Elementary Mathematics for Economists

Juwon Kwak*

Feb 28, 2018

Contents

*juwonkwak@knu.ac.kr. Please do not distribute.

1 Introduction

1.1 How to prepare for class.

1. First of all, this lecture note is very abstract. There is not much explanation. So, if you miss a class, you will have a hard time catching up. Please do not miss any class.

2. If you take a brief look at the composition of the note, there are Definitions and Propositions. Definitions are where we start from and Propositions are where we end up. Thus, memorize all Definitions. They are the most important. By memorizing them, you will understand the rationale behind such definitions. Next try to understand the implications of Propositions. Last, try to follow the proofs of Propositions. This is especially so for linear algebra part.

3. For calculus sections, do a lot of problems. Calculus is basic for natural science, engineering, social science and statistics. Understanding the Propositions of Calculus is important, but not enough. You should be able to differentiate and integrate any functions.

1.2 Set Theory

Definition 1 *Set*

- *Set : A set is a gathering together into a whole of definite, distinct objects* $\{1, 2, 3, ...\}, \{x|0 < x < 1\}$

- *Elements* $x \in A$

- *Subset:* $A \subset B$ *iff* $x \in A \quad \Rightarrow x \in B$

- *Null Set and Universal Set:* $\emptyset, U$

- *Intersection* $A \cap B \equiv \{x|x \in A$ *and* $x \in B\}$

- *Union:* $A \cup B \equiv \{x|x \in A$ *or* $B\}$

- *Complement:* $A^c \equiv \{x \in U|x \notin A\}$

- *Difference:* $A - B \equiv \{x \in A$ *and* $x \notin B\} \equiv A \cap B^c$

- *Power Set:* $2^A \equiv$ *is a set of all subsets of A.*

1.3 Function

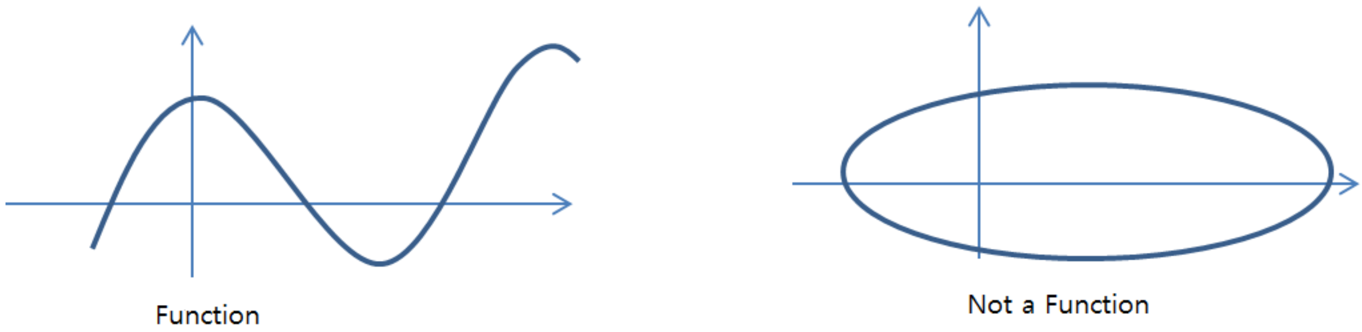


Figure 1: Function

Definition 2 *Function*

- *Function* $f : A \to B$ *is a relation between a set of inputs* (A) *and a set of permissible outputs* (B) *with the property that each input is related to exactly one output.*
- *Onto (surjective) function:* $\forall y \in B, \exists x \in A \quad s.t. \quad f(x) = y$
- *One-to-One (injective) function:* $f(x) = f(y) \Rightarrow x = y$
- *One-to-One Correspondence (bijective function) : injective and surjective.*

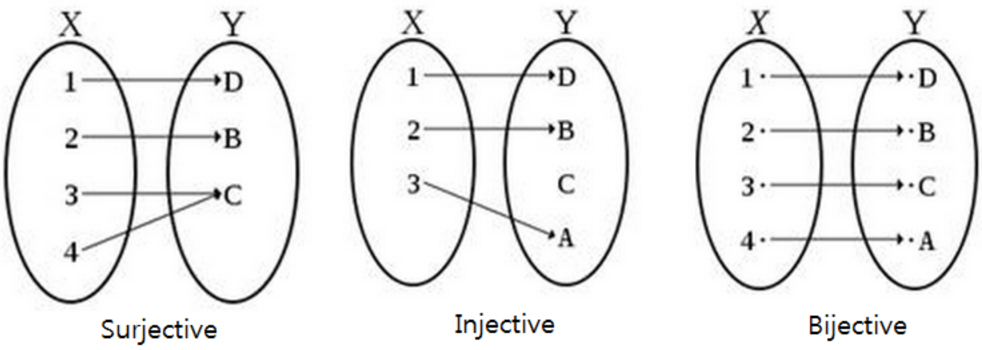


Figure 2: Surjective, Injective and Bijective Functions

- *Inverse Function* $f^{-1} : B \to A$*: if* $f(x) = y$ *then* $x = f^{-1}(y)$
- *Composite Function:* $f \circ g(x) = f(g(x))$
- *Log Function:* $y = a^x \quad \to \quad x = \log_a y$

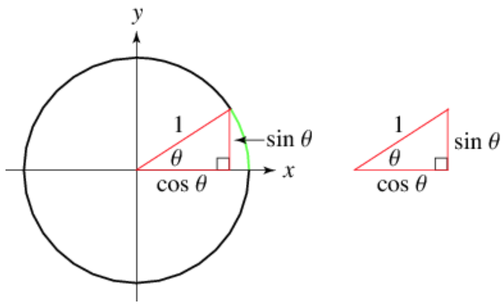


Figure 3: Unit Cirle

1.4 Trigonometry

For a point (x, y) on a circle, the radius of which is 1, on the center,

$$\sin\theta = y$$

$$\cos\theta = x$$

$$\tan\theta = \frac{y}{x} = \frac{\sin\theta}{\cos\theta}$$

Note that the angle θ is the length of the arc, i.e., $90^0 = \frac{\pi}{2}$. $360^0 = 2\pi$

1.5 Methods of Mathematical Proof

H_0: *If P holds, then Q also holds*

We say P is a **sufficient condition** for Q because if P holds then Q is automatically satisfied. Q is a **necessary condition** for P, because it is necessary to be Q to be P, but Q does not automatically mean P. We may also denote H_0 as

$$P \Rightarrow Q$$

Also note that

H_1: *If Q does not hold, then P does not hold.*

If H_0 is true, then H_1 is automatically also true. We say H_1 is the *contraposition* of H_0. We may also denote H_1 as

$$\sim Q \Rightarrow \sim P$$

H_2: *If P holds, then Q also holds, and if Q holds, then P also holds.*

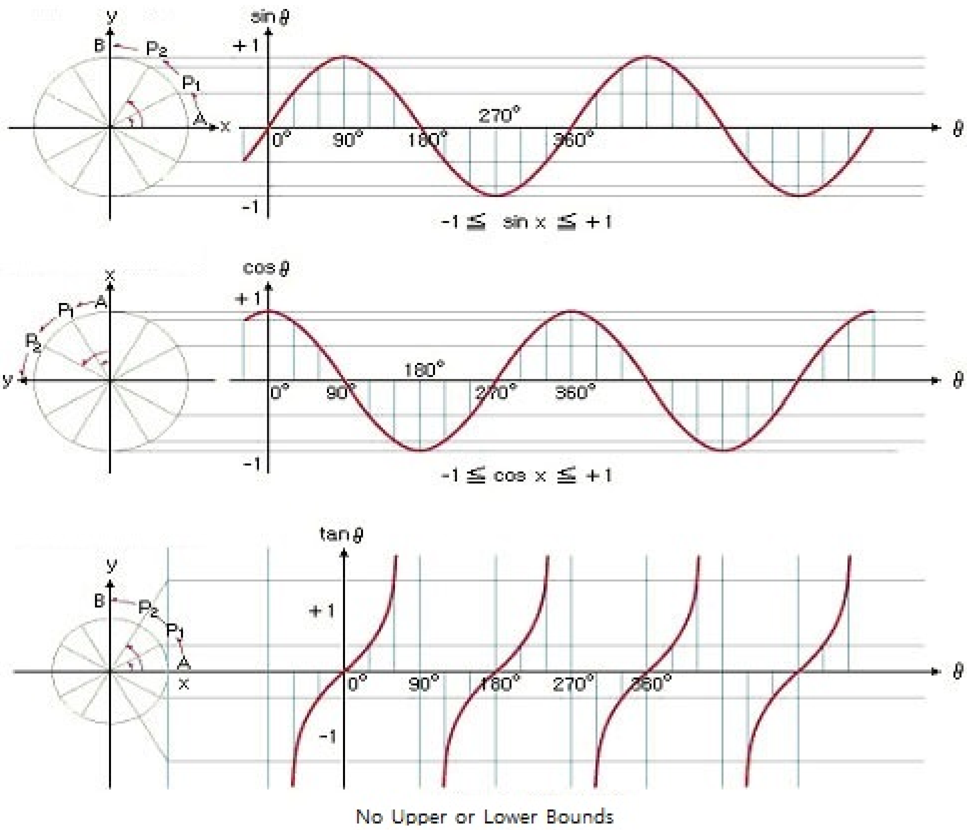


Figure 4: The Graph of Trigonometry

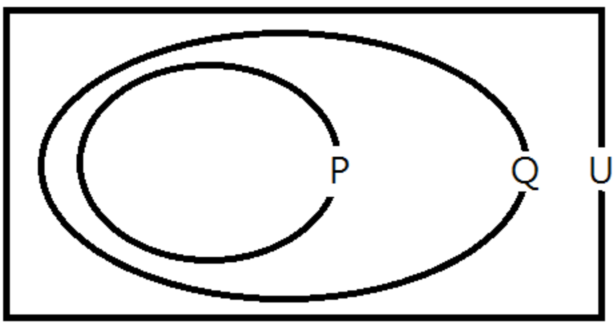


Figure 5: Necessary and Sufficient Conditions

Suppose H_2 is true. Then P and Q are the same conditions, or P and Q are equivalent ($P \Leftrightarrow Q$).

Suppose we want to prove that if P holds, then Q holds too. One way to prove this is by directly proving it. For example, we want to prove that

$$H_3\text{: } \textit{If } x = 2 \textit{ then } x^2 = 4$$

We first assume that $x = 2$ is true. Then, we show that $x^2 = x \times x = 2 \times 2 = 4$. Thus, H_3 is true. We call this method the *constructive proof.* Another way to prove this is by assuming that Q does not hold and showing that P also does not hold. Suppose $x^2 \neq 4$. Then $x^2 > 4$ or $x^2 < 4$. Suppose $x^2 > 4$, then $x > 2$ or $x < -2$. Also if $x^2 < 4$, then $-2 < x < 2$. Thus, $x \neq 2$. We call this method the *contrapositive proof.*

$$H_4\text{: } \textit{If } x = 2 \textit{ then } x^2 \neq 4$$

Another way to prove H_3 is by showing that H_4 is false. If $P \Rightarrow \sim Q$ is false, then $P \Rightarrow Q$ should be true. This is *proof by contradiction*.

$$H_5\text{: } \textit{For } n = 1, 2, 3 \cdots, \textit{ if } A_n \textit{ holds then } B_n \textit{ also holds}$$

To prove H_5, we first show that for $n = 1$, H_5 holds. Second, we assume that when $n = i-1$, H_5 holds. Third, we show that if when $n = i - 1$, H_5 holds, then H_5 holds when $n = i$. Thus, when $n = 1$ holds, so does $n = 2$, and so on. This method of proof is *mathematical induction.*

Problem Set 1

1. Use Venn-Diagram to show Intersection, Union, Complements and Difference

2. What is the number of elements for a power set of a set, the number of elements of which is n?

3. Prove the following:

$$\sin(a + b) = \sin a \cos b + \cos a \sin b$$

$$\sin(a - b) = \sin a \cos b - \cos a \sin b$$

$$\cos(a + b) = \cos a \cos b - \sin a \sin b$$

$$\cos(a - b) = \cos a \cos b + \sin a \sin b$$

2 Linear Algebra

2.1 Vector

Definition 3 *A vector is a quantity that possesses both magnitude and direction, represented by an arrow, the direction of which indicates the direction of the quantity and the length of which is proportional to the magnitude.*

$$(x_1, x_2, \cdots, x_n) \ \ or \ \ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

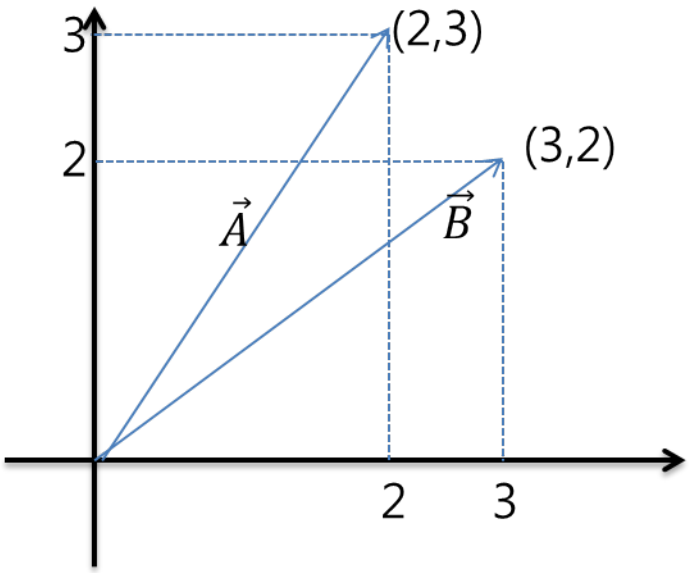


Figure 6: Vector Examples

To distinguish a vector from a number, we place an arrow above it, i.e, $\vec{A} = (2,3)$ and $\vec{B} = (3,2)$ The i-th component of $\vec{A}$ is a_i $\vec{A} = (a_1, a_2, \cdots, a_i, \cdots, a_n)$

Definition 4 *Special Vectors*

- *Null Vector* $\vec{0}$*:* $\vec{0} = (0, 0, \cdots, 0)$
- *Unit Vector* $\vec{e_i}$*: Only the* i*-th component is 1 and all the other components are 0.* $\vec{e_i} = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$

2.1.1 Operations of Vectors

The distance between a point $A = (a_1, \cdots, a_n)$ and a point $B = (b_1, \cdots, b_n)$ is $d(A, B) = \sqrt{\sum_{i=1}^{n} (a_i - b_i)^2}$.

Proposition 1 *Trigonometric Inequalities*

$$d(A, B) \leq d(A, C) + d(C, B)$$

Definition 5 *Operations of Vectors*

- *The magnitude of a vector is the distance from the origin.*

$$|\vec{A}| = \sqrt{\sum_{i=1}^{n} a_i^2}$$

- *Scalar Product of a vector:*

$$k\vec{A} = (ka_1, \cdots, ka_n)$$

- *Addition of vectors:*

$$\vec{A} + \vec{B} = (a_1 + b_1, \cdots, a_n + b_n)$$

- *Linear Combination of vectors:*

$$k\vec{A} + l\vec{B} = (ka_1 + lb_1, \cdots, ka_n + lb_n)$$

- *Inner Product of two vectors:*

$$\vec{A} \cdot \vec{B} = a_1 \cdot b_1 + \cdots + a_n \cdot b_n$$

Proposition 2 *Laws of the Inner Product*

- *Commutative law:* $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$
- $\vec{x} \cdot (a\vec{y}) = a\vec{x} \cdot \vec{y}$
- *Distributive law:* $(\vec{x} + \vec{y}) \cdot \vec{z} = (\vec{x} \cdot \vec{z}) + (\vec{y} \cdot \vec{z})$

Proposition 3 *(Law of cosines) Let a, b, c be the sides of a triangle. If the angle between a and b is θ,*

$$c^2 = a^2 + b^2 - 2ab\cos\theta$$

Proof.

$$\begin{aligned} c^2 &= b^2\sin^2\theta + (a - b\cos\theta)^2 \\ &= b^2\sin^2\theta + a^2 - 2ab\cos\theta + b^2\cos^2\theta \\ &= a^2 + b^2 - 2ab\cos\theta \end{aligned}$$

■

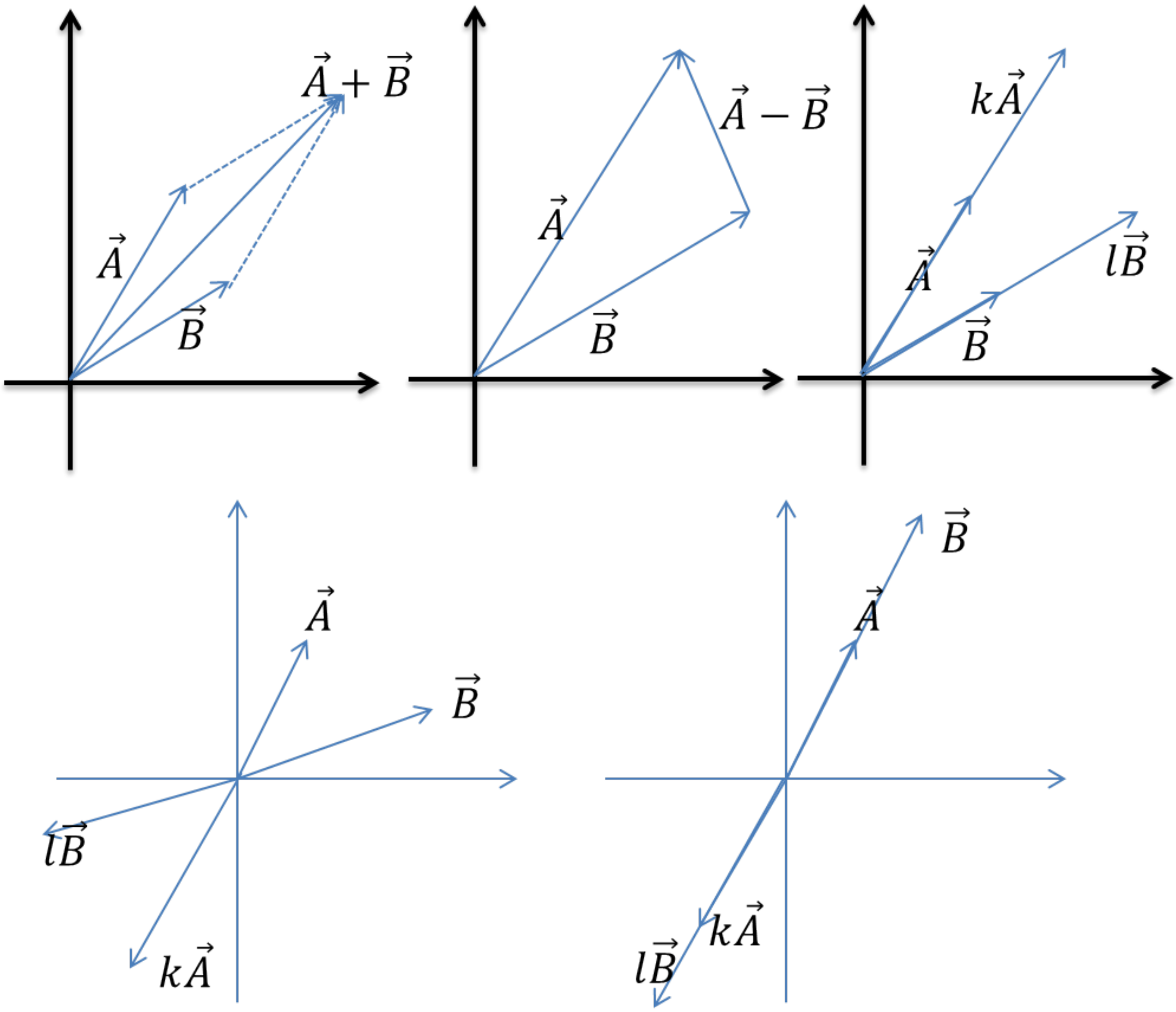


Figure 7: operations of Vectors

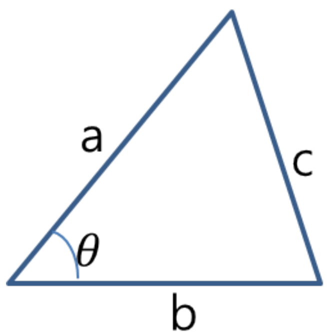


Figure 8: Law of Cosines

Proposition 4 $|\vec{A}|^2 = \vec{A} \cdot \vec{A}$

If θ is the angle between $\vec{A}$ and $\vec{B}$,

Proposition 5 $\vec{A} \cdot \vec{B} = |A||B| \cos\theta$

Proof. By Proposition 4

$$\begin{aligned} |\vec{A} - \vec{B}|^2 &= (\vec{A} - \vec{B}) \cdot (\vec{A} - \vec{B}) \\ &= \vec{A} \cdot \vec{A} + \vec{B}\vec{B} - 2\vec{A} \cdot \vec{B} \\ &= |\vec{A}|^2 + |\vec{B}|^2 - 2\vec{A} \cdot \vec{B} \end{aligned}$$

By the Law of Cosines (Proposition 3) $|\vec{A}|^2 + |\vec{B}|^2 - 2|\vec{A}||\vec{B}| \cos\theta = |\vec{A}|^2 + |\vec{B}|^2 - 2\vec{A} \cdot \vec{B}$. Thus,

$$\vec{A} \cdot \vec{B} = |\vec{A}||\vec{B}| \cos\theta$$

■

Note 1 *In a space, the dimension of which is less or equal to 3, if $\vec{A}$ and $\vec{B}$ are perpendicular (orthogonal)* $\Leftrightarrow \vec{A} \cdot \vec{B} = 0$

Definition 6 *If the dimension is equal to or larger than 4, we define the angle of two vectors $\vec{A}$ and $\vec{B}$, θ as*

$$\cos\theta = \frac{\vec{A} \cdot \vec{B}}{|\vec{A}||\vec{B}|}$$

Definition 7 *In an n-dimensional space, $\vec{A}$ and $\vec{B}$ are* ***orthogonal*** *if and only if* $\Leftrightarrow \vec{A} \cdot \vec{B} = 0$ $\Leftrightarrow \vec{A} \perp \vec{B}$

Example 1 *Let* $\vec{A} = (1, 0, 0, 0)$, $\vec{B} = (0, 0, 1, 0)$.

$$\vec{A} \perp \vec{B} \Leftrightarrow \vec{A} \cdot \vec{B} = 0$$

Proposition 6 (Schwartz Inequality) $|\vec{A} \cdot \vec{B}| \leq |\vec{A}||\vec{B}|$

Because $|\cos\theta| \leq 1$, the Schwartz inequality holds. The equality holds if and only if $\theta = 0$ or π.

2.1.2 Vector Space and Linear Combination

Definition 8 ***Independence and Dependence***

For $\vec{v_1}, \cdots, \vec{v_m}$, if there exists a sequence of real numbers $a_1, \cdots a_m$, at least one which is not 0, so that $\sum_i^m a_i \vec{v_i} = \vec{0}$, then $\vec{v_1}, \cdots, \vec{v_m}$ is ***linearly dependent****. However, if such sequence of real numbers does not exist, these vectors are* ***linearly independent****.*

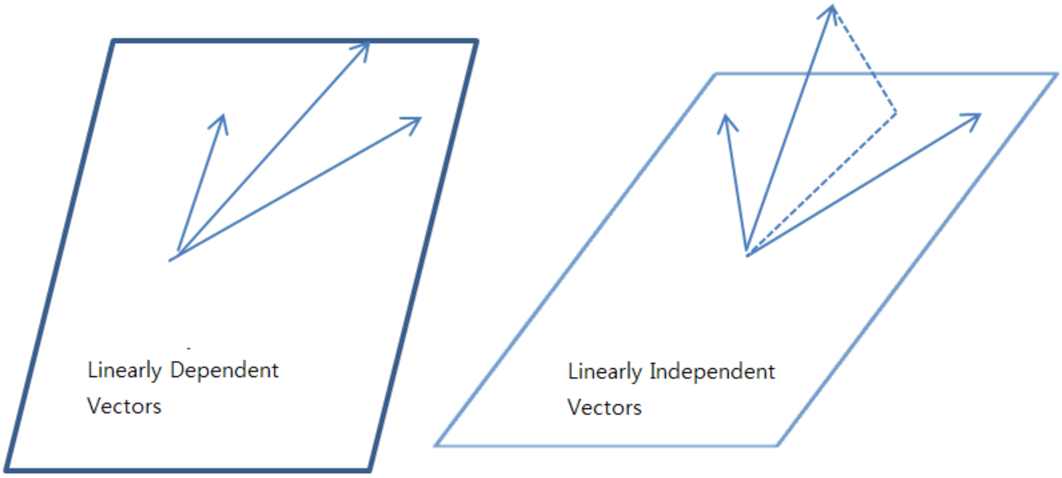


Figure 9: Indepence of Vectors

Example 2 *Examples of dependent and independent vectors.*

1. $(1, 0, 0)$ *and* $(2, 0, 0)$ *are linearly dependent.*

2. $(1, 0, 0)$, $(1, 1, 0)$ *and* $(1, 1, 1)$ *are linearly independent.*

3. $(1, 0, 0)$, $(0, 1, 0)$ *and* $(2, 1, 0)$ *are linearly dependent.*

4. $(1, 0, 0)$, $(0, 1, 0)$, $(0, 0, 1)$ *and* $(3, 2, 0)$ *are linearly dependent.*

Definition 9 ***Spanning of a Vector Space***

Suppose there are m *vectors, i.e.,* $\vec{v_1}, \cdots, \vec{v_m}$. *A vector space,* $\mathbf{V}$, *spanned by* $\vec{v_1}, \cdots, \vec{v_m}$, *is a set of vectors that is a linear combination of* $\vec{v_1}, \cdots, \vec{v_m}$, *i.e., for any real number* $a_1, \cdots, a_m$,

$$\mathbf{V} = \{\vec{v}|\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}\}$$

$\{\vec{v_1}, \cdots, \vec{v_m}\}$ *is the spanning set.*

Note 2 *Two different unit vectors are linearly independent.*

Proposition 7 *If* $\vec{v_1}, \cdots, \vec{v_m}$ *are linearly independent, then vectors of subsets* $\vec{v_1}, \cdots, \vec{v_n}$ *are also linearly independent.*

Proof. Suppose $\{\vec{v_1}, \cdots, \vec{v_n}\} \subset \{\vec{v_1}, \cdots, \vec{v_m}\}$. If $\{\vec{v_1}, \cdots, \vec{v_n}\}$ are linearly dependent, there is a real number sequence $b_1, \cdots, b_n$, at least one of which is not 0, such that

$$b_1\vec{v_1} + \cdots b_n\vec{v_n} = 0$$

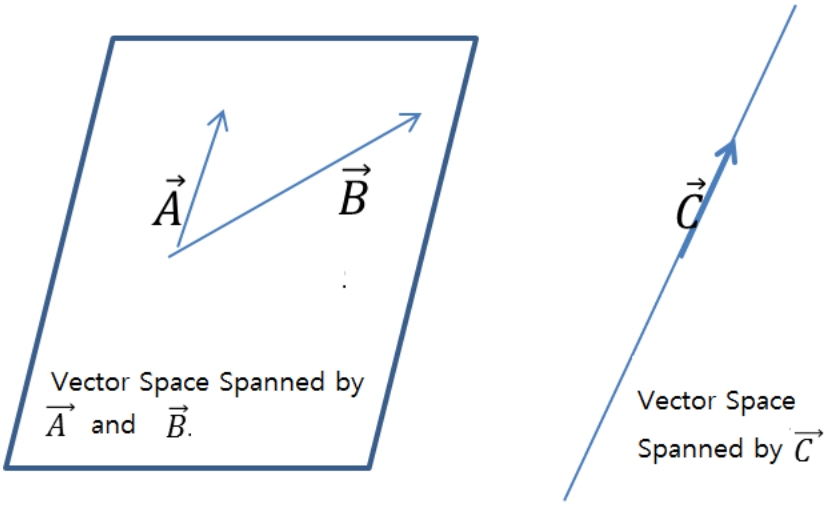


Figure 10: Spanning a Vector Space

Since $\vec{v_1}, \cdots, \vec{v_m}$ are linearly independent, there is a real number sequence $a_1 = \cdots = a_m = 0$, such that

$$a_1\vec{v_1} + \cdots a_m\vec{v_m} = 0$$

Thus,

$$\begin{aligned}
0 &= b_1\vec{v_1} + \cdots b_n\vec{v_n} \\
0 &= b_1\vec{v_1} + \cdots b_n\vec{v_n} + \vec{0} \\
0 &= b_1\vec{v_1} + \cdots b_n\vec{v_n} + a_1\vec{v_1} + \cdots a_m\vec{v_m} \\
0 &= (a_1 + b_1)\vec{v_1} + \cdots + (a_n + b_n)\vec{v_n} + a_{n+1}\vec{v_{n+1}} + \cdots + a_m\vec{v_m}
\end{aligned}$$

Since $\{a_1 + b_1, \cdots, (a_n + b_n), a_{n+1}, \cdots, a_m\} = \{b_1, \cdots, b_n, 0, \cdots, 0\}$ and at least one of $b_1, \cdots, b_n$ is not 0, at least one of $\{b_1, \cdots, b_n, 0, \cdots, 0\}$ is not 0. By Definition 8, $\vec{v_1}, \cdots, \vec{v_m}$ are linearly dependent. Thus , there is a contradiction. ■

Definition 10 ***Basis***

If vectors of a spanning set of a vector space **V** *are linearly independent, we call the spanning set a* ***basis*** *of* **V**.

Note 3 *As there are an infinite number of spanning sets, there are also an infinite number of bases.*

Proposition 8 *If m vectors, $\vec{v_1}, \cdots, \vec{v_m}$, in a vector space* **V** *are orthogonal, i.e., for any $i \neq j \in \{1, 2, \cdots, m\}$, $\vec{v_i} \cdot \vec{v_j} = 0$, $\vec{v_1}, \cdots, \vec{v_m}$ are independent.*

Proof. Suppose they are linearly dependent. Then, for any $\vec{v_i}$, there is at least one sequence of real numbers, at least one of which is not 0, $a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_m$, so that

$$\vec{v_i} = a_1\vec{v_1} + \cdots + a_{i-1}\vec{v_{i-1}} + a_{i+1}\vec{v_{i+1}} + \cdots + a_m\vec{v_m}$$

Take a inner product of $\vec{v_1}$ on both sides. Then,

$$\vec{v_1} \cdot \vec{v_i} = a_1|\vec{v_1}|^2 + 0 + \cdots + 0$$

Thus, $a_1 = 0$. It will be the same for all a_i so

$a_1, \cdots, a_{i-1}, a_{i+1}, \cdots, a_m = 0$. There is a contradiction. ■

Proposition 9 *Let $\vec{v_1}, \cdots, \vec{v_m}$ be the basis of a vector space $\mathbf{V}$. For any vector $\vec{v} \in \mathbf{V}$, there exists a unique real number sequence $a_1, \cdots, a_m$ that satisfies*

$$\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}.$$

and vice versa. Thus, if and only if

$$\vec{v} = b_1\vec{v_1} + \cdots + b_m\vec{v_m}$$

then for all $i \in \{1, \cdots, m\}$, $a_i = b_i$

Proof. (necessary condition) For any vector $v \in \mathbf{V}$, let $\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$ and let $\vec{v} = b_1\vec{v_1} + \cdots + b_m\vec{v_m}$. Then $a_1\vec{v_1} + \cdots + a_m\vec{v_m} = b_1\vec{v_1} + \cdots + b_m\vec{v_m}$ so

$$(a_1 - b_1)\vec{v_1} + \cdots + (a_m - b_m)\vec{v_m} = 0$$

Since $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is the basis, all $\vec{v_i}$s are independent and $a_1 = b_1, \cdots, a_m = b_m$.

(sufficient condition) For any vector $v \in \mathbf{V}$ can be made by linear combination of $\vec{v_1}, \cdots, \vec{v_m}$, $\vec{v_1}, \cdots, \vec{v_m}$ is a spanning set of $\mathbf{V}$. Suppose $a_1\vec{v_1} + \cdots + a_m\vec{v_m} = \vec{0}$. For any vector $\vec{v} = b_1\vec{v_1} + \cdots + b_m\vec{v_m}$,

$$\begin{aligned}
\vec{v} &= b_1\vec{v_1} + \cdots + b_m\vec{v_m} + \vec{0} \\
&= b_1 v_1 + \cdots + b_m\vec{v_m} + a_1\vec{v_1} + \cdots + a_m\vec{v_m} \\
&= (a_1 + b_1)\vec{v_1} + \cdots + (a_m + b_m)\vec{v_m}
\end{aligned}$$

By the assumption of uniqueness, $a_1 = \cdots = a_m = 0$. Thus, $\vec{v_1}, \cdots, \vec{v_m}$ are linearly independent. ■

Proposition 10 *Suppose $\vec{v_1}, \cdots, \vec{v_M}$ spans $\mathbf{V}$. Let $\vec{v_1}, \cdots, \vec{v_m}$ $(m \leq M)$ be the subset of $\vec{v_1}, \cdots, \vec{v_M}$ that are linearly independent and has the maximum number of elements. Then $\vec{v_1}, \cdots, \vec{v_m}$ is basis of $\mathbf{V}$.*

Proof. We have to show that any $\vec{v} \in \mathbf{V}$ is a linear combination of $\vec{v_1}, \cdots \vec{v_m}$.

First, pick any vector $\vec{x} \in \{\vec{v_{m+1}}, \cdots, \vec{v_M}\}$. Then, because $\vec{v_1}, \cdots, \vec{v_m}$ is the subset of independent vectors with a maximum number, $\vec{v_1}, \cdots, \vec{v_m}, \vec{x}$ are linearly dependent. Then there is a real number sequence $a_1, \cdots, a_m, a$, at least one of which is not 0, such that $a_1\vec{v_1} + \cdots + a_m\vec{v_m} + a\vec{x} = 0$. If, $a = 0$, then at least one of $a_1, \cdots, a_m$ is not 0. Thus $\vec{v_1}, \cdots, \vec{v_m}$ is linearly dependent. Thus, $a \neq 0$. Therefore,

$$\vec{x} = b_1\vec{v_1} + \cdots + b_m\vec{v_m}$$

Thus, all $\{\vec{v_{m+1}}, \cdots, \vec{v_M}\}$ is a linear combination of $\vec{v_1}, \cdots, \vec{v_m}$. Since any vector $\vec{v} \in \mathbf{V}$ is a linear combination of $\vec{v_1}, \cdots, \vec{v_M}$, it is also a linear combination of $\vec{v_1}, \cdots, \vec{v_m}$. ■

Proposition 11 *For a basis $\vec{v_1}, \cdots, \vec{v_m}$ of a vector space $\mathbf{V}$, for a $\vec{v} \neq \vec{0}$, there is a real number sequence $a_1, \cdots, a_m$, at least one of which is not 0, such that*

$$\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$$

For $a_i \neq 0$, insert $\vec{v}$ instead of $\vec{v_i}$. Then $\vec{v_1}, \cdots, \vec{v_{i-1}}, \vec{v}, \vec{v_{i+1}}, \cdots, \vec{v_m}$ is also the basis of $\mathbf{V}$.

Proof. We have to prove that 1. $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$ span $\mathbf{V}$ and 2. $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$ are independent.

1. First, show that $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$ span $\mathbf{V}$, i.e., for any $\vec{w} \in \mathbf{V}$, $\vec{w} = b_1\vec{v_1} + \cdots + b_i\vec{v} + \cdots + b_m\vec{v_m}$, there should be $b_1, \cdots, b_m$.

   Since $\vec{w} \in \mathbf{V}$, there is a real number sequence such that

   $$\vec{w} = c_1\vec{v_1} + \cdots + c_i\vec{v_i} + \cdots + c_m\vec{v_m}.$$

   Since there is a real number sequence $a_1, \cdots, a_m$, such that $\vec{v} = a_1\vec{v_1} + \cdots + a_i\vec{v_i} + \cdots + a_m\vec{v_m}$ and $a_i \neq 0$. Thus,

   $$\vec{v_i} = d_1\vec{v_1} + \cdots + d_i\vec{v} + \cdots + d_m\vec{v_m}$$

   $$\begin{aligned}\vec{w} &= c_1\vec{v_1} + \cdots + c_i\vec{v_i} + \cdots + c_m\vec{v_m} \\ &= c_1\vec{v_1} + \cdots + c_i(d_1\vec{v_1} + \cdots + d_i\vec{v} + \cdots + d_m\vec{v_m}) + \cdots + c_m\vec{v_m} \\ &= (c_1 + c_id_1)\vec{v_1} + \cdot + c_id_i\vec{v} + \cdots + (c_m + c_id_m)\vec{v_m}\end{aligned}$$

   Let $b_j = c_j + c_id_j$.

2. Show that $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$ are independent. Suppose $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$ are dependent. Then, there is $e_i \neq 0$ such that,

$$e_1\vec{v_1} + \cdots + e_i\vec{v} + \cdots + e_m\vec{v_m} = 0$$

Since $\vec{v} = a_1\vec{v_1} + \cdots + a_i\vec{v_i} + \cdots + a_m\vec{v_m}$ and $a_i \neq 0$

$$\begin{aligned} 0 &= e_1\vec{v_1} + \cdots + e_i\vec{v} + \cdots + e_m\vec{v_m} \\ &= e_1\vec{v_1} + \cdots + e_i(a_1\vec{v_1} + \cdots + a_i\vec{v_i} + \cdots + a_m\vec{v_m}) + \cdots + e_m\vec{v_m} \\ &= (e_1 + e_ia_1)\vec{v_1} + \cdots + e_ia_i\vec{v_i} + \cdots + (e_m + e_ia_m)\vec{v_m} \end{aligned}$$

Since $e_ia_i \neq 0$ and $\vec{v_1}, \cdots, \vec{v_i}, \cdots, \vec{v_m}$ are linearly dependent, which is a contradiction.

■

Proposition 12 *All bases of a vector space* $\mathbf{V}$ *have the same number of vectors.*

Proof. Let $\vec{v_1}, \cdots, \vec{v_m}$ and $\vec{V_1}, \cdots \vec{V_M}$ are bases of vector space $\mathbf{V}$, and suppose $M > m$. Since $\vec{V_1}, \cdots \vec{V_M} \in \mathbf{V}$, for $1 \leq i \leq m$,

$$v_i = a_1\vec{V_1} + \cdots + a_M\vec{V_M}$$

By Proposition 11, for $a_j \neq 0$, we can replace $\vec{V_j}$ by v_i. By this method, we can replace all $\vec{V_1}, ..., \vec{V_m}$ with $\vec{v_1}, \cdots, \vec{v_m}$, such that $\vec{v_1}, \cdots, \vec{v_m}, \vec{V_{m+1}}, \cdots, \vec{V_M}$ are basis. Since $\vec{v_1}, \cdots \vec{v_m}$ is basis $\vec{V_{m+1}}, \cdots, \vec{V_M}$ is a linear combination of $\vec{v_1}, \cdots \vec{v_m}$, which is a contradiction. ■

Definition 11 *Let the number of vectors of a basis of a vector space* $\mathbf{V}$ *be the* ***dimension*** *of* $\mathbf{V}$*, which we denote by* $\dim \mathbf{V}$.

Proposition 13 *For an* m*-dimensional vector space* $\mathbf{V}$*, these properties always hold.*

1. *For* m *independent vectors in* $\mathbf{V}$ *is always a basis of* $\mathbf{V}$.

2. *A spanning set of* $\mathbf{V}$ *with* m *vectors is always a basis.*

3. *For any* $m + 1$ *vectors in* $\mathbf{V}$ *are always dependent.*

2.2 Matrix

2.2.1 Definition of Matrix

Definition 12 ***A matrix*** *is a combination of vectors with the same number of components so that it is a rectangular array of numbers, symbols, or expressions, arranged in rows and columns.*

Note 4 *View matrix as a combination of vectors rather than numbers.*

Example 3 *An $m \times n$ matrix A_{mn}*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

For $1 \leq i \leq m$, $1 \leq j \leq n$, the i, j element of A is denoted by a_{ij}. We call the $\begin{pmatrix} a_{i1} & \cdots & a_{in} \end{pmatrix}$ the ith row of A, and $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ the jth column of A.

Definition 13 *Special matrix*

1. *Null matrix*

$$\mathbf{0} = \begin{pmatrix} 0 & \cdots & 0 \\ & \cdots & \\ 0 & \cdots & 0 \end{pmatrix}$$

2. *Identity matrix*

$$\mathbf{I}_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 1 \end{pmatrix}$$

3. *Diagonal matrix*

$\forall i \neq j a_{ij} = 0$

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

4. *Symmetric matrix*

$\forall i \& j, \quad a_{ij} = a_{ji}$

5. *Upper triangular matrix*

$i > j \Rightarrow a_{ij} = 0$

6. *Lower triangular matrix*

   $i > j \Rightarrow a_{ji} = 0$

7. *Square matrix*

   A_{nn}

2.2.2 Operations of Matrices

Definition 14 *Scalar product and addition of matrices*

1. *Scalar Product*
   $$\alpha A = \alpha \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} \alpha a_{11} & \cdots & \alpha a_{1n} \\ \vdots & \vdots & \vdots \\ \alpha a_{m1} & \cdots & \alpha a_{mn} \end{pmatrix}$$

2. *Addition*
   $$A + B = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & \vdots & \vdots \\ b_{m1} & \cdots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \cdots & a_{1n} + b_{1n} \\ \vdots & \vdots & \vdots \\ a_{m1} + b_{m1} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$$

3. *Trace*

   For a square matrix $A_{n \times n}$, the sum of all elements on the diagonal line

   $$trA = \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \cdots + n_{nn}$$

   Example 4

   $$tr \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = 5$$

4. *Transpose A'*

   If $B = A'$, $b_{ij} = a_{ji}$.

   Example 5

   $$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}' = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

Proposition 14 *Laws of operation for matrices*

1. *$A + B$ and αA are all matrices*

2. *(commutative law) $A + B = B + A$*

3. *(associative law)* $(A+B)+C = A+(B+C)$
4. $A+0 = 0+A = A$
5. $A+(-A) = 0$
6. *(distributive law)*$\alpha(A+B) = \alpha A + \alpha B$
7. *(distributive law)* $(\alpha+\beta)A = \alpha A + \beta A$
8. *(associative law)* $\alpha(\beta A) = (\alpha\beta)A$
9. $1A = A$
10. $tr(A+B) = trA + trB$
11. $tr(cA) = c(trA)$
12. $trA' = trA$

Matrix Multiplication

One may

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1x + a_2y + a_3z \\ b_1x + b_2y + b_3z \\ c_1x + c_2y + c_3z \end{pmatrix}$$

Although the result is the same, but

$$\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = x \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + y \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} + z \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix}$$

This way, one may view matrix multiplication as linear combinations of vectors (columns).

Similarly, $\begin{pmatrix} x & y & z \end{pmatrix} \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} = x \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix} + y \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} + z \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix}$

Also note that, $\begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{pmatrix} \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \\ z_1 & z_2 \end{pmatrix}$

$$= \left(x_1 \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + y_1 \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} + z_1 \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix} \quad x_2 \begin{pmatrix} a_1 \\ b_1 \\ c_1 \end{pmatrix} + y_2 \begin{pmatrix} a_2 \\ b_2 \\ c_2 \end{pmatrix} + z_2 \begin{pmatrix} a_3 \\ b_3 \\ c_3 \end{pmatrix} \right)$$

$$= \begin{pmatrix} a_1(x_1, x_2) + a_2(y_1, y_2) + a_3(z_1, z_2) \\ b_1(x_1, x_2) + b_2(y_1, y_2) + b_3(z_1, z_2) \\ c_1(x_1, x_2) + c_2(y_1, y_2) + c_3(z_1, z_2) \end{pmatrix}$$

Note 5 *Generally* $AB \neq BA$

Proposition 15 *Laws of matrix multiplications*

1. $A(BC) = (AB)C$
2. $A(B + C) = AB + AC$
3. $(A + B)C = AC + BC$
4. $(\alpha A)B = A(\alpha B) = \alpha(AB)$
5. *For an identity matrix* I, $AI = IA = A$
6. $AA = A^2$
7. $tr(AB) = tr(BA)$
8. $(AB)' = B'A'$

2.2.3 Matrices and Simultaneous Equations

Example 6 *Solve the following simultaneous equations*

First Solution

$$\begin{aligned} 2x + 4y + 6z &= 12 & (1) \\ x + y + 2z &= 4 & (2) \\ x + 2y + 2z &= 5 & (3) \end{aligned}$$

$$(1) - 2 \times (3)$$

$$\begin{aligned} 2x + 4y + 6z &= 12 & (4) \\ x + y + 2z &= 4 & (5) \\ 2z &= 2 & (6) \end{aligned}$$

$$(4) - 2 \times (5)$$

$$\begin{aligned} 2x + 4y + 6z &= 12 & (7) \\ 2y + 2z &= 4 & (8) \\ 2z &= 2 & (9) \end{aligned}$$

Second Solution

You may solve the equations in this way. Note that $\begin{pmatrix} 2 & 4 & 6 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 5 \end{pmatrix}$

$$\begin{pmatrix} 2 & 4 & 6 & 12 \\ 1 & 1 & 2 & 4 \\ 1 & 2 & 2 & 5 \end{pmatrix}$$

We may view $(1) - 2 \times (3)$ as multipling third row by 2 and substracting it from the first row. Replace the third row.

$$\begin{pmatrix} 2 & 4 & 6 & 12 \\ 1 & 1 & 2 & 4 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

We may view $(4) - 2 \times (24)$ asmultipling the second row by 2 and substracting it from the first row. Replace the second row.

$$\begin{pmatrix} 2 & 4 & 6 & 12 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & 2 & 2 \end{pmatrix}$$

We are making a triangular matrix. We call this method *the Gaussian Elimination Method.*

Definition 15 ***Inverse Matrix***

For a $n \times n$ square matrix A, if there is a $n \times n$ matrix X such that $AX = XA = I$, we call X an ***inverse matrix*** *of A and denote it as A^{-1}. An inverse matrix may not exist. If an inverse matrix exist for a matrix A, we call A* ***invertible (or non-singular)*** *and if not* ***non-invertible(or singular)****.*

We may find the inverse matrix of a matrix A by the Gaussian elimination method.

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 1 & 1 & 2 & & 0 & 1 & 0 \\ 1 & 2 & 2 & & 0 & 0 & 1 \end{pmatrix}$$

$$(3) \Leftarrow (1) - 2 \times (3)$$

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 1 & 1 & 2 & & 0 & 1 & 0 \\ 0 & 0 & 2 & & 1 & 0 & -2 \end{pmatrix}$$

$$(3) \Leftarrow (3) \div 2$$

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 1 & 1 & 2 & & 0 & 1 & 0 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(2) \Leftarrow (1) - 2 \times (2)$$

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 0 & 2 & 2 & & 1 & -2 & 0 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(2) \Leftarrow (2) \div 2$$

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 0 & 1 & 1 & & \frac{1}{2} & -1 & 0 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(2) \Leftarrow (2) - (3)$$

$$\begin{pmatrix} 2 & 4 & 6 & & 1 & 0 & 0 \\ 0 & 1 & 0 & & 0 & -1 & 1 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(1) \Leftarrow (1) - 6 \times (3)$$

$$\begin{pmatrix} 2 & 4 & 0 & & -2 & 0 & 6 \\ 0 & 1 & 0 & & 0 & -1 & 1 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(1) \Leftarrow (1) - 4 \times (2)$$

$$\begin{pmatrix} 2 & 0 & 0 & & -2 & 4 & 2 \\ 0 & 1 & 0 & & 0 & -1 & 1 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

$$(1) \Leftarrow (1) \div 2$$

$$\begin{pmatrix} 1 & 0 & 0 & & -1 & 2 & 1 \\ 0 & 1 & 0 & & 0 & -1 & 1 \\ 0 & 0 & 1 & & \frac{1}{2} & 0 & -1 \end{pmatrix}$$

Thus, the inverse matrix of $\begin{pmatrix} 2 & 4 & 6 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$ is $\begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{pmatrix}$

Note 6 *Simulatneous equations can be solved by finding the inverse matrix.*

Suppose we want to solve

$$\begin{aligned} 2x + 4y + 6z &= 12 \\ x + y + 2z &= 4 \\ x + 2y + 2z &= 5 \end{aligned}$$

These simultaneous equations may be expressed as

$$\begin{pmatrix} 2 & 4 & 6 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 5 \end{pmatrix}$$

By multiplying the inverse matrix on both sides

$$\begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 4 & 6 \\ 1 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{pmatrix} \begin{pmatrix} 12 \\ 4 \\ 5 \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 1 \\ \frac{1}{2} & 0 & -1 \end{pmatrix} \begin{pmatrix} 12 \\ 4 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

2.2.4 Row, Column and Null Space of a Matrix and the Fundamental Theorem of Linear Algebra

Definition 16 *Row Space and Column Space of a Matrix*

*A matrix A_{mn} has m rows and n columns. In fact, all columns and rows are vectors, and we call them either column or row vectors. The vector space spanned by the column vectors of A is called the **column space** of A and the vector space spanned by the row vectors is called the **row space** of A. We denote the column space and row space respectively as $Col(A)$ and $Row(A)$.*

Definition 17 *Null Space*

*Of a matrix A_{mn}, the vector space that is spanned by vectors $\vec{x}$ that satisfies $A\vec{x} = 0$ is called the **Null Space** of A and we denote this as $Null(A)$. On the other hand, the vector space that is spanned by vectors $\vec{x}$ that satisfies $A'\vec{x} = 0$ is called the **Left Null Space** of A.*

Proposition 16 *The Fundamental Theorem of Linear Algebra*

For a matrix A_{mn},

1. *the dimensions of the row space and the column space are always equal* $(\dim(Row(A)) = \dim(Col(A)))$*,*
2. *the sum of the dimensions of the row space and the null space is n (or equivalently, the sum of the dimensions of the column space and the left null space is always m).*

First, let us analyze the implication behind the theorem.

Note 7 *The Gaussian elimination method does not change the dimensions of row Space and column Space. For example, a matrix A that has dependent colums which means there is a non-zero vector x such that*

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix} x_2 + \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} \end{pmatrix} x_3 = 0$$

Now suppose we mulitplied c to the second row and added to the third row, which gives

$$\begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} + c \cdot a_{21} \end{pmatrix} x_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} + c \cdot a_{22} \end{pmatrix} x_2 + \begin{pmatrix} a_{13} \\ a_{23} \\ a_{33} + c \cdot a_{23} \end{pmatrix} x_3 = 0$$

By continuing the Gaussian elimination, we can change any A_{mn}, without altering the dimension of the row space or the column space to

$$\begin{pmatrix} a_{11} & \cdots & \cdots & \cdots & a_{1n} \\ 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & a_{kk} & \cdots & a_{kn} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Note that there are k independent row vectors and k independent column vectors. Thus, the dimension of row space and column space are both k and (1) of the Fundamental Theorem holds.

Definition 18 *Let the dimension of the row space of a matrix A be k (thus, $\dim(Row(A)) = \dim(Col(A)) = k$). We call the dimension of the row space (and column space) the **rank** of A and denote it as $r(A) = k$.*

Now let us consider (2) of the Fundamental Theorem. First, consider the meaning of the null space. A vector $\vec{x}$ that is in the null space of A satisfies $A\vec{x} = 0$. Thus, the inner product of $\vec{x}$ and all rows of A are always 0. Thus, $\vec{x}$ is orthogonal to all rows of A by Definition 7. Intuitively, by combining the two subspaces that are orthogonal to each other, we will have the n-dimensional space.

$$\dim(Null(A)) + r(A) = n$$

Also note that for a vector $\vec{y}$ that is in the left null space, $A'\vec{y} = 0$. Thus, the inner product of $\vec{y}$ and all rows of A', which are all columns of A, are always 0. Thus, $\vec{y}$ is orthogonal to all rows of A' (or columns of A) by Definition 7. Intuitively, by combining the two subspaces that are orthogonal to each other, we will have the m-dimensional space.

$$\dim(Null(A')) + r(A) = m$$

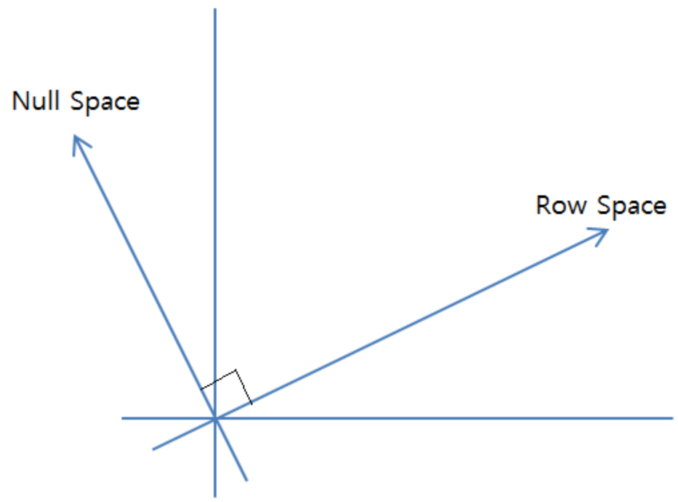


Figure 11: Row Space and Null Space

Because the proof for the Fundamental Theorem of Linear Algebra is out of the scope of this class, we will omit the proof.

Let us return to the previous simultaneous equation. We can solve the simultaneous equations by either the Gaussian method of elimination or by calculating the inverse function. Then, does the Gaussian elimination or the inverse function always gives the solution?

The previous simultaneous equations can also be expressed as

$$x\begin{pmatrix}2\\1\\1\end{pmatrix} + y\begin{pmatrix}4\\1\\2\end{pmatrix} + z\begin{pmatrix}6\\2\\2\end{pmatrix} = \begin{pmatrix}12\\4\\5\end{pmatrix}$$

This shows that the vector on the right hand side $(12, 4, 5)'$ is a linear combination of the column vectors. Thus, if the vector $(12, 4, 5)'$ is in the column space, $Col(A)$, then by Proposition 9, there exists a unique real number sequence (x, y, z) such that

$$x \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 4 \\ 1 \\ 2 \end{pmatrix} + z \begin{pmatrix} 6 \\ 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 12 \\ 4 \\ 5 \end{pmatrix}$$

If the dimension of the column space is n ($\dim(Col(A)) = r(A) = n$), then all vectors of n elements is in the column space, thus there is a unique solution. However, if $r(A) < n$, then there exists no unique solution.

On the other hand, if there always exists a unique solution, the matrix is invertible. Thus, if $r(A) = n$, then the matrix A is invertible.

Example 7 *Compare and contrast these two simultaneous equations*

$$\begin{array}{cc} x + y + z = 2 & x + y + z = 2 \\ x + 2y + z = 3 & x + 2y + z = 3 \\ 2x + 3y + 2z = 5 & 2x + 3y + 2z = 9 \end{array}$$

How many solutions are there for the left simultaneous equations? What about the right one? What do these two simultaneous equations have in common? ? What are their differences?

Proposition 17 *For a $n \times n$ square matrix A, the following statements are equivalent.*

1. *A^{-1} exists, i.e., A is invertible (non-singular).*
2. *A has n number of independent rows and columns.*
3. *$r(A) = n$*
4. *$\dim(Col(A)) = \dim(Row(A)) = n$*
5. *$\dim(Null(A)) = 0$*
6. *For any n component vector b, there is a unique solution, x, to $Ax = b$.*

Proof. $1 \Rightarrow 2$.

Let $\vec{a_1}, \ldots, \vec{a_n}$ be A's column vectors. Then, we only have to prove that the unique real number sequence, $x_1, \cdots, x_n$, that satisfies

$$x_1\vec{a_1} + \ldots + x_n\vec{a_n} = Ax = 0$$

is $x_1, \cdots, x_n = 0$. Since there is A^{-1},

$$A^{-1}Ax = Ix = x = A^{-1}0 = 0.$$

Thus, $\vec{a_1}, \dots \vec{a_n}$ are linearly independent.

$2 \Rightarrow 1$ If $\vec{a_1} \dots \vec{a_n}$ are linearly independent, all n elements vectors are linear combinations of $\vec{a_1} \dots \vec{a_n}$. Thus,

$$\begin{aligned}
(1, 0, \cdots, 0)' &= c_{11}\vec{a_1} + \cdots + c_{n1}\vec{a_n} \\
&\vdots \\
(0, \cdots, 0, 1)' &= c_{1n}\vec{a_1} + \cdots + c_{nn}\vec{a_n}
\end{aligned}$$

Find all c_{ij} then let

$$C = \begin{pmatrix} c_{11} & \cdots & c_{1n} \\ \vdots & \ddots & \vdots \\ c_{n1} & \cdots & c_{nn} \end{pmatrix}$$

Note that $AC = I$. Thus, $C = A^{-1}$.

$2 \Leftrightarrow 3$ It is trivial from the definition of rank. $3 \Leftrightarrow 4 \Leftrightarrow 5$, by the Fundamental Theorem of Linear Algebra.

$2 \Leftrightarrow 6$, by Proposition 9.

■

2.2.5 Determinant

Many of you already know that the determinant is

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$$

Let us extend the definition of determinant to an $n \times n$ matrix.

Definition 19 *For a $n \times n$ matrix A,*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Let

$$A_{ij} = \begin{pmatrix} a_{11} & \cdots & a_{1(j-1)} & a_{1(j+1)} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{(i-1)1} & \cdots & a_{(i-1)(j-1)} & a_{(i-1)(j+1)} & \cdots & a_{(i-1)n} \\ a_{(i+1)1} & \cdots & a_{(i+1)(j-1)} & a_{(i+1)(j+1)} & \cdots & a_{(i+1)n} \\ \cdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{n(j-1)} & a_{n(j+1)} & \cdots & a_{nn} \end{pmatrix}$$

and also let

$$C_{ij} = (-1)^{i+j}|A_{ij}|$$

Then the determinant of a matrix A is

$$|A| = \sum_{k=1}^{n} a_{kj}C_{kj} = \sum_{l=1}^{n} a_{il}C_{il}$$

Proposition 18 *No matter which row or column is chosen, the determinant does not change.*

We will take this proposition as given.

Proposition 19 *Laws of Determinant f*

1. $|I| = 1$

2. *If we interchange a row (column) of A with another row (column) of A and obtain a new matrix B, $|A| = -|B|$*

3. $|\vec{a_1}, \cdots, \vec{a_i}' + \vec{a_i}, \cdots, \vec{a_n}| = |\vec{a_1}, \cdots, \vec{a_i}', \cdots, \vec{a_n}| + |\vec{a_1}, \cdots, \vec{a_i}, \cdots, \vec{a_n}|$

4. $|\vec{a_1}, \cdots, t\vec{a_i}, \cdots, \vec{a_n}| = t|\vec{a_1}, \cdots, \vec{a_i}, \cdots, \vec{a_n}|$

Proof. First, we show that if a is the determinant of A, then three properties hold.

1. $|I| = 1$ is straightforward.

2. If we interchange a row (column) of A with another row (column) of A and obtain a new matrix B, $|A| = -|B|$ It is trivial that it holds for $n = 2$. Suppose this property holds for $n - 1 \geq 2$. For a $n \times n$ matrix B,

$$|B| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|B_{ij}|$$

   Since we assumed that it holds for $n - 1$, $|B_{ij}| = -|A|_{ij}$, thus,

$$|B| = -\sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}| = -|A|$$

3. $|\vec{a_1}, \cdots, \vec{a_i}' + \vec{a_i}, \cdots, \vec{a_n}| = \sum_{j=1}^{n}(a_{ij} + a'_{ij})C_{ij} = \sum_{j=1}^{n} a_{ij}C_{ij} + \sum_{j=1}^{n} a'_{ij}C_{ij} = |\vec{a_1}, \cdots, \vec{a_i}', \cdots, \vec{a_n}| + |\vec{a_1}, \cdots, \vec{a_i}, \cdots, \vec{a_n}|$

4. $|\vec{a_1}, \cdots, t\vec{a_i}, \cdots, \vec{a_n}| = \sum_{j=1}^{n} ta_{ij}C_{ij} = t\sum_{j=1}^{n} a_{ij}C_{ij} = t|\vec{a_1}, \cdots, \vec{a_n}|$

■

Proposition 20 *Properties of Determinants*

1. *If there is a row or a column of 0, then $|A| = 0$.*
2. *If A is a triangular matrix, $|A| = a_{11}a_{22}\cdots a_{nn}$*
3. *If we add a column (row) with a scalar product of another column (row), the determinant does not change.*
4. $|A'| = |A|$
5. $|AB| = |A||B|$
6. $|A^{-1}| = \frac{1}{|A|}$
7. *If there are equivalent columns or rows, $|A| = 0$.*

Proof. Proof to 20.3

It is trivial that it holds when $n = 2$. Suppose this holds for $n - 1 \geq 2$. Let the new matrix be B, the matrix obtained by adding a scalar product of a column or a row,

$$|B| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|B_{ij}|$$

Since $|B_{ij}| = |A_{ij}|$,

$$|B| = \sum_{j=1}^{n} a_{ij}(-1)^{i+j}|A_{ij}| = |A|$$

Proof to 20.5

We take this for given

■

Geometric Interpretation of Determinant

The two row vectors of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $\vec{a_1} = \begin{pmatrix} a \\ b \end{pmatrix}$, $\vec{a_2} = \begin{pmatrix} c \\ d \end{pmatrix}$ make a parallelogram. The area of the parallelogram is $2\Delta OAB$ and

$$\Delta OAB = ad - \frac{1}{2}cd - \frac{1}{2}ab - \frac{(a-c)(d-b)}{2} = \frac{1}{2}(ad - bc)$$

Note 8 *Note that the area is always positive for a parallelogram, but for a determinant if the sequence of the vectors changes, so does the sign of the determinant.*

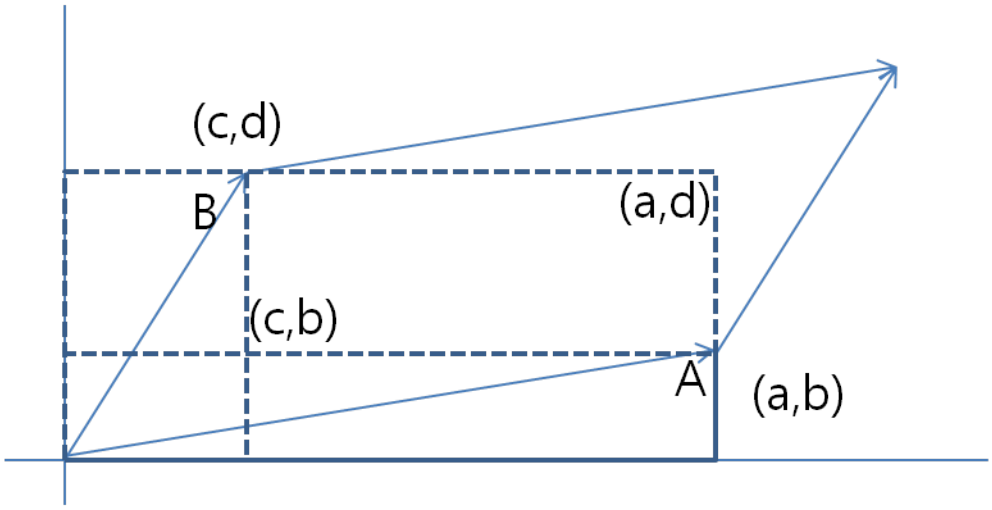


Figure 12: Geometric Interpretation of Determinant

Determinant and Inverse Matrix

Proposition 21 *An $n \times n$ matrix A is invertible $\Leftrightarrow |A| \neq 0$*

Proof. (necessity) If A is invertible, $AA^{-1} = I$ and

$$|I| = 1 = |AA^{-1}| = |A||A^{-1}|$$

Thus, $|A| \neq 0$.

(sufficiency) Suppose A is not invertible. Thus, the column vectors of A are linearly dependent: a column vector A_i is a linear combination of other column vectors. Thus, $A_i = \sum_{j \neq i} \alpha_j A_j$

$$\begin{aligned} |A| &= |A_1, \cdots, A_{i-1}, \sum_{j \neq i} \alpha_j A_j, A_{i+1}, \cdots, A_n| \\ &= \sum_{j \neq i} \alpha_j |A_1, \cdots, A_j, \cdots, A_n| \\ &= 0 \end{aligned}$$

■

2.2.6 Projection

Let $A = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix}$. The column space of A is a line that connects the origin and the point $(2, 1)$. Let $y = (2, 2)$. Then, among the vectors that exist on the column space of A (a vector that is a linear combination of A, i.e., Ax), what is the vector that is closest to y? In the figure below, what is Ax? Let $b = y - Ax$, and b be orthogonal to A. Thus, $A'b = 0$.

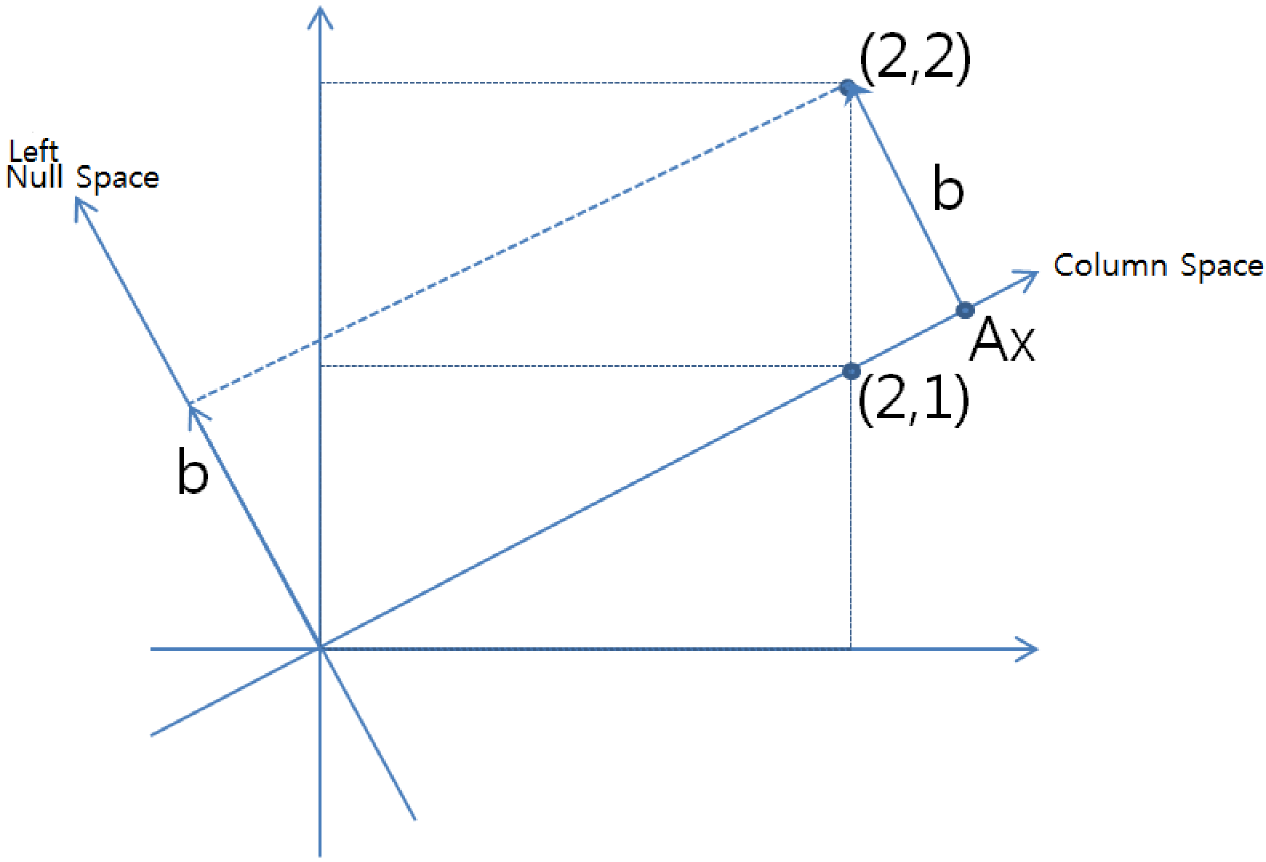


Figure 13: Projection

$$
\begin{aligned}
A'(y - AX) &= 0 \\
A'y - A'AX &= 0 \\
A'y &= A'Ax \\
(A'A)^{-1}A'y &= x \\
A(A'A)^{-1}A'y &= Ax
\end{aligned}
$$

Definition 20 *Projection Matrix*

$$P = A(A'A)^{-1}A'$$

Let $r(A) = n$, i.e., A is non-singular(invertible) matrix. A and A' are both invertible, $(A'A)^{-1} = A^{-1}(A')^{-1}$ and thus, $P = I$. If an $n \times n$ matrix A is invertible, all the column vectors of A are linearly independent; thus, they span the whole n-dimensional space. Therefore, if we project a vector, we obtain the same vector. On the other hand, if A is singular (non-invertible), the projection of y is the closest to Ax.

Last, note that b is orthogonal to the column space of A. Thus, $A'b = 0$. Also since, $Ax = Py$ and $y = Ax + b$, $b = (I - P)y$. Note that

$$A'b = A'(I - A(A'A)^{-1}A')y = (A' - A'A(A'A)^{-1}A')y = (A' - A')y = 0$$

Thus, b **is in the Left Null Space of** A. Ax is in the column space of A. Thus, by combining the columns space and the left null space, we have the whole n-dimensional space.

2.2.7 Eigenvalue and Eigenvector

Definition 21 *The eigenvalue λ and eigenvector $\vec{x}$ of a $n \times n$ matrix A are*

$$A\vec{x} = \lambda x \ \ \textit{and} \ \ \vec{x} \neq 0$$

Let us find the eigenvalue and eigenvectors.

$$Ax = \lambda x \Rightarrow (A - \lambda I)x = 0$$

If any solution except $\vec{0}$ exist, the columns of $(A - \lambda I)$ should be linealy dependent. Thus, $|A - \lambda I| = 0$

Example 8 *Let $A = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}$. Find eigenvalue and eigenvector of A*

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 3-\lambda \end{vmatrix} = 0$$

Thus, $(1-\lambda)(3-\lambda) = 0$ and the eigenvalues are 1 and 3. Let $\lambda_1 = 1, \lambda_2 = 3$. The corresponding eigenvectors are

$$\begin{pmatrix} 0 & 2 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \rightarrow (x, y) = (c, 0)$$

$$\begin{pmatrix} -2 & 2 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0 \rightarrow (x, y) = (c, c)$$

Now let us see why eigenvalues and eigenvectors are important.

Definition 22 *For an $n \times n$ matrix A, there can be n different eigenvalues ($\lambda_1, \cdots, \lambda_n$) and corresponding eigenvectors ($x_1, \cdots, x_n$). Let S be a matrix, the columns of which are eigenvectors of A , and Λ be a matrix of eigenvalues as*

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

.

$$S = \begin{pmatrix} | & \cdots & | \\ x_1 & \cdots & x_n \\ | & \cdots & | \end{pmatrix}$$

We call Λ the eigenvalue matrix and S the eigenvector matrix.

Let us further assume that S is invertible. We then have the following Proposition.

Proposition 22 *Diagonalization*

$$S^{-1}AS = \Lambda$$

Proof.

$$AS = A \begin{pmatrix} | & \cdots & | \\ x_1 & \cdots & x_n \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & \cdots & | \\ \lambda_1 x_1 & \cdots & \lambda_n x_n \\ | & \cdots & | \end{pmatrix}$$

Note that

$$\begin{pmatrix} | & \cdots & | \\ \lambda_1 x_1 & \cdots & \lambda_n x_n \\ | & \cdots & | \end{pmatrix} = \begin{pmatrix} | & \cdots & | \\ x_1 & \cdots & x_n \\ | & \cdots & | \end{pmatrix} \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & \lambda_{n-1} & 0 \\ 0 & \cdots & \cdots & 0 & \lambda_n \end{pmatrix}$$

Thus, $AS = S\Lambda$ or

$$S^{-1}AS = \Lambda$$

■

We call this process the Diagonalization of a matrix, A. However, if S is not invertible then since S^{-1} does not exist, do A cannot be diagonzalized. Although we will not prove this property rigorously, we can intuitively know that

Proposition 23 *An $n \times n$ matrix A with n distinct eigenvalues (or equivalently n independent eigenvectors) can be diagonalized.*

Proposition 24 *The determinant of a matrix A is the product of its eigenvalues.*

Proof. Let $p(\lambda) = |A - \lambda I|$, where λ is an arbitrary real number. for an eigenvalue λ_i, note that $p(\lambda_i) = 0$ by definition. Thus, since $p(\lambda)$ is an n-polynomial of λ and for all n eigenvalues $p(\lambda_i) = 0$,

$$p(\lambda) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$$

Let $\lambda = 0$, then

$$p(0) = \lambda_1 \cdots \lambda_n = |A - 0I| = |A|$$

■

Corollary 1 *If a matrix A is invertible, then it has non-zero eigenvalues.*

Proposition 25 *Let λ be the eigenvalue of A, then λ^{-1} is eigen-value of A^{-1}.*

Proof. $A^{-1}Ax = x = A^{-1}\lambda x = \lambda A^{-1}x$ or $x = \lambda A^{-1}x$. Thus, $A^{-1}x = \lambda^{-1}x$. ■

Example 9 *Let λ be the eigenvalue of A, then show that eigen-value of $A^2 + 2A$ is $\lambda^2 + 2\lambda$.*

2.3 Positive and Negative Definiteness

Definition 23 *An $n \times n$ symmetric matrix A is positive definite if and only if for every nonzero real column vector matrix x,*

$$x^T Ax > 0$$

. An $n \times n$ symmetric matrix$A$ is negative definite if and only if

$$x^T Ax < 0$$

Proposition 26 *The following statements about an $n \times n$ symmetric matrix are equivalent,*

1. *A is positive definite: for every nonzero real column vector matrix x, $x^T Ax > 0$*
2. *All eigenvalues of A are postive*
3. *All the left submatrices A_k have positive determinants, where*

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

   and

$$A_k = \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}$$

   where $k \leq n$

Proof. In this section, we only provide a sketch of the proof.

$1 \to 2$

Let x be an eigenvector of A then, $Ax = \lambda x$, then $x^T Ax = x^T \lambda x = \lambda |x|^2$. Since A is positive definite $x^T Ax > 0$, thus, $\lambda > 0$.

The proof for $2 \to 1$ is out of the scope of this class.

$1 \to 3$

Let x_k be any vector that has zero for $i + 1$-th to the n-th component such that $x_k = (a_1, \cdots, a_k, 0, \cdots, 0)$, $k \le n$

First note that $x_k^T Ax_k = x_k^T A_k x_k$, and also that because A is positive definite, $x_k^T Ax_k > 0$. Thus, A_k are positive definite. In addition, because A_k are positive definite, eigenvalues of A_k are positive and since the determinant of A_k is the product of its eigenvalues,

$$|A_k| > 0$$

■

Corollary 2 *The following statementsabout an $n \times n$ symmetric matrix are equivalent ,*

1. *A is negative definite: for every nonzero real column vector matrix x, $x^T Ax < 0$*
2. *All eigenvalues of A are negative*
3. $(-1)^k |A_k| > 0$

Proposition 27 *Inverse (Transpose) matrix of a positive(negative) defnite matrix is positive(negative) definite.*

Proof. Trnaspose of a positive definite matrix being positive definite is self evident because positive definite matrix is symmetric. If A is positive definite then A is invertible, so define $y = Ax$. Then $y^T A^{-1} y = x^T A^T A^{-1} Ax = x^T Ax > 0$ so is positive definite. ■

Problem Set 2

Problem set on vectors

1. Prove the law of cosines.
2. Let $\vec{A}$, $\vec{B}$ and $\vec{C}$ be n-component vector. Use Schwartz inequality to prove that $|\vec{A} - \vec{C}| \le |\vec{A} - \vec{B}| + |\vec{B} - \vec{C}|$.
3. Show that if a vector of $\vec{v_1}, \cdots, \vec{v_m}$ is a null vector ($\vec{0}$), then the vectors are linearly dependent.

4. If $\mathbf{W}$ is a subspace of $\mathbf{V}$, prove that $\dim \mathbf{W} \leq \dim \mathbf{V}$. Also prove that if the equality holds, then $\mathbf{W} = \mathbf{V}$.

5. Find $\vec{v}, \vec{w}, \vec{u}$ such that $\vec{u} \cdot \vec{v} < 0, \vec{v} \cdot \vec{w} < 0$ and $\vec{w} \cdot \vec{u} < 0$.

Problem set on matrices

1. Prove that if $(a, b) = k(c, d)$, then $(a, c) = l(b, d)$. Thus, if rows of A are linearly dependent, so are its columns.

2. Find an example of $AB \neq BA$.

3. Find a matrix, A such that $A^2 = AA = 0$ and $A \neq 0$.

4. Find all the values of a that makes the following matrices singular.

   (a) $A = \begin{pmatrix} a & 2 & 3 \\ a & a & 4 \\ a & a & a \end{pmatrix}$

   (b) $B = \begin{pmatrix} 2 & a & a \\ a & a & a \\ 3 & 7 & a \end{pmatrix}$

5. Use the Gaussian method of elimination to find inverse matrices of
$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 2 & -1 & -1 \\ -11 & 2 & -11 \\ -1 & -1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

6. Find a matrix A that is not a null matrix and satisfies

   (a) $A^2 = 0$

   (b) $A^2 \neq 0$ but $A^3 = 0$

7. Find the left null space of $\begin{pmatrix} 1 & 0 \\ 1 & 3 \\ 5 & 1 \end{pmatrix}$.

8. Fill in the blanks so that the rank of the matrices is 1,2 and 3, respectively.
$$\begin{pmatrix} 1 & 2 & 4 \\ 2 & & \\ 4 & & \end{pmatrix}$$

$$\begin{pmatrix} 2 & & \\ 1 & & \\ 2 & 6 & -3 \end{pmatrix}$$

9. Solve the following simultaneous equations using Gaussian method of elimination.

$$\begin{aligned} x + 2y + z &= 4 \\ 2x + 4y + 4z + 8t &= 2 \\ 4x + 8y + 6z + 8t &= 10 \end{aligned}$$

10. $A = \begin{pmatrix} 2 & 4 & 6 & 4 \\ 2 & 5 & 7 & 6 \\ 2 & 3 & 5 & 2 \end{pmatrix}$.Find the column space of A. Also find x such that $Ax = \begin{pmatrix} 4 \\ 3 \\ 5 \end{pmatrix}$.

11. If $AB = C$, columns of C are linear combinations of the columns of ______________, so $r(C) \leq r($______________$)$. Also rows of C are linear combinations of the rows of ______________, so $r(C) \leq r($______________$)$.

12. Find the projection matrix, P, of $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$. If we project $\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$ onto the column space of A, what is the projected vector?

13. Find the determinants of
$\begin{vmatrix} 1 & 3 & 2 \\ 2 & -1 & 5 \\ -2 & 4 & -4 \end{vmatrix}, \begin{vmatrix} -3 & 5 & 7 \\ -5 & -4 & 3 \\ 2 & 5 & 6 \end{vmatrix}$

3 Calculus

3.1 Preliminaries

3.1.1 Limit of a Function

For a function $y = f(x)$, if the value of $f(x)$ becomes infinitely close to L as the value of x becomes infinitely close to a, but $x \neq a$, we say that $f(x)$ has a limit L at an input a, and denote it as

$$\lim_{x \to a} f(x) = L$$

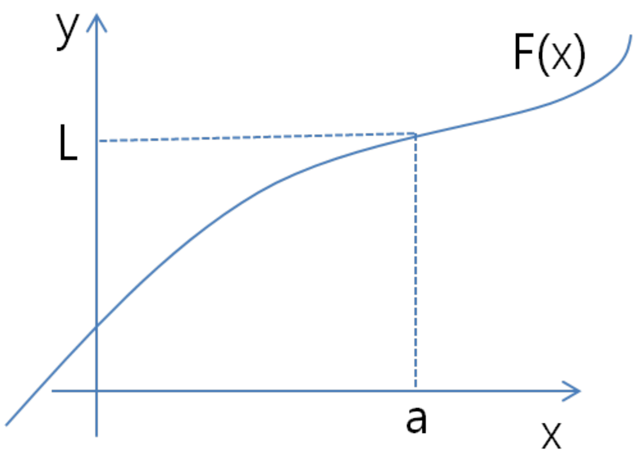


Figure 14: Limit of a Function

If $f(x)$ becomes infinitely large (small) as x becomes infinitely close to a, we say that $f(x)$ diverge to positive (negative) infinity.

$$\text{positive infinity divergence : } \lim_{x \to a} f(x) = \infty$$

$$\text{negative infinity divergence : } \lim_{x \to a} f(x) = -\infty$$

If $f(x)$ becomes infinitely close to a as x diverges to infinity, then we denote it as

$$\lim_{x \to \infty} f(x) = a$$

$$\lim_{x \to -\infty} f(x) = a$$

If $f(x)$ diverges to infinity as x diverge to infinity, we denote it as

$$\lim_{x \to \infty} f(x) = \infty \,, \quad \lim_{x \to \infty} f(x) = -\infty$$

$$\lim_{x \to -\infty} f(x) = \infty \,, \quad \lim_{x \to -\infty} f(x) = -\infty$$

If $f(x)$ becomes infinitely close to α as x becomes infinitely close to a but a little larger (smaller) than a, α is the right(left) limit of $f(x)$ and we denote it as

$$\lim_{x \to a^-} f(x) = \alpha \,, \quad \lim_{x \to a^+} f(x) = \alpha$$

Definition 24 *If the left limit and the right limit at an input a are equal, we say there is a limit at an input a.*

Example 10

$$f(x) = \begin{cases} x < 0, & x + 1 \\ x \geq 0, & x - 1 \end{cases}$$

$$\lim_{x\to 0^-} f(x) = 1\,, \quad \lim_{x\to 0^+} f(x) = -1$$

There is no limit at input 0.

Example 11 *Find limits of the following*

1. $\lim_{x\to 1} \frac{(x-1)(x^2+3x+7)}{(x-1)}$

2. $\lim_{x\to 1} \frac{x+1}{(x^2+5x+10)}$

Definition 25 1. *A function $f(x)$ is continuous at $x = a$ if and only if*

   (a) *At $x = a$, $f(x)$ is definable. $\Leftrightarrow$ $f(a)$ exists.*

   (b) $\lim_{\to a} f(x)$ *exists* $\Leftrightarrow \lim_{x\to a^-} f(x) = \lim_{x\to a^+} f(x)$

   (c) $\lim_{\to a} f(x) = f(a) \Leftrightarrow (a) = (b)$

2. *Function $f(x)$ is said to be continuous on an interval I if and only if $f(x)$ is continuous at each point $x \in I$.*

3.1.2 Exponential Function

Definition 26 ***Euler's Number and Logarithmic Function***

- $e = \lim_{x\to 0}(1+x)^{\frac{1}{x}} = 2.71828\cdots$
- *If $y = e^x$, $x = \ln y$, and if $y = a^x$, $x = \log_a y$.*

Proposition 28 1. *Laws of Exponentials*

- $x^a x^b = x^{a+b}$
- $\dfrac{x^a}{x^b} = x^{a-b}$
- $(xy)^a = x^a y^a$
- $(\dfrac{x}{y})^a = \dfrac{x^a}{y^a}$
- $x^{-a} = \frac{1}{x^a}$
- $x^0 = 1$
- $x^{\frac{m}{n}} = \sqrt[n]{x^m}$
- $(x^a)^b = x^{ab}$

2. *Laws of Log Functions*

- $\log xy = \log x + \log y$
- $\log \dfrac{x}{y} = \log x - \log y$
- $\log x^y = y \cdot \log x$
- $\log 1 = 0$
- $\log_e x = \ln x$
- $\log_y x = \dfrac{\log_a x}{\log_a y} = \dfrac{\ln x}{\ln y}$
- $\ln e = 1$
- $x^y = e^{y \cdot \ln x}$

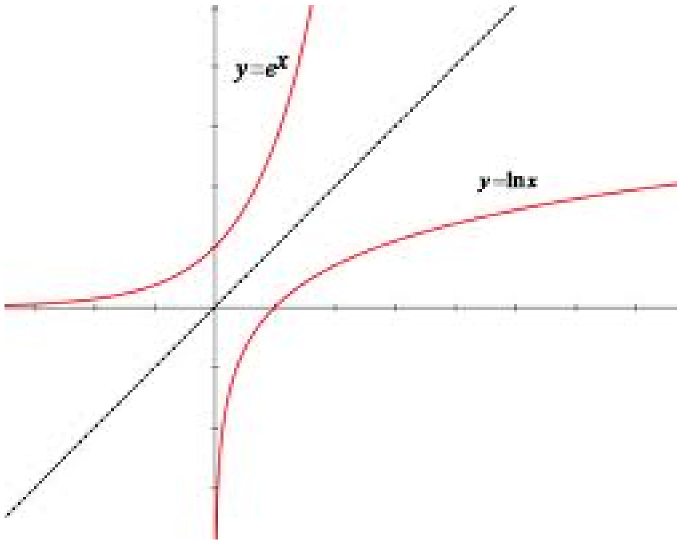


Figure 15: Log and Exponential Functions

3.2 Differentiation

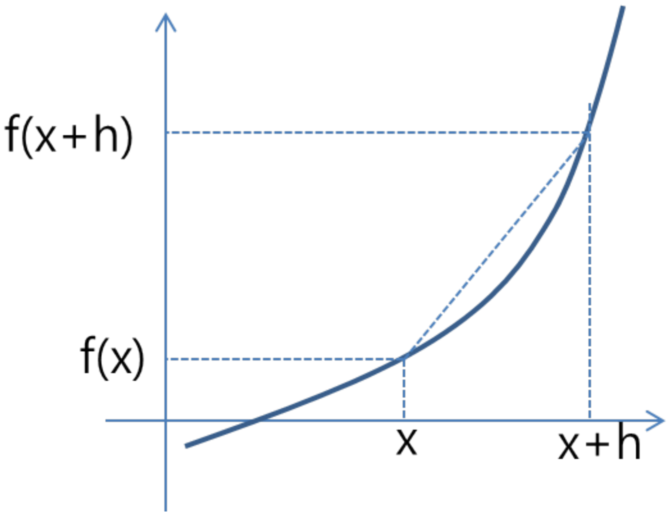


Figure 16: Rate of Change of a Function at x

As x changes to $x+h$, $f(x)$ also changes to $f(x+h)$. The rate of change is

$$\frac{f(x+h)-f(x)}{h}$$

The rate of change at x is

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}$$

We call this the derivative of $f(x)$ with respect to x and denote it as $f'(x)$ or $\frac{\mathrm{d}f(x)}{\mathrm{d}x}$.

Definition 27 *The derivative of $f(x)$ with respect to x is*

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

Proposition 29 *Laws of Derivatives*

1. $\frac{\mathrm{d}[f(x)+g(x)]}{\mathrm{d}x} = \frac{\mathrm{d}f(x)}{\mathrm{d}x} + \frac{dg(x)}{\mathrm{d}x}$
2. $\frac{dcf(x)}{\mathrm{d}x} = c\frac{\mathrm{d}f(x)}{\mathrm{d}x}$
3. $\frac{\mathrm{d}f(x)\cdot g(x)}{\mathrm{d}x} = f'(x)g(x) + f(x)g'(x)$
4. $\frac{d\frac{f(x)}{g(x)}}{\mathrm{d}x} = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}$
5. $\frac{\mathrm{d}f(g(x))}{\mathrm{d}x} = g'(x)f'(g(x))$
6. *If* $g(x) = f(x)^{-1}$, $g'(x) = \frac{1}{f'(x)}$

Proof. Proof (3)

$$\begin{aligned}
\frac{\mathrm{d}f(x)\cdot g(x)}{\mathrm{d}x} &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x+h)g(x)}{h} + \frac{f(x+h)g(x) - f(x)g(x)}{h} \\
&= \lim_{h\to 0} f(x+h)\frac{g(x+h) - g(x)}{h} + g(x)\frac{f(x+h) - f(x)}{h}
\end{aligned}$$

Proof (4)

$$\begin{aligned}
\frac{\mathrm{d}\frac{f(x)}{g(x)}}{\mathrm{d}x} &= \lim_{h\to 0} \frac{\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}}{h} \\
&= \lim_{h\to 0} \frac{\frac{f(x+h)g(x) - f(x)g(x+h)}{g(x+h)g(x)}}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x) - f(x)g(x+h)}{(g(x+h)g(x))\cdot h} \\
&= \lim_{h\to 0} \frac{g(x)(f(x+h) - f(x))}{h}\frac{1}{(g(x+h)g(x))} - \frac{f(x)(g(x+h) - g(x))}{h}\frac{1}{(g(x+h)g(x))}
\end{aligned}$$

Proof (5)

$$\frac{\mathrm{d}f(g(x))}{\mathrm{d}x} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{h} = \lim_{h\to 0} \frac{f(g(x+h)) - f(g(x))}{g(x+h) - g(x)}\frac{g(x+h) - g(x)}{h}$$

■

3.2.1 Left and Right Derivatives and Differentiability

Definition 28 *For $h > 0$, let $f'(x)^- = \lim_{h\to 0} \frac{f(x-h)-f(x)}{h}$ and $f'(x)^+ = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$. If $f'(x)^- = f'(x)^+$, then the left and right limits of the derivative (or the left and right derivatives) are the same: the derivative is continuous. In such a case, we call $f(x)$ differentiable.*

Example 12 *Are the functions below differentiable at $x = a$?*

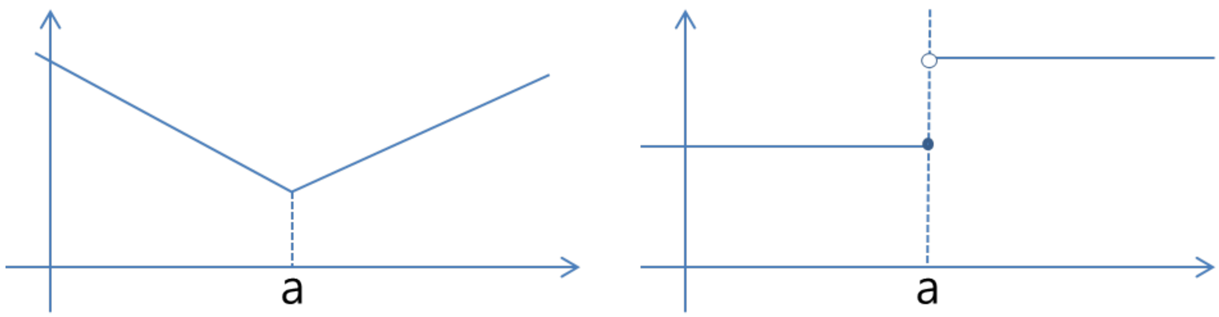


Figure 17: Differentiability

3.2.2 Derivatives of Various Functions

Proposition 30 *Derivatives of Various Functions*

1. $\frac{\mathrm{d}x^n}{\mathrm{d}x} = nx^{n-1}$
2. $\frac{\mathrm{d}a^x}{\mathrm{d}x} = a^x \ln a$
3. $\frac{\mathrm{d}e^x}{\mathrm{d}x} = e^x$

   Hint: By Definition 26, $\lim_{h\to 0} \frac{e^h-1}{h} = 1$.
4. $\frac{\mathrm{d}\ln x}{\mathrm{d}x} = \frac{1}{x}$

   Hint: Use the fact that $e^{f(x)} = e^{\ln x} = x$ and $f'(x)e^{f(x)} = 1$
5. $\frac{\mathrm{d}\sin x}{\mathrm{d}x} = \cos x$
6. $\frac{\mathrm{d}\cos x}{\mathrm{d}x} = -\sin x$
7. $\frac{\mathrm{d}\tan x}{\mathrm{d}x} = \frac{1}{\cos^2 x}$

   Hint: $\sin(a+b) = \sin a \cos b + \cos a \sin b$, $\cos(a+b) = \cos a \cos b - \sin a \sin b$. *Also note that* $\lim_{h\to 0} \frac{\sin h}{h} = 1$.

Example 13 *Find the derivatives of*

1. $x^4 + 2x^3 + \frac{1}{x} + 4$

2. $\tan x$

3. $2x(3x^3 + 1)^2$

4. $\frac{2x+1}{x^2+x-1}$

5. $y = \sqrt{x}$

3.2.3 Maxima, Minima, and Extrema of Functions

Definition 29 ***(Increasing and Decreasing Functions)***

For two arbitrary points $x > x' \in [x_1, x_2]$, if $f(x) \geq (\leq) f(x')$, then $f(x)$ is increasing (decreasing) in the interval $[x_1, x_2]$. If $f(x) > (<) f(x')$ holds without equalities, then $f(x)$ is strictly increasing (strictly decreasing).

Proposition 31 *If a differentiable function $f(x)$ satisfies $f'(x_0) > (<) 0$, then f is increasing (decreasing) in the neighborhood of x_0.*

Derivatives are the rate of change. If a rate of change is positive, it is increasing; if a rate of change is negative, it is decreasing.

Definition 30 *At a point x_0, if there exists an arbitrary real number $\delta > 0$ such that for all $x \in [x_0 - \delta, x_0 + \delta]$, $f(x_0) \geq (\leq) f(x)$, then $f(x)$ has a local maximum (minimum) at point $x = x_0$. Among the local maxima (minima), the largest local maximum (smallest local minimum) is called the global maximum (minimum)*

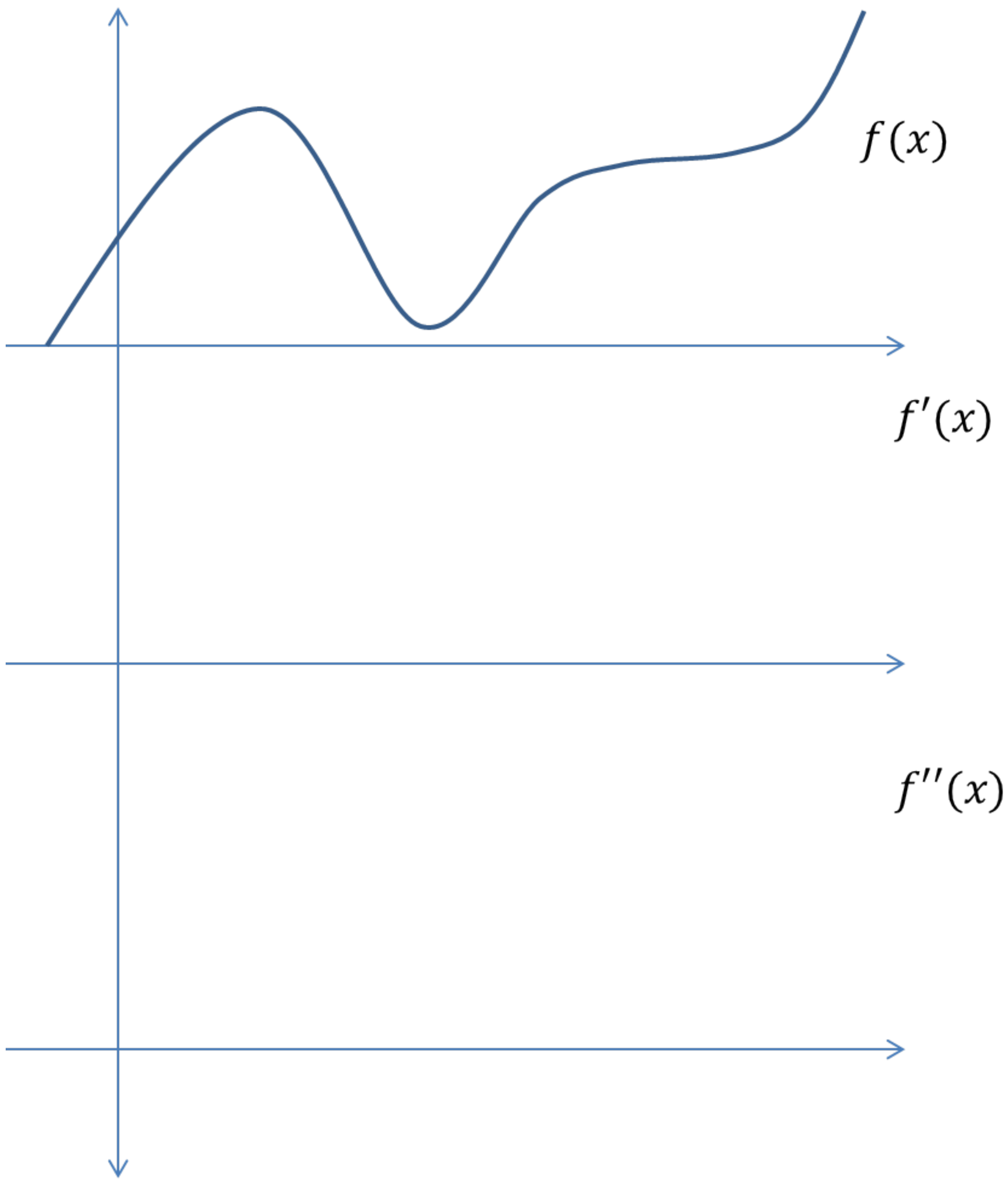


Figure 18: First Derivative, Second Derivative Functions

Proposition 32 *If a differentiable function $f(x)$ is a local minimum or maximum at point c, then $f'(c) = 0$.*

Proof. Suppose f is a local maximum at c.

For a sufficiently small $|h|$, $f(c+h) - f(c) \leq 0$

If $h < 0$, $\frac{f(c+h)-f(c)}{h} \geq 0$.

If $h > 0$, $\frac{f(c+h)-f(c)}{h} \leq 0$.

Thus, $f'(c) = \lim_{h \to 0} \frac{f(c+h)-f(c)}{h} = 0$. ■

Definition 31 ***Convex and Concave Functions***

- *For all $\lambda \in [0, 1]$, if $f[\lambda x_1 + (1-\lambda)x_2] \leq (<)[\lambda f(x_1) + (1-\lambda)f(x_2)]$, $f(x)$ is (strictly) convex in $[x_1, x_2]$.*

- *For all $\lambda \in [0, 1]$, if $f[\lambda x_1 + (1-\lambda)x_2] \geq (>)[\lambda f(x_1) + (1-\lambda)f(x_2)]$, $f(x)$ is (strictly) concave in $[x_1, x_2]$.*

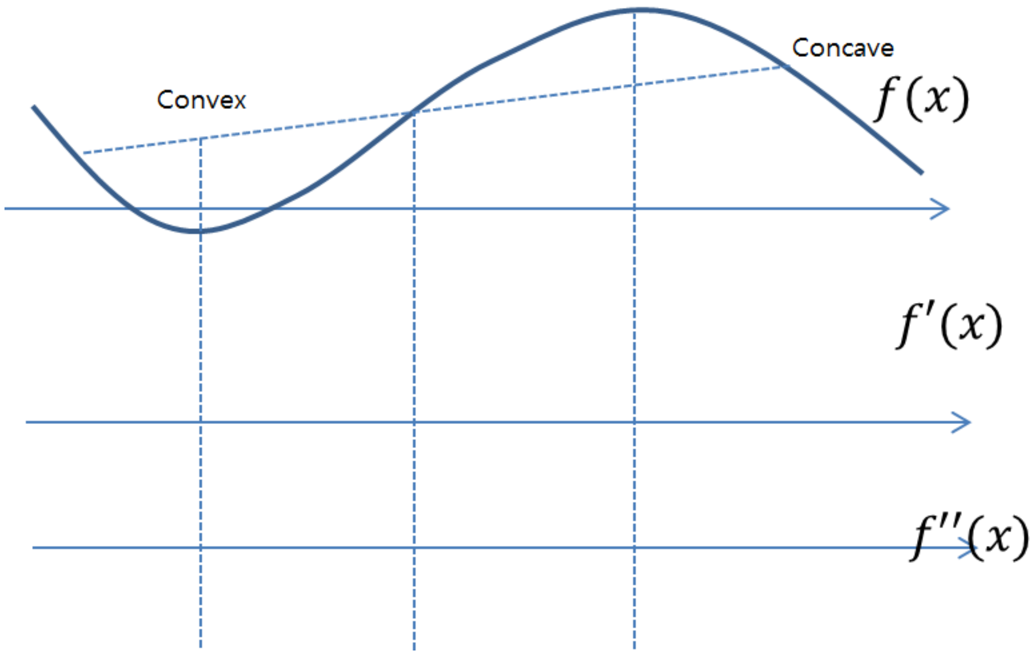


Figure 19: Convex and Concave Functions

Proposition 33 *If $f'(x_0) = 0$ and $f''(x_0) < 0 (> 0)$, then f is a local maximum (minimum) at $x = x_0$.*

3.2.4 Mean Value Theorem and L'Hopital's Rule

Proposition 34 ***Rolle's Theorem***

Suppose $f(x)$ is differentiable and continuous in $[a, b]$ and $f(a) = f(b)$. There always is a point or points $c \in (a, b)$ such that $f'(c) = 0$.

Proof. When $f(x)$ is defined on a closed interval such as $[a, b]$, it is well known that $f(x)$ has both minimum and maximum inside the interval. We will not prove this theorem, which is referred as the **Extreme Value Theorem**, but take it as given.

Suppose the maximum is obtained at $c \in (a, b)$. Then by definition, for a sufficiently small $|h|$, $f(c + h) - f(c) \leq 0$

If $h < 0$, $\frac{f(c+h)-f(c)}{h} \geq 0$.

If $h > 0$, $\frac{f(c+h)-f(c)}{h} \leq 0$.

Thus, $f'(c) = \lim_{h\to 0} \frac{f(c+h)-f(c)}{h} = 0$.

Similarly, if the minimum is obtained at $c \in (a, b)$, then by similar reasoning, $f'(c) = 0$.

Suppose the maximum and minimum are both obtained at $x = a = b$. Then $f(x)$ is constant so that $f'(x) = 0$ for all $x \in [a, b]$

■

Proposition 35 ***Mean Value Theorem***

Suppose $f(x)$ is differentiable and continuous in $[a, b]$. There always is a point or points $c \in (a, b)$ such that $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Proof. Let $g(x) = \frac{f(b)-f(a)}{b-a}x - f(x)$. Note that $g(a) = \frac{f(b)-f(a)}{b-a}a - f(a) = \frac{af(b)-bf(a)}{b-a} = g(b) = \frac{f(b)-f(a)}{b-a}b - f(b) = \frac{af(b)-bf(a)}{b-a}$. Thus, by Rolle's theorem, there is a $c \in (a, b)$ such that $g'(c) = 0$, where $g'(x) = \frac{f(b)-f(a)}{b-a} - f'(x)$. $g'(c) = \frac{f(b)-f(a)}{b-a} - f'(c) = 0$. Thus, $f'(c) = \frac{f(b)-f(a)}{b-a}$

■

Proposition 36 ***Cauchy's Mean-Value Theorem***

If $g(x)$ and $f(x)$ are differentiable and continuous in $[a, b]$, there always is a point or points $c \in (a, b)$ such that $\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}$.

Proof. Let $h(x) = f(x) - \frac{f(b)-f(a)}{g(b)-g(a)}g(x)$. Note that $h(a) = h(b)$. By Rolle's Theorem, there is a $c \in (a, b)$, such that $h'(c) = f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)}g'(c) = 0$. Note that

$$\frac{f'(c)}{g'(c)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

■

Proposition 37 ***L'Hopital's Rule***

When $x \to c \in [-\infty, \infty]$, if $f(x) \to 0$ and $g(x) \to 0$, or $f(x) \to \infty$ and $g(x) \to \infty$ and $g'(x) \neq 0$, then $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$.

Proof. This is only a sketch of the proof. A formal proof is out of the scope of this class.

1. As $x \to c$, $f(x) \to 0$ and $g(x) \to 0$

   First, redefine $f(c) = g(c) = 0$ so that $f(x)$ and $g(x)$ is continuous at $x = c$. Then there is $\epsilon \in (c, x)$, such that by Cauchy's Mean-Value Theorem,
   $$\frac{f'(\epsilon)}{g'(\epsilon)} = \frac{f(c) - f(x)}{g(c) - g(x)} = \frac{f(x)}{g(x)}$$
   As $x \to c$, $\epsilon \to c$. Thus, and
   $$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(\epsilon)}{g'(\epsilon)} = lim_{x\to c} \frac{f'(x)}{g'(x)}$$

2. As $x \to c$, $f(x) \to \infty$ and $g(x) \to \infty$

   Let $F(x) = \frac{1}{f(x)}$ and $G(x) = \frac{1}{g(x)}$. Then we know that as $x \to c$, $F(x) \to 0$ and $G(x) \to 0$. Thus,
   $$\lim_{x\to c} \frac{g(x)}{f(x)} = \lim_{x\to c} \frac{F(x)}{G(x)} = \lim_{x\to c} \frac{F'(x)}{G'(x)} = \lim_{x\to c} \frac{\frac{f'(x)}{f(x)^2}}{\frac{g'(x)}{g(x)^2}}$$
   By simplification, we have
   $$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$$

■

3.2.5 Multivariate Calculus

Let $z = f(x, y)$ and x, y are exogeneous (independent) variables and z is an endogeneous (dependent) variable. A partial derivative of a function of several variables is its derivative with respect to one of those variables, with the others held constant (as opposed to the total derivative, in which all variables are allowed to vary). In the previous example, z is a function of both x and y. To know the effect of x on z alone, we assume that y is a constant and differentiate z with respect to x. We will denote the partial derivative as $\frac{\partial z}{\partial x}$ or f_x.

Definition 32 ***Partial Derivatives***

A partial derivative of a function of several variables is its derivative with respect to one of those variables, with the others held constant (as opposed to the total derivative, in which all variables are allowed to vary).

Proposition 38 ***Rules of Partial Derivatives***

- *If* $z = g(x,y) \cdot h(x,y)$, $\frac{\partial z}{\partial x} = g(x,y) \cdot \frac{\partial h}{\partial x} + h(x,y) \cdot \frac{\partial g}{\partial x}$
- *If* $z = \frac{g(x,y)}{h(x,y)}$, $\frac{\partial z}{\partial x} = \frac{h(x,y) \cdot \frac{\partial g}{\partial x} - g(x,y) \cdot \frac{\partial h}{\partial x}}{[h(x,y)]^2}$
- *If* $z = [g(x,y)]^n$, $\frac{\partial z}{\partial x} = n[g(x,y)]^{n-1} \cdot \frac{\partial g}{\partial x}$
- *(Young's Theorem)* $f_{xy} = f_{yx}$

Suppose $z = f(x,y)$ where $y = g(x)$. Then

$$\frac{\mathrm{d}z}{\mathrm{d}x} = \frac{\partial f}{\partial x} + \frac{\partial y}{\partial x}\frac{\partial f}{\partial y}$$

Example 14 *Find the partial derivatives of* z *with respect to* x *and* y*.* $z = x + y$ *,* $z = xy$*,* $z = x^2y^3 + 2x + 4y$*,* $z = 3x^2(8x - 7y)$*,* $z = \frac{9y^3}{x-y}$

3.2.6 Implicit Function Theorem

A derivative ($\frac{\mathrm{d}y}{\mathrm{d}x}$) is a rate of change and we may interpret $\mathrm{d}y$ as the change of y and $\mathrm{d}x$ as the change of x. Thus, the rate of change $f'(x)$ is the change of y divided by the change of x and $\mathrm{d}y = f'(x)\mathrm{d}x$ holds.

The change of F caused by the change of y is $\mathrm{d}y \cdot F_y$ and the change of F caused by the change of x is $\mathrm{d}x \cdot F_x$.

$$\mathrm{d}F = \mathrm{d}x \cdot F_x + \mathrm{d}y \cdot F_y$$

However, since $F = 0$ in any case (thus, F does not change), $\mathrm{d}F = 0$ holds.

$$\mathrm{d}x \cdot F_x + \mathrm{d}y \cdot F_y = 0$$

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{F_x}{F_y}$$

Proposition 39 ***Implicit Function Theorem***

If $F(x,y) = 0$, $\frac{dy}{dx} = -\frac{F_x}{F_y}$.

Example 15 *Find* $\frac{\mathrm{d}y}{\mathrm{d}x}$ *of the following.*

$y - 6x + 7 = 0$, $3x^2 + 2xy + 4y^3 = 7$, $7x^2 + 2xy^2 + 9y^4 = 9$, $x^2y^3 + z^2 + xyz = 3$, $3x^2y^3 + xz^2y^2 + y^3zx^4 + y^2z = 3$

3.2.7 Extrema of Bivariate Functions

Suppose $z = f(x, y)$ (a bivariate function). Let us find the local extrema of z. By the implicit function theorem, we know that

$$\mathrm{d}z = \mathrm{d}x \cdot f_x + \mathrm{d}y \cdot f_y$$

The change of z is caused by the change of x, which is $f_x\mathrm{d}x$, and the change of y, which is $f_y\mathrm{d}y$.

Now let us find the change of z at a specific point (x_0, y_0). For z to be an extremum at $(x, y) = (x_0, y_0)$,

$$\mathrm{d}z = \mathrm{d}x \cdot f_x + \mathrm{d}y \cdot f_y = 0$$

Thus, for any directionsto which x, y move (thus for all combinations of $\mathrm{d}x$ and $\mathrm{d}y$), the change of z ($\mathrm{d}z = 0$) should be 0. The sufficient conditions are $f_x = 0$ and $f_y = 0$. Thus, at point (x_0, y_0), if $f_x(x_0) = 0$ and $f_y(y_0) = 0$, then it is an extremum.

Then let us find whether the point (x_0, y_0), where $f_x(x_0) = 0$ and $f_y(y_0) = 0$, is a local minimum, local maximum, or something else.

For (x_0, y_0) to be a local maximum, $\mathrm{d}z$ should be decreasing, which means $\mathrm{d}(\mathrm{d}z) = d^2z < 0$.

$$\mathrm{d}^2z = \mathrm{d}(\mathrm{d}z) = \frac{\partial \mathrm{d}z}{\partial x}\mathrm{d}x + \frac{\partial \mathrm{d}z}{\partial y}\mathrm{d}y$$

Since $\frac{\partial \mathrm{d}z}{\partial x} = f_{xx}\mathrm{d}x + f_{yx}\mathrm{d}y$, $\frac{\partial \mathrm{d}z}{\partial y} = f_{xy}\mathrm{d}x + f_{yy}\mathrm{d}y$ and $f_{xy} = f_{yx}$

$$\mathrm{d}^2z = f_{xx}\mathrm{d}x^2 + 2f_{xy}\mathrm{d}x\mathrm{d}y + f_{yy}\mathrm{d}y^2$$

For (x_0, y_0) to be a local maximum, for $\mathrm{d}x$ and $\mathrm{d}y$ sufficiently near 0, $\mathrm{d}^2z < 0$ should hold. Thus, $f_{xx} < 0$ and $f_{xy}^2 - f_{xx}fyy < 0$ should hold because of the following Lemma.

Lemma 1 *For all x and y, the sufficient and requisite conditions for $ax^2 + 2bxy + cy^2 < 0$ to hold are $a < 0$ and $b^2 - ac < 0$.*

Also if $f_{xx} > 0$ and $f_{xy}^2 - f_{xx}f_{yy} < 0$, it is a local minimum. If $f_{xy}^2 - f_{xx}f_{yy} \geq 0$, it is a saddle point.

Proposition 40 ***Extrema of a Bivariate Function***

Let $z = f(x, y)$.

1. *For a point to be a local minimum,*

$$f_x = f_y = 0 \ \textit{ and } \ f_{xx} > 0 \ \textit{ and } \ f_{xy}^2 - f_{xx}f_{yy} < 0$$

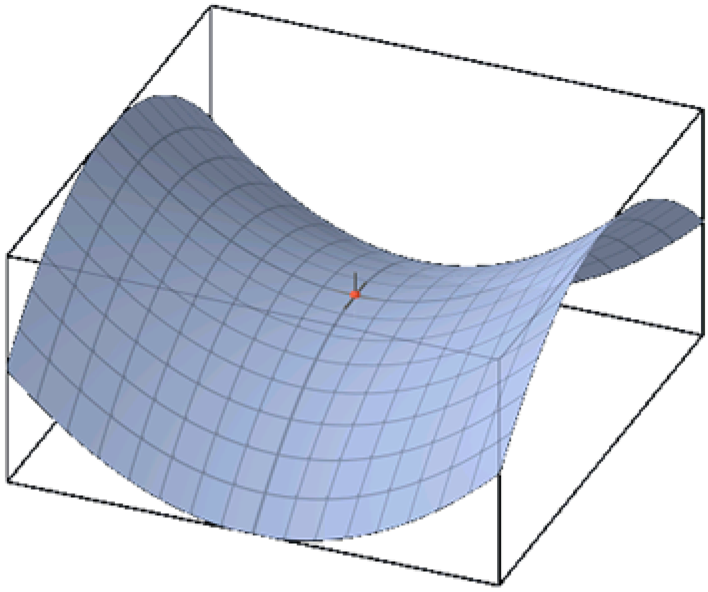


Figure 20: Saddle Point

2. *For a point to be a local maximum,*

$$f_x = f_y = 0 \ \text{ and } \ f_{xx} < 0 \ \text{ and } \ f_{xy}^2 - f_{xx}f_{yy} < 0$$

3. *For a point to be a saddle point,*

$$f_x = f_y = 0 \ \text{ and } \ f_{xy}^2 - f_{xx}f_{yy} \geq 0$$

3.2.8 Extrema of a Multivariate Function

Now let us consider extrema of a multivariate function such as $f(x_1, \cdots, x_n)$. Before we begin our anlalysis, we first define the gradient of a function.

$$(a,b)\cdot(1,\frac{\mathrm{d}y}{\mathrm{d}x}) = 0 \Leftrightarrow a + b\frac{\mathrm{d}y}{\mathrm{d}x} = 0 \Leftrightarrow -\frac{a}{b} = \frac{\mathrm{d}y}{\mathrm{d}x} = -\frac{\frac{\partial f(x,y)}{\partial x}}{\frac{\partial f(x,y)}{\partial y}}$$

The last equality holds because of the implicit function theorem.

$$(a,b) = (\frac{\partial f(x,y)}{\partial x}, \frac{\partial f(x,y)}{\partial y})$$

We call (a,b) the *gradient* of f.

Definition 33 *The* ***Gradient*** *of a function* $f(x_1, x_2, \cdots, x_n)$ *is*

$$\nabla f(x_1, x_2, \cdots, x_n) = \Big(\frac{f(x_1, x_2, \cdots, x_n)}{\partial x_1}, \cdots, \frac{f(x_1, x_2, \cdots, x_n)}{\partial x_n}\Big)$$

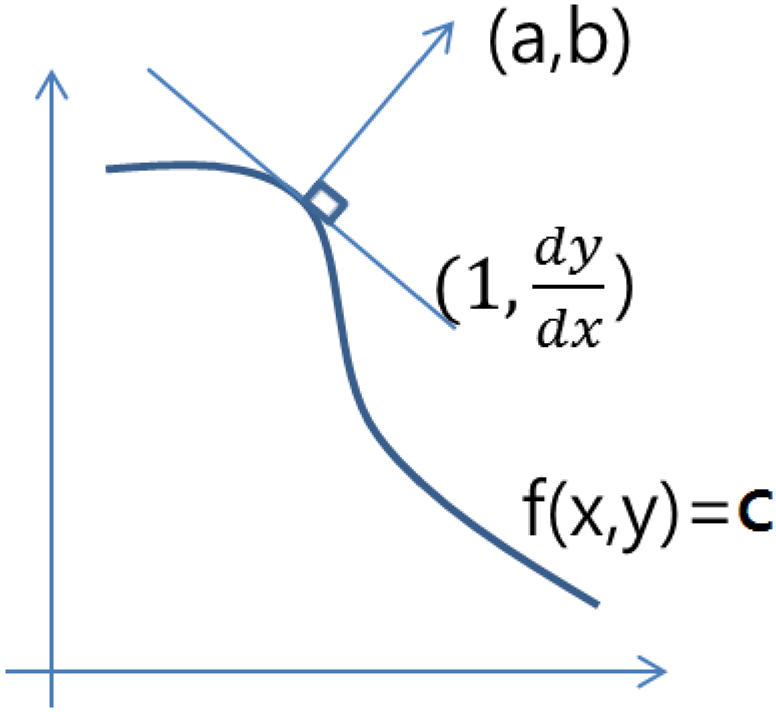


Figure 21: Vector Orthogonal to the Targent Line

It is trivial that the minimum or maximum of a function $(x_1^*, \cdots, x_n^*)$ should satisfy $f_{x_1}(x_1^*, \cdots, x_n^*) = \cdots = f_{x_n}(x_1^*, \cdots, x_n^*) = 0$, or equivalently $\nabla f(x_1^*, \cdots, x_n^*) = 0$. Let us formally prove this.

Proposition 41 *A local minimum* $x^* = (x_1^*, \cdots, x_n^*)$ *should satisfy*

$$\nabla f(x_1^*, \cdots, x_n^*) = 0$$

Proof. Let $g(t) = f(x^* + tz)$ for an arbitrary vector z. $x^* + tz$ is simply a vector that is different from x^* for non-zero t and z; g coincides with some value of f. For f to reach a local minimum at x^*, g must also reach a local minimum at $t = 0$. Thus, $g'(0) = 0$ Using the chain rule, $g'(t) = \sum_i^n \frac{\partial f(x^*+tz)}{\partial x_i + tz_i} z_i$. Thus,

$$g'(0) = \sum_i^n \frac{\partial f(x^*)}{\partial x_i} z_i = \nabla f(x^*) z = 0$$

Since this should hold for all z,

$$\nabla f(x^*) = 0$$

■

Now let us consider the second order condition. Is x^* a minimum or maximum? Before we answer this question, we first define a Hessian matrix of a function $f(x_1, \cdots, x_n)$.

Definition 34 *A Hessian matrix of f, H is*

$$H = \begin{pmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \cdots & \vdots \\ f_{n1} & \cdots & f_{nn} \end{pmatrix}$$

Note that a Hessian matrix is a symmetric matrix because of Young's Theorem.

Then we derive the second order condition.

Proposition 42 *If x^* that satisfies $\nabla f(x^*) = 0$ is*

1. *a local minimum, then $H(x^*)$ is positive definite.*

2. *a local maximum, then $H(x^*)$ is negative definite.*

Proof. We again define $g(t) = f(x^* + tz)$ for an arbitrary vector z. $x^* + tz$ is simply a vector that is different from x^* for non-zero t and z; g coincides with some value of f.

$$g''(t) = \sum_i^n \sum_j^n \frac{\partial^2 f(x + tz)}{\partial x_i \partial x_j} z_i z_j$$

and

$$g''(0) = \sum_i^n \sum_j^n \frac{\partial^2 f(x)}{\partial x_i \partial x_j} z_i z_j = z^T H(x) z$$

For f to reach a local minimum at x^*, g must also reach a local minimum at $t = 0$. Thus, $g''(0) > 0$ Thus, $H(x^*)$ must be positive definite.

(2) The second order condition for a local maximum can be proved similarly. ■

3.3 Matrix Differentiation

Let $A_{m\times n}$ be a $m \times n$ matrix and x be a $n \times 1$ column vector and y a $m \times 1$ column vector.

Definition 35

$$\begin{pmatrix} \frac{\partial y_1}{\partial x_1} & \frac{\partial y_1}{\partial x_2} & \cdots & \frac{\partial y_1}{\partial x_n} \\ \frac{\partial y_2}{\partial x_1} & \frac{\partial y_2}{\partial x_2} & \cdots & \frac{\partial y_2}{\partial x_n} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial y_m}{\partial x_1} & \frac{\partial y_m}{\partial x_2} & \cdots & \frac{\partial y_m}{\partial x_n} \end{pmatrix}$$

Let $Ax = y$ then

$$\frac{\partial y}{\partial x} = A$$

Furthermore

$$\frac{\partial y}{\partial z} = A \frac{\partial x}{\partial z}$$

Let a scalar $\alpha = y^T A x$ then

$$\frac{\partial \alpha}{\partial x} = y^T A$$

$$\frac{\partial \alpha}{\partial y} = x^T A^T$$

$$\frac{\partial \alpha}{\partial z} = x^T A^T \frac{\partial y}{\partial z} + y^T A \frac{\partial x}{\partial z}$$

Let a scalar $\alpha = x^T A x$ where A is $n \times n$ matrix then

$$\frac{\partial \alpha}{\partial x} = x^T (A + A^T)$$

Let a scalar $\alpha = y^T x$ then

$$\frac{\partial \alpha}{\partial z} = x^T \frac{\partial y}{\partial z} + y^T \frac{\partial x}{\partial z}$$

Problems

1. Prove Proposition 29. 6.

2. Prove Proposition 30

3. Prove Lemma 1.

4. Find derivatives of the following functions.

   a) $f(x) = \frac{1}{(3x^4+2)^{20}}$

   b) $f(x) = \frac{2}{\sqrt{5x+x^2}}$

   c) $f(x) = \sqrt{3x^2 + x + 1}$

   d) $f(x) = 4x^5 + x^3 + 3x$

   e) $f(x) = \frac{\ln x}{x} + \frac{x}{\ln x}$

   f) $f(x) = 2e^x(x^3 + 7)$

   g) $f(x) = \frac{e^x}{1+e^x}$

   h) $f(x) = e^{x^{\frac{1}{2}x}}$

   i) $f(x) = x^{x+1}$

   j) $f(x) = x^{\ln x}$

5. Find extrema of the following functions.

   a) $f(x) = (x^2 - 4x - 12)^2$

   b) $f(x) = (x + 7)^3(2x + 5)^2$

   c) $f(x) = x^3 + 3x^2 + 3x + 4$

d) $f(x) = \frac{x^2}{x^2+1}$

e) $f(x) = \frac{x^2}{x-2}$

6. Find minima and maxima of the following functions.

a) $f(x) = x^2 - 2x + 1$, given that $0 \leq x \leq 3$

b) $f(x) = x^3 - 3x^2$, given that $-1 \leq x \leq 1$

c) $f(x) = 4x + \frac{16}{x^2}$, given that $1 \leq x \leq 4$

d) $f(x) = \begin{cases} 100 \leq x \leq 200, & 1200x - 2x^2 \\ 200 < x \leq 300, & 800x \end{cases}$

7. If the price of an orange is 150 and the total cost to produce x oranges is $TC(x) = \frac{1}{6}x^3 - 8x^2 + 142x + 200$, what is the profit maximizing number of orange production?

8. Is $f(x) = 3x^4 - 4x^3$ convex or concave?

9. Represent $2^x, 2^{-|x|}, (\frac{1}{2})^{2x}$ in terms of Euler's number.

10. Suppose we invested \$ 3000 and the interest rate is r,

a) What is the value of the investment after t period of time?

b) When does the value double?

11. Find the partial derivatives of the following functions.

a) $f(x, y) = x^5y + 3x^4y^3 - x + 7$

b) $f(x, y) = \ln(2xy - x^3)$

c) $f(x, y) = x^2e^{\sqrt{xy}}$

d) $f(x, y) = \frac{e^{x-y}}{2x+y}$

e) $f(x, y) = x^5y + 3x^4y^3 - x + 7$

f) $f(x, y) = (3xy^6 - y + 1)^7$

g) $f(x, y) = e^{3x2y}$

h) $f(x, y) = \frac{e^{x-y}}{2x-z}$

12. Find extremas of the following functions.

a) $f(x, y) = 3y^2 - x^2$

b) $f(x, y) = xy - \frac{x}{y}$

c) $f(x, y) = 2x^3 + 3x^2 - 12x + y^2 - y + 5$

d) $f(x,y) = 3x - x^2 - y^2$

e) $f(x,y) = 3x^2 - y^2 - 3xy - 15x + 8y$

f) $f(x,y) = e^{x^2+y^2}$

g) $f(x,y) = x^2 + y\ln y + y$

h) $f(x,y,z) = 2x^2 + y^2 + 3z^2 + 2xz + 3yz - x - y - z$

13. Find $\frac{\mathrm{d}y}{\mathrm{d}x}$.

a) $\sqrt{x} + \sqrt{y} = x$

b) $\frac{1}{x^2} + \frac{1}{y^2} = 3$

c) $x^2 e^y = 4x$

d) $\ln(x^2 - y^2) = 7$

14. Find $\frac{\mathrm{d}z}{\mathrm{d}t}$

a) $z = x^2 - 8xy - y^3$ given that $x = 3t$, $y = 1 - t$

b) $z = 3u + 2vt$ given that $u = 2t^2$, $v = t + 1$

15. If you invest x in A and y in B, the annual profit is $1.5x + 2.2y - \frac{1}{100}x^2 - \frac{2}{100}y^2$. How much should you invest in A and B?

16. The demand for apples is $x = \frac{2500}{P_o^2}$ and the demand for grapes is $y = \frac{3000}{P_g^2}$. P_i is the price respectively. The total cost of production is $C = 5x + 3y + 200$. How many oranges and grapes should you produce?

17. Determine whether the following functions are convex or concave.

a) $f(x,y) = 3x - x^2 - y^2$

b) $f(x,y) = 3x^2 - y^2 - 3xy - 15x + 8y$

c) $f(x,y) = x^2 + y\ln y + y$

3.4 Integral

3.4.1 Indefinite Integral

Let $F'(x) = f(x)$. Integration is finding $F(x)$, the reverse of differentiation. Thus, the integral of $f(x)$ is $F(x)$.

Definition 36 *An indefinite integral is the opposite operation of differentiation: Let $F'(x) = f(x)$, then the integral of $f(x)$ is $F(x)$ and we denote it as*

$$F(x) = \int f(x)\mathrm{d}x$$

$\int$ is the integral sign and $f(x)$ is the integrand. For example, the derivative of x^2 is $2x$. Thus, integral of $2x$ might be x^2, but also $x^2 + 1$. Thus, in general, we will write as $x^2 + C$, where C represents a constant.

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = f(x) \;\;\Rightarrow\;\; F(x) = \int f(x)\mathrm{d}x + C$$

Proposition 43 *The following rules of integrals hold.*

1. $\int x^n \mathrm{d}x = \frac{1}{n+1}x^{n+1} + C$
2. $\int e^x \mathrm{d}x = e^x + C$
3. $\int \frac{1}{x}\mathrm{d}x = \ln x + C$
4. $\int f'(x)g'(f(x))\mathrm{d}x = g(f(x)) + C$

   Note. Let $t = f(x)$. $\mathrm{d}t = f'(x)\mathrm{d}x$, thus, $\int f'(x)g'(f(x))\mathrm{d}x = \int g'(t)\mathrm{d}t = g(t) = g(f(x))$. We call this integration by substitution.
5. $\int f(x) + g(x)\mathrm{d}x = \int f(x)\mathrm{d}x + \int g(x)\mathrm{d}x$
6. *Let k be a constant.* $\int kf(x)\mathrm{d}x = k\int f(x)\mathrm{d}x$.
7. $\int f'(x)g(x)\mathrm{d}x = f(x)g(x) - \int f(x)g'(x)\mathrm{d}x$. *We call this integration by parts.*

3.4.2 Definite Integral: Riemann Integral

Definition 37 *Riemann Sum*

Let $y = f(x)$ be a real valued function on $[a, b]$. Let Δ be an arbitrary partition on the interval $[a, b]$ such that it divides the interval as $a = x_0 < x_1 < \cdots < x_n = b$. Let $\Delta x_i = x_i - x_{i-1}$ and $|\Delta| = \max\{\Delta x_i : i = 1, 2, \cdots, n\}$. The Riemann sum of $f(x)$ over $[a, b]$ with partition Δ is

$$\mathbf{R}(\Delta, f) \equiv \sum_{i=1}^{n} f(\epsilon_i)\Delta x_i$$

where, $\epsilon_i \in [x_{i-1}, x_i]$.

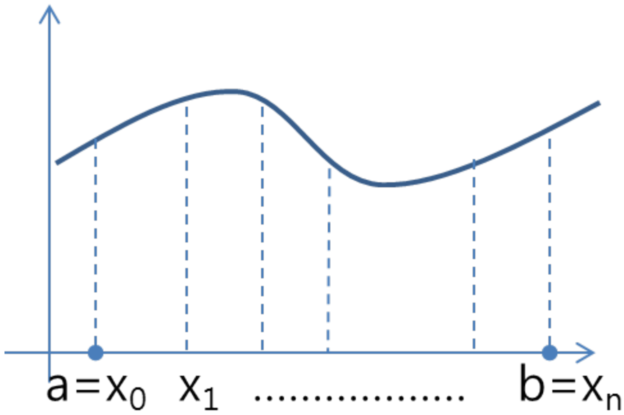


Figure 22: Riemann Sum

Definition 38 *Upper Riemann Sum and Lower Riemann Sum*

On the interval $[x_{i-1}, x_i]$*, let* x_i', x_i'' *be a point of* x *that minimizes and maximizes* $f(x)$*, respectively. The Lower Riemann Sum is*

$$\mathbf{L}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i')\Delta x_i$$

and the Upper Riemann Sum is

$$\mathbf{U}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i'')\Delta x_i$$

Definition 39 *When* $|\Delta| = \max\{\Delta x_i : i = 1, 2, \cdots, n\} \to 0$*, if* $\mathbf{L}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i')\Delta x_i = \mathbf{U}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i'')\Delta x_i$ *holds, we call* $y = f(x)$ *is* ***integrable.*** *We also denote the limit of* $\mathbf{L}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i')\Delta x_i = \mathbf{U}(\Delta, f) \equiv \sum_{i=1}^{n} f(x_i'')\Delta x_i$ *as*

$$\int_a^b f(x)\mathrm{d}x = \lim_{|\Delta| \to 0} \sum_{i=1}^{n} f(\epsilon_i)\Delta x_i$$

Note that $\int_a^b f(x)\mathrm{d}x$ is the area of $f(x)$ in the interval $[a, b]$.

Proposition 44 ***Fundamental Theorem of Calculus***

Let $F'(x) = f(x)$*. Then*

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$$

Proof. Let $a = x_0 < x_1 < \cdots < x_n = b$. Then,

$$F(b) - F(a) = F(x_n) - F(x_{n-1}) + F(x_{n-1}) - \cdots - F(x_1) + F(x_1) - F(x_0)$$

By the Mean Value Theorem, there exists $\epsilon_i \in (x_{i-1}, x_i)$ such that

$$F(x_i) - F(x_{i-1}) = F'(\epsilon_i)(x_i - x_{i-1}) = f(\epsilon_i)(x_i - x_{i-1})$$

Thus,

$$\begin{aligned} F(b) - F(a) &= F(x_n) - F(x_{n-1}) + F(x_{n-1}) - \cdots - F(x_1) + F(x_1) - F(x_0) \\ &= f(\epsilon_n)(x_n - x_{n-1}) + \cdots + f(\epsilon_1)(x_1 - x_0) = \sum_{i=1}^{n} f(\epsilon_i)\Delta x_i \end{aligned}$$

When $|\Delta| = \max\{\Delta x_i : i = 1, 2, \cdots, n\} \to 0$,

$$F(b) - F(a) = \lim_{|\Delta| \to 0} \sum_{i=1}^{n} f(\epsilon_i)\Delta x_i = \int_a^b f(x)\mathrm{d}x$$

■

Corollary 3 *The following rules hold.*

1. $\int_a^b f(x)\mathrm{d}x = -\int_b^a f(x)\mathrm{d}x$
2. $\int_a^a f(x)\mathrm{d}x = 0$
3. $\int_a^c f(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_b^c f(x)\mathrm{d}x$
4. $\int_a^b k \cdot f(x)\mathrm{d}x = k\int_a^b f(x)\mathrm{d}x$
5. $\int_a^b f(x) + g(x)\mathrm{d}x = \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x$
6. $\int_a^b f'(x)g(x)\mathrm{d}x = (f(b)g(b) - f(a)g(a)) - \int_a^b f(x)g'(x)\mathrm{d}x$ *This is also called the integral by parts.*

Example 16 *Solve* $\int_0^2 xe^{-2x}\mathrm{d}x$ *and* $\int_1^3 \frac{\ln x}{x}\mathrm{d}x$.

Definition 40 *An improper integral is:*

$$\int_a^\infty f(x)\mathrm{d}x \equiv \lim_{b \to \infty} \int_a^b f(x)\mathrm{d}x$$

Example 17 *Solve*

1. $\int_0^{-\infty} e^x \mathrm{d}x$
2. *Let* $f(x) = \begin{cases} x \geq 2, & 4x^{-3} \\ x < 2, & 0 \end{cases}$

   Find $\int_{-\infty}^{\infty} f(x)\mathrm{d}x$

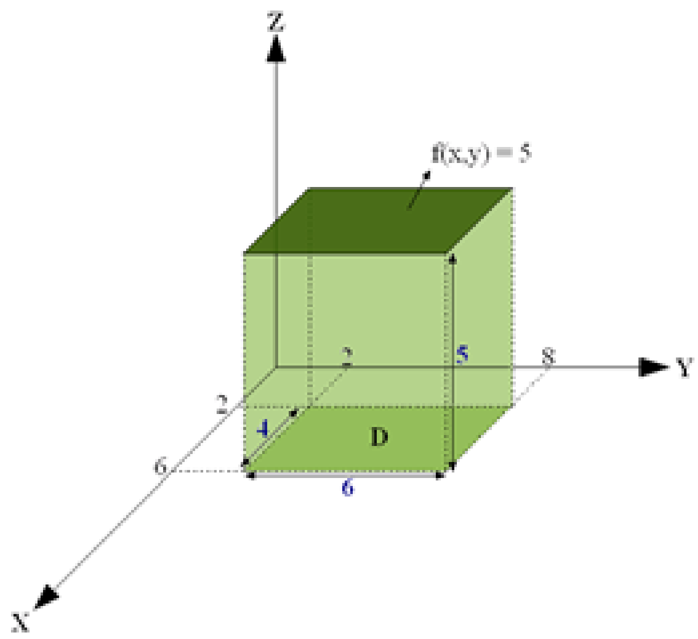


Figure 23: A Rectangular Cuboid

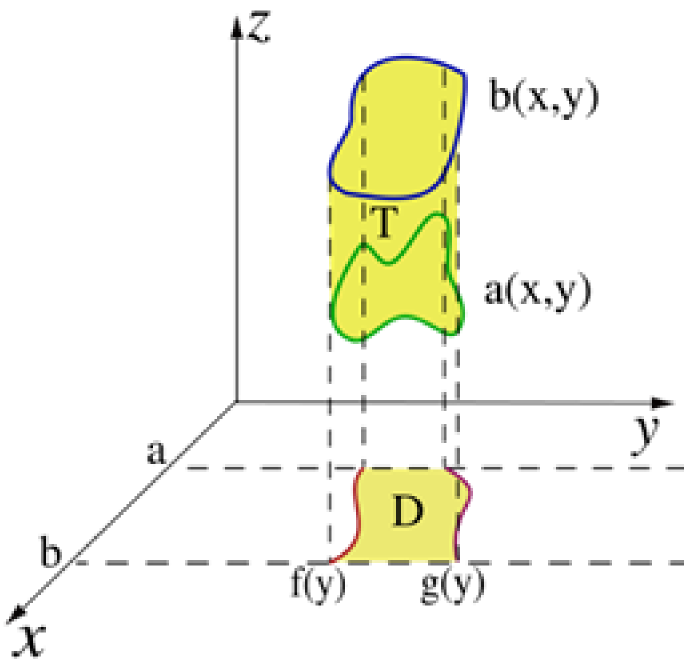


Figure 24: Figure by a Function

3.4.3 Double Integral

We know that the volume of the cuboid is $2 \times 4 \times 6 = 48$. What about the volume of the figure in Figure 24, which is defined by $a(x, y)$ and D? First we finely partition D and name each sub-partition as ΔD_i. We also choose a point $(x_i^*, y_i^*) \in \Delta D_i$. Then the approximation of the volume is

$$\sum_{i=1}^{n} A(x_i^*, y_i^*)|\Delta D_i|$$

where $A(x, y) = b(x, y) - a(x, y)$ and $|\Delta D_i|$ is the area of ΔD_i.

Let $P = \max(|\Delta D_i|)$ Then the volume of the figure is

$$\lim_{P\to 0} \sum_i A(x_i^*, y_i^*)|\Delta D_i|$$

Note that the previous equation is the same as the Riemann Sum.

Example 18 *Find the volume of the figure surrounded by a rectangle* $[0, 1] \times [1, 3]$ *and* x^2y.

$$\int_0^1 \int_1^3 x^2 y \mathrm{d}y \mathrm{d}x = \int_0^1 x^2 \left[\frac{y^2}{2}\right]_1^3 \mathrm{d}x = \int_0^1 4x^2 \mathrm{d}x = \frac{4}{3}$$

Example 19

$$\begin{aligned} \int_0^1 \int_1^{\frac{\pi}{2}} e^y + \cos x \mathrm{d}x \mathrm{d}y &= \int_0^1 \left[xe^y + \sin x\right]_0^{\frac{\pi}{2}} \mathrm{d}y \\ &= \int_0^1 \frac{\pi}{2} e^y + 1 \mathrm{d}y = \left[\frac{\pi}{2} e^y + y\right]_0^1 = \frac{\pi(e-1)}{2} + 1 \end{aligned}$$

Example 20 *Find the volume of a figure surrounded by* $y = x^2$, $y = x$ *and* $z = xy$.

$$\int_0^1 \int_{x^2}^{x} xy \mathrm{d}y \mathrm{d}x = \int_0^1 \left[\frac{xy^2}{2}\right]_{x^2}^{x} \mathrm{d}x = \int_0^1 \frac{x^3}{2} - \frac{x^5}{2} \mathrm{d}x = \left[\frac{x^4}{8} - \frac{x^6}{12}\right]_0^1 = \frac{1}{24}$$

Problems

1. Prove Proposition 43.7.

2. Do the following integrals.

(1) $\int x^3 \mathrm{d}x$	(2) $\int 3^x \mathrm{d}x$	(3) $\int \frac{2}{x^7} \mathrm{d}x$
(4) $\int 2e^{-x} \mathrm{d}x$	(5) $\int 3p^2 - p + 1 \mathrm{d}p$	(6) $\int 2\sqrt{t} - \frac{2}{\sqrt{t}} \mathrm{d}t$
(7) $\int x^{\frac{2}{5}} + 8x^{-\frac{9}{5}} \mathrm{d}x$	(8) $\int 12x^5 - 2x^3 - x + \frac{1}{x} \mathrm{d}x$	(9) $\int \frac{x^4 - 2x^2 + 3}{x} \mathrm{d}x$
(10) $\int x(x^2 - 4)^8 \mathrm{d}x$	(11) $\int 12x^3(x^4 + 1)^5 \mathrm{d}x$	(12) $\int (t^3 - t)^{-\frac{1}{8}}(6t^2 - 2) \mathrm{d}t$
(13) $\int (6 - y)^{-\frac{2}{3}} \mathrm{d}y$	(14) $\int \frac{\ln x}{x} \mathrm{d}x$	

(16) $\int \frac{1}{x\ln 2x}\mathrm{d}x$ (17) $\int \frac{e^{\sqrt{x}}}{\sqrt{x}}\mathrm{d}x$ (18) $\int \frac{e^x-e^{-x}}{e^x+e^{-x}}\mathrm{d}x$
(19) $\int 2x^2(5+x)^{\frac{3}{2}}\mathrm{d}x$ (20) $\int (x^2+1)\sqrt{x-2}\mathrm{d}x$ (21) $\int \frac{x-1}{\sqrt{2x-4}}\mathrm{d}x$
(22) $\int (x-2)^2\ln x\mathrm{d}x$ (23) $\int 2x^3(\ln x)^2\mathrm{d}x$ (24) $\int xe^{2x+1}\mathrm{d}x$
(25) $\int xe^{x^2}\mathrm{d}x$ (26) $\int x\sqrt{x+1}\mathrm{d}x$

3. Suppose $f''(x) = x^2 - 10x + 1$, $f'(3) = -23$ and $f(0) = -2$. Find $f(x)$.

4. Suppose the marginal cost is $MC(Q) = \frac{1}{50}Q^2 - 2Q + 107$ and the fixed cost is 2,000.

   (a) Find the total cost.

   (b) What is the increase in the cost of producing the 31st product. Compare it with $MC(30)$.

5. The marginal revenue is $MR(Q) = 150 - Q$, marginal cost is $MC(Q) = \frac{1}{10}Q^2 - 4Q + 110$, and the total cost of producing 30 products is 3,000.

   (a) Find the fixed cost.

   (b) What is the profit function?

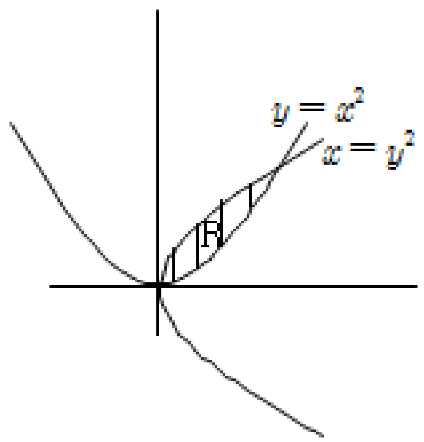
   (c) What is the profit maximizing Q?

6. Do the following integrals.

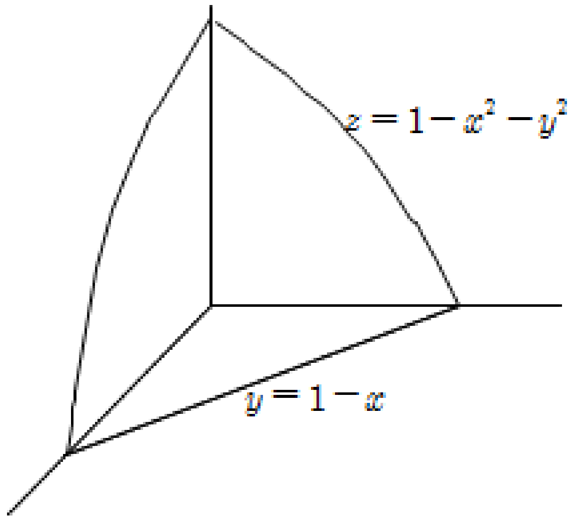
(1) $\int_1^5 \frac{2}{u} + \frac{u}{2}\mathrm{d}u$ (2) $\int_8^{27} 4x^{\frac{1}{3}}\mathrm{d}x$ (3) $\int_2^8 3e^{\frac{x}{2}}\mathrm{d}x$
(4) $\int_0^1 e^x + e^{-x}\mathrm{d}x$ (5) $\int_1^3 \frac{\ln x}{x}\mathrm{d}x$ (6) $\int_2^6 x\sqrt{2x^2-8}\mathrm{d}x$
(7) $\int_1^5 \frac{3}{(2x-1)^{\frac{3}{2}}}\mathrm{d}x$ (8) $\int_0^2 xe^{-2x}\mathrm{d}x$ (9) $\int_1^4 \sqrt{x}\ln x\mathrm{d}x$
(10) $\int_1^\infty \frac{1}{x^4}\mathrm{d}x$ (11) $\int_{-\infty}^{-2} \frac{1}{x^2}\mathrm{d}x$

7. Find the volume ofthe followidlng figures.

   (a) The figure surrounded by R in the following figure and xy^2.



   (b) The figure surrounded by $z = 1 - x^2 - y^2$, $x = 0$, $y = 0$, $z = 0$ and $x + y = 1$.



(c) Find the area surrounded by $x = 2y^2$ and $y^2 = x - 4$.

(d) Find the volume of the cylinder, the radius of which is 1, cut by the xy-plane and $x + y + z = 1$

(e) The figure surrounded by $z = x^2$, $z = 2x^2$, $y = x^2$ and $y = 8 - x^2$.

4 Optimization with Constraints

4.1 Optimization with one constraint

The following equations allow you to control x, y to maximize $f(x, y)$. However, x, y should satisfy $g(x, y) \leq c$. We call $f(x, y)$, the objective function, $g(x, y) \leq c$ the constraints, and x, y the control variables.

$$\begin{aligned} &\operatorname*{Max}_{x,y} \quad z = f(x, y) \\ &\text{s.t.} \quad\quad g(x, y) \leq c \end{aligned}$$

The x^*, y^* that solve this optimization problem will have a geometric quality as in Figure 25. (x, y) that satisfy $g(x, y) \leq c$ is the shaded area. If $f(x, y)$ increases as (x, y) moves to the north-east direction, it will maximize at the point z^*, where the gradients of f and g are linearly dependent, which means that $\nabla f(x^*, y^*) = \lambda \nabla g(x^*, y^*)$. Thus,

$$\frac{\partial f(x^*, y^*)}{\partial x} = \lambda \frac{\partial g(x^*, y^*)}{\partial x}, \frac{\partial f(x^*, y^*)}{\partial y} = \lambda \frac{\partial g(x^*, y^*)}{\partial y} \text{ and } g(x^*, y^*) = c$$

where λ is a constant.

The previous conditions can be rewritten as the following. Let $\mathcal{L} = f(x, y) - \lambda(g(x, y) - c)$. Then the previous three conditions are

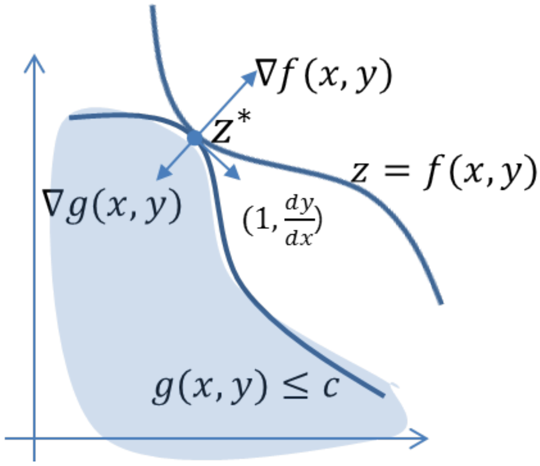


Figure 25: Optimization Problem

$$\frac{\partial \mathcal{L}}{\partial x} = 0, \quad \frac{\partial \mathcal{L}}{\partial y} = 0, \quad \frac{\partial \mathcal{L}}{\partial \lambda} = 0$$

$\mathcal{L}$ is the Lagrangian function, λ is the Lagrangian multiplier, and the optimization by finding the first order condition of the Lagrangian function is called the Lagrangian method.

4.2 Optimization with many constraints

Now let us solve an optimization problem with m constraints and n control variable, which can be written as

$$\begin{aligned} \underset{x_1,\cdots,x_n}{\text{Max}} \quad & f(x_1,\cdots,x_n) \\ \text{s.t.} \quad & g_1(x_1,\cdots,x_n) \leq \quad c_1 \\ & \vdots \\ & g_m(x_1,\cdots,x_n) \leq \quad c_m \end{aligned}$$

The control variables are $x_1,\cdots,x_n$ and the constraints are $g_1(x_1,\cdots,x_n) \leq c_1,\cdots,g_m(x_1,\cdots,x_n) \leq c_m$. Let $z^* = (x_1^*,\cdots,x_m^*)$ be the solution to the problem as it is shown in Figure 26.

From Figure 26, we can easily see that $\nabla f(x_1^*,\cdots,x_n^*)$ is a linear combination of $\nabla g_1(x_1^*,\cdots,x_n^*)$, $\cdots,\nabla g_m(x_1^*,\cdots,x_n^*)$. Thus,

$$\nabla f(x_1^*,\cdots,x_n^*) = \lambda_1 \nabla g_1(x_1^*,\cdots,x_n^*) + \cdots + \lambda_m \nabla g_m(x_1^*,\cdots,x_n^*)$$

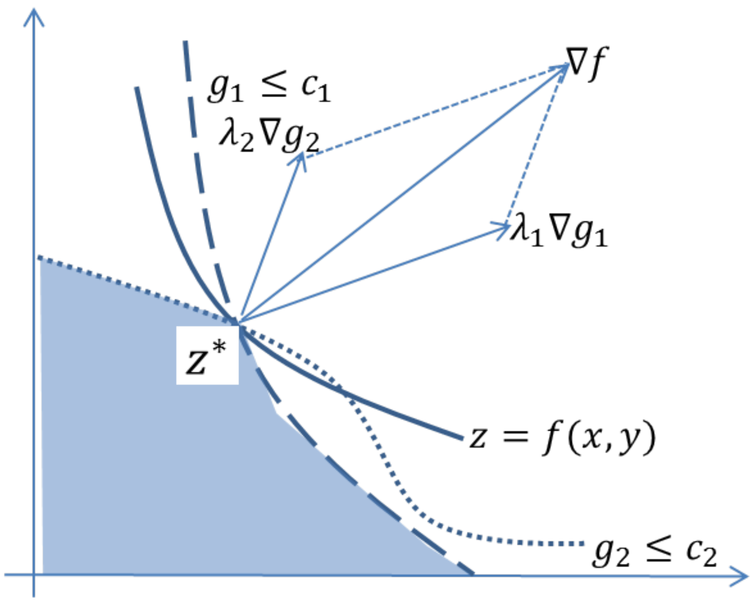


Figure 26: Optimization with Many Constraints

This can be generalized as

$$\begin{aligned}
\frac{\partial f(x_1^*, \cdots, x_n^*)}{\partial x_1} &= \lambda_1 \frac{\partial g_1(x_1^*, \cdots, x_n^*)}{\partial x_1} + \cdots + \lambda_m \frac{\partial g_m(x_1^*, \cdots, x_n^*)}{\partial x_1} \\
&\vdots \\
\frac{\partial f(x_1^*, \cdots, x_n^*)}{\partial x_n} &= \lambda_1 \frac{\partial g_1(x_1^*, \cdots, x_n^*)}{\partial x_n} + \cdots + \lambda_m \frac{\partial g_m(x_1^*, \cdots, x_n^*)}{\partial x_n} \\
g_1(x_1^*, \cdots, x_n^*) &= c_1 \\
&\vdots \\
g_m(x_1^*, \cdots, x_n^*) &= c_m
\end{aligned}$$

This also can be solved by the Lagrangian method. First, we define the Lagrangian function as

$$\mathcal{L} = f(x_1, \cdots, x_n) - \lambda_1 \cdot (g_1(x_1, \cdots, x_n) - c_1) - \cdots - \lambda_m \cdot (g_m(x_1, \cdots, x_n) - c_m)$$

The problem solving conditions are

$$\frac{\partial \mathcal{L}}{\partial x_1} = 0, \cdots, \frac{\partial \mathcal{L}}{\partial x_n} = 0$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = 0, \cdots, \frac{\partial \mathcal{L}}{\partial \lambda_m} = 0$$

Now, let us focus on the optimization problem as in Figure 27, where there is a non-binding constraint. The constraint that $g_3 \leq c_3$ has no effect on the optimization problem, i.e., it is non-binding. This means that ∇f is a linear combination of only ∇g_1 and ∇g_2, but not ∇g_3. Therefore, from $\nabla f(x_1^*, \cdots, x_n^*) = \lambda_1 \nabla g_1(x_1^*, \cdots, x_n^*) + \cdots + \lambda_m \nabla g_m(x_1^*, \cdots, x_n^*)$, we know that $\lambda_3 = 0$. Also note the direction of $\lambda_i \nabla g_i$, where $\lambda \neq 0$, is north-east. Thus, we know that $\lambda_i \geq 0$.

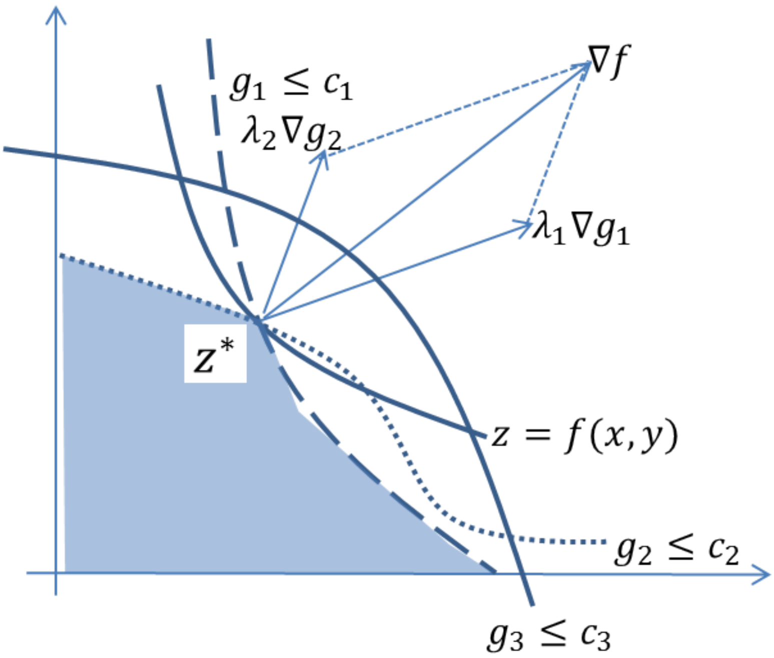


Figure 27: Optimization with Non-Binding Constraints

Proposition 45 ***Kuhn-Tucker Theorem***

$$\begin{aligned} \underset{x_1, \cdots, x_n}{\mathit{Max}} \quad & f(x_1, \cdots, x_n) \\ \text{s.t.} \quad & g_1(x_1, \cdots, x_n) \leq c_1 \\ & \vdots \\ & g_m(x_1, \cdots, x_n) \leq c_m \end{aligned}$$

The solution of the previous optimization problem $(x_1^*, \cdots, x_n^*)$ *has the following properties.*

$$\mathcal{L} = f(x_1, \cdots, x_n) - \lambda_1 \cdot (g_1(x_1, \cdots, x_n) - c_1) - \cdots - \lambda_m \cdot (g_m(x_1, \cdots, x_n) - c_m)$$

then,

$$\left.\frac{\partial \mathcal{L}}{\partial x_1}\right|_{(x_1^*, \cdots, x_n^*)} = 0, \cdots, \left.\frac{\partial \mathcal{L}}{\partial x_n}\right|_{(x_1^*, \cdots, x_n^*)} = 0$$

$$\left.\frac{\partial \mathcal{L}}{\partial \lambda_1}\right|_{(x_1^*, \cdots, x_n^*)} \times \lambda_1 = 0, \cdots, \left.\frac{\partial \mathcal{L}}{\partial \lambda_m}\right|_{(x_1^*, \cdots, x_n^*)} \times \lambda_m = 0$$

and,

$$\lambda_i \geq 0$$

If $\lambda_i = 0$*, the* i*-th constraint is non-binding.*

Finally, we will study the effect of change of parameter (not control variable) on the optimization problem. Parameter is a variable in the optimization problem that cannot be controlled. First, we define Value Function.

Definition 41 ***Value Function***

$$(x_1^*, \cdots, x_n^*) \in \underset{x_1,\cdots,x_n}{\textit{argmax}}\, f(x_1, \cdots, x_n, r)$$

$$\begin{aligned} \textit{s.t.}\quad & g_1(x_1, \cdots, x_n, r) \leq c_1 \\ & \vdots \\ & g_m(x_1, \cdots, x_n, r) \leq c_m \end{aligned}$$

A value function is

$$v(r) = f(x_1^*, \cdots, x_n^*, r)$$

Now, let us analyze the effect of change of r on the value function. The first order derivative of the value function with respect to r is

$$\frac{\mathrm{d}v(r)}{\mathrm{d}r} = \frac{\partial f(x_1^*, \cdots, x_n^*, r)}{\partial r} + \sum_{i=1}^{n} \frac{\partial f(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)} \times \frac{\mathrm{d}x_i^*}{\mathrm{d}r}$$

By Proposition 45,

$$\frac{\partial f(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)} = \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)}$$

Thus,

$$\begin{aligned} \frac{\mathrm{d}v(r)}{\mathrm{d}r} &= \frac{\partial f(x_1^*, \cdots, x_n^*, r)}{\partial r} + \sum_{i=1}^{n} \frac{\mathrm{d}x_i^*}{\mathrm{d}r} \times \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*|r)} \\ &= \frac{\partial f(x_1^*, \cdots, x_n^*, r)}{\partial r} + \sum_{j=1}^{m} \lambda_j \sum_{i=1}^{n} \frac{\mathrm{d}x_i^*}{\mathrm{d}r} \frac{\partial g_j(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)} \end{aligned} \tag{10}$$

If, for j that is $\lambda_j = 0$, $\lambda_j \sum_{i=1}^{n} \frac{\mathrm{d}x_i^*}{\mathrm{d}r} \frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)} = 0 = \lambda_j \frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial r}$ holds. If $\lambda_j \neq 0$, by Proposition 45., $g_j = c_j$. By the implicit function theorem,

$$\frac{\partial g_j(x_1, \cdots, x_n, r)}{\partial r} + \sum_{i=1}^{n} \frac{\mathrm{d}x_i^*}{\mathrm{d}r} \frac{\partial g_j(x_1, \cdots, x_n, r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)} = 0$$

In summary we have

$$\frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial r} = -\sum_{i=1}^{n}\frac{\mathrm{d}x_i^*}{\mathrm{d}r}\frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial x_i}\Big|_{(x_1^*,\cdots,x_n^*)}$$

Insert this result into Equation (10), then

$$\frac{\mathrm{d}v(r)}{\mathrm{d}r} = \frac{\partial f(x_1^*,\cdots,x_n^*,r)}{\partial r} - \sum_{j=1}^{m}\lambda_j\frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial r}\Big|_{(x_1^*,\cdots,x_n^*)}$$

Proposition 46 ***Envelope Theorem***

Let $v(r)$ be a value function for an optimization problem, where $f(x_1,\cdots,x_n)$ is the objective function, $g_1,\cdots,g_m$ are constraints, and $\lambda_1,\cdots,\lambda_m$ are Lagrangian multipliers.

$$\frac{\mathrm{d}v(r)}{\mathrm{d}r} = \frac{\partial f(x_1^*,\cdots,x_n^*,r)}{\partial r} - \sum_{j=1}^{m}\lambda_j\frac{\partial g_j(x_1,\cdots,x_n,r)}{\partial r}\Big|_{(x_1^*,\cdots,x_n^*)}$$

The implication of the Envelope Theorem is that if we want to evaluate the effect of a change of parameter on the optimization problem, we only need to evaluate the direct effect of the change of parameter ($\frac{\partial f}{\partial r} - \sum_j \lambda_j \frac{\partial g}{\partial r}\Big|_{(x_1^*,\cdots,x_n^*)}$), but do not need to take account of the indirect effect of the change of parameter because of the first order condition ($\frac{\partial f}{\partial x} = \lambda_j \frac{\partial g}{\partial r}\Big|_{(x_1^*,\cdots,x_n^*)}$ so that $\frac{\mathrm{d}x}{\mathrm{d}r}\frac{\partial f}{\partial x} = \sum_j \lambda_j \frac{\partial g}{\partial r}$).

Now we study the economic meaning of the Lagrangian multiplier. Each Lagrangian multiplier corresponds to each constraint one-to-one. The Lagrangian multiplier denotes how much the change of the corresponding constraint affects the value function. The larger the multiplier, the larger is the effect of the change of the constraint on the value function.

For example, let us consider a profit maximization problem of a firm. We want to study how the change of the cost (c) may affect the maximized profit (Π^*). Thus, $\frac{\mathrm{d}\Pi^*}{\mathrm{d}c}$. Then, according to the Envelope Theorem, we only need to evaluate the direct effect of the change of the cost on the profit ($\frac{\partial \Pi^*}{\partial c}$), but need not consider the effect on the behavior of the firm (q^*) ($\frac{\mathrm{d}q^*}{\mathrm{d}c}$), and need not study the effect of the change of behavior on the profit.

$$\frac{\mathrm{d}\Pi^*}{\mathrm{d}c} = \frac{\partial \Pi^*}{\partial c} + \frac{\partial \Pi^*}{\partial q^*}\frac{\mathrm{d}q^*}{\mathrm{d}c} = \frac{\partial \Pi^*}{\partial c}$$

Example 21 ***Profit Maximization***

We want to find the profit maximizing quantity (q) of a firm. The price of the product is p and total cost is $c\cdot q^2$. The profit is $\Pi = p\cdot q - c\cdot q^2$. By solving the profit maximization problem, we know that $q^* = \frac{p}{2c}$, $\Pi^* = \frac{p^2}{4c}$. The effect of the rise of the cost (c) on profit is

$$\frac{\mathrm{d}\Pi^*}{\mathrm{d}c} = \frac{\partial \Pi^*}{\partial c} + \frac{\mathrm{d}q^*}{\mathrm{d}c}\frac{\partial \Pi^*}{\partial q^*}$$

By the first order condition, $\frac{\partial \Pi^*}{\partial q^*} = 0$*, thus,*

$$\frac{\mathrm{d}\Pi^*}{\mathrm{d}c} = \frac{\partial \Pi^*}{\partial c} = -q^{*2}$$

Note that $\Pi^* = \frac{p^2}{4c}$*,*

$$\frac{\mathrm{d}\Pi^*}{\mathrm{d}c} = -\frac{p^2}{4c^2} = -q^{*2}$$

Example 22 ***Utility Maximization***

An individual consumes goods X and Y, and her utility from consumption is $U(X, Y)$, where $U_X > 0$, $U_Y > 0$, $U_{XX} < 0$, $U_{YY} < 0$. The prices of goods are P_X, P_Y, and she has w amount of budget.

$$\begin{aligned} \underset{X,Y}{\mathit{Max}} \quad & U(X, Y) \\ \mathit{s.t.} \quad & P_X X + P_Y Y \leq w \end{aligned}$$

Let the value function be

$$v(w, P_X, P_Y) = U(X^*, Y^*)$$

Using the Envelope Theorem,

$$\begin{aligned} \frac{\mathrm{d}v(w, P_X, P_Y)}{\mathrm{d}w} &= \frac{\partial U(X^*, Y^*)}{\partial w} - \lambda \frac{\partial P_X X^* + P_Y Y^* - w}{\partial w} = \lambda \\ \frac{\mathrm{d}v(w, P_X, P_Y)}{\mathrm{d}P_X} &= \frac{\partial U(X^*, Y^*)}{\partial P_X} - \lambda \frac{\partial P_X X^* + P_Y Y^* - w}{\partial P_X} = -\lambda X^* \\ \frac{\mathrm{d}v(w, P_X, P_Y)}{\mathrm{d}P_Y} &= \frac{\partial U(X^*, Y^*)}{\partial P_Y} - \lambda \frac{\partial P_X X^* + P_Y Y^* - w}{\partial P_Y} = -\lambda Y^* \end{aligned}$$

λ denotes the effect of a change of budget on the utility, i.e., if the budget increases by 1, the utility increases by λ. It also shows the effect of a price change. If the price of X increases by 1, it is equivalent to the budget decreasing by X^*. Thus, with respect to the effect on the utility, increasing the price by 1 is equivalent to decreasing the budget by X^*. Thus, the utility decreases by $-\lambda X^*$. Thus, λ is the price one pays in terms of utility, and it is often referred to as the *Shadow Price*.

4.3 Second Order Condition for Constrained Optimization

Suppose we are faced withsuch a maximization problem as

$$\begin{aligned}
\underset{x_1,\cdots,x_n}{\text{Max}} \quad & f(x_1,\cdots,x_n) \\
\text{s.t.} \quad & g^1(x_1,\cdots,x_n) \leq \quad c_1 \\
& \vdots \\
& g^m(x_1,\cdots,x_n) \leq \quad c_m
\end{aligned}$$

Let us further assume that all constraints are binding, such that $g_i = c_i$. Then solving this maximization problem may be equivalent to solving an unconstrained maximization problem such as

$$\mathbf{L} = \underset{x_1,\cdots,x_n}{\text{Max}} \quad f(x_1,\cdots,x_n) - \lambda_1(g^1 - c_1) - \cdots - \lambda_m(g^m - c_m)$$

First, note that the first order conditions of λ_is will guarantee that the constraints are met. Second, the first order conditions of x_is will find the maximum value of f while the constraints are met. Thus, the Lagrangian method is actually changing the constrained maximization problem to an unconstrained maximization problem.

Thus, the second order condition should also hold. First, note that the Hessian matrix for $\mathbf{L}$ is

$$H = \begin{pmatrix} \mathbf{L}_{11} & \cdots & \mathbf{L}_{1n} \\ \mathbf{L}_{n1} & \cdots & \mathbf{L}_{nn} \end{pmatrix}$$

Since (x^*, λ^*) solves maximization (minimization) it should be negative (positive) definite.

Lemma 2 *By Proposition 42, we see that*

1. *a local minimum, then H is positive definite.*

2. *a local maximum, then H is negative definite.*

However, alternatively, we have Bordered Hessian Matrix of Lagragian which is

$$B = \begin{pmatrix} 0 & \cdots & 0 & g_1^1 & \cdots & g_n^1 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & g_1^m & \cdots & g_n^m \\ g_1^1 & \cdots & g_1^m & \mathbf{L}_{11} & \cdots & \mathbf{L}_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ g_n^1 & \cdots & g_n^n & \mathbf{L}_{n1} & \cdots & \mathbf{L}_{nn} \end{pmatrix}$$

With little calculation, we can easily see that

Proposition 47 *If the x^* that satisfies the Kuhn-Tucker Theorem is*

1. *a local maximum, the bordered-preserving principal minor of order k (determinant of the left submatrix) has the sign $(-1)^k$ for $k = 2, 3, ..., n$*

2. *a local minimum, then all the determinants of submatrices of B is negative.*

Note that The bordered-preserving principal minor of order 1 is the determinant of

$$B = \begin{pmatrix} 0 & \cdots & 0 & g_1^1 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & g_1^m \\ g_1^1 & \cdots & g_1^m & \mathbf{L}_{11} \end{pmatrix}$$

4.4 Economic Application of Maximization Method

4.4.1 Perfect Competition Economy with Two Goods

Market Demand

There are two goods, X and Y. An individual consumer's utility function is $X^\alpha Y^{1-\alpha}$ and the prices are P_X and P_Y. The total mass of all consumer is 1. This means that if there are n number of consumers and each consumer consumes 1 unit, then we normalize the number of consumers to 1 and the amount of consumption to $\frac{1}{n}$. If $n \to \infty$, the mass of each individual converges to 0, which is precisely the assumption of perfect competition. Each consumer has a different preference α, where α is uniformly distributed over $[0, 1]$ ($\alpha \sim U[0, 1]$). We also assume that each consumer's budget is w. The utility maximization problem of a consumer with preference α is

$$\begin{aligned} \underset{X,Y}{\text{Max}} \quad & X^\alpha Y^{1-\alpha} \\ \text{s.t.} \quad & P_X X + P_Y Y \leq w \end{aligned}$$

$$\mathcal{L} = X^\alpha Y^{1-\alpha} - \lambda(P_X X + P_Y Y - w)$$

The first order conditions are, $\lambda P_X = \alpha \left(\frac{Y}{X}\right)^{1-\alpha}$, $\lambda P_Y = (1-\alpha)\left(\frac{X}{Y}\right)^\alpha$ and $P_X X + P_Y Y = w$. The individual's demand functions are

$$\begin{aligned} X &= \frac{\alpha}{P_X} w \\ Y &= \frac{1-\alpha}{P_Y} w \end{aligned}$$

Thus, the market demands are

$$\begin{aligned} X^D &= \int_0^1 \frac{\alpha}{P_X} w \mathrm{d}\alpha = \frac{w}{2P_X} \\ Y^D &= \int \frac{1-\alpha}{P_Y} w \mathrm{d}\alpha = \frac{w}{2P_Y} \end{aligned}$$

Market Supply

Now let us focus on the market supply. Each firm can either produce X or Y. We also normalize the mass of firms to 1, thus an individual produces 1 unit, and then the amount of production is normalized to $\frac{1}{n}$. An individual firm's production functions are $X = \sqrt{C_X}$ and $Y = \sqrt{C_Y}$, where C_X and C_Y are the total cost. Thus, the total costs of producing each good are $C_X = X^2$, $C_Y = Y^2$. Let λ be the mass of the firms that decide to produce X and $1-\lambda$ be the mass of the firms that decided to produce Y. The profit maximization problem of the X producing firm is

$$\underset{X}{\text{Max}} \ P_X X - X^2$$

and the first order condition is $X = \frac{P_X}{2}$, and similarly we have $Y = \frac{P_Y}{2}$. Thus, the market supplies of X and Y are

$$X^S = \lambda \frac{P_X}{2}$$

$$Y^S = (1-\lambda)\frac{P_Y}{2}$$

Market Clearance

For the market to clear, the demand and supply should meet and the profit of X-producing firm should be equal to the profit of the Y-producing firm (We call this the 'no arbitrage condition'). If supply and demand meet, $\frac{w}{2P_X} = X^D = X^S = \lambda \frac{P_X}{2}$ and $\frac{w}{2P_Y} = Y^D = Y^S = (1-\lambda)\frac{P_Y}{2}$. Thus the equilibrium prices and quantities are

$$X = \frac{\sqrt{\lambda w}}{2},\ P_X = \sqrt{\frac{w}{\lambda}},\ Y = \frac{\sqrt{(1-\lambda)w}}{2},\ P_Y = \sqrt{\frac{w}{1-\lambda}}$$

By tje no arbitrage condition, the profit of the X-producing firm should be equal to the profit of the Y-producing firm. Thus,

$$P_X X - X^2 = P_Y Y - Y^2$$

which leads to $\lambda = \frac{1}{2}$. Therefore,

$$X = Y = \frac{\sqrt{w}}{2\sqrt{2}}$$

$$P_X = P_Y = \sqrt{2w}$$

4.4.2 Perfect Competition/Monopoly Economy with 1 Goods

Perfect Competition

We normalize the total mass of consumers to 1, where the individual's mass is infinitely divisible ($\frac{1}{n} \to 0$). Each consumer obtains utility of α when she consumes 1 unit of the goods, where α is uniformly distributed on the interval $[0, 1]$. She can only consume a unit. If the price of the goods is p, the consumers with $\alpha > p$ will consume. Thus, the market demand function is

$$Q = \int_p^1 1 \mathrm{d}\alpha = 1 - p$$

We also normalize the total mass of firms. Each firm's production cost is c where c is uniformly distributed on the interval $[0, 1]$. Each firm can only produce one unit. Thus, the firm with $c < p$ will produce and the market supply is

$$Q = \int_0 p 1 \mathrm{d}c = p$$

Thus, the market clearance condition is $1 - p = p \ \leftrightarrow p = \frac{1}{2}$. Therefore, the market price and quantity are

$$p = Q = \frac{1}{2}$$

Monopoly

In this case, we assume that consumers are the same, but that all the firms merge into a single firm, i.e., there is only one firm that produces as many units as it wants. The total cost is $\frac{1}{2}q^2$. Thus, the profit of the firm is,

$$\pi = p \cdot q - \frac{1}{2}q^2 = (1 - q)q - \frac{1}{2}q^2$$

The second equality holds because the market demand is $p = 1 - q$. The equilibrium price and quantity are

$$q^* = \frac{1}{3} \quad p^* = \frac{2}{3}$$

5 Real Analysis focused on Theorem of Maximum and Fixed Point Theorem

5.1 Basics

Definition 42 *Preliminiaries*

*Given a vector $x \in R^n$ and a positive real number e, we define an **open ball of radius** e at x as $B_e(x) = \{y \in R^n||y - x| < e\}$. A set of points A is a **open set** if for every x in A there is some $B_e(x)$ which is contained in A. If $x \in A$ and there exists an $e > 0$ such that $B_e(x)$ is in A, then x is said to be in the **interior** of A.*

*The **complement** of a set A in R^n consists of all the points in R^n that are not in A; it is denoted by $R^n \setminus A$. A set is a **closed set** if $R^n \setminus A$ is an open set. A set A is **bounded** if there is some x in A and some $e > 0$ such that A is contained in $B_e(x)$.*

*An infinite **sequence** in R^n, $(x_i) = (x_1, x_2, ...)$ is just an infinite set of points, one point for each positive integer. A sequence (x_i) is said to converge to a point x^* if for every $e > 0$, there is an integer m such that, for all $i > m$, x_i is in $B_e(x^*)$. We sometimes say that x_i gets arbitrarily close to x^*. We also say that x^* is the limit of the sequence (x_i) and write $\lim_{i\to\infty} x_i = x^*$. If a sequence converges to a point, we call it a **convergent sequence**.*

$x_0 \in A$ is a limit point of A, if and only if every open ball $B_e(x_0)$ contains a point in A.

Definition 43 *If A is a compact set, then every sequence in A has a subsequence converging to a point in A.*

Proposition 48 *Heine–Borel theorem*

Every compact set in R^n is closed and bounded.

Proof. Suppose compact set is not closed. Then there is a limit point X_0 not in A. Let x_n be a sequence in A converging to x_0. Then all subsequence also converges to x_0. Thus, contradiction.

Suppose compact set is not bounded. First, define $(x_i) \in A$. Choose any $x_1 \in A$. Then pick x_i so that $d(x_i, x_1) > i$. Then there should be subsequence y_j of x_i that converges to a point $y_0 \in A$, which means there is an integer N such that for $n > N$, $d(y_n, y_0) < 1$. Thus, $d(yn, x_1) \leq d(y_n, y_0)+d(y_0, x_1) < 1+d(y_0, x_1)$. However, since y_n is at least the $n-th$ member of x_i, it follows that $d(y_n, x_1) \geq d(x_n, x_1) > n$. for large enough n, it is a contradiction. ■

Proposition 49 *Every closed and bounded set in R^n is also compact.*

Definition 44 *A function $f(x)$ is **continuous** at x^* if for every sequence (x_i) that converges to x^*, we have the sequence ($f(x_i)$) converging to $f(x^*)$. A function that is continuous at every point in its domain is called a **continuous function**.*

Definition 45 *A set of points A in R^n is convex if $x \in A$ and $y \in A$ implies $tx + (1 - t)y \in A$ for all t such that $0 \leq t \leq 1$. A set of points in A is strictly convex if $tx + (1 - t)y$ is in the interior of A for all t such that $0 < t < 1$.*

Proposition 50 *Separating hyperplane theorem.*

If A and B are two nonempty, disjoint, convex sets in R^n, then there exists $\vec{p}$ such that $\vec{p} \cdot \vec{x} \geq \vec{p} \cdot \vec{y}$ for all $x \in A$ and $y \in B$.

Similarly

Proposition 51 *Supporting Hyperplane theorem. If $x \in A$ where A is nonempty, convex sets in R^n and $y \notin A$ then there exists $\vec{p}$ such that $\vec{p} \cdot \vec{x} \geq \vec{p} \cdot \vec{y}$ for all $x \in A$ and $y \notin A$.*

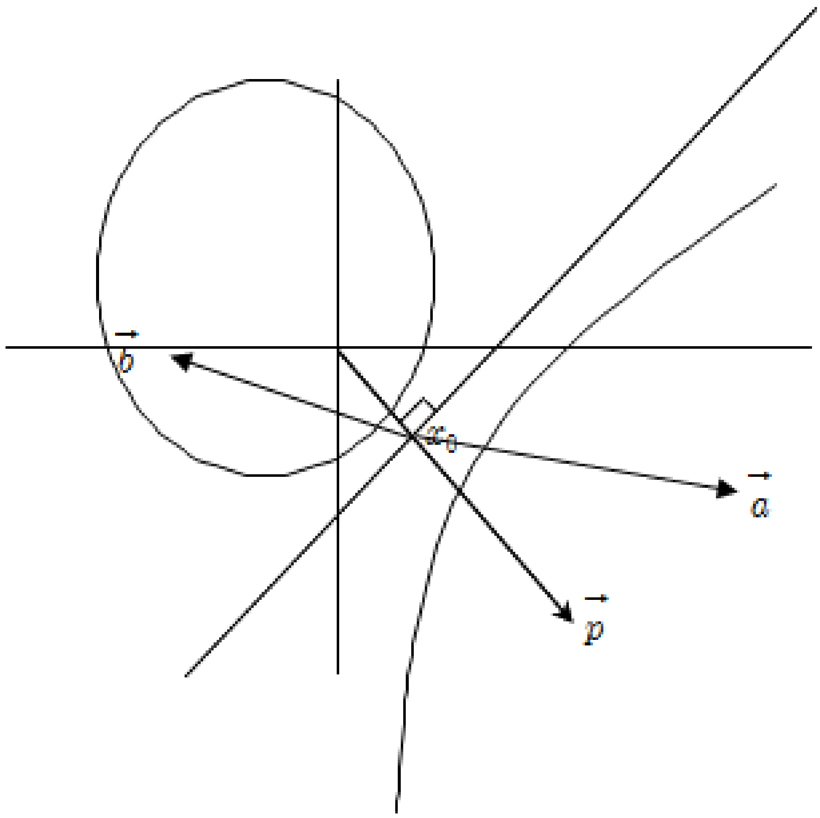


Figure 28: Separating Hyperplane

Definition 46 *A function $f : R^n \to R$ is* ***concave*** *if*

$$f(tx + (1-t)y) \geq tf(x) + (1-t)f(y)$$

for all x and y and all $0 \leq t \leq 1$.

Definition 47 *A function $f : R^n \to R$ is* ***quasi-concave*** *if the upper contour sets of the function are convex sets. In other words, sets of the form $P_a = \{x \in R^n | f(x) \geq a\}$ are convex for all values of a. A function $f(x)$ is quasiconvex if $-f(x)$ is quasiconcave.*

Proposition 52 *A concave function is quasiconcave. A convex function is quasiconvex.*

Proof. Let f be a concave function. Denote the upper contour set $P_a : x \in P_a$ and $y \in P_a$. We need to show that P_a is convex. The concavity of f implies that

$f((1`t)x + ty)e(1`t)f(x) + tf(y) \geq a$ since $f(x)ea$ and $f(y)ea$. ■

The converse of this result is not true: **a quasiconcave function may not be concave**. Consider, for example, the function $f(x, y) = xy$ defined on the set of pairs of nonnegative real numbers. This function is quasiconcave (its upper level sets are the sets of points above rectangular hyperbolae), but is not concave (for example, $f(0, 0) = 0$, $f(1, 1) = 1$, and $f(2, 2) = 4$, so that $f((1/2)(0, 0) + (1/2)(2, 2)) = f(1, 1) = 1 < 2 = (1/2)f(0, 0) + (1/2)f(2, 2))$.

5.2 Equilibrium

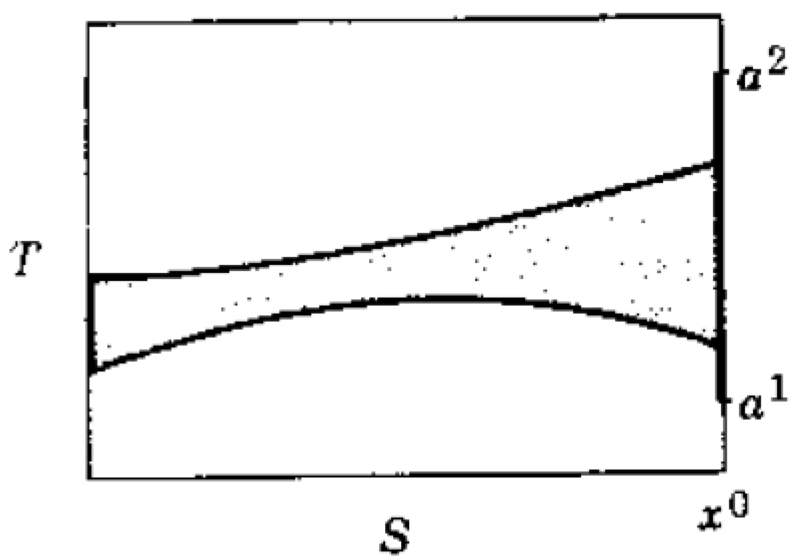
Definition 48 *Given a set in $A \subset R^n$, a correspondence $g : A \to R^k$ is a rule that assigns a set $g(x) \subset R^k$ to every $x \in A$.*

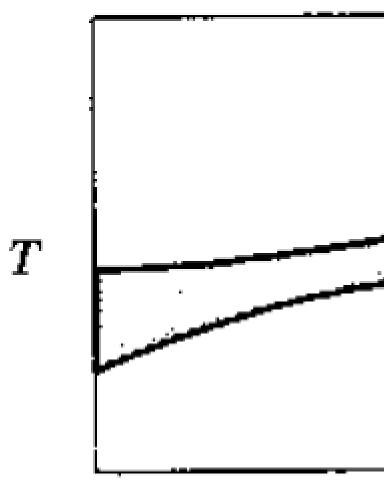
Definition 49 *A correspondence $g(x)$ is*

- *upper-hemi(semi) continuous at a if $g(a)$ is nonempty and if, for every sequence $a_n \to a$ and every sequence b_n such that $b_n \in g(a_n)$ for all n, there exists a convergent subsequence of b_n whose limit point b is in $g(a)$.*

- *lowe-hemi(semi) continuous at a if $g(a)$ is nonempty and if, for every $b \in g(a)$ and every sequence $a_n \to a$, there exists $N \geq 1$ and a sequence b_n from $n = N$ to ∞ such that $b_n \to b$ and $b_n \in g(a_n)$ for all neN.*

- *continuous if and only if both u.h.c and l.h.c.*





Proposition 53 *Extreme Value Theorem*

if a real-valued function f is continuous on the closed interval [a,b], then f must attain a maximum and a minimum, each at least once.

Proposition 54 *Maximum Theorem or Theorem of Maximum*

Consider a parameterized maximization problem of the form

$$v(a) = \max_{x} f(x, a)$$
$$\textit{subject to } \; x \in G(a)$$

or equivalently

$$v(a) = \max\{f(x, a)|x \in G(a)\}$$

$$x^*(a) = \arg\max\{f(x, a)|x \in G(a)\} = \{x \in G(a)|f(x, a) = v(a)\}$$

Let $f(x, a)$ be a continuous function with a compact range and suppose that the constraint set $G(a)$ is a nonempty, compact-valued, continuous correspondence of a. Then

1. *$v(a)$ is a continuous in a function and*

2. *$x * (a)$ is a non-empty upper-semicontinuous correspondence.*

Corollary 4 *If in addition to the conditions above,*

1. *if f is quasiconcave and $G(a)$ is convex, then $x(a)$ is also convex.*

2. *if f is strictly quasiconcave and $G(a)$ is convex, then $x(a)$ is single-valued, and thus is a continuous function rather than a correspondence.*

3. *if f is concave and $G(a)$ is convex, then $v(a)$ is concave and $x(a)$ is convex.*

4. *if f is strictly concave and $G(a)$ is convex, then $x(a)$ is a continuous function.*

Proposition 55 *Brouser's Fixed Point Theroem*

Supppose that a $A \subset R^n$ is a non-empty, compact, convex set and $f : A \rightarrow A$ is a continuous function from A into itself. Then $f(\cdot)$ had a fixed point $x \in A$ such that $f(x) = x$.

Proposition 56 *Kakutani's Fixed Point Theorem*

Supppose that a $A \subset R^n$ is a non-empty, compact, convex set and $f : A \rightarrow A$ is a upper-hemi continuous correspondence from A into itself. Then $f(\cdot)$ had a fixed point $x \in A$ such that $f(x) \in x$.

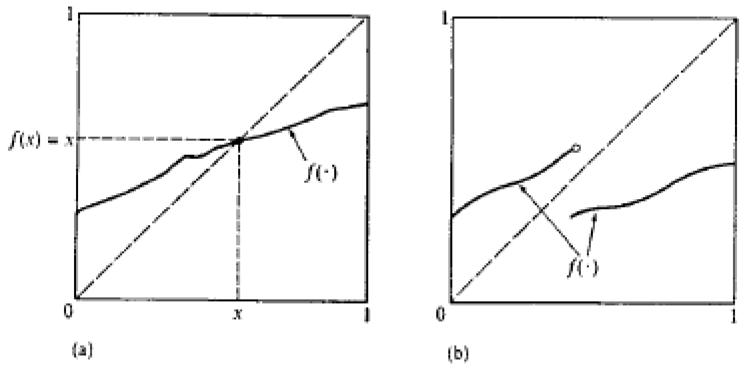


Figure M.I.1
Brouwer's fixed point theorem.
(a) A continuous function from [0, 1] to [0, 1] has a fixed point.
(b) The continuity assumption is indispensable.

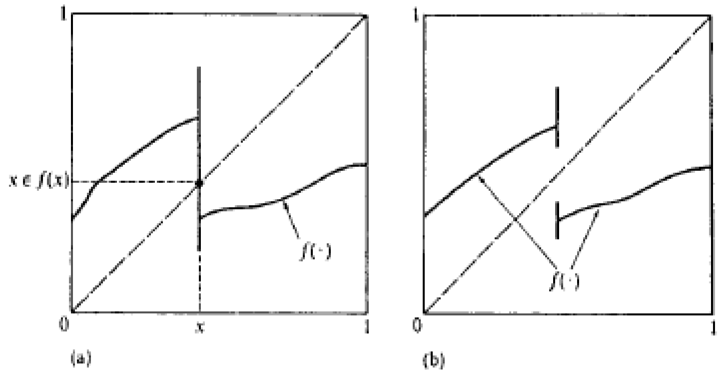


Figure M.I.2
Kakutani's fixed point theorem.
(a) A fixed point exists.
(b) The convex-valuedness assumption is indispensable.

Example 23 *Existence of a Nash Equilibrium and Fixed Point Theorem*

For each i, let $BR_i : S_{-i} \to S_i$ be defined as follows:

$$BR_i(s_{-i}) = \{s + i \in S_i | u_i(s_i; s_{-i}) \geq u_i(t; s_{-i}, \forall t \in S + i\}$$

this function describes $i's$ best response to the other players strategy s_{-i}. Define the mapping $BR : S \to S$ as

$$BR(s) = BR_1(s1) \times \cdots BR_I(s_I)$$

is a Nash equilibrium if and only if

$$s^* = BR(s^*)$$

which is a fixed point.

Past Exams-Jockbo

수리경제학 Quiz 1 (2014년도 1학기)

이름: ______________________________ 학번: ______________________________

1. 행렬 A와 행렬 B는 서로 다른 행렬이지만, 같은 column space를 가지고 있다. 다음의 문장이 참인지 거짓인지를 밝히고, 그 이유를 답하라.

   (a) A 와 B는 독립인 column의 수가 같다.

   (b) A와 B의 left null space는 같다.

   (c) A와 B의 projection matrix들을 각각 P_A와 P_B라 하자. 그렇다면, 모든 벡터 $\vec{y}$에 대해, $P_A\vec{y} = P_B\vec{y}$이다.

   (d) rank(AB)=max(rank(A),rank(B)) 단, max(X,Y)는 X와 Y중 큰 값을 주는 함수.

2. $\vec{v_1}, \cdots, \vec{v_m}$가 기저인 벡터공간 V에서 속한 $\vec{0}$가 아닌 벡터 $\vec{v}$는 다음과 같이 쓸 수 있다.

$$\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$$

   이 중, $a_i = 0$인 $1 \leq i \leq m$를 택해서, $\vec{v_i}$ 대신 $\vec{v}$를 삽입한다면, $\vec{v_1}, \cdots, \vec{v}, \cdots, \vec{v_m}$는

   (a) V의 생성집합이 아님을 증명하시오. (3점)

   (b) 서로 독립이 아님을 증명하시오. (1점)

   * Hint: 만약 생성집합 또는 독립이라고 가정한다면 어떤 모순이 발생하는지를 보이면 된다.

   * Caution : 수업 노트의 Corollary 1과 혼돈하지 마시오

3. $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 7 & 11 \\ 2 & 6 & 8 \end{pmatrix}$이다. 다음에 답하여라.

   (a) A의 역행렬을 구하여라.(만약 존재한다면) (1점)

   (b) A의 행렬식을 구하여라. (1점)

   (c) $Ax = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$ 의 해를 구하여라. (1점)

   (d) $B = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & 9 \\ 2 & 4 & 6 \end{pmatrix}$이다. B의 Null Space의 기저를 구하여라. (2점)

수리경제학 Quiz 2

이름: ______________________________ 학번: ______________________________

1. 다음 함수의 그래프를 그리시오.

   (a) $y = x^{\frac{1}{x}}$ 그리고 극대값, 극소값, 최대값 또는 최소값이 존재하면, 이를 구하시오. 단, $x > 0$ (4점)

   (b) $y = x - \ln x$. 그리고 극대값, 극소값, 최대값 또는 최소값이 존재하면 이를 구하시오. 단, $x > 0$. 이때 $\lim_{x\to\infty} x - \ln x$는 얼마인가? 이를 어떻게 증명할 수 있는가? (4점)

2. 가정용 정수기 수요가 $x = e^{-P_x}$, 사무실용 정수기 수요가 $y = e^{-P_y}$이다. 단, x, y는 각각 가정용 및 사무실용 정수기의 수요량을, P_x와 P_y는 각 정수기의 가격을 의미한다. 총생산비용이 $5x + 3y$이라면, 각각 얼마나 생산해야 이윤이 극대화 되는가? 그리고 그때의 가격은 얼마인가? (3점)

3. $f(x, y) = 3x + 6y - x^3 - y^3$이다. $x - y$ 좌표 위에 $f(x, y)$가 오목함수인 영역, 볼록함수인 영역 그리고 알 수 없는 영역을 표시하시오. (4점)

수리경제학 Quiz 3

이름: ______________________________ 학번: ______________________________

1. 미적분학의 기본정리를 적고, 이를 증명하시오.(3점)

2. 다음 적분 문제를 푸시오

   (a) $\int_1^e \ln x \mathrm{d}x$ (2점)

   (b) $\int_2^\infty \frac{1}{(1-x)^2} \mathrm{d}x$ (2점)

   (c) $\int_2^6 x\sqrt{2x^2 - 8} \mathrm{d}x$ (2점)

   (d) $\int \frac{1}{x \ln 2x} \mathrm{d}x$ (2점)

3. $y = e^x$와 $x = 1$ 그리고 x축 y축에 둘러싸인 영역과 높이 xy^2에 둘러싸인 영역의 부피를 구하라 (4점)

수리경제학 기말고사

이름: ________________________________ 학번: ________________________________

1. 다음 문장에 대해 참/거짓을 밝히고 그 이유를 적으시오. (각 2점)

   (a) $AB = I$이다. 여기서 A와 B는 임의의 행렬이고, I는 단위행렬이다. 그렇다면, B는 A의 역행렬이다.

   (b) $rank(AB) < \max(rank(A), rank(B))$는 모든 행렬 A와 B에 대해 성립한다.

   (c) 어떤 행렬 A에 대해, $A\vec{x} = \vec{y}$를 만족시키는 두 벡터가 있다고 하자. 여기서 $\vec{y}$는 영벡터가 아니다. 행렬 A의 column space에 대한 projection matrix가 P라 할 때, $P\vec{y} = \vec{y}$이다.

   (d) $A\vec{x} = 0$를 만족시키는 영벡터가 아닌 $\vec{x}$가 존재한다고 하자. 단, 여기서 0은 영벡터를 의미한다. 행렬 A의 determinant는 0이 아니다.

2. $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 3 \end{pmatrix}$이다. 다음에 답하여라.

   (a) A의 역행렬을 구하여라. (1점)

   (b) A의 Determinant를 구하여라. (1점)

   (c) A의 eigen-value를 구하여라. (2점)

3. 다음 함수의 그래프를 그리시오. 단, 극대값, 극소값, 최대값 또는 최소값이 존재하면, 이를 구하시오. (각 3점)

   (a) $y = e^{\frac{x}{3}} - \ln x$ for $x > 0$. 단, a는 $a = e^{\frac{1}{a}}$를 만족시킨다. 이때 극값을 a로 표시하라. 단, $e < \ln \frac{3}{a}$.

   (b) $y = x^{-x}$ for $x > 0$.

4. $z = f(x, y)$이다.

   (a) $\mathrm{d}z$와 $\mathrm{d}^2 z$를 적으시오. (2점)

   (b) 모든 $\mathrm{d}x$, $\mathrm{d}y$에 대해, $\mathrm{d}^2 z < 0$이기 위한 충분조건을 유도하시오.(2점)

5. 다음 적분을 하시오.

   (a) $\int_4^9 \frac{1}{\sqrt{x}-1}\mathrm{d}x$ (3점)

   (b) $\int \frac{1}{1+x^2}\mathrm{d}x = \arctan x$이다. 이를 이용하여 $\int \ln(1 + x^2)\mathrm{d}x$를 구하시오. (3점)

6. 다음 극대화 문제를 푸시오. (각 4점)

(a)

$$\max_{x,y} \quad z = \ln x + \ln y$$

$$\text{s.t.} \qquad x + y = c$$

(b)

$$\max_{x,y} \quad z = x^a y^{1-a}$$

$$\text{s.t.} \qquad x + y = c$$

7. 1기와 2기의 두 기간 동안의 소비자 효용함수가 $\ln c_1 + \beta \ln c_2$로 주어져 있다. 여기서 β는 time discount factor(시간 할인율)이고, c_1은 1기의 소비이고, c_2는 2기의 소비이다.

(a) 소비자 1의 소득이 1기에는 0이고, 2기에는 W 이다. 그의 효용 함수는 $\ln c_1 + \beta \ln c_2$이다. 따라서, 이 소비자는 1기에 돈을 빌려 생활하고 2기에 이자와 함께 상환해야 한다. 이자율이 r일 때 효용을 극대화하는 소비수준을 구하여라. (3점)

(힌트 : $c_1(1+r) + c_2 = W$가 제약 조건식이다.)

(b) 소비자 2의 소득은 1기에는 W이고, 2기에는 0이다. 그의 효용함수는 $\ln d_1 + \beta \ln d_2$ 이다. 따라서, 이 소비자는 1기에는 돈을 빌려주고 2기에 원금과 이자를 받아 생활한다. 은행이자율이 r 일 때 효용을 극대화하는 소비수준을 구하여라.(3점)

(힌트 :$d_1 + \frac{d_2}{1+r} = W$가 제약 조건 식이다.)

(c) 전체 경제에 소비자 1과 소비자 2만 존재하여, 소비자 1이 소비자 2에게 돈을 빌려 주는 경우를 생각해보자. 그런 경우, 자금 공급과 자금 수요가 일치하기 위해서는 이자율이 얼마여야 하는가? (3점)

(힌트 : $c_1(1+r) = d_2$가 성립하여야 한다.)

Mathematical Economics Final Exam

Name: ________________________ Student Id. : ________________________

1. State whether following statements are true or false. Do not give reasons for your choice. Each are 2 points. However, if you get the answer wrong, minus 2 points.

   (a) Suppose matrix A and B are both $n \times m$ matrices, where $A \neq B$. Suppose a set of vectors $\{v_1, \cdots, v_k\}$ is basis of both $col(A)$ and $col(B)$. Then $\vec{v} \notin col(A)$ if and only if $\vec{v} \notin col(B)$.

   True

   (b) Suppose $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is spanning set of a vector space $\mathbf{V}$ and $\vec{v_1} = \alpha \vec{v_2}$. Then $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is basis of $\mathbf{V}$.

   False

2. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 0 \\ 1 & 2 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

   (a) Find inverse matrix of A. (2 points)

   $$A^{-1} = \begin{pmatrix} 2 & 0 & -1 \\ 0 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

   (b) Find determinant of A. (2 points)

   1

   (c) Find eigen-values and eigen-vectors of of B. (2 points)

   - $\lambda = 1$ $c(1, 0, 0)$
   - $\lambda = 2$ $c(2, 1, 0)$
   - $\lambda = 3$ $c(5, 2, 2)$

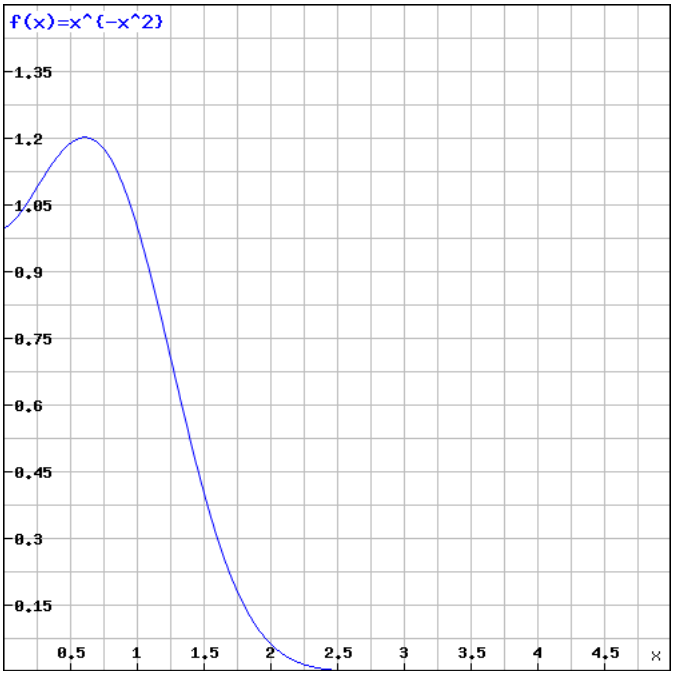
3. Draw the graph of x^{-x^2} where $x > 0$. Find all extremas and saddle points. (8 points)

   minimum at $x = 0,\ y = 1$ and maximum at $x = e^{-0.5},\ y = (e^{-0.5})^{-e^{-1}} = e^{0.5e^{-1}}$

   Definition 50 ***Value Function***

$$(x_1^*, \cdots, x_n^*) \in \underset{x_1, \cdots, x_n}{argmax}\ f(x_1, \cdots, x_n | r)$$

$$\begin{aligned} s.t. \quad & g_1(x_1, \cdots, x_n) \le c_1 \\ & \vdots \\ & g_m(x_1, \cdots, x_n) \le c_m \end{aligned}$$



A value function is

$$v(r) = f(x_1^*, \cdots, x_n^*|r)$$

Proposition 57 ***Envelope Theorem*** *Let $v(r)$ be a value function for an optimization problem, where $f(x_1, \cdots, x_n)$ is the objective function, $g_1, \cdots, g_m$ are constraints, and $\lambda_1, \cdots, \lambda_m$ are Lagrangian multipliers.*

$$\frac{\mathrm{d}v(r)}{\mathrm{d}r} = \frac{\partial f(x_1^*, \cdots, x_n^*|r)}{\partial r} - \sum_{j=1}^{m} \lambda_j \frac{\partial g_j(x_1, \cdots, x_n|r)}{\partial r}\Big|_{(x_1^*, \cdots, x_n^*)}$$

4. Prove the Envelope Theorem. (8 pts)

   See lecture note.

5. Find the volume/area of following figures

   (a) Figure surrounded by $y = x^2$, $y = x$ and $z = x + y$ (8 pts)

   3/20

   (b) Figure surrounded by $y = x^3$, $y = x^2$, and $z = xy$ (8 pts)

   1/48

6. Suppose the utility function of Suji is $u(x_1, x_2) = \alpha \ln x_1 + \beta \ln x_2$. Suppose the price of x_1 and x_2 are p_1 and p_2 respectively. Suji has y amount of money.

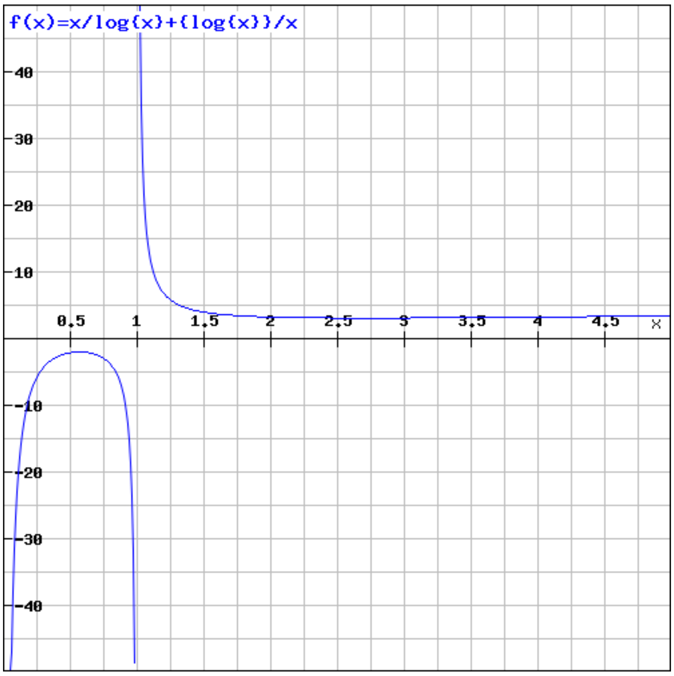
   (a) How much x_1 and x_2 would Suji consume? Assume $0 < \alpha, \beta < 1$ (9 pts) $x_1 = \frac{\alpha}{\lambda p_1}$ and $x_2 = \frac{\beta}{\lambda p_2}$

(b) Now suppose the goods are sold in a bundle such that Suji can only buy same amount of x_1 and x_2, i.e., $x_1 = x_2$. What is the optimal consumption level for Suji? Let the bundle price be p. (9 pts) x_1과 x_2는 묶음으로 팔리고 있다. 즉, x_1과 x_2는 같은 양을 소비하여야 한다. 그렇다면, 총 몇개를 소비하겠는가? 단, 그 묶음 상품의 가격은 p이다.

$x = x_1 = x_2 = y/p$

7. Stat/Math/Engineering/Science major students, answer the following question.

(a) Draw of the graph of $y = \frac{x}{\ln x} + \frac{\ln x}{x}$. Find all extremas and saddle points, where $x > 0$. (6 pts)



(b) Draw of the graph of $y = \frac{x^2}{e^x}$. Find all extremas and saddle points. (6 pts)

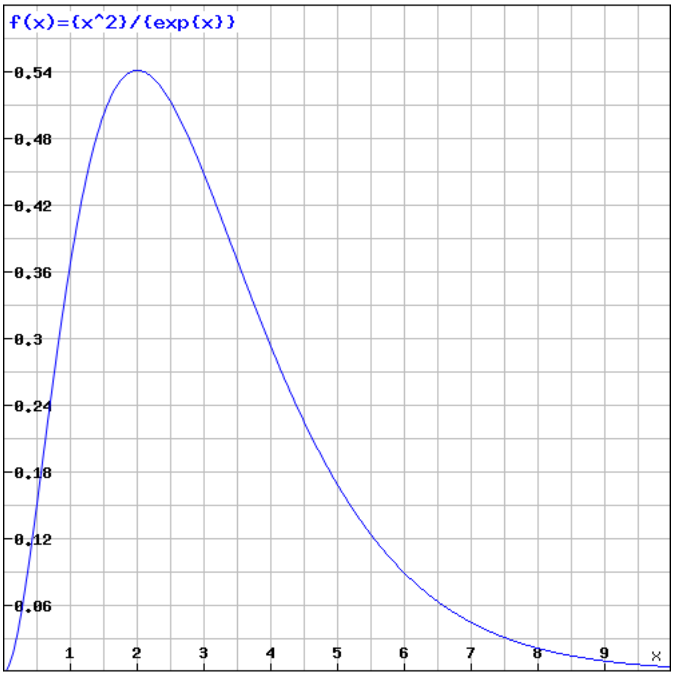
(c) Suppose the utility function of Suji is $u(x_1, x_2) = (x_1^\rho + x_2^\rho)^{\frac{1}{\rho}}$. Suppose the price of x_1 and x_2 are p_1 and p_2 respectively. Suji has y amount of money. How much x_1, and x_2 would Suji consume? Assume $0 < \rho < 1$ (6 pts)

$x_1 = \frac{p_1^{r-1} y}{p_1^r + p_2^r}$, $x_2 = \frac{p_2^{r-1} y}{p_1^r + p_2^r}$, where $r = \frac{\rho}{\rho - 1}$

☐ End of Exam

Let X_i be the points you earned in question i. Let X be the final score of your mid-term exam.

- (Econ/Bussiness and other non-math oriented major students) $X = \sum_{j=1}^{7} X_j$. If you solve 8, you will be counted as a Math major student.
- (Math/Stat/Engineering/Science students) Your final score is $X = X_8 + 0.7 \times \sum_{j=1}^{7} X_j$



Mathematical Economics Mid-term Exam

Name: ______________________ Student Id. : ______________________

1. State whether following statements are true or false. Do not give reasons for your choice. Each are 2 points. However, if you get the answer wrong, minus 3 points.

   (a) Suppose matrix A and B are both $n \times m$ matrices, where $A \neq B$. Suppose a set of vectors $\{v_1, \cdots, v_k\}$ is basis of both $col(A)$ and $col(B)$. Then $dim(Null(A)) = dim(Null(B))$. T

   (b) Suppose matrix A and B are both $n \times m$ matrices, where $A \neq B$. Suppose a set of vectors $\{v_1, \cdots, v_k\}$ is basis of $col(A)$ and $col(B)$. Then $LeftNull(A) = LeftNull(B)$. T

   (c) Suppose matrix A and B are both $n \times m$ matrices, where $A \neq B$. Suppose a set of vectors $\{v_1, \cdots, v_k\}$ is basis of $col(A)$ and $col(B)$. Then $rank(A) = rank(B)$. T

   (d) Suppose A and B are two different matrices, where $row(A) \subset row(B)$. Then $rank(A) > rank(B)$ F

   (e) Suppose $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is spanning set of a vector space $\mathbf{V}$ and $\vec{v_i} \cdot \vec{v_j} = 0$ for any $i \leq m$ and $j \leq m$. Then $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is basis of $\mathbf{V}$. T

   (f) Suppose $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is spanning set of a vector space $\mathbf{V}$. Suppose $\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$. Then real number sequence $a_1, \cdots, a_m$ is always unique. F

(g) Suppose A and B are both $n \times n$ matrices. Let $rank(A) < rank(B)$. Then A is invertible. F

(h) Suppose A is an $n \times n$ matrices, and any n-components vector $\vec{v} \in col(A)$. Then A is invertible. T

(i) Suppose A is an $n \times n$ matrices and $col(A) = row(A)$. Then A is invertible. F

(j) Suppose $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is basis of a $col(A)$ and $col(B)$. Then for any x $P_A x = P_B x$, where P_A and P_B are projection matrix of A and B respectively. T

2. Let $A = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \\ 1 & 1 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

(a) Find inverse matrix of A. (3 points) $\begin{pmatrix} 2 & -3 & -1 \\ 0 & 1 & 0 \\ -1 & 1 & 1 \end{pmatrix}$

(b) Find determinant of A. (3 points)

1

(c) Find eigen-values and eigen-vectors of of B. (4 points) eigen value is 1 and eigenvector is $(x_1, 0, 0)$

eigen value is 2 and eigenvector is $(2x_2, x_2, 0)$

eigen value is 3 and eigenvector is $(\frac{5}{2}x_3, x_3, x_3)$

Proposition 58 *For a basis $\vec{v_1}, \cdots, \vec{v_m}$ of a vector space $\mathbf{V}$, for a $\vec{v} \neq \vec{0}$, there is a real number sequence $a_1, \cdots, a_m$, at least one of which is not 0, such that*

$$\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$$

For $a_i \neq 0$, insert $\vec{v}$ instead of $\vec{v_i}$. Then $\vec{v_1}, \cdots, \vec{v_{i-1}}, \vec{v}, \vec{v_{i+1}}, \cdots, \vec{v_m}$ is also the basis of $\mathbf{V}$.

3. Prove that all bases of a vector space $\mathbf{V}$ have the same number of vectors. Take Proposition 61 for given. (5 points)

Proof. Let $\vec{v_1}, \cdots, \vec{v_m}$ and $\vec{V_1}, \cdots \vec{V_M}$ are bases of vector space $\mathbf{V}$, and suppose $M > m$. Since $\vec{V_1}, \cdots \vec{V_M} \in \mathbf{V}$, for $1 \leq i \leq m$,

$$v_i = a_1\vec{V_1} + \cdots + a_M\vec{V_m}$$

By Proposition 58, for $a_j \neq 0$, we can replace $\vec{V_j}$ by v_i. By this method, we can replace all $\vec{V_1}, ..., \vec{V_m}$ with $\vec{v_1}, \cdots, \vec{v_m}$, such that $\vec{v_1}, \cdots, \vec{v_m}, \vec{V_{m+1}}, \cdots, \vec{V_M}$ are basis. Since $\vec{v_1}, \cdots \vec{v_m}$ is basis $\vec{V_{m+1}}, \cdots, \vec{V_M}$ is a linear combination of $\vec{v_1}, \cdots \vec{v_m}$, which is a contradiction. ■

4. P_A is projection matrix on $col(A)$. Suppose Q_A is a projection matrix on $row(A)$. Find Q_A. (5 points)

   (Hint: Let y be any vector. Then $Q_A y$ is a vector in $row(A)$ such that it is closest to y. Note that a vector in $row(A)$ is $A'x$.)

$$\begin{aligned}(y - A'X) &= 0\\ A'y - A'AX &= 0\\ A'y &= A'Ax\\ (A'A)^{-1}A'y &= x\\ A(A'A)^{-1}A'y &= Ax\end{aligned}$$

$$Q_A = A(A'A)^{-1}A'$$

5. Stat/Math/Engineering/Science major students, answer the following questions.

   (a) Suppose $\vec{v_1}, \cdots, \vec{v_m}$ are dependent and $\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$. Prove that there are infinite number of real number sequences $(a_1, \cdots a_m)$, which means there are infinite ways to linearly combine $\vec{v_1}, \cdots, \vec{v_m}$ to make $\vec{v}$. (6 points)

   Proof. Since $\vec{v_1}, \cdots, \vec{v_m}$ are dependent, $\vec{v_1} = b_2\vec{v_2} + \cdots + b_m\vec{v_m}$. Then $\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m} = r \cdot a_1(b_2\vec{v_2} + \cdots + b_m\vec{v_m}) + (1-r)a_1\vec{v_1} + \cdots + a_m\vec{v_m}$ for $\forall r \in \mathbf{R}$ ■

   Proposition 59 *For an m-dimensional vector space $\mathbf{V}$, these properties always hold.*

   i. For m independent vectors in $\mathbf{V}$ is always a basis of $\mathbf{V}$.

   ii. A spanning set of $\mathbf{V}$ with m vectors is always a basis.

   Proposition 60 *Any subset of independent vectors are also independent.*

   (b) Prove that any $m+1$ vectors in m-dimensional vector space $\mathbf{V}$ are always dependent. Take Proposition 2 and 3 for given. (6 points)

   Proof. Suppose $\vec{v_1}, \cdots, \vec{v_{m+1}} \in \mathbf{V}$ are independent. By Proposition 3, $\vec{v_1}, \cdots, \vec{v_m}$ are also independent. By Proposition 2.i., $\{\vec{v_1}, \cdots, \vec{v_m}\}$ is a basis. Thus, one can linearly combine $\vec{v_1}, \cdots, \vec{v_m}$ to make v_{m+1}. Thus, $\vec{v_1}, \cdots, \vec{v_{m+1}}$ are not independent. Thus, contradiction. ■

☐ End of Exam

Let X_i be the points you earned in question i. Let X be the final score of your mid-term exam.

- (Econ/Bussiness and other non-math oriented major students) $X = \sum_{j=1}^{4} X_j$. If you solve 5, you will be counted as a Math major student.
- (Math/Stat/Engineering/Science students) Your final score is $X = X_5 + 0.7 \times \sum_{j=1}^{5} X_j$

2016 Mid-term

1. State whether following statements are true or false. Do not give reasons for your choice. Each are 2 points. However, if you get the answer wrong, minus 3 points.

   (a) Let A be a $n \times n$ square matrix and y be a $n \times 1$ matrix. Also suppose there are infinite number solution x to $Ax = y$. Then the column vector $y \in Col(A)$ T

   (b) Let A be a $n \times n$ square matrix and y be a $n \times 1$ matrix. Also suppose there are infinite number solution x to $Ax = y$. The $dim(LeftNull(A) \geq 1$. T

   (c) Let A be a $n \times n$ square matrix. Suppose all columns are orthogonal to each others. Then all row vector with n components vector is in $Row(A)$. T

   (d) Let A be a $n \times n$ square matrix. Let P_A be the projection matrix of A on $Col(A)$. Suppose for any column vector x, $P_A x = x$. Then projection matrix Q_A of A on $Row(A)$, $Q_A x = x$ always holds. T

   (e) Suppose A and B are two different matrices, where $row(A) \subset row(B)$. Then $rank(A) \leq rank(B)$ T

2. Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 1 & 2 & 2 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$

   (a) Find inverse matrix of A. (3 points) $A^{-1} = \begin{pmatrix} 2 & 0 & -1 \\ 2 & -1 & 0 \\ -3 & 1 & 1 \end{pmatrix}$

   (b) Find determinant of A. (3 points) -1

   (c) Find eigen-values and eigen-vectors of of B. (6 points)

      i. $\lambda_1 = 1$ and $x_1 = (c, 0, 0)'$

      ii. $\lambda_2 = 2$ and $x_2 = (2c, c, 0)'$

      iii. $\lambda_3 = 3$ and $x_3 = (\frac{3}{2}c, c, c)'$

3. Prove that if m vectors, $\vec{v_1}, \cdots, \vec{v_m}$, in a vector space V are orthogonal, i.e., for any $i \neq j \in \{1, \cdots, m\}$, $\vec{v_i} \cdot \vec{v_j} = 0$, then $\vec{v_1}, \cdots, \vec{v_m}$ are independent. (6 points)

4. Suppose $\vec{a_i}$ be a column vector with n elements. Prove that $|\vec{a_1}, \cdots, \vec{a_i}' + \vec{a_i}, \cdots, \vec{a_n}| = |\vec{a_1}, \cdots, \vec{a_i}', \cdots, \vec{a_n}| + |\vec{a_1}, \cdots, \vec{a_i}, \cdots, \vec{a_n}|$. (6 points)

5. Suppose for two different matrices $|AB| = |A||B|$. Prove that an $n \times n$ matrix A is invertible if and only if $|A| \neq 0$.(6 points)

6. Stat/Math/Engineering/Science major students, answer the following question.

   (a) Let $\vec{v_1}, \cdots, \vec{v_m}$ are independent and $\vec{v} = a\vec{v_i}$ where a is a non-zero real number. Prove that $\vec{v}, \vec{v_2}, \cdots, \vec{v_m}$ are independent. (6 points) **Proof.** Suppose $\vec{v}, \vec{v_2}, \cdots, \vec{v_m}$ are dependent. Then there is non-zero real number sequence such that $b_1\vec{v} + b_2\vec{v_2} + \cdots + b_m\vec{v_m} = 0$. Then $b_1\frac{1}{a}\vec{v_1} + b_2\vec{v_2} + \cdots + b_m\vec{v_m} = 0$. Thus, $\vec{v_1}, \vec{v_2}, \cdots, \vec{v_m}$ are dependent. ■

   Proposition 61 *For a basis $\vec{v_1}, \cdots, \vec{v_m}$ of a vector space $\mathbf{V}$, for a $\vec{v} \neq \vec{0}$, there is a real number sequence $a_1, \cdots, a_m$, at least one of which is not 0, such that*

   $$\vec{v} = a_1\vec{v_1} + \cdots + a_m\vec{v_m}$$

   For $a_i \neq 0$, insert $\vec{v}$ instead of $\vec{v_i}$. Then $\vec{v_1}, \cdots, \vec{v_{i-1}}, \vec{v}, \vec{v_{i+1}}, \cdots, \vec{v_m}$ is also the basis of $\mathbf{V}$.

   (b) Suppose Proposition 61 holds. For an m-dimensional vector space V, a set of m independent vectors, $\{\vec{v_1}, \cdots, \vec{v_m}\} \subset V$ is always a basis of V. (6 points)

   Proof.

   Suppose $\{\vec{w_1}, \cdots, \vec{w_m}\}$ is a basis of V. Then $\vec{v_1}$ is a linear combination of $\{\vec{w_1}, \cdots, \vec{w_m}\}$. Thus, by Proposition 61, $\vec{v_1}$ can replace one of $\{\vec{w_1}, \cdots, \vec{w_m}\}$ and still be a basis. We can keep doing this until all w_is are replaced. ■

2016 final

1. State whether following statements are true or false. Do not give reasons for your choice. Each are 3 points. However, if you get the answer wrong, minus 3 points.

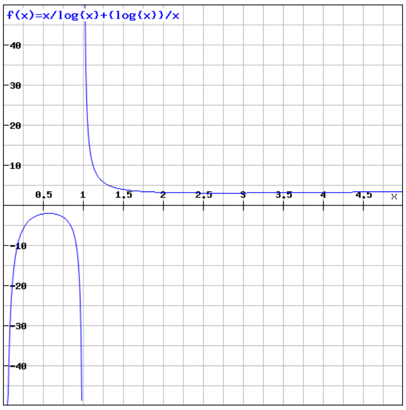
   (a) Any vector in a vector space, $\vec{v} \in V$, is a linear combination of column vectors of a matrix A. Then $V \subset Col(A)$. True

   (b) Any vector which is a linear combination of vector column vectors of matrix A is in the vector space V. Then $V = Col(A)$. False

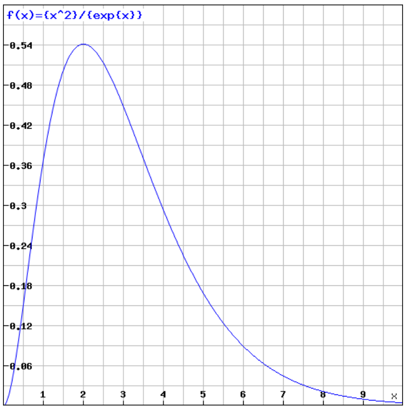
   (c) Suppose $Row(A) = Row(B)$ where A and B are different matrices. Then $Null(A) = Null(B)$. True

2. Draw graphs of following functions

   (a) $y = x^{\frac{1}{x^2}}$ where $x > 0$ (6 points)



   (b) $y = x^{-x^2}$ where $x > 0$. Note that when $x = e^{-\frac{1}{2}}$, $y > 1$. (6 points)



3. Let $z = f(x, y)$ be a bivariate function. Then $\mathrm{d}z = \mathrm{d}x \cdot f_x + \mathrm{d}y \cdot f_y$. Find $\mathrm{d}^2 z$ and conditions for a point (x, y) to be a local minimum, local maximum and saddle point? (6 points)

4. Do the following integration

(a) $\int a^x \mathrm{d}x$ for $a > 0$ (3 points) $\frac{a^x}{\ln(a)} + c$

(b) $\int_0^1 2x^3 \ln(x^2+1)\mathrm{d}x$ (6 points) $= \frac{1}{2}x^4 \ln(x^2+1)|_0^1 - \int_0^1 \frac{x^5}{x^2+1}\mathrm{d}x = \frac{1}{2}\ln 2 - \int_0^1 x^3 - x - \frac{x}{x^2+1}\mathrm{d}x = \frac{1}{4}$

5. Find the volume of a figure sorrounded by $x-y$ plane, $y = \sqrt{x}$, $y = x^3$ and $z = 4xy - y^3$. (9 points) 55/156

6. Suppose market demand function is $p = a - q$ where p is the price chosen and q is the quantity produced. The total cost function is cq. What is the profit maximizing quantity, q? (6 points)

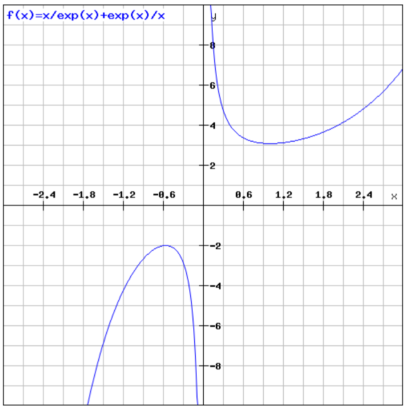
   $\pi = (a-q)q - cq = (a-c-q)q$. FOC is $a - c - q - q = 0$ Thus, $q = \frac{a-c}{2}$.

7. Suppose utility function of a consumer is $U = X^\alpha Y^\beta$. The prices of X and Y are P_X and P_Y. Moreover, she only has Z amount of money. Find the utility maximizing amount of X and Y. (9 points)

   $X = \frac{\alpha}{\alpha+\beta}\frac{Z}{P_X}$ and $Y = \frac{\beta}{\alpha+\beta}\frac{Z}{P_Y}$

8. Stat/Math/Engineering/Science major students, answer the following question.

(a) Draw graph of $y = \frac{x}{e^x} + \frac{e^x}{x}$ How many extrema are there? (9 points).



(b) Find the volume of the figure enclosed by the planes $4x + 2y + z = 10$, $y = 3x$,$z = 0$ and $x = 0$. (9 points) 25/3