Introduction to Convexity

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

1. Introduction
   - Notation and Conventions

Subsection 1

Notation and Conventions

Sets

- The empty set is denoted by $\emptyset$.
- A set consisting of a single element is called a **singleton**.
- We use the symbol $\{x_1,\ldots,x_m\}_{\neq}$ to denote the set consisting of *distinct elements* $x_1,\ldots,x_m$.
- If A and B are sets, then $A \backslash B$ is used to denote the set consisting of those elements belonging to A but not to B.
- Two sets are said to **meet** if they have a non-empty intersection, otherwise they are said to be **disjoint**.
- If A is a subset of some universal set X, then the set $X \backslash A$ is called the **complement** of A in X and is denoted by A^c.

Families

- The intersection (union) of an empty family of subsets of the universal set X is taken to be X ($\varnothing$).
- Let $(A_i : i \in I)$ be a family of subsets of X indexed by some index set I.
- Then the family is said to be **pairwise disjoint** if A_i and A_j are disjoint whenever $i, j \in I$ with $i \neq j$.
- De Morgan's complementation laws assert that

$$\begin{aligned} \left(\bigcup(A_i : i \in I)\right)^c &= \bigcap(A_i^c : i \in I), \\ \left(\bigcap(A_i : i \in I)\right)^c &= \bigcup(A_i^c : i \in I). \end{aligned}$$

Injective, Surjective and Bijective Mappings

- Consider a mapping $f : X \to Y$, where X and Y are non-empty sets.
- Then f is said to be **injective** if $f(x) = f(x')$, where $x, x' \in X$, implies that $x = x'$;
- it is said to be **surjective** if, for each $y \in Y$, there exists $x \in X$ such that $f(x) = y$.
- A mapping which is both injective and surjective is said to be **bijective**.
- An injective (surjective, bijective) mapping is called an **injection** (**surjection**, **bijection**).
- Each bijection $f : X \to Y$ gives rise to an inverse mapping $f^{-1} : Y \to X$ defined by the condition that $f^{-1}(y) = x$ if and only if $f(x) = y$, where $x \in X$, $y \in Y$.

Image, Inverse Image and Composition

- Let $f : X \to Y$, $A \subseteq X$ and $B \subseteq Y$.
- The **image** $f(A)$ of A under f is the subset $\{f(a) : a \in A\}$ of Y.
- The **inverse image** $f^{-1}(B)$ of B under f is the subset $\{x \in X : f(x) \in B\}$ of X.
- If $f : X \to Y$ and $g : Y \to Z$ are mappings, then the composite mapping $g \circ f : X \to Z$ is defined by the equation

$$(g \circ f)(x) = g(f(x)), \text{ for } x \in X.$$

Bolzano-Weierstrass Theorem and Infinite Products

- We assume the rudiments of real analysis, including sequences, series, and the continuity, differentiability and integration of functions.
- One important result is the **Bolzano-Weierstrass Theorem**:
  - Every bounded sequence of real numbers contains a convergent subsequence.
- The infinite product notation is used in the following sense:

  If the real sequence $a_1, a_2, \ldots, a_k, \ldots$ is such that the sequence of partial products $a_1, a_1 a_2, \ldots, a_1 a_2 \cdots a_k, \ldots$ converges to the real number a, then this is indicated by writing

$$a = \prod_{k=1}^{\infty} a_k.$$

First Mean Value Theorem and Side Derivatives

- The **First Mean Value Theorem** (of the differential calculus) states:

  If the function $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$ and differentiable on (a,b), then there exists c in (a,b), such that
  $$f'(c) = \frac{f(b) - f(a)}{b - a}.$$
- If the real function f is such that the one-sided limit $\lim_{x \to a^+} \frac{f(x)-f(a)}{x-a}$ exists, then its value is denoted by $f'_+(a)$ and is called the **right derivative** of f at a.
- Similar remarks apply to the **left derivative** $f'_-(a)$ of f at a.
- If both $f'_-(a)$ and $f'_+(a)$ exist, then f is continuous at a.
- If, in addition, $f'_-(a) = f'_+(a)$, then f is differentiable at a.
- A superficial knowledge of the differentiability of a real function of n real variables is assumed, since the chain rule is used.

Vector Spaces and the Dimension Theorem

- We also assume a basic understanding of elementary linear algebra, including a knowledge of vector spaces, bases, dimension, linear transformations and eigenvectors.
- The only vector space considered here is the real n-dimensional space $\mathbb{R}^n$ whose points are real n-tuples $\boldsymbol{x} = (x_1, \ldots, x_n)$, and in which addition and scalar multiplication are defined coordinatewise.
- One result that will be needed is the **Dimension Theorem**:

  If A and B are finite dimensional subspaces of a vector space, then

  $$\dim(A+B) + \dim(A \cap B) = \dim A + \dim B,$$

  where $\dim A$ etc. denotes the dimension of A.

Inner Products and Norms

- The **inner product** $\boldsymbol{x} \cdot \boldsymbol{y}$ of vectors $\boldsymbol{x} = (x_1, \ldots, x_n)$ and $\boldsymbol{y} = (y_1, \ldots, y_n)$ in $\mathbb{R}^n$ is the real number defined by the equation

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + \cdots + x_n y_n.$$

- The **norm** or **length** $\|\boldsymbol{x}\|$ of a vector $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$ is the nonnegative number defined by the equations

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Cauchy-Schwarz and Triangle Inequalities

- Two important inequalities relating to the inner product and the norm are the Cauchy-Schwarz and the Triangle Inequalities.
- The **Cauchy-Schwarz Inequality**:

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \leq \|\boldsymbol{x}\| \|\boldsymbol{y}\|.$$

- The **Triangle Inequality**:

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|.$$

Orthogonal Complement and Unique Decomposition

- The orthogonal complement $A^\perp$ of a subspace A of $\mathbb{R}^n$ is the subspace of $\mathbb{R}^n$ defined by the equation

$$A^\perp = \{x \in \mathbb{R}^n : \boldsymbol{x} \cdot \boldsymbol{a} = 0 \text{ for all } \boldsymbol{a} \in A\}.$$

- We have

$$\dim A + \dim A^\perp = n.$$

- Each point $\boldsymbol{x}$ of $\mathbb{R}^n$ can be expressed uniquely in the form

$$\boldsymbol{x} = \boldsymbol{a} + \boldsymbol{b},$$

  where $\boldsymbol{a} \in A$ and $\boldsymbol{b} \in A^\perp$.

Matrices

- All matrices considered in the book are real.
- They are denoted by bold, upper case letters and expressed using square brackets.
- Thus we write $\boldsymbol{A}$ is the $m \times n$ matrix $[a_{ij}]$ to indicate that $\boldsymbol{A}$ is the $m \times n$ matrix having the element a_{ij} in its ith row and jth column.
- A zero matrix, i.e., one all of whose elements are zero, is denoted by $\mathbf{0}$, its size being determined by the context.
- The identity matrix $\boldsymbol{I}_n$ is the $n \times n$ matrix having ones along its leading diagonal and zeros elsewhere.
- The transpose of a matrix $\boldsymbol{A}$ is denoted by $\boldsymbol{A}^T$.
- The determinant of a square matrix $\boldsymbol{A}$ is denoted by $\det \boldsymbol{A}$.
- A square matrix $\boldsymbol{A}$ that has a non-zero determinant is said to be **non-singular** and has an inverse matrix $\boldsymbol{A}^{-1}$.

Linear Equations and Transformations

- Matrices can be used to represent both systems of linear equations and linear transformations.
- A general system of m linear equations in n unknowns $x_1, \ldots, x_n$ can be represented by the matrix form $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, where $\boldsymbol{A}$ is an $m \times n$ matrix, $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{b} \in \mathbb{R}^m$, $\boldsymbol{x}$ and $\boldsymbol{b}$ being regarded as column matrices.
- When $\boldsymbol{b} = \mathbf{0}$, the system is said to be **homogeneous**.
- Such a homogeneous system for which $m < n$ always has a non-trivial solution.
- A transformation $f : \mathbb{R}^n \to \mathbb{R}^m$ is linear if and only if it can be expressed in the form $f(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, where $\boldsymbol{A}$ is an $m \times n$ matrix and $\boldsymbol{x} \in \mathbb{R}^n$.

Rank of a Matrix

- The set of all linear combinations of the columns of an $m \times n$ matrix is a subspace of $\mathbb{R}^m$ called the **column space** of $\boldsymbol{A}$.

  Its dimension is called the **column rank** of $\boldsymbol{A}$.
- The set of all linear combinations of the rows of an $m \times n$ matrix is a subspace of $\mathbb{R}^n$ called the **row space** of $\boldsymbol{A}$.

  Its dimension is called the **row rank** of $\boldsymbol{A}$.
- The column rank and the row rank of a matrix are always equal, and are referred to simply as the **rank** of the matrix.

Symmetric Matrices and Quadratic Forms

- Let $\boldsymbol{A}$ be a symmetric $n \times n$ matrix.
- Then the eigenvalues of $\boldsymbol{A}$ are all real.
- Moreover, there exists an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $\boldsymbol{A}$.
- The matrix $\boldsymbol{A}$ gives rise to a quadratic function $q : \mathbb{R}^n \to \mathbb{R}$ of n real variables $x_1, \ldots, x_n$ defined by the equation

$$q(\boldsymbol{x}) = \boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x},$$

  where $\boldsymbol{x} = (x_1, \ldots, x_n)$ and the 1×1 matrix $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x}$ is identified with the real number defining it.
- We say that $\boldsymbol{A}$ is **non-negative semi-definite** if $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} \geq 0$ for all $\boldsymbol{x}$, and **positive definite** if $\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{x} > 0$ when $\boldsymbol{x} \neq \boldsymbol{0}$.
- If $\boldsymbol{A}$ is non-negative semi-definite (positive definite), then its eigenvalues are non-negative (positive).