Notes on Math Camp for Economics at UChicago

陈科宇

** My personal notes, Sept 5 - 15, 2023. **

Table of Contents

1

Basic Language of Math

1.1 Sets

Let us begin with the *sets*. A **set** is any well-specified collection of elements/members. There are two ways of denoting sets:

1. In style of large brackets, e.g., $\{1,2,3\}, \mathbb{N} = \{1,2,3,4,\cdots\}$;
2. In style of $\{a|P\}$, where $|$ reads "such that", a as a element, belongs to P as a set, e.g., $\{x|x^2 - 5x - 6 = 0\}$

There are some notations indicating the relations between sets and elements: $a \in A$, $a \notin A$, $A = B$, $A \subset B$[1.1].

Property 1.1. Given that (1)$A \subset A$; (2)$A \subset B$ and $B \subset A \Leftrightarrow A = B$.
(3) $A \subset B$, $B \subset C \Leftrightarrow A \subset C$.

Example. $N = \{1,2,3,\cdots\}, \mathbb{Z} = \{\cdots,-3,-2,-1,0,1,2,\cdots\}, \mathbb{R} =$ the set of real numbers, while $\varnothing$ indicates empty set.

Given the existence of a **universal set** U, where other sets are embedded in this universal set U. Now we introduce notations of set operations:

1. The **union** between two such sets A, B embedded in U is denoted by $A \cup B = \{x|x \in A \text{ or } x \in B\}$;
2. The **intersection** between A, B is denoted by $A \cap B = \{x|x \in A \text{ and } x \in B\}$.
3. The **complement** of <u>*B relative to A*</u> is denoted by $A - B = \{x|x \in A \,\text{but}\, x \notin B\}$.
4. Now let A be the universal set, then we have $A - B = B^C$.

Property 1.2. **(De Morgan's Laws)** The followings are the properties applicable to set operations:

1.1. Any element of A is an element of B. We say that, A is a subset of B, or B is a superset of A: $B \supset A$.

1. $C-(A\cup B)=(C-A)\cap(C-B)$. If $C=U$, then we have $(A\cup B)^C=A^C\cap B^C$.
2. $C-(A\cap B)=(C-A)\cup(C-B)$.
3. *(Association Rule)* $(A\cup B)\cup C=A\cup(B\cup C)$

De Morgan's Laws extends to something more general. Denote that $\bigcup_{i=1}^{n}A_i=A_1\cup\cdots\cup A_n$:

4. $(\bigcup_{i\in I}A_i)^C=\bigcap_{i\in I}A_i^C$; $(\bigcap_{i\in I}A_i)^C=\bigcup_{i\in I}A_i^C$.

Let A be a set. **Power set** is the set of all subsets of A, denoted by $P(A)=\{B|B\subset A\}$. The **cardinality** of a power set for a set of n elements is 2^n.

1.2 Functions

Definition 1.1. (Cartesian Product)
Let A and B be sets. Then the **Cartesian product** of A and B is represented by:

$$A\times B=\{(a,b)|a\in A,b\in B\}$$

Definition 1.2. (Function)
A function or map from X to Y, denoted by $f\colon X\to Y$, is a subset of $X\times Y$, s.t., for each $x\in X$, there exists a unique $y\in Y$ with $(x,y)\in f$.

X is denoted by $D(f)$, the **domain** of f; while Y is denoted by the **range** of f, such that $R(f)=\{y\in Y\,|\exists x\in X, s.t., f(x)=y\}$.

Definition 1.3. (Images)
Let X,Y be two sets, and the mapping of $f\colon X\to Y.W\subset X$ is the **direct image** of W under f, such that

$$f(W)=\{f(x)|x\in W|\}\subseteq Y$$

Then let $Z\subset Y$, the **inverse image** of Z under f is:

$$f^{-1}(Z)=\{x\in X\,|\,f(x\}\in Z\}\subset X$$

Example. $f(x)=x^2.W=\{x|0\le x\le 2\}$, we have the direct image of W under f be $f(W)=\{y|0\le y\le 4\}$. Let $Z=\{y|0\le y\le 4\}$, then we have the inverse image of Z under f be $f^{-1}(Z)=\{x|-2\le y\le 2\}$.

Specially note that, in general[1.2]:

$$f^{-1}(f(W)) \neq W$$

Definition 1.4. (Injective, Surjective, and Bijective)

Given $f: X \to Y$, f is **injective** or **one-to-one** if whenever $x_1 \neq x_2$, then $f(x_1) \neq f(x_2)$; f is **surjective** or **onto** if $\forall y \in Y$, $\exists x \in X$, s.t., $f(x) = y$.

Finally, if f is both **injective** and **surjective**, it is called **bijective**.

Definition 1.5. (Inverse Function)

Let $f: X \to Y$, given that f is bijective, then $f^{-1} = \{(y, x) \in Y \times X | (x, y) \in f$ is the **inverse function** of f.

Example. $D(f^{-1}) = R(f); R^{-1}(f) = D(f)$.

Property 1.3. Given $f: X \to Y$, then f is well-defined $\Leftrightarrow$ f is a function.

Definition 1.6. (Composite Function)

Given $f: X \to Y$, $g: Y \to Z$. Then **composite function** $g \circ f: X \to Z$, s.t., $(g \circ f)(x) = g(f(x)), \forall x \in X$.

Proposition 1.1. $f: X \to Y; g: Y \to Z$. Given that $A \subset Z$, then

$$(g \circ f)^{-1}(A) = f^{-1}(g^{-1})(A)$$

1.3 Cardinality

For finite sets, one could directly work out the amount of elements; while for infinite sets, it becomes important to define the ways of *counting*.

Definition 1.7. (Cardinality)

Let two sets A, B. If $\exists$ bijective map between A and B, we say A and B have the same **cardinal number** or the same **cardinality**.

Definition 1.8. (Finity)

1. A is **finite** if A has the same cardinality as $\{1, \cdots, n\}$ for some $n \in \mathbb{N}$;
2. A is **infinite** if A is not finite:
   - A is **countably infinite** if it has the same cardinality as $\mathbb{N}$, e.g., the set of even natural numbers, the set of squared numbers;
   - A is **uncountable** if it is neither finite nor countable.

Proposition 1.2. Every infinite subset of a countably infinite set A is countably infinite.[1.3]

Remark. The cardinality of real numbers equals the cardinality of the power set of natural numbers.

1.2. What is the exact relation then?

1.3. The proof is left to references. The cardinality of such infinite subset is strictly larger than that of $\mathbb{N}$.

2

Linear Algebra

2.1 Vector Space

Definition 2.1. (Field)

A field K is a set with operation: $+$ and $\cdot$. For every pair $a, b \in K$, there are unique elements $a+b \in K$, $a \cdot b \in K$.

1. Communicative. $a+b=b+a, a \cdot b=b \cdot a$;
2. Associative. $(a+b)+c=a+(b+c)$;
3. $\exists 0 \in K$, $1 \in K$, s.t., $0+a=a; 1 \cdot a=a$;
4. $\forall a \in K$, $\forall b \in K, b \neq 0, \exists c, d \in K$, $c+a=0, b \cdot d=1$;
5. $a \cdot (b+c)=a \cdot b+a \cdot c$.

Definition 2.2. (Vector Space)

Let K be a field whose elements will be called **scalars**. A vector space V is a non-empty set whose elements we will referred to as *vectors*. There are two operations: addition and scalar multiplication:

$$\boldsymbol{u}, \boldsymbol{v} \in V, c \in \mathbb{R} \Rightarrow \boldsymbol{u}+\boldsymbol{v} \in V, c\boldsymbol{u} \in V$$

Property 2.1. There are several properties of vectors:

1. $u+(v+w)=(u+v)+w$;
2. $u+v=v+u$;
3. $\exists 0 \in V$, s.t., $0+u=u$;
4. $\forall u \in V, \exists -u \in V$, s.t., $u+(-u)=0$;
5. On scalar multiplication:
   - $\alpha(u+v)=\alpha u+\alpha v$;
   - $(a+\beta)u=\alpha u+\beta u$;
   - $(\alpha\beta)u=\alpha(\beta u)$;

- $1 \cdot u = u$.

Example. (1) $V = K^n = \{(\alpha_1, \cdots, \alpha_n) | \alpha_i \in K\}$. We have the following properties satisfied:

$(\alpha_1, \cdots, \alpha_n) + (\beta_1, \cdots, \beta_n) = (\alpha_1 + \beta_1, \cdots, \alpha_n + \beta_n); \lambda(\alpha_1, \cdots, \alpha_n) = (\lambda\alpha_1, \cdots, \lambda\alpha_n)$

(2) $V = \{\{a_n\}_{n=1}^{\infty}\}$, then $\{a_n\}_{n=1}^{\infty} + \{b_n\}_{n=1}^{\infty} = \{a_n + b_n\}_{n=1}^{\infty}$, $\lambda\{a_n\}_{n=1}^{\infty} = \{\lambda a_n\}_{n=1}^{\infty}$;

(3) $V = \{f : K \to K\}$, $f, g \in V$, $\alpha \in K$, therefore $(f + g)(x) = f(x) + g(x)$; $(\alpha f)(x) = \alpha f(x)$;

(4) (Polynomial) $\mathbb{R}[x] = \{\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n | n \in \mathbb{N}, \alpha_i \in \mathbb{R}\}$.

Definition 2.3. (Linear Combination)

Given scalars $a_1, \cdots, a_n \in \mathbb{R}$ and $v_1, \cdots, v_n \in V$, then $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$ is called a **linear combination** of vectors $v_1, \cdots, v_n$.

Definition 2.4. (Linear Independence)

$v_1, \cdots, v_n \in V$ are **linearly dependent** if $\exists a_1, \cdots, a_n \in K$,s.t., at least one of them is non-zero and $a_1 v_1 + \cdots + a_n v_n = 0$, otherwise, they are **linearly independent**.

Definition 2.5. (Spanning Set)

$\mathcal{S} \subset V$. The subset spanned by $\mathcal{S}$ is:

$$\mathrm{Span}(\mathcal{S}) = \{a_1 v_1 + a_2 v_2 + \cdots a_n v_n | a_n \in K, v_n \in S\}.$$

If $\mathrm{Span}(\mathcal{S}) = V$, we say $\mathcal{S}$ **spans** V.

Theorem 2.1. $\mathcal{S} \subset V$, $\mathcal{S}$ is linearly independent $\Leftrightarrow$ every vector in $\mathrm{span}(\mathcal{S})$ has a unique expression as a linear combination of vectors in $\mathcal{S}$.[2.1]

Definition 2.6. (Basis)

$\mathcal{S} \subset V$ is **basis** of V if $\mathcal{S}$ is linearly independent and $\mathcal{S}$ spans V.

2.1. See the proof. Prove by contradiction.

Theorem 2.2. $\mathcal{S}=\{v_1,\cdots,v_n\}\in V$, $\mathcal{S}$ is a basis of $V \Leftrightarrow \forall v\in V, \exists$ a unique set of scalars $a_1,\cdots,a_n$, s.t.,$a_1v_1+\cdots+a_nv_n=v$.[2.2]

Definition 2.7. (Coordinates)

$\mathcal{S}=\{v_1,\cdots,v_n\}$ be a basis of V. For a given $v\in V$, the unique scalars $a_1,\cdots,a_n\in K$, s.t., $v=a_1v_1+\cdots+a_nv_n$ are called **coordinates** of V w.r.t. the basis $\{v_1,\cdots,v_n\}$.

Example. Consider $V=\mathbb{R}^2$.

1. $e_1=(1,0), e_2=(0,1)$, $(\alpha_1,\alpha_2)\in\mathbb{R}^2, (\alpha_1,\alpha_2)=\alpha_1e_1+\alpha_2e_2$;
2. $e_1=(1,0), e_2=(1,1), (\alpha_1,\alpha_2)\in\mathbb{R}^2, (\alpha_1,\alpha_2)=(\alpha_1-\alpha_2)e_1+\alpha_2e_2$.

Theorem 2.3. In a finite dimensional vector space V, any two of its bases of vectors $\{v_1,\cdots,v_m\},\{u_1,\cdots,u_n\}$ must have $m=n$.

Definition 2.8. (Dimension) The number n of vectors in a basis of a finite dimensional vector space V is called **dimension of V**, denoted by $\dim(V)$.

Proposition 2.1. Suppose $\{v_1,\cdots,v_r\}\subset V$ spans V[2.3], then there is a subset of $\{v_1,\cdots,v_r\}$ which is a basis of V.[2.4]

Proposition 2.2. Suppose V has a finite spanning set. Suppose $\{v_1,\cdots,v_r\}\subset V$ are linearly independent[2.5]. Then we can expand $\{v_1,\cdots,v_r\}$ to $\{v_1,\cdots,v_n\}$, $n\geq r$, s.t., $\{v_1,\cdots,v_n\}$ is a basis.

Proposition 2.3. Suppose $\{v_1,\cdots,v_n\}$ span V and $\{w_1,\cdots,w_m\}$ are linearly independent, then $n\geq m$.

Corollary 2.1. Let $\dim(V)=n$.

1. Any n vectors which span V form a basis of V, and no $n-1$ vectors can span V;
2. Any n linearly independent vectots form a basis of V, and no $n+1$ vectors can be linearly independent.

2.2. *ditto*.

2.3. Note that, we have not guaranteed that the vectors in $\{v_1,\cdots,v_r\}$ are linearly independent. We may reduce them to a set of vectors that are linearly independent.

2.4. The proof is done by **sifting**.

2.5. This set of vectors do not necessarily *span* the vector space V.

Definition 2.9. (Subspace)

A **subspace** $S \subset V$ that is a vector space in its own right using the operations of addition(+) and scalar multiplication($\cdot$) obtained by restricting the addition and scalar multiplication from V to S.

Theorem 2.4. A nonempty set $S \subset V$ is a **subspace** iff S is closed under addition and scalar multiplication, s.t., $\forall \alpha, \beta \in K, \forall u, v \in S \Rightarrow \alpha u + \beta v \in S$.

Example. Let $V = \mathbb{R}^2$, $W = \{(\alpha, \beta) | \beta = 2\alpha\}$, $W \subset V$, and W is a subspace of V. This is verified by checking both addition and scalar multiplication.

Example. Let V be the vector space with $v_1, \cdots, v_n \in V$, then $\text{span}\{v_1, \cdots, v_n\}$ is a subspace of V.

2.2 Linear Transformation

Definition 2.10. (Linear Transformation)

Let $\mathcal{U}, \mathcal{V}$ be vector spaces. Then $T: \mathcal{U} \to \mathcal{V}$ is a **linear transformation** or linear map if:

1. $T(u_1 + u_2) = T(u_1) + T(u_2), \forall u_1, u_2 \in \mathcal{U}$;
2. $T(\alpha u) = \alpha T(u), \forall \alpha \in K, u \in \mathcal{U}$.

The above conditions can be condensed into:

$$\forall u_1, u_2 \in \mathcal{U}, T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2)$$

Property 2.2. The followings are the properties of linear transformation:

1. $T(0_U) = 0_V$;

   Example. Think about geometric transformations. One noteworthy case is ***translation***, which is not linear transformation, since the mapping of 0 results in non-zero result.

2. $T(-u) = -T(u)$;

Example.

1. $T: \mathbb{R}^3 \to \mathbb{R}^2. T(\alpha, \beta, \gamma) = (\alpha, \beta)$;
2. $V = C^\infty(\mathbb{R}), \mathcal{D}: C^\infty(\mathbb{R}) \to C^\infty(\mathbb{R})$, where

$$(\mathcal{D}f)(x) = \frac{df}{dx}(x)$$

3. *Identity map*: $I: U \to U$, $I(u) = u$;
4. *Zero map:* $O_{U,V}: U \to V$, $O_{U,V}(u) = 0_V$.

Theorem 2.5. Let $\mathcal{U}, \mathcal{V}$ be two vector spaces. Let $\{e_1, \cdots, e_n\}$ be a basis of $\mathcal{U}$. For any vectors $v_1, \cdots, v_n \in \mathcal{V}$, there is a unique linear map $T: \mathcal{U} \to \mathcal{V}$, s.t., $T(e_i) = v_i$.

Proof. From previous result, $\forall u \in \mathcal{U}$, $\exists$ unique set of coordinates $(\alpha_1, \cdots, \alpha_n)$, s.t.,

$$u = \alpha_1 e_1 + \cdots + \alpha_n e_n$$

Therefore, we have:

$$T(u) = T(\alpha_1 e_1 + \cdots + \alpha_n e_n) = \alpha_1 T(e_1) + \cdots + \alpha_n T(e_n) = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

Example. $T: \mathbb{R}^3 \to \mathbb{R}^2, T(\alpha, \beta, \gamma) = T(\alpha, \beta)$, that is, $T(e_1) = (1, 0), T(e_2) = (0, 1), T(e_3) = (0, 0)$.We have $u = \alpha e_1 + \beta e_2 + \gamma e_3$, then

$$T(u) = T(\alpha e_1 + \beta e_2 + \gamma e_3) = \alpha T(e_1) + \beta T(e_2) + \gamma T(e_3) = (\alpha, \beta)$$

Definition 2.11. (Kernels and Images)
Given the following linear transformation $T: U \to V$:

1. $\text{image}(T) = \{T(u) | u \in U\} \subset V$;
2. $\ker(T) = \{u \in U | T(u) = 0_V\} \subset U$.

If $T(u_1) = 0 = T(u_2), T(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 T(u_1) + \alpha_2 T(u_2) = 0$.

Both kernels and images are subspaces, then we could discuss on their dimensions.

Definition 2.12. (Rank and Nullity)
$T: \mathcal{U} \to \mathcal{V}$, $\text{rank}(T)$ is the dimension of $\text{im}(T)$, while $\text{null}(T)$ is the dimension of $\ker(T)$.

Example. Given the transformation $T: \mathbb{R}^3 \to \mathbb{R}^2, T(\alpha, \beta, \gamma) = (\alpha, \beta)$. Then we have $\text{image}(T) = \mathbb{R}^2, \ker(T) = \{(0, 0, \gamma) | \gamma \in \mathbb{R}\}$. Hence $\text{rank}(T) = 2$, $\text{null}(T) = 1$, $\dim(U) = 3$.

Example. $U = \mathbb{R}[x]_{\leq n} = \{\alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_n x^n | \alpha_i \in \mathbb{R}\}$. Let $\mathcal{D}: \mathbb{R}[x]_{\leq n} \to \mathbb{R}[x]_{\leq n}$. Then we have $\text{image}(\mathcal{D}) = \mathbb{R}[x]_{\leq n-1}$; $\ker(\mathcal{D}) = \mathbb{R}[x]_0$; $\text{rank}(\mathcal{D}) = n$; $\text{null}(\mathcal{D}) = 1$.

Theorem 2.6. (Rank-Nullity Theorem) [2.6]

Let U, V be finite dimensional vector spaces, and let $T: U \rightarrow V$ be a linear map, then we have:

$$\operatorname{rank}(T) + \operatorname{null}(T) = \dim(U)$$

Theorem 2.7. Let $T: U \rightarrow V$ be a linear transformation. Suppose $\dim(U) = \dim(V) = n$, then the following statements are equivalent:

1. T is surjective(onto);
2. $\operatorname{rank}(T) = n$;
3. $\operatorname{null}(T) = 0$;
4. T is injective(one-to-one);
5. T is bijective.

Proof. Statement (1) states that $\operatorname{im}(T) = V$, which naturally gives (2). [2.7]

Definition 2.13. (Singularity)

$T: U \rightarrow V$. T is **non-singular** if (1) or (2) or (5) hold. Otherwise, T is **singular**. Note that, this is defined only for U and V have the same dimension.

2.3 Matrices

We start with some denotations:

1. $K^{m,n}$: the space of all $m \times n$ matrices. $K^{1,n}$: the space of all row vectors. $K^{n \times 1}$: the space of all column matrices.
2. The denotation of matrix:

$$A = [\alpha_{ij}] = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} \end{bmatrix}$$

3. Identity matrix I_n; zero matrix $0_{m \times n}$,$m, n \in \mathbb{N}$.

Let U, V be two matrices, such that $\dim(U) = n, \dim(V) = m$. Define $T: U \rightarrow V$. Then fix the bases $\{e_1, \cdots, e_n\}$ of U and $\{f_1, \cdots, f_m\}$ of V.

2.6. Check the proof in the notes.

2.7. The following proof to be continued by yourself.

We have the following system:

$$\begin{aligned} T(e_1) &= \alpha_{11}f_1 + \alpha_{21}f_2 + \cdots + \alpha_{m1}f_m \\ T(e_2) &= \alpha_{12}f_1 + \alpha_{22}f_2 + \cdots + \alpha_{m2}f_m \\ &\vdots \\ T(e_n) &= \alpha_{1n}f_1 + \alpha_{2n}f_2 + \cdots + \alpha_{mn}f_m \end{aligned}$$

Where

$$A = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{bmatrix}_{m\times n}$$

Theorem 2.8. $\dim(U) = n, \dim(V) = m$. For given bases of U and V, there is a bijective correspondence between $K^{m,n}$ and $\mathrm{Hom}(U, V)$[2.8].

Example. $T: \mathbb{R}^3 \to \mathbb{R}^2$, $T(\alpha, \beta, \gamma) = (\alpha, \beta), \{e_1, e_2, e_3\} \subset U, \{f_1, f_2\} \subset V, e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$; $f_1 = (1, 0)$, $f_2 = (0, 1)$.

Therefore, $T(e_1) = (1, 0) = 1f_1 + 0f_2; T(e_2) = 0f_1 + 1f_2; T(e_3) = 0f_1 + 0f_2$. Hence, we have

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

We fix the basis $\{e_1, \cdots, e_n\}$ of U, $u \in U$, and we have $u = \lambda_1 e_1 + \cdots \lambda_n e_n$; similarly, we fix the basis $\{f_1, \cdots, f_m\}$ of V, $v \in V$, we have $v = \mu_1 f_1 + \cdots \mu_n f_m$, then we denote:

$$\tilde{u} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix}, \tilde{v} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{pmatrix}$$

Proposition 2.4. (Linear Map)

Given a linear map $T: U \to V$. Fix the bases $\{e_1, \cdots, e_n\}$ of U and $\{f_1, \cdots, f_m\}$ of V, and let A be the matrix corresponding to the transformation, therefore we have:

$$T(u) = v \Leftrightarrow A\tilde{u} = \tilde{v}$$

2.8. homomorphism.

Theorem 2.9. (Operations of Linear Transformation)

1. Given the linear transformations $T_1, T_2: U \to V$ with matrices A_1, A_2. $T_1 + T_2$ has the corresponding matrix $A_1 + A_2$;
2. Given the linear transformations $T_1, T_2: U \to V$ with matrices A_1, A_2. λT_1 has the corresponding matrix λA_1;
3. Given the linear transformations $T_1: V \to W$ with A_1; $T_2: U \to V$; under the premise of feasible matrix multiplication, the corresponding matrix to $T_1 \circ T_2$ is $A_1 A_2$.

Now we focus on the conceptions related *rank of a matrix*.

Definition 2.14. (Row and column spaces)

Let $K^{m \times n}$ be the space of all $m \times n$ matrices, which is composed of row vectors $\{r_1, \cdots, r_m\} \in K^{1,n}$, and alternatively, column vectors $\{c_1, \cdots, c_n\} \in K^{m \times 1}$.

1. Row space of A is the subspace of $K^{1,n}$ spanned by $\{r_1, \cdots, r_m\}$, while the row rank of $A = \dim(\text{row}(A))$;
2. Column space of A is the subspace of $K^{m,1}$ spanned by $\{c_1, \cdots, c_n\}$, while the column rank of $A = \dim(\text{col}(A))$.

Theorem 2.10. Given the linear mapping $T: \mathcal{U} \to \mathcal{V}$, by fixing some bases $\{u_1, \cdots, u_n\}$ and $\{v_1, \cdots, v_m\}$, the matrix representation of T is A. Also, $\text{rank}(T) = \text{col}(A) = \text{row}(A) = \text{rank}(A)$.[2.9]

2.4 Invertibility and Isomorphism

Definition 2.15. (Inverse Matrix)

Let $\mathcal{U}, \mathcal{V}$ be two vector spaces, and $T: \mathcal{U} \to \mathcal{V}$ be a linear transformation. Then the function T^{-1} is said to be the **inverse** of T if $TT^{-1} = I_V, TT^{-1} = I_U$. We then say that T is invertible.

2.9. Proof is left to the references.

Proposition 2.5. If $T:\mathcal{U}\rightarrow\mathcal{V}$ is linear and invertible, then $\dim(\mathcal{U})=\dim(\mathcal{V})$.

Proposition 2.6. Suppose the mapping $T:\mathcal{U}\rightarrow\mathcal{V}$ is linear and invertible, then T^{-1} is also linear.

Definition 2.16. (Invertibility)
Let $A\in K^{n,n}$, then A is **invertible** if $\exists A^{-1}\in K^{n,n}$, s.t., $A^{-1}A=AA^{-1}=I_n$.

Proposition 2.7. Given the linear transformation $T:\mathcal{U}\rightarrow\mathcal{V}$, and A is the matrix form of T. Then T is invertible if and only if A is invertible.

For invertible maps $T, T_1, T_2: U\rightarrow V$ and their corresponding matrices A, A_1, A_2 , the following properties may easily be checked:

1. T^{-1}, A^{-1} are unique;
2. $(T_1T_2)^{-1}=T_2^{-1}T_1^{-1}$ and $(A_1A_2)^{-1}=A_2^{-1}A_1^{-1}$;
3. $(T^{-1})^{-1}=T$ and $(A^{-1})^{-1}=A$.

Definition 2.17. (Isomorphism)
Given two vector spaces $\mathcal{U}$ and $\mathcal{V}$, $\mathcal{U}$ is said to be **isomorphic** to $\mathcal{V}$ if $\exists T$: $U\rightarrow V$ where T is linear and invertible.

Theorem 2.11. Let $\mathcal{U}$ and $\mathcal{V}$ be finite vector spaces, then $\mathcal{U}$ is **isomorphic** to V if and only if $\dim(\mathcal{U})=\dim(\mathcal{V})$.

Inverse of Linear Maps and Matrices.

The study of linear algebra is largely motivated by the study of the system of equations:

$$\begin{gathered}\alpha_{11}x_1+\cdots+\alpha_{1n}x_n=\beta_1\\ \alpha_{21}x_1+\cdots+\alpha_{2n}x_n=\beta_2\\ \vdots\\ \alpha_{m1}x_1+\cdots+\alpha_{mn}x_n=\beta_m\end{gathered}$$

Start with the linear map $T:\mathcal{U}\rightarrow\mathcal{V}$. Fix a basis $\{e_1,\cdots,e_n\}$ of $\mathcal{U}$ and $\{f_1,\cdots,f_m\}$ of $\mathcal{V}$, and we have A be the corresponding matrix to T. Then we have:

$$T(u)=v\Leftrightarrow A\tilde{u}=\tilde{v}$$

The $\Leftarrow$ side is known as **inverse image problem**.[2.10]

Homogeneous Problem.

$$b=0, Ax=0 \Leftrightarrow v=0, T(u)=0.$$

Correspondingly, $\ker(T)=\{u \in \mathcal{U} | T(u)=0\}$, and $\text{null}(A)=\{x \in K^{n,1} | Ax=0\}$.

Proposition 2.8. Take $A \in K^{m,n}, b \in K^{m\times 1}$, if $x^* \in K^{n,1}$ is a **particular solution** to $Ax=b$, then the full set of solutions to $Ax=b$ is

$$x^* + \text{null}(A) = \{x^* + y | Ay = 0\}$$

Proof. $\{x \in K^{n,1} | Ax = b\} = \{x^* + y | Ay = 0\}$.

"$\supset$": $x = x^* + y$ where $Ay = 0$, then $Ax = (Ax^* + y) = Ax^* + Ay = Ax^* = b$;
"$\subset$": $\bar{x}$, where $A\bar{x} = b$, define $y = \bar{x} - x^*$, therefore $Ay = A\bar{x} - Ax^* = b - b = 0$. Hereby $\bar{x} = x^* + y$. This build up their equivalence.

A more general case is for $T(u) = v$ as for $u^* + \ker(T)$.

Example. $V = C^\infty$, with

$$a\frac{d^2y}{dx^2}(x) + b\frac{dy}{dx}(x) + cy(x) = f(x)$$

where a, b, c are known real numbers, while $f(x)$ is a known function. Find $y(x)$.

Solution. $\mathcal{L}y = a\frac{d^2y}{dx^2} + b\frac{dy}{dx} + c$, such that $\mathcal{L}: C^\infty \to C^\infty$. Then $\mathcal{L}y = f \Leftrightarrow (\mathcal{L}y)(x) = f(x)$, $\forall x$, satisfying that $\mathcal{L}(y_1 + y_2) = \mathcal{L}(y_1) + \mathcal{L}(y_2)$, while $\mathcal{L}(\alpha y) = \alpha\mathcal{L}(y)$.

Then the following holds:

$$y = y_P + y_H, \mathcal{L}y_P = f, \mathcal{L}y_H = 0, \mathcal{L}(y_P + y_H) = \mathcal{L}(y_P) + \mathcal{L}(y_H) = f$$

Theorem 2.12. Let $A \in K^{n,n}$

1. The homogeneous system $Ax = 0$ has a non-zero solution if and only if A is singular,s.t., $\text{null}(A) \neq \{0\}$;
2. The non-homogeneous system $Ax = b$ has a unique solution if and only if A is non-singular, s.t., $\text{null}(A) = \{0\}$.[2.11]

2.10. If $v \notin \text{Im}(T)$, then there is no solution to the inverse image problem.

2.11. Such that, the linear map is a bijection.

2.5 Determinants

For $A \in K^{n,n}$, we denote the determinant of A by $\det(A)$ or $|A|$.

Property 2.3. The following properties hold for determinants:

1. $\det(I_n) = 1$;
2. If A has two equal rows, then $\det(A) = 0$;
3. If $A_{n\times n}$ is an upper triangular matrix, then $\det(A) = \prod_{i=1}^{n} a_{ii}$
4. $\det(A^T) = \det(A)$;
5. $\det(AB) = \det(A)\det(B)$;
6. $\det(A) = 0$ if and only if A is non-singular.

2.6 Change of Basis

Let $\mathcal{U}$ be a vector space of dimension n with "old" basis $\{e_1, \cdots, e_n\}$, and "new" basis $\{e'_1, \cdots, e'_n\}$.

The matrix P corresponding to the identity map $I_U : U \to U$ using the basis $\{e_1, \cdots, e_n\}$ in the domain and $\{e'_1, \cdots, e'_n\}$ in the range is called the **change of basis matrix** from the basis of $e_i s$ to the basis of e'_is.

Let $P = (\sigma_{ij})$, then

$$I_U(e_j) = e_j = \sum_{i=1}^{n} \sigma_{ij} e'_i$$

i.e. the jth column of the change of basis matrix P is the coordinates of the *"old"* basis vector e_j.

Example. Let $\mathcal{U} = \mathbb{R}^3$, let $e'_1 = (1,0,0), e'_2 = (0,1,0), e'_3 = (0,0,1), e_1 = (0,2,1)$, $e_2 = (1,1,0), e_3 = (1,0,0)$.

Solution. $e_1 = 0e'_1 + 2e'_2 + 1e'_3$; $e_2 = 1e'_1 + 1e'_2 + 0e'_3$; $e_3 = 1e'_1 + 0e'_2 + 0e'_3$.

Therefore, we have:

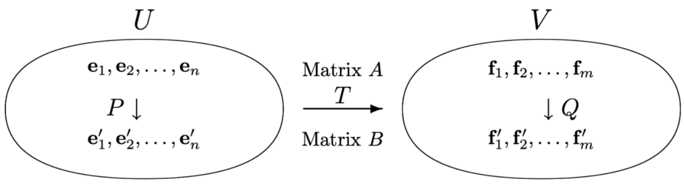
$$P = \begin{pmatrix} 0 & 1 & 1 \\ 2 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

Proposition 2.9. Choose $u \in \mathcal{U}$, let $\tilde{u}$ be coordinates of u using $\{e_i\}_i$, let $\tilde{u}'$ be coordinates of u using $\{e'_i\}_i$, then $P\tilde{u} = \tilde{u}'$.

Proposition 2.10. If $\{e_i\} \to^P \{e_i'\}$, then there exists Q, s.t., $\{e_i'\} \to^Q \{e_i\}$, and we must have $QP = I$.

$T: \mathcal{U} \to \mathcal{V}$, such that $\dim(\mathcal{U}) = \dim(\mathcal{V}) = n$. Suppose $\{e_i\} \subset \mathcal{U}, \{f_i\} \subset \mathcal{V}$, and $T(e_j) = \sum_{i=1}^m \alpha_{ij} f_i$; $\{e_i'\} \subset \mathcal{U}, \{f_i'\} \subset \mathcal{V}$, and $T(e_j') = \sum_{i=1}^m \beta_{ij} f_i'$. Let $A = (\alpha_{ij})$, $B = (\beta_{ij})$. Let $\{e_i\} \to^P \{e_i'\}$, $\{f_i\} \to^Q \{f_i'\}$.

Theorem 2.13. We have $BP = QA$, or equivalently, $B = QAP^{-1}$.[2.12]



Definition 2.18. (Equivalence of Matrices)

Two matrices $A, B \in K^{m,n}$ are said to be **equivalent** if there exists invertible P and Q with $B = QAP^{-1}$.

Theorem 2.14. Let $A, B \in K^{m,n}$. The following conditions are equivalent:

1. A and B are equivalent;
2. A and B represent the same linear map with respect to different bases;
3. A and B has the same rank.

Now consider a specific case, such that $\mathcal{U} = \mathcal{V}$. For simplicity, we write $T: \mathcal{V} \to \mathcal{V}$ as a linear map, there will be a matrix representing T with respect to that basis. Let $\{e_1, \cdots, e_n\}$ and $\{e_1', \cdots, e_n'\}$ be two bases of V, and let $A = (\alpha_{ij})$ and $B = (\beta_{ij})$ be the matrices of T with respect to $\{e_i\}$ and $\{e_i'\}$, respectively. Let $P = (\sigma_{ij})$ be the change of basis matrix from $\{e_i'\}$ to $\{e_i\}$.

Theorem 2.15. With the above notation, $B = P^{-1}AP$.

Definition 2.19. (Similar Matrices)

Two $A, B \in K^{m,n}$ matrices are said to be **similar** if $\exists$ an invertible matrix $P \in K^{m,n}$ with $B = P^{-1}AP$.

So two matrices are similar if and only if they represent the same linear map $T: V \to V$ with respect to different bases of V .

2.12. The proof is to be seen on the references.

2.7 Eigenstuff

Definition 2.20. **(Diagonalizable)**
A matrix which is similar to a diagonal matrix is said to be **diagonalizable**.

It turns out that the possible entries on the diagonal of a matrix similar to A can be calculated directly from A. They are called **eigenvalues** of A and depend only on the linear map to which A corresponds, and not on the particular choice of basis.

Definition 2.21. (Eigenpair)
Let $T: V \to V$ be a linear map, where V is a vector space over K. Suppose that for some non-zero vector $v \in V$ and some scalar $\lambda \in K$, we have $T(v) = \lambda v$. Then v is called an **eigenvector** of T, and λ is called the **eigenvalue** of T corresponding to v.

Example. $T: \mathbb{R}^2 \to \mathbb{R}^2, T(\alpha_1, \alpha_2) = (2\alpha_1, 0)$. Therefore $T(1,0) = (2,0)$. Correspondingly, we have eigenpair $(2,(1,0))$. Also, $T(0,1) = (0,0)$, so we have another eigenpair $(0,(0,1))$.

Let $\{e_1, \cdots, e_n\}$ be a basis of V, and let $A = (\alpha_{ij})$ be the matrix of T with respect to this basis. To each vector $v \in V$, where $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$, we associate its column vector of coordinates,

$$\tilde{v} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix}$$

Then, for $u, v \in V$, we have $T(u) = v$ if and only if $A\tilde{u} = \tilde{v}$, and in particular,

$$T(v) = \lambda v \Leftrightarrow A\tilde{v} = \lambda \tilde{v}$$

Also note that, similar matrices have the same eigenvalues.

Theorem 2.16. Let $A \in K^{n,n}$. Then λ is an eigenvalue of A if and only if $\det(A - \lambda I_n) = 0$, and

$$p(\lambda) = \det(A - \lambda I_n)$$

is known as the **characteristic polynomial** of A.

Corollary 2.2. Suppose A (or linear transformation $T{:}\,\mathcal{V}\to\mathcal{V}$)have n distinct eigenvalues, then A is **diagonalizable**.

Example. Find the eigenvalues of the following matrix:

$$A=\begin{bmatrix} 4 & 5 & 2 \\ -6 & -9 & -4 \\ 6 & 9 & 4 \end{bmatrix}$$

$\lambda_1=0, \lambda_2=1, \lambda_3=-2,\ \tilde{v_1}=[1,-2,3]^T; \tilde{v_2}=[1,-1,1]^T; \tilde{v_3}=[1,-2,2]^T$, and

$$P=\begin{bmatrix} 1 & 1 & 1 \\ -2 & -1 & -2 \\ 3 & 1 & 2 \end{bmatrix}$$

P is the change of basis from standard basis to a basis consisting of eigenvectors of A, and we could arrive at the following result: a diagonal matrix whose entries are eigenvalues.

$$D=P^{-1}AP=\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

Application. Coupled System of ODE

$$\begin{aligned} \dot{x_1}(t)&=\alpha_{11}x_1(t)+\alpha_{12}x_2(t) \\ \dot{x_2}(t)&=\alpha_{21}x_1(t)+\alpha_{22}x_2(t) \end{aligned}$$

Denote $\boldsymbol{x}(t)=(x_1(t),x_2(t))$, and $\dot{\boldsymbol{x}(t)}=(\dot{x_1}(t),\dot{x_2}(t))$, therefore, we have

$$\dot{\boldsymbol{x}}(t)=A\boldsymbol{x}(t)$$

with initial value $\boldsymbol{x}(0)=x_0$. By factorization, we have $P^{-1}AP=D$.

Therefore $\boldsymbol{y}(t)=P^{-1}\boldsymbol{x}(t)\Leftrightarrow\boldsymbol{x}(t)=P\boldsymbol{y}(t)$, where P is **decoupled**. Following the above steps, we have:

$$[P\dot{\boldsymbol{y}(t)}]=AP\boldsymbol{y}(t)\Rightarrow\dot{\boldsymbol{y}(t)}=\underbrace{P^{-1}AP}_{D}\boldsymbol{y}(t)=D\dot{\boldsymbol{y}(t)}$$

Resulting that

$$\begin{bmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{bmatrix}=\underbrace{\begin{bmatrix} e^{\lambda_1 t} & & & \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ & & & e^{\lambda_n t} \end{bmatrix}}_{E(t)}\begin{bmatrix} y_1(0) \\ y_2(0) \\ \vdots \\ y_n(0) \end{bmatrix}$$

Therefore, we have $\boldsymbol{y}(t) = E(t)\boldsymbol{y}(0)$, by plugging back, we have:

$$P^{-1}\boldsymbol{x}(t) = E(t)P^{-1}\boldsymbol{x}(0)$$

Thus, we have:

$$\boldsymbol{x}(t) = PE(t)P^{-1}\boldsymbol{x}(0)$$

2.8 Inner Product, Symmetricity and Orthogonality

Definition 2.22. (Inner Product)

Let V be a vector space over $\mathbb{R}$. An **inner product** on V is a function that assigns, to every pair of vectors (u, v) in $V \times V$, a scalar in $\mathbb{R}$, such that for all $u, v, w \in V$ and $\alpha \in K$, the following hold:

1. $\langle u + w, v\rangle = \langle u, v\rangle + \langle w, v\rangle$;
2. $\langle \alpha u, v\rangle = \alpha \langle u, v\rangle$;
3. $\langle u, v\rangle = \langle v, u\rangle$;
4. If $u \neq 0$, then $\langle u, u\rangle > 0$.

Example.

1. Let $V = \mathbb{R}^n$ and let $u, v \in \mathbb{R}^n$ with $u = (\alpha_1, \cdots, \alpha_n), v = (\beta_1, \cdots, \beta_n)$. A commonly used inner product is,
$$\langle u, v\rangle = \sum_{i=1}^{n} \alpha_i \beta_i$$
2. Let $V = C[0, 1]$ be the space of all continuous functions on $[0, 1]$. For f, $g \in V$, the following is an inner product space:
$$\langle f, g\rangle = \int_0^1 f(x) g(x)\, dx$$

Definition 2.23. (Norm)

Let V be an inner product space. For $v \in V$, we define the **norm** or **length** of v by $\|v\| = \sqrt{\langle v, v\rangle}$.

Definition 2.24. (Orthogonality)

$u, v \in V$ are **orthogonal** or perpendicular if $\langle u, v\rangle = 0$. A subset $S \subset V$ is said to be **orthogonal** if any two distinct vectors in S are orthogonal.

A vector u in V is a **unit vector** if $\|u\|=1$.

Definition 2.25. (Orthonormality)
A subset S of V is **orthonormal** if S is orthogonal and consists entirely of unit vectors.

Definition 2.26. (Symmetricity)
We say $A \in K^{n,n}$ is **symmetric** if $A = A^T$.

Definition 2.27. (Orthogonal Matrix)
We say $A \in K^{n,n}$ is **orthogonal** if $A^{-1} = A^T$,s.t.,

$$\langle Au, Av\rangle = (Au)^T Av = u^T A^T A v = u^T v = \langle u, v\rangle$$

Theorem 2.17. Let $A \in K^{n,n}$ be a real symmetric matrix. Then there exists a real orthogonal matrix P with $P^{-1}AP = P^T AP$ is diagonal.

3

Exercises I

These exercises are worked out or excerpted by Tom Hierons.

3.1 Basic Language of Math

Exercise 3.1. Consider $f(x) = x^2$.

a) $f\colon \mathbb{R} \to \mathbb{R}$; Neither injective, nor surjective.
b) $f\colon \mathbb{R}_+ \to \mathbb{R}$; Injection
c) $f\colon \mathbb{R} \to \mathbb{R}_+$; Surjection
d) $f\colon \mathbb{R}_- \to \mathbb{R}_+ (\mathbb{R}_- \equiv \{x \in \mathbb{R} \colon x \leq 0\})$ Bijection

Exercise 3.2. **(Cardinality)** (a) Q; (b) $\bigcup_{n=1}^{\infty} \{n, n+1, n+2\}$; (c)$\bigcap_{n=1}^{\infty} \{n, n+1, n+2\}$;(d)$\bigcup_{n=1}^{\infty} \left[1, 1 + \frac{1}{n}\right]$;(e)$\bigcap_{n=1}^{1+n} \left[1, 1 + \frac{1}{n}\right]$;(f)$\{X \colon X \subseteq \mathbb{N}\}$.

Check out which of these are finite, countably infinite, and uncountable.

(a) (Rational Numbers) countably infinite;
(b) $= \mathbb{N}$, countably infinite;
(c) $= \varnothing$, finite;
(d) $= [1, 2]$, uncountably infinite, by mapping;
(e) $= \{1\}$, finite;

(f) Power set of natural number, uncountably infinite.

3.2 Linear Algebra

Exercise 3.3. (Definiteness and Cholesky Factorization)

Let A be an $n \times n$ real matrix. A is said to be positive semi-definite if it is symmetric and $x^T A x \geq 0$ for any $x \in \mathbb{R}$. A is said to be positive definite if it is symmetric and $x^T A x > 0$ for $x \in \mathbb{R} \backslash \{0\}$.

If we can write $A = LL^T$ for a lower triangular matrix L then this is known as the **Cholesky factorization** of A.

1. Show that if A is positive definite then it is invertible.

Proof. Proof by contradiction. Given that $x^T A x > 0$ $\forall x \in \mathbb{R} \backslash \{0\}$ while A is not invertible, then we have $Ax = 0$ for some $x \in \mathbb{R} \backslash \{0\}$, a contradiction.

2. Given an example of a positive semi-definite matrix that is not invertible.

Solution.
$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

3. Show that if A is positive semi-definite but not positive definite then it is not invertible.

Proof. Solve by using eigenvalue. There exists a zero eigenvalue corresponds to the matrix A, indicating A does not have full rank.

4. Show that if A has a Cholesky factorization then it is positive semidefinite.

Proof. If A has a Cholesky factorization, then for $x \in \mathbb{R}^n$, we have $A = LL^T$, hence $x^T A x = (L^T x)^T (L^T x) \geq 0$.

5. Show that A is positive semidefinite if and only if it has a Cholesky decomposition. Show that A is positive definite if and only if it has a unique Cholesky decomposition.

Proof. (1)$\Leftarrow$: From 4; $\Rightarrow$:Given that A is positive semidefinite, then $\forall x \in \mathbb{R}^n$, we have

$$x^T A x \geq 0$$

$\Rightarrow$: Proof by induction. When $n = 1$, $A = (a)$. Obviously, when A is positive semidefinite, i.e., $a \geq 0$, we have $L = (\pm\sqrt{a}) = L^T$. Now suppose when $n = k$, A_k can be factorized in style of $A_k = LL^T$, we are now to induce that, when $n = k + 1$, A_{k+1} can also be factorized in form of $A_{k+1} = L*(L*)^T$.

Now we have A' be factorized into

$$A' = \begin{pmatrix} c & \boldsymbol{v}^T \\ \boldsymbol{v} & A \end{pmatrix} = \begin{pmatrix} \sqrt{c} & 0 \\ \frac{\boldsymbol{v}}{\sqrt{c}} & I_k \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \bar{A} \end{pmatrix}\begin{pmatrix} \sqrt{c} & \frac{\boldsymbol{v}^T}{\sqrt{c}} \\ 0 & I_k \end{pmatrix}$$

where $\boldsymbol{v} \in \mathbb{R}^k$, and c is a scalar, and $\bar{A} = A - \frac{\boldsymbol{v}\boldsymbol{v}^T}{c}$. It is obvious that $\bar{A}$ is also symmetric, and it is also positive semidefinite since for any nonzero vector $\boldsymbol{x}$ of length k, we have

$$\boldsymbol{x}^T\bar{A}\boldsymbol{x} = \boldsymbol{x}^T\left(A - \frac{\boldsymbol{v}\boldsymbol{v}^T}{c}\right)\boldsymbol{x} = \begin{pmatrix} -\frac{\boldsymbol{x}^T\boldsymbol{v}}{c} & \boldsymbol{x}^T \end{pmatrix}\begin{pmatrix} c & \boldsymbol{v}^T \\ \boldsymbol{v} & A \end{pmatrix}\begin{pmatrix} -\frac{\boldsymbol{x}^T\boldsymbol{v}}{c} \\ \boldsymbol{x}^T \end{pmatrix} \geq 0$$

By the previous induction, $\bar{A}$ has a Cholesky decomposition, denoted by $\bar{A} = \bar{L}\bar{L}^T$, therefore, we have

$$A' = \begin{pmatrix} \sqrt{c} & 0 \\ \frac{\boldsymbol{v}}{\sqrt{c}} & I_k \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \bar{L} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \bar{L}^T \end{pmatrix}\begin{pmatrix} \sqrt{c} & \frac{\boldsymbol{v}^T}{\sqrt{c}} \\ 0 & I_k \end{pmatrix} = L * (L*)^T$$

(2) $\Leftarrow$: Trivial and following the steps of (1).

$\Rightarrow$:

Claim. If there exists $L_1, L_2, L_1 \neq L_2$,s.t.,

$$A = L_1 L_1^T = L_2 L_2^T$$

Then we have $L_2 = D L_1$, where D is a diagonal matrix.

Suppose $\exists L_1, L_2, L_1 \neq L_2$, s.t.,

$$A = L_1 L_1^T = L_2 L_2^T$$

This gives that

$$I = L_1^{-1} L_2 L_2^T (L_1^{-1})^T = (L_1^{-1} L_2)(L_1^{-1} L_2)^T \Rightarrow (L_1^{-1} L_2) = ((L_1^{-1} L_2)^T)^{-1}$$

This means that $L_1^{-1} L_2$ is both upper and lower triangular, s.t., $L_1^{-1} L_2$ is diagonal. Moreover, by the above equation, we have $L_1^{-1} L_2$ has all its diagonal entries be ± 1, and since $L_1 \neq L_2$, therefore the two factorizations differ from each other by the signs of their columns.

Therefore, after adding the restrictions on "uni-sign", we have the proposition proved.

Exercise 3.4. (Matrix Powers)
Let $A = \begin{bmatrix} 7 & -12 \\ 4 & -7 \end{bmatrix}$.

a) Write A in the form of $P\mathrm{DP}^{-1}$ where D is diagonal.

$$A = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$$

b) Find A^{100} and A^{101}.

$$A^{100} = I; A^{101} = A$$

Exercise 3.5. (The Trace and More on Determinants) The trace of an $n \times n$ matrix $A = [a_{ij}]$ is defined by

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}$$

a) Show that the trace is a linear function.

Proof. We need to show the following: $\mathrm{tr}(A+B) = \mathrm{tr}(A) + \mathrm{tr}(B); \mathrm{tr}(kA) = k\,\mathrm{tr}(A); \mathrm{tr}(O) = 0$, which are obviously true.

b) For $n \times n$ matrices A, B, show that $\mathrm{tr(AB)} = \mathrm{tr(BA)}$.

Proof.

$$\mathrm{tr(AB)} = \sum_{j=1}^{n} \sum_{i=1}^{n} a_{ji} b_{ij}; \mathrm{tr(BA)} = \sum_{i=1}^{n} \sum_{j=1}^{n} b_{ij} a_{ji}.$$

c) Suppose A is diagonalizable. Show that the trace of A is given by the sum of its eigenvalues.

Proof. Since A is diagonalizable, therefore we write $A = P^T D P$, by results of b), we have:

$$\mathrm{tr}(A) = \mathrm{tr}(P^T D P) = \mathrm{tr}(D P P^T) = \mathrm{tr}(D) = \sum_{i=1}^{n} \lambda_i$$

d) Suppose A is diagonalizable. Show that the determinant of A is given by the product of its eigenvalues.

Proof. Since A is diagonalizable, therefore we write $A = P^T D P$, for P is invertible, and therefore

$$\det(A) = \det(P^T D P) = \det(P^T)\det(D)\det(P) = \det(D)$$

Of which D is triangular and hence $\det(D) = \det(A) = \prod_{i=1}^{n} \lambda_i$.

Exercise 3.6. (Eigenvalues of Some Special Matrices) Characterize the eigenvalues of the following:

a) If A is **idempotent**, i.e., $A^2 = A$, show its eigenvalues are 0 or 1.

Proof. There are two cases, either $A = I$ or $A = O$, then correspondingly, we have eigenvalues of I be 1, and those of O be 0.

b) If A is **nilpotent**, i.e., $A^k = 0, \exists k$, show its eigenvalues are 0.

Proof. We have $\det(A^k) = [\det(A)]^k = 0$, hence $\det(A) = 0$, then by old results, the eigenvalues of A are 0.

c) If A is **involutory**, i.e., $A^2 = I$, show its eigenvalues are ± 1.

Proof. Again, we have $\det(I) = \det(A)\det(A) = 1$, therefore either $\det(A) = 1$ or $\det(A) = -1$, indicating that eigenvalues of A are either 1 or -1.

d) If A is **Hermitian**, i.e.,$A^H = A$, show its eigenvalues are in $\mathbb{R}$.

Proof. Let (λ, v) be an eigenpair of A, then $<v^H, v> = \sum_{i=1}^n |v_i|^2 \geq 0$. Since v is an eigenvector, therefore $\exists v_i \neq 0$, hence $<v^H, v> = 0$.

Thus,

$$\begin{aligned} v^H A v &= v^H(\lambda v) = \lambda v^* v \\ &= v^H A^H v = (Av)^H v = (\lambda v)^H v = \bar{\lambda} v^* v \end{aligned}$$

Therefore, we have $\lambda v^* v = \bar{\lambda} v^* v \Rightarrow (\lambda - \bar{\lambda}) v^* v = 0$, hence $\lambda = \bar{\lambda}$, indicating that its eigenvalues are real.

e) If A is **unitary**, i.e., $A^H A = I$, show its eigenvalues are unit complex numbers in $\mathbb{C}$, or, have modulus 1.

Proof. First we have the following two relations: $Av = \lambda v$, and by taking its conjugate transpose $(Av)^H = v^H A^H = \lambda^H v^H$. Hence we multiply our two relations to obtain:

$$\begin{aligned} v^H A^H A v &= \lambda^H v^H \lambda v \\ v^H I v &= (\lambda^H \lambda) v^H v \\ v^H v &= (\lambda^H \lambda) v^H v \\ \|v\|^2 &= |\lambda|^2 \|v\|^2 \\ 1 &= |\lambda| \end{aligned}$$

3.3 Differential Equations

Exercise 3.7.

a) **(First-order, linear, constant coefficients)**

$$\dot{m(t)} = \bar{r}m(t) - \bar{c}, m(0) = m_0$$

Solve for $m(t)$ as a function of $\bar{r}, \bar{c}, m_0$.

General Recipe. (First-order,linear ODE)

1. Write $\dot{x(t)} + a(t)x(t) = b(t)$;
2. Integrating factor: $e^{\int a(t)dt}[\dot{x(t)} + a(t)x(t)] = e^{\int a(t)dt}b(t)$
3. Differential form:
$$\frac{d}{dt}[e^{\int a(t)dt}x(t)] = e^{\int a(t)\,dt}b(t)$$
4. Use FTC to integrate
$$e^{\int a(t)d(t)}x(t) = \int e^{\int a(t)dt}b(t)\,dt + C$$
5. General solution.
$$x(t) = e^{-\int a(t)dt}\left[\int e^{\int a(t)dt}b(t)dt + C\right]$$
6. Use $x(t_0) = x_0$ to solve for a particular solution:
$$x(t) = x_0 e^{-\int_{t_0}^{t} a(s)ds} + \int_{t_0}^{t} e^{-\int_{s}^{t} a(z)dz}b(s)ds$$

[Economic Intepretation] The stock problem.

$$\dot{m(t)} - \bar{r}m(t) = -\bar{c} \Rightarrow e^{\int -\bar{r}dt}[\dot{m(t)} - \bar{r}m(t)] = e^{\int -\bar{r}dt}(-\bar{c})$$

Then by differentiating on both sides:

$$\frac{d}{dt}[e^{\int -\bar{r}dt}m(t)] = \frac{d}{dt}[e^{\int -\bar{r}dt}(-\bar{c})] \Rightarrow m(t)e^{-\bar{r}t} = \frac{\bar{c}}{\bar{r}}e^{-\bar{r}t} + C$$

By integration, we have the general solution

$$m(t) = \frac{\bar{c}}{\bar{r}} + Ce^{\bar{r}t}$$

with specific solution:

$$m(0) = m_0 = \frac{\bar{c}}{\bar{r}} + C \rightarrow C = m_0 - \frac{\bar{c}}{\bar{r}}$$

Therefore,

$$m(t) = \frac{\bar{c}}{\bar{r}} + \left(m_0 - \frac{\bar{c}}{\bar{r}}\right)e^{\bar{r}t}$$

b) **(First-order, linear, variable coefficients)**

$$\dot{m(t)} = r(t)m(t) - c(t), m(0) = m_0$$

Solve for $w(t)$ as a function $r(\cdot), c(\cdot), m_0$.

$$\dot{m(t)} - r(t)m(t) = -c(t) \Rightarrow e^{\int -r(t)dt}[\dot{m(t)} - r(t)m(t)] = e^{\int -r(t)dt}(-c(t))$$

Then by differentiating on both sides and integrate:

$$\frac{d}{dt}[e^{\int -r(t)dt}m(t)] = \frac{d}{dt}[e^{\int -r(t)dt}(-c(t))] \Rightarrow m(t)e^{-r(t)t} = \frac{c(t)}{r(t)}e^{-r(t)} + C$$

By integeration, we have the general solution

$$m(t) = \frac{c(t)}{r(t)} + Ce^{r(t)t}$$

with specific solution:

$$m(0) = m_0 = \frac{c(t)}{r(t)} + C \rightarrow C = m_0 - \frac{c(t)}{r(t)}$$

Therefore,

$$m(t) = \frac{c(t)}{r(t)} + \left(m_0 - \frac{c(t)}{r(t)}\right)e^{\bar{r}t}$$

c) **(Separable)**

$$Y(t) = A(t)K(t)^\alpha$$
$$\dot{K(t)} = sY(t)$$
$$\dot{A(t)} = g$$
$$A(0) = A_0, K(0) = K_0$$

The recipe. Solve for $Y(t)$ as a function of A_0, K_0, α, g, s.

1. Rearrange the function

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

2. Differentiation

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

3. Integration

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

4. Solve and check for zeros of $g(\cdot)$

We have:

$$\frac{d}{dt}A(t) = gA(t) \Rightarrow A(t) = A_0 e^{gt} \ldots [\text{Malthus Growth Model}]$$

$$\dot{K(t)} = sA_0 e^{gt} K(t)^{\alpha} \Rightarrow \frac{dK(t)}{dt} K(t)^{-\alpha} = sA_0 e^{gt}$$

Hence, we have

$$\frac{1}{1-\alpha} K(t)^{1-\alpha} = \frac{sA_0}{g} e^{gt} + C$$

Then, by using $K(0) = K_0$, we have:

$$C = \frac{1}{1-\alpha} K_0^{1-\alpha} - \frac{sA_0}{g}$$

Therefore,

$$K(t) = \left[\frac{(1-\alpha)sA_0}{g} e^{gt} + (1-\alpha)C\right]^{\frac{1}{1-\alpha}}; Y(t) = A_0 e^{gt} K(t)^{\alpha}$$

Exercise 3.8. (Linear Systems and Stability)

a) Find the solution to the system of differential equations:

$$\underbrace{\begin{bmatrix} \dot{y_1(t)} \\ \dot{y_2(t)} \end{bmatrix}}_{\dot{\boldsymbol{y}}(t)} = \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}}_{D} \underbrace{\begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix}}_{\boldsymbol{y}(t)}$$

Given $\boldsymbol{y}(0) = \boldsymbol{y}_0$. Throughout, we assume that λ_1 and λ_2 are nonzero.

D has two eigenvalues, λ_1 and λ_2. Then eigenvector corresponding to λ_1 is $(1,0)^T$, and that of λ_2 is $(0,1)^T$. Therefre we have the basic solution matrix as:

$$\Phi(t) = \begin{bmatrix} e^{\lambda_1 t} & 0 \\ 0 & e^{\lambda_2 t} \end{bmatrix}$$

and general solution:

$$y(t) = C_1 \begin{bmatrix} e^{\lambda_1 t} \\ 0 \end{bmatrix} y_1 + C_2 \begin{bmatrix} 0 \\ e^{\lambda_2 t} \end{bmatrix} y_2$$ [3.1]

b) Suppose $\lambda_1, \lambda_2 \in \mathbb{R}$, under what conditions will $\boldsymbol{y}(t) \to \mathbf{0}$ as $t \to \infty$ for any value of $\boldsymbol{y}_0$. For $\lambda_1, \lambda_2 < 0$.

3.1. Tom's answer form: $y(t) = \begin{bmatrix} y_{10} e^{\lambda_1 t} \\ y_{20} e^{\lambda_2 t} \end{bmatrix}$

c) Repeat b) but $\lambda_1, \lambda_2 \in \mathbb{C}$.(By using Euler[3.2])

We have $\lambda = a + b\,i$, and therefore we have $y_{10}e^{\lambda_1 t} = y_{10}e^{(a+bi)t} = y_{10}e^{at}e^{i(bt)}$, where $e^{i(bt)} = \cos(bt) + i\,\mathrm{sin}t(bt)$, which oscillates within the span of a unit circle. Hence, we only need the real parts of $\lambda_1, \lambda_2 < 0$.

d) Let

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

And suppose that $A = PDP^{-1}$ where D is the matrix given in a) consisting of the eigenvalues of A and P is an invertible matrix consisting of the corresponding eigenvectors. Let $\boldsymbol{x}(t) = P\boldsymbol{y}(t)$. 1) Show that $\dot{\boldsymbol{x}(t)} = A\boldsymbol{x}(t)$. 2) for what values of $\boldsymbol{x}(t)$ is $\dot{\boldsymbol{x}(t)} = 0$.

$$\dot{\boldsymbol{x}(t)} = \frac{d}{dt}(P\boldsymbol{y}(t)) = P\dot{\boldsymbol{y}(t)} = PD\,\boldsymbol{y}(t) = PDP^{-1}\boldsymbol{x}(t) = A\,\boldsymbol{x}(t)$$

For A has non-zero eigenvalues, hence only if $\boldsymbol{x}(t) = 0$, we have $\dot{\boldsymbol{x}(t)} = 0$.

e) Using the previous parts, give conditions on A so that $\boldsymbol{x}(t) \to 0$ as $t \to \infty$ for any value of $\boldsymbol{x}_0$.

From d), we just need $\boldsymbol{y}(t) \to 0$. Therefore, in general, we just need the real parts of the eigenvalues of A are strictly negative.

f) Suppose that

$$\dot{\boldsymbol{w}(t)} = A\boldsymbol{w}(t) + \boldsymbol{b}$$

For some $\boldsymbol{b} \in \mathbb{R}^2$. Find the steady state of the system (where $\dot{\boldsymbol{w}(t)} = 0$) and suggest a transformation $\boldsymbol{x}(t) = g(\boldsymbol{w}(t))$ so that $\boldsymbol{x}(t)$ satisfies $\dot{\boldsymbol{x}(t)} = A\boldsymbol{x}(t)$. Then interpret.

$$A\boldsymbol{w}^* + b = 0 \Rightarrow \boldsymbol{w}^* = -A^{-1}b$$

Take $\boldsymbol{x}(t) = \boldsymbol{w}(t) - \boldsymbol{w}^* = \boldsymbol{w}(t) + A^{-1}b$, then

$$\dot{\boldsymbol{x}(t)} = \dot{\boldsymbol{w}(t)} = A\boldsymbol{w}(t) + b = A(\boldsymbol{x}(t) - A^{-1}b) + b = A\,\boldsymbol{x}(t)$$

3.2. Euler's equation: $e^{it} = \cos(t) + i\sin(t)$.

3.4 Calculus

Exercise 3.9. (Norms, Metrics, and Topologies)
Let X be a vector space over $\mathbb{R}$.

a) Let $\|\cdot\| = d(x,0)$ be a norm on X. Construct a metric $d(\cdot,\cdot)$ on X using the norm and show that it satisfies the definition of a metric.
Define $d'(x,y) = \|x-y\|$. Therefore we have:
- $d'(x,y) \geq 0$ and $''='' $ iff $x = y$;
- $d(x,y) = d(y,x) = \|x-y\|$;
- $d(x,z) = \|x-z\|$; $d(x,y) + d(y,z) = \|x-y\| + \|y-z\| \geq \|(x-y) + (y-z)\|$ by triangular inequality.

b) Let $d(\cdot,\cdot)$ be a metric on X. Construct a *topology* on X and show that it satisfies the definition of a *topology*.

Definition. (Topology) X is a set. A topology on X is a set of subsets τ of X with the following properties:
- Whenever $(U_i)_{i\in I}$ is a family of subsets of X, s.t., $U_i \in \tau, \forall i \in I$, then $\cup_{i\in I} U_i \in \tau$;
- Whenver $U_1, U_2 \in \tau$, then $U_1 \cap U_2 \in \tau$;
- $\varnothing \in \tau$ and $X \in \tau$.

Consider any open ball on metric space X in the following format:

$$B = \{y \in X : d(x,y) < r\}$$

Let $\tau = \{U \subseteq X \,|\, \forall x \in U : \exists B \in \mathcal{B}, s.t., x \in B \subset U\}$, where $\mathcal{B}$ is the family of all open balls in (X,d). Therefore,
- $\forall x \in X, \exists B \in \mathcal{B}$, s.t., $x \in B$
- $\forall\ B_1, B_2 \in \mathcal{B}$, if $x \in B_1 \cap B_2$, then $\exists B_3 \in \mathcal{B}$, s.t., $x \in B_3 \subseteq B_1 \cap B_2 \in \mathcal{B}$.

Then we have the properties of the definition be immediate.

Exercise 3.10. (Homogeneous Functions)
Let $f: \mathbb{R}^n \to \mathbb{R}$ and assume throughout that it is C^2, f is said to be *homogeneous of degree* $k \in \mathbb{R}$ if for any $\lambda \in \mathbb{R}$ and $\boldsymbol{x} \in \mathbb{R}^n$, $f(\lambda \boldsymbol{x}) = \lambda^k f(\boldsymbol{x})$.

Cobb-Douglas function $f(x,y) = x^\alpha y^{1-\alpha}$, $f(\beta x, \beta y) = \beta f(x,y)$.

a) **(Euler's Theorem)** Show that if f is homogeneous of degree k, then

$$k f(\boldsymbol{x}) = \sum_{i=1} x_i \frac{\partial f}{\partial x_i}(\boldsymbol{x})$$

$$\frac{\partial\left[\lambda^k f(\boldsymbol{x})\right]}{\partial\lambda} = k\lambda^{k-1} f(\boldsymbol{x}) = \frac{\partial f}{\partial(\lambda x_1)}\frac{\partial(\lambda x_1)}{\partial\lambda} + \cdots + \frac{\partial f}{\partial(\lambda x_n)}\frac{\partial(\lambda x_n)}{\partial\lambda}$$
$$= x_1\frac{\partial f}{\partial(\lambda x_1)} + \cdots + x_n\frac{\partial f}{\partial(\lambda x_n)}$$

Let $\lambda = 1$, we have:

$$k f(\boldsymbol{x}) = \sum_{i=1} x_i \frac{\partial f}{\partial x_i}(\boldsymbol{x})$$

b) Show that if f is homogeneous of degree k, then for $i = 1, \cdots, b, \frac{\partial f}{\partial x_i}(\boldsymbol{x})$ is homogeneous of degree $k - 1$.

$$\frac{\partial f}{\partial x_i}(\lambda\boldsymbol{x}) = \lim_{\boldsymbol{h}\to 0}\frac{f(\lambda(\boldsymbol{x}+\boldsymbol{h})) - f(\lambda\boldsymbol{x})}{\lambda\boldsymbol{h}}$$
$$= \lambda^{k-1}\lim_{\boldsymbol{h}\to 0}\frac{f(\boldsymbol{x}+\boldsymbol{h}) - f(\boldsymbol{x})}{\boldsymbol{h}}$$
$$= \lambda^{k-1}\frac{\partial f}{\partial x_i}(\boldsymbol{x})$$

c) **(Wicksell's law)** Suppose that $f: \mathbb{R}_+^2 \to \mathbb{R}$ is homogeneous of degree 1, and that $\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}) < 0$ for any $\boldsymbol{x} \in \mathbb{R}_+^2$. Show that $\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}) > 0$ for any $\boldsymbol{x} \in \mathbb{R}_+^2$.

As what have been given, we have

$$f(\lambda x_1, \lambda x_2) = \lambda f(x_1, x_2)$$

and by using

$$k f(\boldsymbol{x}) = \sum_{i=1} x_i \frac{\partial f}{\partial x_i}(\boldsymbol{x})$$

we have:

$$\frac{\partial f}{\partial x_1} = \frac{\partial^2 f}{\partial x_1 \partial x_2} x_1 + \frac{\partial^2 f}{\partial x_1^2} x_1 + \cdot \frac{\partial f}{\partial x_1}$$

given $k = 1$. Then the result is obvious.

Exercise 3.11. Calculus problems.

a) **(L'Hopital's)** CES preferences are often parameterized as $u: \mathbb{R}_+^n \to \mathbb{R}$ where

$$u(\boldsymbol{x}) = \left[\sum_{i=1}^{n} (\beta_i x_i)^\rho\right]^{\frac{1}{\rho}}$$

with $\beta_i > 0$ and $\rho \in (-\infty, 1]$. For a given x derive this limit a $\rho \to -\infty$.

L'Hopital's Rule. f and g are *almost everywhere differentiable* and

$$\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0 \text{ or } \pm\infty$$

where $g'(x) \neq 0, \forall x \in I \setminus \{0\}$, then we have:

$$\lim_{x \to c} \frac{f(x)}{g(x)} = \lim_{x \to c} \frac{f'(x)}{g'(x)}$$

First let $\bar{L} = [\min\{\beta_1 x_1, \cdots, \beta_n x_n\}]$,

$$u(\boldsymbol{x}) \equiv \bar{L}\left[\sum_{i=1}^{n} \left(\frac{\beta_i x_i}{\bar{L}}\right)^{\rho}\right]^{\frac{1}{\rho}}$$

Then take the logarithm form:

$$\log(u) = \log(\bar{L}) + \frac{1}{\rho}\log\left(\sum_{i=1}^{n} \left(\frac{\beta_i x_i}{\bar{L}}\right)^{\rho}\right)$$

Hence, we have $u(\boldsymbol{x}) = \bar{L}$ when $\rho \to -\infty$.

b) **(Leibniz's)** In continuous time, the present discounted value of an asset paying a continuous flow of income at time t of $y(t)$ from time 0 to T and paying 0 thereafter, given a discount rate ρ, is given by:

$$V(\rho, T) = \int_0^T y(t) e^{-\rho t}\, dt$$

Find $\frac{\partial V}{\partial \rho}(\rho, T)$ and $\frac{\partial V}{\partial T}(\rho, T)$ stating any assumptions needed on $y(t)$.

(Leibniz Integral Rule)

$$\begin{aligned} &\frac{d}{dx}\left(\int_{a(x)}^{b(x)} f(x,t) dt\right) \\ = \;& f(x, b(x)) \cdot \frac{d}{dx} b(x) - f(x, a(x)) \cdot \frac{d}{dx} a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x} f(x,t)\, \mathrm{dt} \end{aligned}$$

$$\begin{cases} \frac{\partial V}{\partial \rho}(\rho, T) = \int_0^T \frac{\partial}{\partial \rho} y(t) e^{-\rho t}\, dt = [-t y(t) e^{-\rho t}]_0^T = -t[y(T) e^{-\rho T} - y(0)] \\ \frac{\partial V}{\partial T}(\rho, T) = y(T) e^{-\rho T} \end{cases}$$

$y(t)$ need to be continuous, differentiable, and has fixed integral limits.

c) **(Log-linearization)** Take a first order Taylor expansion of $f(K, L)$ around $(\bar{K}, \bar{L})$ in logs. Under what conditions on f does this imply that the approximation can be written in the form $\ln f(K, L) \simeq \ln A + \alpha \ln K + (1 - \alpha) \ln L$? Interpret.

$$\begin{aligned} \ln f(K, L) &= \ln f(e^{\ln K}, e^{\ln L}) \\ &\approx \ln f(\bar{K}, \bar{L}) + \frac{\partial \ln f}{\partial \ln K}(\ln K - \ln \bar{K}) + \frac{\partial \ln f}{\partial \ln L}(\ln L - \ln \bar{L}) \end{aligned}$$

Letting $\alpha = \frac{\partial \ln f}{\partial \ln K}, 1-\alpha = \frac{\partial \ln f}{\partial \ln L}$, s.t., when $f(K,L)$ follows a Cobb-Douglas Function, where α represents output elasticity, then we have

$$\ln A = \ln f(\bar{K}, \bar{L}) - \alpha \ln \bar{K} - (1-\alpha) \ln \bar{L}.$$

d) **(Example)** Consider the closed economy accounting identity

$$y_t = c_t + i_t$$

1. Take logs and compute the first order Taylor series approximation around y^*, c^*, i^*, simplify;

$$\ln(y_t) = \ln(c_t + i_t)$$

Therefore, by using first order Taylor series approximation, we have:

$$\ln y^* + \frac{1}{y^*}(y_t - y^*) = \ln(c_t + i_t) + \frac{1}{c_t^* + i_t^*}[(c_t - c^*) + (i_t - i^*)]$$

Then we have:

$$\frac{1}{y^*}(y_t - y^*) = \frac{1}{c_t^* + i_t^*}[(c_t - c^*) + (i_t - i^*)]$$

2. Assume that y^*, c^*, i^*, correspond to the steady-state of the system. Compute deviations from the steady state.

By using $\tilde{y}_t = y_t - y^*, \tilde{c}_t = c_t - c, \tilde{i}_t = i_t - i$, following the last result, we have:

$$\frac{1}{y^*}(y_t - y^*) = \frac{c^*}{c_t^* + i_t^*} \cdot \frac{(c_t - c^*)}{c^*} + \frac{i^*}{c_t^* + i_t^*} \cdot \frac{(i_t - i^*)}{i^*}$$

This gives that:

$$\tilde{y}_t = \frac{c^*}{y^*}\tilde{c}_t + \frac{i^*}{y^*}\tilde{i}_t$$

4

Real Analysis

4.1 Basic Topology

Definition 4.1. (Metric Spaces) A set X, whose elements we shall call **points**, is said to be a **metric space**, if with any two points $p, q \in X$, there is an associated real number $d(p, q)$, called the **distance** between p and q, such that,

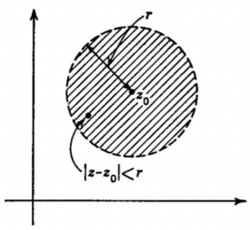
1. $d(p, q) > 0$ if $p \neq q$; $d(p, p) = 0$.
2. $d(p, q) = d(q, p)$.
3. (Triangle Inequality) $d(p, q) \leq d(p, r) + d(r, q), \forall r \in X$

Any function with these three properties is called a **distance** function or a **metric**.

Example. $X = \mathbb{R}$, $d(x, y) = |x - y|$; $x = \mathbb{R}^2, d(x, y) = \|x - y\|$.

Definition 4.2. (Open Ball) Let (X, d) be a metric space. Let $x_0 \in X$ and $r > 0$. The **open ball** centered at x_0 with radius $r > 0$ is a subset of X defined by,

$$B_r(x_0) = \{x \in X \,|\, d(x, x_0) < r\}$$



Definition 4.3. (Open Set) A subset $G \subset X$ of a metric space is called an **open** set if for all $x \in G$, there exists an $r > 0$, such that $B_r(x_0) \subset G$.

Example. $X = \mathbb{R}, A = [a, b)$ is not open; $A = (a, b)$ is open.

Proposition 4.1. In any metric space, $B_r(x)$ is an open set.

Proof. Let $B_r(x_0)$ be an open ball and $x \in B_r(x_0)$. We have to show that there exists an $r_1 > 0$ such that $B_{r_1}(x) \subset B_r(x_0)$. Let $r_1 = r - d(x, x_0)$. Consider the point $y \in B_{r_1}(x)$. Then by construction,$\forall y \in B_r(x_0)$

$$d(x_0, y) \leq d(x_0, x) + d(x, y) < d(x_0, x) + r_1 = r - d(x, x_0) + d(x, x_0) = r$$

Proposition 4.2. Let (X, d) be a metric space. Let $\{G_\alpha\}_{\alpha \in I}$ be a collection of open sets, s.t., $\forall \alpha \in I$, $G_\alpha \subseteq X$ is open.

1. $\bigcup_{\alpha \in I} G_\alpha$ is open;
2. $\bigcap_{\alpha=1}^{N} G_\alpha$ is also open.

Proof.
(1) Let $\{G_i\}$ be an arbitrary class of open sets in X, and let $G = \bigcup_\alpha G_\alpha$. We consider the following cases:

1. If $\{G_\alpha\}$ is empty, then G is empty and hence open;
2. If $\{G_\alpha\}$ is not-empty, therefore every point in $\{G_\alpha\}$ is the center of an open ball centered around it, and similarly, $G = \bigcup_\alpha G_\alpha$ has the same property. Therefore G is also open.

(2) Let $\{G_\alpha\}$ be a finite class of open set. The case in which $\{G_\alpha\}$ is empty is trivial, then suppose the family of $\{G_\alpha\}$ is non-empty, and equalling $\{G_1, G_2, \cdots, G_n\}$. Then if G is empty, it is open. Now suppose the intersection of these open sets is non-empty, and let $x \in G$, then $x \in G_\alpha$,$\forall \alpha$. Since each G_α is open, $\exists r_i > 0$,$\forall i \in \{1, \cdots, n\}$, s.t., $B_{r_\alpha}(x) \subset G_\alpha$. Let $r = \min_{i \in \{1, \cdots, n\}} r_i$. Then $B_r(x) \subset B_{r_i}(x) \subset G_i$, $\forall i \in \{1, \cdots, n\}$. Since $G = \bigcap_{i=1}^{n} G_i$, we therefore have $B_r(x) \subset G$ and therefore G is open.

Definition 4.4. (Interior Point) Let (X, d) be a metric space $A \subseteq X$. A point in A is called can **interior point** of A if it is the center of some open sets contained in A. The **interior** of A, denoted by int(A), is the set of all its interior points, such that:

$$\operatorname{int}(A) = \{x | x \in A, \exists r > 0, s.t., B_r(x) \in A\}$$

Definition 4.5. (Limit Point) Let (X, d) be a metric space and $A \subset X$. A point $x \in X$ is called a **limit point** of A, denoted by A', if each open set centered on x contains at least one point of A that is different from x:

$$B_\varepsilon(x) \backslash \{x\} \cap F \neq \varnothing$$

Definition 4.6. (Closed set) A set $F \subset X$ of a metric space (X, d) is **closed** if it contains all of its limit points, s.t., $\forall \varepsilon > 0, B_\varepsilon(x) \cap F \neq \varnothing$.

Proposition 4.3. In any metric space X, the empty set O and the full space X are closed sets.

Proposition 4.4. Let (X, d) be a metric space. A subset F of X is closed if and oly if its complement F^c is closed.

Proof. $\Rightarrow$: Assume $F \subset X$ is closed, except for the trivial case that F^c is empty, let $x \in F^c$. Since F is closed and $x \notin F$, such that x is not a limit point of F, it must be that there exists an open set centered around x, such that $B_r(x)$ is disjoint from F, indicating that $B_r(x) \in F^c$.

$\Leftarrow$: Assume F^c is open, still ignore the trivial case. Now prove by contradiction. Suppose there exists a limit point of F that is not in F but in F^c, then we have

$$B_\varepsilon(x) \backslash \{x\} \cap F \neq \varnothing, \forall \varepsilon > 0$$

a condtradiction to that $B_\varepsilon(x) \subset F^c$, a contradiction to the fact that F^c is open.

Definition 4.7. (Closure) Let (X, d) be a metric space, and let $A \subset X$. The **closure** of A, denoted by $\bar{A}$, is the union of A and the set of all its limit points.

$\bar{A}$ is the smallest closed set, s.t., $A \subset \bar{A}$.

Example. $\overline{B_\varepsilon(x)} = \{y \in X | d(x, y\} \leq \varepsilon$.

Theorem 4.1. Let (X, d) be a metric space, and $\{F_\alpha\}_{\alpha \in I}$ be a collection of closed sets.

1. $\bigcap_{\alpha \in I} F_\alpha$ is also closed;
2. $\bigcup_{\alpha=1}^{N} F_\alpha$ is also closed.

This can be proved by using DeMorgan's laws.

Remark. Singletons are closed sets.

Example. $\{0\} = [(-\infty, 0) \cup (0, \infty)]^C = (-\infty, 0)^c \cap (0, \infty)^c = \{0\}$.

Definition 4.8. (Bounded Set) Let (X, d) be a metric space and A be a subset of X. We say that A is a bounded set if $\exists x \in X, \varepsilon > 0$, s.t., $A \subseteq B_\varepsilon(x)$.

Definition 4.9. (Compact Set) Let $X = \mathbb{R}^n$ with any metric. A **compact set** in X is a closed and bounded set.

4.2 Sequences

Let (X, d) be a metric space with metric and let, $\{x_n\} = \{x_1, x_2, x_{3,} \cdots\}$ be a **sequence** of points in X.

Definition 4.10. (Sequence) A sequence is $f: \mathbb{N} \rightarrow X$ and for notation, (1) $f(n) = x_n$; (2) $\{x_n\}_{n \in \mathbb{N}} = \text{Image}(f)$. "The values of f, that is, the elements x_n, are called the terms of the sequence.[4.1]

Definition 4.11. (Convergence of Sequence) A sequence $\{x_n\}_n$ is convergent to $x \in X$, if $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, $d(x_n, x) < \varepsilon$.

Theorem 4.2. $\{x_n\}_n$ converges to $x \in X$ if and only if for every $\varepsilon > 0$, all but finite terms of the sequence are contained in $B_\varepsilon(x)$.

Proof. $\Rightarrow$: By definition (by *tautology*);

$\Leftarrow$: Suppose every nbbd of x contains all but finite points of $\{x_n\}_n$. Fix $\varepsilon > 0$, and by assumption the set

$$V_\varepsilon(x) = \{n \in \mathbb{N} | d(x_n, x) \geq \varepsilon\}$$

is finite. Let $N = \max V_\varepsilon(x)$. Then $\forall n > N$, it is true that $d(x_n, x) < \varepsilon$. Then $x_n \rightarrow x$.

Theorem 4.3. (Uniqueness of Limit)

4.1. Rudin **Definition 2.7**.

Let $\{x_n\}_n$ be a sequence in a metric space X, with $x_n \to x$, $x_n \to y$ as $n \to \infty$, then $x = y$.

Proof. Fix $\varepsilon > 0$ and assume that x_n converges to both x and y. Then by definition there exists N_x and N_y, such that

$$d(x_n, x) < \frac{\varepsilon}{2}, \forall n \geq N_x; d(x_n, y) < \frac{\varepsilon}{2}, \forall n \geq N_y$$

Let $N = \max\{N_x, N_y\}$ and by using the triangle inequality,

$$d(x, y) \leq d(x, x_n) + d(x_n, y) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for any $n \geq N$. Thus, $d(x, y) < \varepsilon, \forall \varepsilon > 0$. Hence $d(x, y) = 0$ and thus $x = y$.

Theorem 4.4. Let A be a non-empty closed set, and $x \in A'$. Then there is a sequence $\{x_n\}_n \subseteq A'$, such that $x_n \to x$.

Proof. Fix A and $x \in A'$. Pick $x_n \in B_{\frac{1}{n}}(x) \backslash \{x\} \bigcap A$. Since x is a limit point of A, then x_n always exists. Fix $\varepsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, $\frac{1}{n} < \varepsilon$. Then, for every $n \geq N$, it follows that

$$d(x_n, x) \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon$$

Hence, $x_n \to x$ and by construction $\{x_n\}_n \subset A$.

Theorem 4.5. Let $\{x_n\}_n$ be a convergent sequence, then it is also bounded.

Proof. Given that as $n \to \infty$, $x_n \to x$, then there exists $N \in \mathbb{N}$, s.t., $\forall n \geq N$, $d(x, x_n) < \varepsilon$. Then let δ be defined as

$$\delta = \max\{\varepsilon, d(x_1, x), \cdots, d(x_{N-1}, x)\}$$

Then $\{x_n\}_n \subset B_\delta(x)$, implying that the sequence is bounded.

Theorem 4.6. Let $\{x_n\}_n$ and $\{y_n\}_n$ be convergent sequences in the metric space (X, d), s.t., $x_n \to x, y_n \to y$, therefore we have:

1. $x_n + y_n \to x + y$;

2. $\alpha x_n \rightarrow \alpha x$, for any $\alpha \in \mathbb{R}$;

   The first two identities can be generalized to any vector space.

3. $x_n y_n \rightarrow x y$; (In $\mathbb{R}^n$ and $\mathbb{C}$)
4. $\frac{1}{x_n} \rightarrow \frac{1}{x}$, whenever $x_n \neq 0, x \neq 0$.

Proof. (1) Given $\varepsilon > 0$, $\exists\, N_1, N_2 \in \mathbb{N}$, s.t., $n \geq N_1, |x_n - x| < \frac{\varepsilon}{2}$; $n \geq N_2, |y_n - y| < \frac{\varepsilon}{2}$. If $N = \max\{N_1, N_2\}$, then $n \geq N$ implies

$$|(x_n + y_n) - (x + y)| \leq |x_n - x| + |y_n - y| < \varepsilon$$

(2) Fix $\varepsilon > 0$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N, |x_n - x| < \frac{\varepsilon}{|\alpha| + 1}$, then by using the following identity, for $\alpha \in \mathbb{R}$, we have

$$||\alpha x_n - \alpha x|| = \|\alpha(x_n - x)\| = |\alpha| \cdot \|x_n - x\| < \frac{|\alpha|}{(|\alpha| + 1)} \varepsilon$$

Definition 4.12. (Subsequence) Let $\{x_n\}_n$ be a sequence, consider a sequence of natural numbers $\{n_k\}_k \subset \mathbb{N}$, such that $n_1 < n_2 < \cdots < n_k < \cdots$, in other terms:

1. $\forall k \in \mathbb{N}, n_k \in N$;
2. $\forall k, k' \in \mathbb{N}, s.t., k > k'$, then $n_k > n_{k'}$.

Then the sequence $\{x_{n_k}\}_k$ is called a subsequence of $\{x_n\}_n$. If $\{x_{n_k}\}_k$ converges then its limit is called a *subsequential limit*.

Theorem 4.7. Let $\{x_n\}_n$ be a sequence in a metric space X. $\{x_n\}_n$ converges to x if and only if every subsequence of $\{x_n\}_n$ converges to x.[4.2]

Proof. $\Rightarrow$: Fix $\varepsilon > 0$, and let $\{x_{n_k}\}$ be a subseqeunce of $\{x_n\}$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, $d(x_n, x) < \varepsilon$. $d(x_{n_k}, x) < \varepsilon$, for every $n_k > N_k$. s.t., $N_k = \min_k \{n_k | n_k \geq N\}$.

$\Leftarrow$: Proof by contradiction. Suppose $\lim_{n\rightarrow\infty} x_n \neq x$, s.t., $\exists \varepsilon > 0, \forall n \in \mathbb{N}$ we have $d(x_n, x) > \varepsilon$; while for arbitrary choice of subsequence $\{x_{n_k}\}$, we have $d(x_{n_k}, x) < \varepsilon$, while $\{x_{n_k}\} \subseteq \{x_n\}$, a contradiction.

Theorem 4.8. (Bolzano-Weierstrass Theorem) Let $\{x_n\}_n$ be a sequence and $\{x_n\}_n \subseteq K$, s.t., K is compact, then $\{x_n\}_n$ has a convergent subsequence.

4.2. Rudin **Definition 3.5**.

Proof. Every sequence in a closed and bounded subset is bounded, so it has a convergent subsequence, which converges to a point in the set, because the set is closed.[4.3]

Definition 4.13. (Cauchy Sequence) A sequence $\{x_n\}$ is called a **Cauchy sequence** if for every $\varepsilon > 0$, there exists an $N \in \mathbb{N}$, such that, for all $m, n > N$, we have $d(x_n, x_m) < \varepsilon$.

Theorem 4.9. Any sequence $\{x_n\}_{n \in \mathbb{N}}$ that is convergent is also Cauchy.

Proof. Let $\varepsilon > 0$, since $x_n \to x$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, we have $d(x_n, x) < \frac{\varepsilon}{2}$. Let $m, n \geq N$, then by triangular inequality,

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Remark. The converse does not hold, by checking $\{3, 3.1, 3.14, 3.142, \cdots\}$ converges to π, where all elements of the sequence are in the space of rational number($\mathbb{Q}$), but π is irrational. The sequence is Cauchy, but not convergent.

Definition 4.14. (Complete Metric Space) A **complete** metic space is a metric space in which every *Cauchy* sequence is convergent.

Theorem 4.10. $\mathbb{R}^n$ is a complete metric space, $\forall n$.

4.3 Continuous Functions

Definition 4.15. (Continuous Function)

Let (X, d_X) and (Y, d_Y) be two metric spaces. Let $f: X \to Y$, we say f is *continuous* at x, if $\forall \varepsilon > 0$, $\exists \delta > 0$, s.t., $\forall z \in X$ satisfying $d_x(x, z) < \delta$, $d_y(f(x), f(z)) < \varepsilon$. Alternatively:

$$z \in B_\delta(x) \Rightarrow f(z) \in B_\varepsilon(f(x))$$

Or:

For each open ball B_ε centered on $f(x)$, there exists an open ball $B_\delta(x_0)$ centered on x_0, such that $f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$.

(Not Continuous) We say that f is *not continuous* at x, if is $\exists \varepsilon > 0$, s.t., $\forall \delta > 0$, s.t., $\exists z \in X$, satisfying $d_x(x, z) < \delta$, $d_y(f(x), f(z)) \geq \varepsilon$.

4.3. Check the workout of *sequential compactness* in Xuefeng's HW6 ex4.

Theorem 4.11. Let (X, d_X) and (Y, d_Y) be metric spaces and $f: X \to Y$. Then f is continuous at x_0 if and only if for all sequences $x_n \to x_0 \Rightarrow f(x_n) \to f(x_0)$.

Proof. $\Rightarrow$: Assuming that f is continuous at x_0. Let $\varepsilon > 0$, by continuity, $\exists \delta$, $N > 0$, s.t., $x_n \in B_\delta(x_0), \forall n \geq N$, we have $f(x_n) \in B_\varepsilon(f(x_0))$.

Moreover, since $x_n \to x_0$, then $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, s.t. $x_n \in B_\delta(x_0)$.

Thus, we have $x_n \in B_\delta(x_0)$ and $f(x_n) \in f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$, that is, $f(x_n) \to f(x_0)$.

$\Leftarrow$: Proof by contradiction. Given $f(x_n) \to f(x_0)$, suppose f is not continuous at x_0, then $\exists \varepsilon > 0, \forall \delta, n > 0$, s.t., $f(B_\delta(x_0))$ is not a subset of $B_\varepsilon(f(x_0))$. Then construct $\{x_n\}$ by choosing $x_n \in B_{\frac{1}{n}}(x_0)$ and $f(x_n) \notin B_\varepsilon(f(x_0))$, then clearly $x_n \to x_0$ but $f(x_n)$ does not converge to $f(x_0)$.

Continuity preserves limit.

Example. (1) Let $X = Y = \mathbb{R}$, let $d(x, z) = |x - z|$, and $f(x) = x$.

Let $\varepsilon > 0$. Pick $\delta = \varepsilon$, then for every $z \in B_\delta(x)$, it follows that,

$$|f(x) - f(z)| = |x - z| < \delta = \varepsilon$$

(2) Let $X = Y = \mathbb{R}$, let $d(x, z) = |x - z|$, and $f(x) = x^2$.

Let $\varepsilon > 0$. For every $z \in B_\delta(x)$, it follows that,

$$|f(x) - f(z)| = |(x + z)(x - z)| = |x + z||x - z| < (2|x| + \delta) \cdot \delta < \varepsilon$$

That is, our choice of δ should satisfy:

$$(2|x| + \delta) \cdot \delta < \varepsilon$$

By condition, due to its existence, we further restrict $\delta < 1$, so as to make that,

$$(2|x| + 1) \cdot \delta < \varepsilon \Rightarrow \delta < \frac{\varepsilon}{2|x| + 1}$$

and thus,

$$\delta = \min\left\{1, \frac{\varepsilon}{2|x| + 1}\right\}$$

for arbitrary δ, x that satisfy.

Theorem 4.12. Let (X, d_X) and (Y, d_Y) be metric spaces and $f: X \to Y$. Then f is continuous if and only if $f^{-1}(G)$ is open in X whenever G is open in Y.

Proof. $\Rightarrow$: Assume f is continuous. If G is an open set in Y, we must show that $f^{-1}(G)$ is open in X. Suppose $f^{-1}(G)$ is non-empty. Let $x \in f^{-1}(G)$. Then $f(x) \in G$. Since G is open, there exists a $\varepsilon > 0$ such that $B_\varepsilon(f(x)) \subset G$. Furthermore, from the continuity, $\exists \delta > 0$, s.t., $f(B_\delta(x)) \subset B_\varepsilon(f(x))$, and therefore, $f(B_\delta(x)) \subset G$, implying that $B_\delta(x) \subset f^{-1}(G)$. Thus we have an open ball centered on x that is contained in $f^{-1}(G)$, showing that $f^{-1}(G)$ is open.

$\Leftarrow$: We assume that $f^{-1}(G)$ is open whenever G is open in Y. Since $f^{-1}(G)$ is open, then let $x \in X$ and consider $B_\varepsilon(f(x))$, which is open by nature. By assumption, $f^{-1}(B_\varepsilon(f(x)))$ is open too. Since $f^{-1}(B_\varepsilon(f(x)))$ is open and contains x, there exists an open ball centered on x, s.t., $B_\delta(x) \subset f^{-1}(B_\varepsilon(f(x)))$. It is clear then $f(B_\delta(x)) \subset B_\varepsilon(f(x))$, which proves the continuity of f.

Corollary 4.1. Let (X, d_X) and (Y, d_Y) be metric spaces and $f: X \to Y$. Then f is continuous if and only if $f^{-1}(G)$ is closed in X whenever G is closed in Y.

Proof. "The complement."[4.4]

Proposition 4.5. Let (X, d_X) and (Y, d_Y) be metric spaces and $f, g: X \to Y$ be continuous functions. Then $f + g$, fg, f / g are also continuous functions.

Proof.

(1) Fix $\varepsilon > 0$ and $x \in X$. Since f is continuous, there $\exists \delta_1 > 0$ s.t. $f(B_{\delta_1}(x)) \subseteq B_{\frac{\varepsilon}{2}}(f(x))$; $\exists \delta_2 > 0$ s.t. $g(B_{\delta_2}(x)) \subseteq B_{\frac{\varepsilon}{2}}(g(x))$. Take $\delta = \min\{\delta_1, \delta_2\}$, therefore for $x \in B_\delta(x')$, we have:

$$|f(x) + g(x) - (f(x') + g(x'))| \leq |f(x) - f(x')| + |g(x) - g(x')| < \varepsilon$$

(2) Fix $\varepsilon > 0$ and $x \in X$. $\forall x' \in B_\delta(x)$, we have:

$$\begin{aligned} |f(x)g(x) - f(x')g(x')| &= |f(x)g(x) - f(x)g(x') + f(x)g(x') - f(x')g(x')| \\ &\leq |f(x)g(x) - f(x)g(x')| + |f(x)g(x') - f(x')g(x')| \\ &= |f(x)|\,|g(x) - g(x')| + |g(x')|\,|f(x) - f(x')| \\ &\leq |f(x)|\frac{\varepsilon}{2|f(x)|} + |g(x')|\frac{\varepsilon}{2\,\mathrm{argmax}_{x'}|g(x')|} \end{aligned}$$

4.4. Rudin **Theorem 4.8**.

Theorem 4.13. Let $f_1, \cdots, f_k$ be real-valued functions on the metric space X and let $\boldsymbol{f}: \mathbb{R} \to \mathbb{R}^n$ be defined as,

$$\boldsymbol{f}(x) = (f_1(x), f_2(x), \cdots, f_n(x))$$

Then $\boldsymbol{f}$ is continuous if it is component-wise continuous.

Proof. Fix $\varepsilon > 0$. Since every f_k is continuous, therefore $\exists \delta_k > 0$, s.t., $\forall x' \in B_{\delta_k(x)} \Rightarrow \boldsymbol{f}(x) \in B_{\frac{\varepsilon}{\sqrt{n}}}(\boldsymbol{f}(x))$. Now finx $\delta = \min_k \delta_k$, we have:

$$\|\boldsymbol{f}(x) - \boldsymbol{f}(x')\| \leq \sqrt{\sum_{k=1}^{n} \underbrace{|f_k(x) - f_k(x')|^2}_{\leq \left(\frac{\varepsilon}{\sqrt{n}}\right)^2}} \leq \varepsilon$$

Theorem 4.14. Let (X, d_X) and (Y, d_Y) be metric space and $f: X \to Y$ be continous. Then $K \subseteq X$ if is compact, then $f[K]$ is compact.

Proof. **Closedness.** Let $y \in f[K]'$, $\exists \{y_n\}_{n \in N} \subseteq f[K]$, s.t., $y_n \to y$. Consider the sequence $\{x_n\}_{n \in \mathbb{N}} \subseteq K$, s.t., $f(x_n) = y_n$.

By Bolzano-Weierstrass theorem, $\exists \{x_{n_k}\}_{k \in \mathbb{N}}$ to be a subsequence of $\{x_n\}_{n \in \mathbb{N}}$, s.t., $x_{n_k} \to x$ for some $x \in K$. Given the continuity of $f()$, we have:

$$f(x) = f\left(\lim_{k \to \infty} x_{n_k}\right) = \lim_{k \to \infty} f(x_{n_k}) = \lim_{k \to \infty} y_{n_k} = y$$

Boundedness. Proof by contradiction. Given that $f[K]$ is not bounded, want to show K is also not bounded. Then $\exists \{y_n\}_n \in f[K]$ that is not Cauchy. By the relation of $f(x_n) = y_n$, there exists a convergent subsequence of $\{x_n\}_n \in K$, which is Cauchy, which notwithstands the definition of continuity.

Alternatively, assuming that $f[K]$ is not bounded, $\forall \varepsilon > 0$, $y \in f[K], s.t.$, $\|y\| \geq \varepsilon$. Pick $y_n \in f[K]$, s.t., $\|y_n\| \geq n$. Consider $\{x_n\}_{n \in N} \subseteq K$, s.t., $f(x_n) = y_n$. By Bolzano-Weierstrass, $\exists \{x_{n_k}\}_{k \in N} \in K$, $x_{n_k} \to x$ for some element of K, that is,

$$\|f(x)\| = \left\|f\left(\lim_{k \to \infty} x_{n_k}\right)\right\| = \left\|\lim_{k \to \infty} f(x_{n_k})\right\| \geq \left\|\lim_{k \to \infty} n_k\right\|$$

a contradiction.

Theorem 4.15. If $f: X \to \mathbb{R}$ is continous and $K \subseteq X$ is compact, then $\exists \bar{x}$ and $\underline{x} \in K$, s.t., $f(\bar{x}) = \max_{x \in K} f(x)$, and $f(\underline{x}) = \min_{x \in K} f(x)$.

Proof. Since K is compact, while f is continuous by last theorem, $f[K]$ is compact. Ley $\alpha = \sup f[K]$, s.t., $\alpha < \infty, \alpha \in f[K]$, hence $\exists \bar{x} \in K$, s.t., $f(\bar{x}) = \alpha$.

Definition 4.16. (Supremum and Infimum)

Let $A \subseteq \mathbb{R}$. $\alpha = \sup A$; $\beta = \inf A$ if $\forall \varepsilon > 0$

1. $\forall x \in A$, $\alpha + \varepsilon > x$; $\exists x \in A$, $x > \alpha - \varepsilon$.
2. $\forall x \in A$, $\beta - \varepsilon < x$; $\exists x \in A$, $x < \beta + \varepsilon$.

Observation. α, β are limit points of A.

4.4 Connected Sets

Definition 4.17. (Separated Sets) Let (X, d) be a metric space, with A, $B \subseteq X$, A and B are **separated** if

1. $\bar{A} \cap B = \varnothing$;
2. $A \cap \bar{B} = \varnothing$.

A set is otherwise *connected* if it is not *separated* by two sets.

Theorem 4.16. *Let $X = \mathbb{R}$. Suppose $C \subseteq \mathbb{R}$ is connected, and $x_1, x_2 \in C$. WLOG, $x_1 < x_2$, $\forall z \in [x_1, x_2]$, $z \in C$.*

Proof. $\exists z \in (x_1, x_2)$, s.t., $z \notin C$, and $A = (-\infty, z) \cap C$, while $B = (z, \infty) \cap C$, then we have $\bar{A} = [\inf C, z]$; $B = (z, \sup C)$, where $\bar{A} \cap B = \varnothing$, a contradiction.

Theorem 4.17. Given $f: [a, b] \to \mathbb{R}$ is continuous, $f([a, b]) = [c, d]$. $\forall z \in [c, d]$, s.t., $\exists x \in [a, b]$, s.t., $f(x) = z$.

4.5 Fixed Point Theorem

Definition 4.18. (Contraction Mapping) Let (X, d) be a metric space. A mapping T of X into itself is said to be a **contraction**, or **contraction mapping**, if there exists a real number α, $0 < \alpha < 1$, s.t.,

$$d(Tx, Ty) \leq \alpha d(x, y)$$

$\forall x, y \in X$.

Theorem 4.18. (Banach Fixed Point Theorem) Let $T: X \to X$ be a continuous contraction. Then $\exists x \in X$, s.t., $T(x) = x$, which is unique.

Proof. Let $x_0 \in X$, and let $\{x_n\}$ be a sequence iteratively defined as $x_{n+1} = Tx_n$ for $n = 0, 1, 2, \cdots$. We shall first prove that $\{x_n\}$ is Cauchy. For $p = 1, 2, \cdots$, by the definition of contraction mapping, we have:

$$d(x_{p+1}, x_p) = d(Tx_p, Tx_{p-1}) \leq \alpha d(x_p, x_{p-1}) \leq \cdots \leq \alpha^n d(x_1, x_0)$$

Consider m, n be positive integers, WLOG, let $m > n$, therefore, by triangular inequality,

$$\begin{aligned} d(x_m, x_n) &\leq d(x_m, x_{m-1}) + d(x_{m-1}, x_{m-2}) + \cdots + d(x_{n+1}, x_n) \\ &\leq (\alpha^{m-1} + \alpha^{m-2} + \cdots + \alpha^n) d(x_1, x_0) \\ &\leq \frac{\alpha^n}{1-\alpha} d(x_1, x_0) \end{aligned}$$

Giving that the sequence $\{x_n\}$ is Cauchy in the complete space (X, d), hence $\{x_n\}$ is convergent. Let y be the limit of $\{x_n\}$, since continuity preserves limits, it follows that

$$Ty = T\Big(\lim_{n\to\infty} x_n\Big) = \lim_{n\to\infty} Tx_n = \lim_{n\to\infty} x_{n+1} = y$$

Hence, y is a fixed point of T.

If $y \neq z$, but $Ty = y$, $Tz = z$, therefore $d(y, z) = d(Ty, Tz) \leq \alpha d(y, z) < d(y, z)$, implying that $d(y, z) = 0 \Rightarrow y = z$, thus prove for uniqueness.

4.6 Differentiability

Definition 4.19. (Derivative) Let $f: \mathbb{R} \to \mathbb{R}$ defined on $[a, b]$. For any $x \in [a, b]$, define $f'(x)$ to be the derivative of f at x as

$$f'(x) = \lim_{t\to x} \frac{f(x) - f(t)}{x - t}$$

where $t \in [a, b]$.

Theorem 4.19. Consider a real-value function f on $[a, b]$ and assume for $x \in [a, b]$, the function f is differentiable. Then, f is also continuous at x.

Proof. Fix $\varepsilon < 0$. Note that,

$$f'(x) = \frac{f(x) - f(t)}{x - t} + \eta(t)$$

where $\eta(t) \to 0$ as $t \to x$. Observe that $\eta(t)$ is continuous, therefore, $\exists \delta > 0$, s.t., $|\eta(t)| < \delta$, hence, we have

$$\begin{aligned} |f(x) - f(t)| &= \frac{|f(x) - f(t)|}{|x - t|} |x - t| = \left| \frac{f(x) - f(t)}{x - t} \right| |x - t| \\ &= |f'(x) - \eta(t)|\, |x - t| \\ &\leq (|f'(x)| + |\eta(t)|) |x - t| \\ &\leq (|f'(x)| + \delta) |x - t| \\ &< (|f'(x)| + \delta) \delta' = \varepsilon \end{aligned}$$

Theorem 4.20. Suppose f and g are real and differentiable, then we have:

1. $(f+g)'(x) = f'(x) + g'(x)$;
2. $(fg)'(x) = f'(x)g(x) + f(x)g'(x)$;
3. $$\left(\frac{f}{g}\right)'(x) = \frac{g(x)f'(x) - g'(x)f(x)}{g^2(x)}, g(x) \neq 0$$

Proof. (2) Let $h(x) = f(x)g(x)$, therefore

$$\begin{aligned} h(x) - h(t) &= f(x)g(x) - f(t)g(t) \\ &= f(x)g(x) - f(t)g(t) - f(x)g(t) + f(x)g(t) \\ &= f(x)[g(x) - g(t)] + g(t)[f(x) - f(t)] \end{aligned}$$

Hence we have:

$$\frac{h(x) - h(t)}{x - t} = \frac{f(x)[g(x) - g(t)] + g(t)[f(x) - f(t)]}{x - t}$$

By taking limits as $t \to x$, we have the results hold.

(3) Let $h(x) = \frac{f(x)}{g(x)}$, $g(x) \neq 0$, following the above manner, we have:

$$h(x) - h(t) = \frac{f(x)}{g(x)} - \frac{f(t)}{g(t)} = \frac{1}{g(x)g(t)}[f(x)g(t) - f(t)g(x)]$$

The result is then obvious.

Theorem 4.21. Suppose f is continuous on $[a, b]$, $f'(x)$ exists at some point $x \in [a, b]$, g is defined on an interval I which contains the range of f, and g is differentiable at the point $f(x)$. If $h(t) = g(f(t))$, $a \leq t \leq b$, then h is differentiable at x, and

$$h'(x) = g'(f(x))f'(x)$$

Proof. Let $y = g(x)$, $s = g(t)$, therefore

$$f(y) - f(s) = [f'(y) + \varepsilon(s)](y - s); g(x) - g(t) = [g'(x) + \eta(t)](x - t)$$

$\varepsilon(s), \eta(t) \to 0$ as $s \to y, t \to x$. Then we have,

$$\begin{aligned} f(g(x)) - f(g(t)) &= f(y) - f(s) \\ &= [f'(y) + \varepsilon(s)](y - s) \\ &= [f'(g(x)) + \varepsilon(s)][g(x) - g(t)] \\ &= [f'(g(x)) + \varepsilon(s)][g'(x) + \eta(t)](x - t) \\ &\Longrightarrow \\ \lim_{t \to x} \frac{f(g(x)) - f(g(t))}{x - t} &= f'(g(x))g'(x) \end{aligned}$$

Theorem 4.22. (Generalized Mean Value Theorem)
Consider two functions f, g to be real and differentiable on $[a, b]$, then $\exists x \in (a, b)$,s.t.,

$$g'(x)(f(b) - f(a)) = f'(x)(g(b) - g(a))$$

A special case is done by setting $g(x) = x_0$,

$$f'(x)(b - a) = f(b) - f(a)$$

and this is known by **Mean Value Theorem**.

Definition 4.20. (Local Maxima) Consider a function f in an interval $[a, b]$, $x \in [a, b]$ is a **local maxima** if $\exists \delta > 0$, s.t., $\forall y \in B_\delta(x) \cap [a, b]$, s.t. $f(y) \leq f(x)$.

Lemma 4.1. Consider f be continuous on $[a, b]$ and be differentiable at (a, b). If $x \in (a, b)$, then x is a local maxima if $f'(x) = 0$.

Proof. $\exists \delta > 0$, s.t., $\forall y \in B_\delta(x)$, s.t., $f(y) \leq f(x)$, $B_\delta(x) \subseteq (a, b)$. Let $t \in (x - \delta,$ $x)$, note that

$$\frac{f(x) - f(t)}{x - t} > 0$$

Pick $t \in (x, x + \delta)$, we have

$$\frac{f(x) - f(t)}{x - t} \leq 0$$

Now pick a sequence of $\{\underline{t_n}\}_n$, s.t., $t_n \to x$, and

$$h(t) = \frac{f(x) - f(t)}{x - t}$$

Where $h(\underline{t_n}) \geq 0$, and $h(\underline{t_n}) \to f'(x)$. By taking a similar argument on the another side, we have $\forall n \in N$, $h(\underline{t_n}) > f'(x) \geq h\{\bar{t_n}\}$. Therefore, we should have $f'(x) = 0$.

Now continue to prove the generalized mean value theorem:

Proof. The claim is equivalent to:

$$g'(x)(f(b) - f(a)) - f'(x)(g(b) - g(a)) = 0$$

Then by FTC, let

$$h(x) = g(x)(f(b) - f(a)) - f(x)(g(b) - g(a))$$

where $h(a) = h(b)$, and x is a local optimum. There are three cases to discuss about.

1. When $h \equiv C$, then trivially done.
2. If $\exists t, h(t) > h(a), t \in (a, b)$, by the above lemma, $x = \max f(x), s.t., h'(x) = 0$;
3. If $\exists t, h(t) < h(a), t \in (a, b)$, similarly done.

4.7 Sequence of Functions

Definition 4.21. (Pointwise Convergent) Let $\{f_n\}_n$ be a sequence of functions defined in a metric space X. Suppose that $\forall x \in X$, the sequence of points $\{f_n(x)\}_n$ converges. Then define

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

In this case we say that the sequence $\{f_n\}_n$ converges **pointwise** to f.

Note that, properties of the sequence can be lost in the limit.

Example.

(1) Consider $f_n(x) = x^n$, $x \in [0, 1]$ $\forall n \in \mathbb{N}$, f_n is continuous.

1. $0 \leq x < 1$, $f_n(x) = x^n \to 0$;
2. $x = 1, f_n(x) = 1$.

(2)

$$f_n(x) = \frac{\sin(nx)}{\sqrt{n}}$$

Fix $x \in \mathbb{R}$, $|f_n(x)| = \left|\frac{\sin(nx)}{\sqrt{n}}\right| \leq \frac{1}{\sqrt{n}}$. Then as $n \to \infty$, $f_n(x) = 0$. Thus $f(x) = 0$.

Now, differentiating the sequence and the limit it yields $f'(x) = 0$, we have

$$f_n'(x) = \sqrt{n}\cos(nx)$$

Where $f_n'(0) = \sqrt{n}$, which does not converge to 0. Therefore, the sequence $\{f_n'\}_n$ does not converge pointwise to f'.

Definition 4.22. (Uniform Convergence) Let $\{f_n\}_n$ be a sequence of functions on a metric space X. We say that f_n converges uniformly to f, if $\forall \varepsilon, \exists N \in \mathbb{N}$, s.t., $\forall n \geq N$,

$$d(f_n(x), f(x)) < \varepsilon$$

$\forall x \in X$.

Example.
$$f_n(x) = \frac{1}{nx+1}$$

$f_n(x)$ converges pointwise to $f(x)=0$ for $x>0$, but it is not uniformly convergent.

Theorem 4.23. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of functions, s.t., $f_n \to^{c.u} f$, then $\forall \varepsilon>0$, $\exists N \in \mathbb{N}$, s.t., $\forall n, m \geq N$, we have $|f_n(x) - f_m(x)| < \varepsilon$.

Proof. Since $f_n(x)$ converges uniformly to $f(x)$ as $n \to \infty$. Since $\forall \varepsilon > 0$,$\exists N$, s.t., $\forall$ $m, n > N$, $\forall x$, we have, $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$,$|f_m(x) - f(x)| < \frac{\varepsilon}{2}$,

$$|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f_m(x) - f(x)| < \varepsilon$$

Corollary 4.2. f_n converges uniformly on I iff $\sup_{x\in X} |f_n(x) - f(x)| \to 0$ as $n \to \infty$.

Theorem 4.24. Consider the sequence $\{f_n\}_{n\in\mathbb{N}}$, s.t., $f_n \to^{c.u} f$. Let $E \subset X$, and $x \in E'$. Denote $y_n = \lim_{t\to x} f_n(t), t \in E$, therefore, we have:

$$\lim_{t\to x}\left(\lim_{n\to\infty} f_n(t)\right) = \lim_{n\to\infty} \underbrace{\left(\lim_{t\to x} f_n(t)\right)}_{y_n}$$

Proof. Since $f_n(t)$ uniformly converges, then $\forall \varepsilon$, $\exists N$, s.t., $|f_n(t) - f(t)| < \frac{\varepsilon}{2}$, $|f_m(t) - f(t)| < \frac{\varepsilon}{2}$, for $n, m \geq N$. Therefore, we have:

$$\lim_{t\to x} |f_n(t) - f_m(t)| = \left|\lim_{t\to x} f_n(t) - \lim_{t\to x} f_m(t)\right| < \varepsilon$$

Thefore $\lim_{t\to x} f_n(x)$ is Cauchy. Let $y_n = \lim_{t\to x} f_n(t)$, then by the claim in the theorem, we have $\lim_{n\to\infty} y_n = y$.

Thus, $\exists \delta > 0$, s.t., if $0 < |t - x| < \delta$, we have $|f_n(t) - y_n| < \varepsilon$, thus

$$|f(t) - y| \leq |f(t) - f_n(t)| + |f_n(t) - y_n| + |y_n - y| = 3\varepsilon$$

Corollary 4.3. $f_n \to^{c.u.} f$ and $\forall n \in N$, f_n is continuous, then we have f be continuous.

Theorem 4.25. $f_n \to^{p.t.} f$ and $\forall n \in N$, where f_n is monotonically increasing and continuous, while f is continuous. f_n's and f are on a compact set K.

Proof. Let $\varepsilon > 0$. Define $g_n(x) = f_n(x) - f(x)$, where $g_n(x) \geq 0$ and is decreasing in n, while $g_n(x) \to^{p.t.} 0$.

Let

$$K_n = \{x \in K \,|\, g_n(x) \geq \varepsilon\}$$

The set K suffices the following conditions:

1. $\forall n \in \mathbb{N}, K_n \subseteq K$;
2. $K_{n+1} \subseteq K_n$;
3. $\forall n \in \mathbb{N}$, K_n is closed, for g_n is continuous, and its image is closed.

Also, since $K_n \subseteq K$, K_n is compact.

Fix $t \in K$. Then $g_n(t) \to 0$, $\exists N \in \mathbb{N}$, s.t., $\forall n \geq N$, $|g_n(t)| < \varepsilon \Rightarrow g_n(t) < \varepsilon$. This infers that $\forall n \geq N, t \notin K_n$, therefore $t \notin \cap_{n \in \mathbb{N}} K_n$. Therefore, $\exists N \in \mathbb{N}$, s.t., $K_N = \varnothing \Rightarrow \forall n \geq N, K_n = \varnothing$.

Theorem 4.26. Suppose $\{f_n\}$ is a sequence of functions, differentiable on $[a, b]$, and s.t., $\{f_n(x_0)\}$ converges for some point x_0 on $[a, b]$. If $\{f_n'\}$ converges uniformly on $[a, b]$, then $\{f_n\}$ converges uniformly on $[a, b]$, to a function f, and

$$f'(x) = \left(\lim_{n \to \infty} f_n(x)\right)'$$

for $x \in [a, b]$.[4.5]

4.5. See proof on Chapter 6, MAT2050.

5

Optimization

Definition 5.1. (Partial Derviatives) Let $f\colon \mathbb{R}^n \to R$. We define

$$\frac{\partial f}{\partial x_k} = \lim_{t \to 0} \frac{f(x_1, x_2, \cdots, x_k + t, \cdots, x_n) - f(x_1, x_2, \cdots, x_k, \cdots, x_n)}{t}$$

Definition 5.2. (**Gradient**) If $f\colon \mathbb{R}^n \to \mathbb{R}$ we define the gradient

$$\nabla f(\boldsymbol{x}) = \left[\frac{\partial f(\boldsymbol{x})}{\partial x_1}, \cdots, \frac{\partial f(\boldsymbol{x})}{\partial x_n}\right]$$

We also have the **tangent line** at $f(\boldsymbol{x})$ be:

$$T_{\boldsymbol{x}}(f) = \{\boldsymbol{y} \in \mathbb{R}^n | \nabla f(\boldsymbol{x})(\boldsymbol{x} - \boldsymbol{y}) = 0\}$$

Definition 5.3. (Concavity)
A function is **concave(convex)** if $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n, \alpha \in [0, 1]$,

$$f(\alpha \boldsymbol{x} + (1 - \alpha)\boldsymbol{y}) \geq (\leq) \alpha f(\boldsymbol{x}) + (1 - \alpha) f(\boldsymbol{x})$$

Proposition 5.1. If $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable and $\boldsymbol{x}^* \in \mathbb{R}^n$ is a critical point, s.t., $\nabla f(\boldsymbol{x}) = \boldsymbol{0}$:

1. If f is concave, then $\boldsymbol{x}^*$ is a local maximum;
2. If f is convex, then $\boldsymbol{x}^*$ is a local minimum.

Definition 5.4. (Directional Derivative) Let $f\colon \mathbb{R}^n \to \mathbb{R}$ and $\boldsymbol{v} \in \mathbb{R}$, then

$$f'_{\boldsymbol{v}}(\boldsymbol{x}) = \frac{f(\boldsymbol{x} + \boldsymbol{v}h) - f(\boldsymbol{x})}{h}$$

And

$$f'_{\boldsymbol{v}}(\boldsymbol{x}) = \nabla f(\boldsymbol{x}) \cdot \boldsymbol{v}$$

Proof. By definition,

$$\begin{aligned}
f(\alpha\boldsymbol{y} + (1-\alpha)\boldsymbol{x}^*) &\geq \alpha f(\boldsymbol{y}) + (1-\alpha) f(\boldsymbol{x}^*) \\
f(\boldsymbol{x}^* + \alpha(\boldsymbol{y} - \boldsymbol{x}^*)) &\geq f(\boldsymbol{x}^*) + \alpha(f(\boldsymbol{y}) - f(\boldsymbol{x}^*)) \\
\lim_{\alpha \to 0} \frac{f(\boldsymbol{x}^* + \alpha(\boldsymbol{y} - \boldsymbol{x}^*)) - f(\boldsymbol{x}^*)}{\alpha} &\geq f(\boldsymbol{y}) - f(\boldsymbol{x}^*) \\
\nabla f(\boldsymbol{x}^*)(\boldsymbol{y} - \boldsymbol{x}^*) &\geq f(\boldsymbol{y}) - f(\boldsymbol{x}^*)
\end{aligned}$$

Theorem 5.1. (Envelope Theorem) In the problem $\max_{x \in S} f(x, \theta)$, suppose that there is a maximum point $x^*(\theta)$ in S. Furthermore, assume that the functions $\theta \longmapsto f(x^*(\theta^*), \theta)$ and $\theta \longmapsto f^*(\theta)$ are differentiable at θ^*. Then

$$\frac{\partial f^*(\theta^*)}{\partial \theta_j} = \left[\frac{\partial f(x, \theta)}{\partial \theta_j}\right]_{(x = x^*(\theta^*), \theta = \theta^*)}$$

We have the following formula to work out:

$$\frac{\partial f^*(\theta)}{\partial \theta_j} = \frac{\partial f^*(x^*(\theta), \theta)}{\partial \theta_j} + \underbrace{\sum_{k=1}^{n} \frac{\partial f(x^*(\theta), \theta)}{\partial x_k} \frac{\partial x^*(\theta)}{\partial \theta_j}}_{=0, \text{interior solution}}$$

Constrained Optimization.

$$f^*(\Theta) = \max_{x \in K} \{f(\boldsymbol{x}; \Theta), s.t., \boldsymbol{g}(\boldsymbol{x}; \Theta) = b\}$$

To solve the above formulation, we need less constraints than variables, and we have the Lagrangian:

$$\mathcal{L}(\boldsymbol{x} | \Theta) = f(\boldsymbol{x}; \Theta) + \boldsymbol{\lambda}(\boldsymbol{b} - \boldsymbol{g}(\boldsymbol{x}; \Theta))$$

6

Exercises II

These exercises are worked out or excerpted by Jeanne Sorin.

Exercise 6.1. (Convexity and Concavity)

1. Show that the function
$$f(x_1, \cdots, x_n) = \prod_{i=1}^{n} x_i^{\alpha_i}, \sum_{i=1}^{n} \alpha_i = 1$$
is strictly concave.
$$\frac{\partial f}{\partial x_i} = \frac{\alpha_i f}{x_i}, \frac{\partial^2 f}{\partial x_i^2} = \frac{\alpha_i(\alpha_i - 1) f^2}{x_i^2}, \frac{\partial^2 f}{\partial x_i x_j} = \frac{\alpha_i \alpha_j f^2}{x_i x_j}$$
Then we check its corresponding Hessian matrix..
$D_3 = f_{11}f_{22}f_{33} - f_{12}^2 f_{33} - f_{13}^2 f_{22} - f_{23}^2 f_{11}$; $D_4 = f_{11}f_{22}f_{33}f_{44} - f_{12}^2 f_{33}f_{44} - f_{13}^2 f_{22}f_{44} - f_{14}^2 f_{22}f_{33} - f_{23}^2 f_{11}f_{44} - f_{24}^2 f_{11}f_{33} - f_{34}^2 f_{11}f_{22}$
In conclusion, we have:
$$D_k = \prod_{i=1}^{k} f_{ii} - \sum_{i \neq j}^{k} f_{ij}^2 \prod_{l \neq i,j} f_{ll}$$
2. Show that the function
$$f(x_1, \cdots, x_n) = \left(\sum_{i=1}^{n} x_i^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \sigma > 1$$
is strictly concave.

We first prove the following lemma:

Lemma. *If $f(\cdot), g(\cdot)$ are both strictly concave and $f(\cdot)$ is increasing, then $f(g(\cdot))$ is increasing.*

Proof. By the definition of concavity, and given the fact that f is increasing, we have:

$$\begin{aligned}(1)\ f(g(\alpha x+(1-\alpha)y)) &> f(\alpha g(x)+(1-\alpha)g(y))\\ (2)\ f(\alpha g\,(x)+(1-\alpha)g(y)) &> \alpha f(g(x))+(1-\alpha)f(g(y))\end{aligned}$$

Therefore, we have

$$f(g(\alpha x+(1-\alpha)y))>\alpha f(g(x))+(1-\alpha)f(g(y))$$

For $h(x)=x^{\frac{\sigma}{\sigma-1}}$, when $\sigma>1$, $f(x)$ is increasing in x, with $f''(x)<0$.

For $g(x)=\sum_{i=1}^{n}x_i^{\frac{\sigma-1}{\sigma}}$, we have:

$$\begin{cases}\dfrac{\partial g^2(x)}{\partial x_i^2}=-\dfrac{\sigma-1}{\sigma^2}x_i^{-\frac{\sigma+1}{\sigma}}<0\\ \dfrac{\partial g^2(x)}{\partial x_i x_j}=0, i\neq j\end{cases}$$

Then obviously we find $g(\boldsymbol{x})$ to be ND, thus strictly concave, by the above lemma $f(\boldsymbol{x})=h(g(\boldsymbol{x}))$ is also strictly concave.

3. Find the solution of the following problem by solving the constraints for x and y:

$$\min x^2+(y-1)^2+z^2, s.t.\begin{cases}x+y=\sqrt{2}\\ x^2+y^2=1\end{cases}$$

There is only one point satisfying the binding constraints, $(x,y)=\left(\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\right)$. Therefore the minimum is at $\frac{1}{2}+\left(\frac{\sqrt{2}}{2}-1\right)^2=1+1-\sqrt{2}=2-\sqrt{2}$.

Exercise 6.2. (Envelope Theorem)

Let $f(\boldsymbol{x},d)$ and $\boldsymbol{g}(\boldsymbol{x},d)$ be real-valued, continuously differentiable functions on $\mathbb{R}^{m+\ell}$, where $X\in\mathbb{R}^m$ we choose variables and α parameters in $\mathbb{R}^\ell$, and we try to solve the following problem:

$$\max_{\boldsymbol{x}} f(\boldsymbol{x},\boldsymbol{\alpha}), s.t., \boldsymbol{g}(\boldsymbol{x},\boldsymbol{\alpha})\geq 0$$

evaluated at $(\boldsymbol{x}^*(\boldsymbol{\alpha}),\boldsymbol{\lambda}^*(\boldsymbol{\alpha}))$, with the Lagrangian

$$\mathcal{L}^*(\boldsymbol{\alpha})=f(\boldsymbol{x})(\boldsymbol{x}(\boldsymbol{\alpha})^*,\boldsymbol{\alpha})+\boldsymbol{\lambda}^*(\boldsymbol{\alpha})\boldsymbol{g}(\boldsymbol{x}(\boldsymbol{\alpha})^*,\boldsymbol{\alpha})$$

Define $V(\boldsymbol{\alpha}) = f(\boldsymbol{x}^*(\alpha), \alpha)$. Then

$$\frac{\partial V(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_k} = \frac{\partial \mathcal{L}(\boldsymbol{x}^*(\boldsymbol{\alpha}), \boldsymbol{\lambda}^*(\boldsymbol{\alpha}), \boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}_k} = \frac{\partial f(\boldsymbol{x}^*)}{\partial \alpha_k} + \boldsymbol{\lambda}^* \frac{\partial \boldsymbol{g}(\boldsymbol{x}^*)}{\partial \alpha_k}$$

<u>Elicit.</u> $\boldsymbol{U(C, L) = c^{\alpha} l^{1-\alpha}, s.t., cp_c + lp_l = w.}$

We have the Lagrangian:

$$\mathcal{L}(c, \ell; \lambda) = c^{\alpha} l^{1-\alpha} + \lambda(w - cp_c - lp_l)$$

And per its corresponding value function and subsequently, envelope theorem, we have:

$$\frac{\partial V(p_c, p_\ell)}{\partial p_c} = -\lambda^* c^*; \frac{\partial V(p_c, p_\ell)}{\partial p_\ell} = -\lambda^* \ell^*$$

Let an individual's utility maximization problem be like:

$$\begin{aligned} & \max{}_{X,Y} U(X, Y) = \log(X) + 2Y \\ s.t. \quad & Xp_x + Yp_y \leq W \\ & X, Y \geq 0 \end{aligned}$$

1. Solve for the choice variables as a function of the parameters , and for the corresponding indirect utility function V.

   $$\mathcal{L}(X, Y, \lambda) = \log(X) + 2Y + \lambda(W - Xp_x - Yp_y)$$

   By taking Lagrangian, we have:

   $$\frac{\partial \mathcal{L}}{\partial X} = \frac{1}{X} - \lambda p_x = 0$$
   $$\frac{\partial \mathcal{L}}{\partial X} = 2 - \lambda p_y = 0$$
   $$\frac{\partial \mathcal{L}}{\partial \lambda} = W - Xp_x - Yp_y$$

   which gives their optimum, by applying these, we then get its indirect utility function V:

   $$V(W, p_x, p_y) = \log\left(\frac{p_y}{2p_x}\right) + 2\left(\frac{W}{p_y} - \frac{1}{2}\right)$$

2. Totally differentiate V w.r.t W.

   $$\frac{\partial V}{\partial W} = \frac{2}{p_y}$$

3. Simplify using the FOCs. Think carefully about what holds when at an interior solution ($e.g. W = 10$, $p_y = 3$, $p_x = 1$) vs a corner solution ($e.g. W = 1, p_y = 3, p_x = 1$). Check that it satisfies the envelope theorem.

Interior solution: $X^* = \frac{3}{2}; Y^* = \frac{17}{6};$
Corner Soluton: $X^* = \min\left\{\frac{3}{2}, \frac{W}{p_x} = 1\right\};\ Y^* = 0.$

Exercise 6.3. (Roy's Identity)

Let $u: \mathbb{R}^{\ell}_{++} \to \mathbb{R}$ be a differentiable and strictly quasi-concave utility function. Then the demand function $\hat{x}(\cdot)$ and the indirect utility function $v(\cdot)$ satisfy the equations:

$$\hat{x}_k(\boldsymbol{p}, w) = -\frac{\frac{\partial v}{\partial p_k}(\boldsymbol{p}, w)}{\frac{\partial v}{\partial w}(\boldsymbol{p}, w)}, k = 1, \cdots, \ell$$

By using Lagrangian and Envelope theorem, we have:

$$\mathcal{L}(\boldsymbol{x}; \boldsymbol{\lambda}) = u(\boldsymbol{x}) + \boldsymbol{\lambda}(\boldsymbol{g}(\boldsymbol{x}, \boldsymbol{p}, w)), g(\boldsymbol{x}) = w - \sum_k p_k x_k$$

Therefore, we have:

$$\frac{\partial v}{\partial p_k} = \frac{\partial \mathcal{L}}{\partial p_k} = \boldsymbol{\lambda}\frac{\partial \boldsymbol{g}}{\partial p_k} = -\lambda \hat{x_k}; \frac{\partial v}{\partial w} = \frac{\partial \mathcal{L}}{\partial w} = \boldsymbol{\lambda}\frac{\partial \boldsymbol{g}}{\partial w} = \lambda$$

Therefore we arrive at our conclusion.