)\|x - x^*\|^3$$

$$= f(x^*) + (x - x^*) + \frac{1}{2}f''(x^*)(x - x^*)^2 + \frac{1}{6}f'''(x^*)(x - x^*)^3 + R_3(x, x^*)\|x - x^*\|^3 \tag{41}$$

1. If $f''(x^*) \neq 0$, then we consider two cases.
   (a) $f''(x^*) > 0$. In this case, $f'(x)$ is increasing in a neighborhood $(x^* - \delta, x^* + \delta)$ of x^*. Since $f'(x^*) = 1$, we have $f'(x) > 1 = f'(x^*)$ $\forall x \in (x^*, x^* + \delta)$. By adopting the argument in the proof of part 2 of Theorem 24, we can prove that x^* is unstable.
   (b) $f''(x^*) < 0$. Using the same argument, we have that x^* is unstable.
2. If $f''(x^*) = 0$ and $f'''(x^*) > 0$, then we have
$$f(x) - f(x^*) = (x - x^*)\left(1 + \frac{1}{6}f'''(x^*)(x - x^*)^2\right) + R_3(x, x^*)\|x - x^*\|^3.$$
   By adopting the argument in the proof of part 2 of Theorem 24, we can prove that x^* is unstable.
3. If $f''(x^*) = 0$, and $f'''(x^*) < 0$, then (41) implies that
$$f(x) - f(x^*) = (x - x^*)\left(1 + \frac{1}{6}f'''(x^*)(x - x^*)^2\right) + R_3(x, x^*)\|x - x^*\|^3$$
   By adopting the argument in the proof of part 1 of Theorem 24, we can prove that x^* is asymptotically stable.

□

5.2.6 Application: Solow growth models

We consider a model à la Solow.

$$\textbf{Solow Model:} \quad \begin{aligned} & c_t + S_t = Y_t \\ & I_t = S_t \\ & k_{t+1} = k_t(1-\delta) + I_t \\ & S_t = sY_t \\ & Y_t = A_t k_t^\alpha L_t^{1-\alpha}, \alpha \in (0,1) \\ & A_t = A \\ & L_t = 1 \end{aligned}$$

where c_t, S_t, I_t are consumption, saving, investment at date t ($t = 0, 1, \dots, +\infty$), $s \in (0,1)$ is the exogenous saving rate, k_t is the physical capital stock at date t ($k_0 > 0$ is given), $\delta \in [0,1]$ is the capital depreciation rate, Y_t is the output.

Proposition 20. *Consider the above Solow model.*

1. *Prove that, for any $t \geq 0$,*

$$\begin{aligned} Y_t &= Ak_t^\alpha \\ \frac{Y_{t+1}}{Y_t} &= \Big(\frac{k_{t+1}}{k_t}\Big)^\alpha \\ k_{t+1} &= k_t(1-\delta) + sAk_t^\alpha \end{aligned}$$

2. *Prove that k_t and Y_t converge. Find $k^* \equiv \lim_{t\to\infty} k_t$ and $Y^* \equiv \lim_{t\to\infty} Y_t$. How k^* and Y^* depend on A, s?*

3. *Is k^* asymptotically stable? Why? Illustrate your arguments by diagrams.*

The long-term rate of growth g of the output depends strongly on the rate of growth of the TFP A. The higher A, the higher the rate of growth g.

Proof. Left to the readers. □

We now consider a more general model à la Solow.

$$\textbf{Solow Model:} \quad \begin{aligned} & c_t + S_t = Y_t \\ & I_t = S_t \\ & k_{t+1} = k_t(1-\delta) + I_t \\ & S_t = sY_t \\ & Y_t = A_t k_t^\alpha L_t^{1-\alpha}, \alpha \in (0,1) \\ & A_t = a(1+\gamma)^t \\ & L_t = L_0(1+n)^t \end{aligned}$$

Here $\gamma > -1$ is the rate of growth of the TFP A_t, $n > -1$ is the rate of growth of the labor force. Both of them are assumed to be exogenous.

Proposition 21. *Consider the above Solow model.*

1. *Prove that, for any $t \geq 0$,*

$$\begin{aligned} Y_t &= a(1+\gamma)^t k_t^{\alpha} L_t^{1-\alpha} \\ \frac{Y_{t+1}}{Y_t} &= (1+\gamma)(1+n)^{1-\alpha}\Big(\frac{k_{t+1}}{k_t}\Big)^{\alpha} \\ k_{t+1} &= k_t(1-\delta) + sa(1+\gamma)^t k_t^{\alpha} L_t^{1-\alpha} \end{aligned}$$

2. *Prove that $\frac{\Delta Y_t}{Y_t} \to g$ where $\Delta Y_t \equiv Y_{t+1} - Y_t$ and g satisfies*

$$1 + g = (1+n)(1+\gamma)^{\frac{1}{1-\alpha}}$$

The long-term rate of growth g of the output depends strongly on the rate of growth of the TFP A_t. The higher γ, the higher the rate of growth g.

Proof. Left to the readers. □

5.2.7 Application: a Malthusian growth model

In 1798, Thomas Malthus wrote: "Through the animal and vegetable kingdoms, nature has scattered the seeds of life abroad with the most profuse and liberal hand. ... The germs of existence contained in this spot of earth, with ample food, and ample room to expand in, would fill millions of worlds in the course of a few thousand years. Necessity, that imperious all pervading law of nature, restrains them within the prescribed bounds. The race of plants, and the race of animals shrink under this great restrictive law. And the race of man cannot, by any efforts of reason, escape from it. Among plants and animals its effects are waste of seed, sickness, and premature death. Among mankind, misery and vice."[18]

Thomas Malthus hypothesized that population growth is an inverse function of income per capita. Inspired by this idea, we assume that

$$\frac{N_{t+1} - N_t}{N_t} = n - \frac{b}{w_t}$$

where w_t is income per capita, N_t is population in period t while a, b are positive constants.

Assume that w_t is given by

$$w_t = \frac{Y_t}{N_t}$$

where Y_t is aggregate output in the economy. Assume that

$$Y_t = N_t^{\alpha}$$

where $\alpha \in (0, 1)$.

[18]See Thomas Malthus, 1798. An Essay on the Principle of Population. Chapter I.

From this, we get that

$$w_t = N_t^{\alpha-1}.$$

Hence, the income is a decreasing function of the population. Rearranging equations, we obtain a nonlinear, first-order difference equation:

$$N_{t+1} = N_t(1 + n - bN_t^{1-\alpha})$$

Exercise 32. *1. Find the positive steady states of this dynamical system.*

2. Are they asymptotically stable or unstable? Why?

5.3 Multiple-dimensional, first-order systems

We now consider a multiple-dimensional, first-order system: $x_{t+1} = f(x_t)$ where $f : \mathbb{R}^n \to \mathbb{R}^n$. Explicitly, we write that

$$\begin{aligned} x_{1,t+1} &= f_1(x_{1,t}, \cdots, x_{n,t}) \\ &\cdots \\ x_{n,t+1} &= f_n(x_{1,t}, \cdots, x_{n,t}). \end{aligned}$$

In general, it is difficult to deal with this system.

Let us look at a steady states x^* determined by $f(x^*) = x^*$. Notice that it is not easy to compute x^*.

Assume that, for any i, the function f_i is in C^1 (continuously differentiable). Applying the Taylor's theorem, in a neighborhood of x^*, we can approximate f_i by an affine function

$$\begin{aligned} f_i(x_{1,t}, \dots, x_{n,t}) =& f_i(x_1^*, \dots, x_n^*) + \sum_{k=1}^{n} \frac{\partial f_i(x^*)}{\partial x_k}(x_{k,t} - x_k^*) \\ &+ R_i(x_t, x^*)\|x_t - x^*\| \end{aligned}$$

where $R_i(x^*, x^*) = 0$ and $R_1(x_t, x^*) \to 0$ as $x_t \to x^*$.

This can be rewritten in a matrix form

$$\begin{bmatrix} x_{1,t+1} \\ \vdots \\ x_{n,t+1} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(x^*) & \cdots & \frac{\partial f_1}{\partial x_n}(x^*) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x^*) & \cdots & \frac{\partial f_n}{\partial x_n}(x^*) \end{bmatrix} \begin{bmatrix} x_{1,t} \\ \vdots \\ x_{n,t} \end{bmatrix} + \begin{bmatrix} b_1^* \\ \vdots \\ b_n^* \end{bmatrix} + \begin{bmatrix} R_1(x_t, x^*)\|x_t - x^*\| \\ \vdots \\ R_n(x_t, x^*)\|x_t - x^*\| \end{bmatrix}$$

where $b_i^* \equiv f_i(x_1^*, \dots, x_n^*) - \sum_{k=1}^{n} \frac{\partial f_i(x^*)}{\partial x_k} x_k^*$.

So, in the following we focus on a linear dynamical system given by $x_{t+1} = Ax_t + b$, i.e.,

$$\begin{bmatrix} x_{1,t+1} \\ \vdots \\ x_{n,t+1} \end{bmatrix} = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{n1} & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} x_{1,t} \\ \vdots \\ x_{n,t} \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}$$

when $b = 0$, the system is said to be homogeneous.

Let x^* be a fixed point $x^* = Ax^* + b$. x^* exists and equals $(I - A)^{-1}b$ if the determinant $det(I - A) \neq 0$ where I is the $n \times n$ identity matrix. We have that

$$x_{t+1} - x^* = A(x_t - x^*)$$
$$\text{hence } x_t - x^* = A^t(x_0 - x^*)\forall t \geq 0$$

It remains to compute A^t. This task is not easy. Linear Algebra helps us to deal with this problem. Assume that we can diagonalize the matrix A in the sense that

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

where the matrix V is convertible and D is a diagonal matrix

$$D = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}$$

and $\lambda_i \in \mathbb{R}, \lambda_i \neq \lambda_j \ \forall i, j$. This happens if the characteristic polynomial $det(A - \lambda I)$ has n different real roots. So, under this condition, we obtain $A^t = VD^tV^{-1} \ \forall t$ and hence

$$x_t - x^* = VD^tV^{-1}(x_0 - x^*)\forall t \geq 0$$
$$D^t = \begin{bmatrix} \lambda_1^t & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n^t \end{bmatrix}$$

To sump up, x_t converges to the steady state x^* if the polynomial $det(A - \lambda I)$ has n different real roots $(\lambda_i)_{i=1}^n$ and $|\lambda_i| < 1 \ \forall i$.[19]

[19] See Simon and Blume (1994) among others for more details.

References

Bosi, S., & Ragot L. (2011). Discrete time dynamics: an introduction. *Bologna: CLUEB.*

Boyd, S., and Vandenberghe, L., 2004. *Convex Optimization*. Cambridge University Press.

Florenzano, M., Le Van, C., 2001. *Finite dimensional convexity and optimization*, Springer, 2001.

Michael Hoy, John Livernois, Chris McKenna, Ray Rees, Thanasis Stengos, 2001. *Mathematics for Economics*, 2nd Edition, MIT Press.

Jehle, G. A., and Reny, P. J., 2011. *Advanced microeconomic theory*. Financial Times Press, 2011.

Mas-Colell, A., Whinston, M.D., and Green, J. R., 1995 *Microeconomic theory*. Oxford University Press.

Elaydi, S., 2007. *Discrete Chaos, With Applications in Science and Engineering*, second edition, CRC Press.

Rudin, W., 1976. *Principles of Mathematical Analysis (International Series in Pure and Applied Mathematics)*, 3rd ed. McGraw-Hill.

Simon, C. P, Blume, L. E., 1994. *Mathematics for Economists*, W. W. Norton & Company.

Varian, H. R., 2014. *Intermediate Microeconomics: A Modern Approach*, Ninth Edition. W. W. Norton & Company.