Notes for Operations Research & More

Lan Peng, PhD Student

Department of Industrial and Systems Engineering
University at Buffalo, SUNY
lanpeng@buffalo.edu

November 10, 2019

November 10, 2019

Contents

Part I

Preliminary Topics

Chapter 1

Introduction to Optimization

1.1 Optimization Model

The following is the basic forms of terminology:

$$\begin{aligned}
(\text{P}) \quad \min \quad & f(x) && (1.1)\\
\text{s.t.} \quad & g_i(x) \le 0, \quad i = 1, 2, ..., m && (1.2)\\
& h_j(x) = 0, \quad j = 1, 2, ..., l && (1.3)\\
& x \in X && (1.4)
\end{aligned}$$

We have

- - $x \in R^n \to X \subseteq R^m$
- $g_i(x)$ are called inequality constraints
- $h_j(x)$ are called equality constraints
- X is the domain of the variables (e.g. cone, polygon, $\{0,1\}^n$, etc.)
- Let F be the feasible region of (P):
  - x^0 is a feasible solution iff $x^0 \in F$
  - x^* is an optimized solution iff $x^* \in F$ and $f(x^*) \le f(x^0), \forall x^0 \in F$ (for minimized problem)

Notice: Not every (P) has a feasible region, we can have $F = \emptyset$. Even if $F \neq \emptyset$, there might not be an solution to P, e.g. unbounded. If (P) has optimized solution(s), it could be 1) Unique 2) Infinite number of solution 3) Finite number of solution

Types of Optimization Problem

- $m = l = 0$, $x \in R^n$, unconstrained problem
- $m + l > 0$, constrained problem
- $f(x), g_i(x), h_j(x)$ are linear, Linear Optimization
  - If $X = R^n$, Linear Programming
  - If X is discrete, Discrete Optimization
  - If $X \subseteq Z^n$, Integer Programming
  - If $X \in \{0,1\}^n$, Binary Programming
  - If $X \in Z^n \times R^m$, Mixed Integer Programming

Chapter 2

Review of Linear Algebra

2.1 Field

Definition 2.1.1 (Field)**.** Let F denote either the set of real numbers or the set of complex numbers.

- Addition is commutative: $x + y = y + x, \forall x, y \in F$
- Addition is associative: $x + (y + z) = (x + y) + z, \forall x, y, z \in F$
- Element 0 exists and unique: $\exists 0, x + 0 = x, \forall x \in F$
- To each $x \in F$ there corresponds a unique element $(-x) \in F$ such that $x + (-x) = 0$
- Multiplication is commutative: $xy = yx, \forall x, y \in F$
- Multiplication is associative: $x(yz) = (xy)z, \forall x, y, z \in F$
- Element 1 exists and unique: $\exists 1, x1 = x, \forall x \in F$
- To each $x \neq 0 \in F$ there corresponds a unique element $x^{-1} \in F$ that $xx^{-1} = 1$
- Multiplication distributes over addition: $x(y + z) = xy + xz, \forall x, y, z \in F$

Suppose one has a set F of objects $x, y, z, ...$ and two operations on the elements of F as follows. The first operation, called addition, associates with each pair of elements $x, y \in F$ an element $(x + y) \in F$; the second operation, called multiplication, associates with each pair x, y an element $xy \in F$; and these two operations satisfy all conditions above. The set F, together with these two operations, is then called a **field**.

Definition 2.1.2 (Subfield)**.** A **subfield** of the field C is a set F of complex numbers which itself is a field.

Example. The set of integers is not a field.

Example. The set of rational numbers is a field.

Example. The set of all complex numbers of the form $x + y\sqrt{2}$ where x and y are rational, is a subfield of $\mathbb{C}$.

Notice: In this note, we (...Lan) assume that the field involved is a subfield of the complex numbers $\mathbb{C}$. More generally, if F is a field, it may be possible to add the unit 1 to itself a finite number of times and obtain 0, which does not happen in the subfield of $\mathbb{C}$. If it does happen in F, the least n such that the sum of n 1's is 0 is called **characteristic** of the field F. If it does not happen, then F is called a field of **characteristic zero**.

2.2 Real Vector Spaces

2.3 Linear, Conic, Affine, and Convex Combinations

2.4 Determinants

2.5 Inner Products

Definition 2.5.1 (Inner Product)**.** Let F be the field of real numbers or the field of complex numbers, and V a vector space over F. An **inner product** on V is a function which assigns to each ordered pair of vectors α, β in V a scalar $< \alpha|\beta >$ in F in such a way that $\forall \alpha, \beta, \gamma \in V, c \in \mathbb{R}$ that

- $< \alpha + \beta|\gamma > = < \alpha|\gamma > + < \beta|\gamma >$
- $< c\alpha|\beta > = c < \alpha|\beta >$
- $< \alpha|\beta > = \overline{< \beta|\alpha >}$
- $< \alpha|\alpha > \geq 0$, $< \alpha|\alpha > = 0$ iff $\alpha = \mathbf{0}$

Furthermore, the above properties imply that

- $< \alpha|c\beta + \gamma > = \bar{c} < \alpha|\beta > + < \alpha|\gamma >$

Definition 2.5.2. On F^n there is an inner product which we call the **standard inner product**. It is defined on $\alpha = (x_1, x_2, ..., x_n)$ and $\beta = (y_1, y_2, ..., y_n)$ by

$$< \alpha|\beta > = \sum_j x_j \bar{y_j} \tag{2.1}$$

For $F = \mathbb{R}^n$

$$< \alpha|\beta >= \sum_j x_j y_j \tag{2.2}$$

In the real case, the standard inner product is often called the dot product and denoted by $\alpha \cdot \beta$

Example. For $\alpha = (x_1, x_2)$ and $\beta = (y_1, y_2)$ in $\mathbb{R}^2$, the following is an inner product.

$$< \alpha|\beta >= x_1y_1 - x_2y_1 - x_1y_2 + 4x_2y_2 \tag{2.3}$$

Example. For $\mathbb{C}^{n\times n}$,

$$< \mathbf{A}|\mathbf{B} >= trace(\mathbf{B}^*\mathbf{A}) \tag{2.4}$$

is an inner product, where

$$\mathbf{A}^*_{ij} = \bar{\mathbf{A}}_{ji} \quad (\textbf{conjugate transpose}) \tag{2.5}$$

For $\mathbb{R}^{n\times n}$,

$$< \mathbf{A}|\mathbf{B} >= trace(\mathbf{B}^T\mathbf{A}) = \sum_j (AB^T)_{jj} = \sum_j \sum_k A_{jk}B_{jk} \tag{2.6}$$

2.6 Norms

Definition 2.6.1 (Norms)**.** A **norm** on a vector space $\mathcal{V}$ is a function $\|\|: \mathcal{V} \to \mathbb{R}$ for which the following three properties hold for all point $\mathbf{x}, \mathbf{y} \in \mathcal{V}$ and scalars $\lambda \in \mathbb{R}$

- (Absolute homogeneity) $\|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\|$
- (Triangle inequality) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$
- (Positivity) Equality $\|\mathbf{x}\| = 0$ holds iff $\mathbf{x} = 0$

Definition 2.6.2 (L_p-norms)**.** Let $p \geq 1$ be a real number. We define the p-norm of vector $\mathbf{v} \in \mathbb{R}^n$ as:

$$\|\mathbf{x}\|_p = (\sum_{i=1}^n |v_i|^p)^{\frac{1}{p}} \tag{2.7}$$

Particularly

$$\|\mathbf{v}\|_1 = \sum_{i=1}^n |v_i| \tag{2.8}$$

$$\|\mathbf{v}\|_2 = \sqrt{\sum_{i=1}^n v_i^2} \tag{2.9}$$

$$\|\mathbf{v}\|_\infty = \max_{i=1}^n |v_i| \tag{2.10}$$

Definition 2.6.3 (Dual norm)**.** For an arbitrary norm $\|\cdot\|$ on $\mathbf{E}$, the **dual norm** $\|\cdot\|^*$ on $\mathbf{E}$ is defined by

$$\|\mathbf{v}\|^* = \max\{< \mathbf{v}|\mathbf{x} > | \|\mathbf{x}\| \leq 1\} \tag{2.11}$$

For $p, q \in [1, \infty]$, the l_p and l_q norms on $\mathbb{R}^n$ are dual to each other whenever $\frac{1}{p} + \frac{1}{q} = 1$.

2.7 Eigenvectors and Eigenvalues

Definition 2.7.1. If $\mathbf{A}$ is an $n\times n$ matrix, then a nonzero vector $\mathbf{x} \in \mathbb{R}^n$ is called an **eigenvector** of $\mathbf{A}$ if $\mathbf{Ax}$ is a scalar multiple of $\mathbf{x}$, i.e.

$$\mathbf{Ax} = \lambda\mathbf{x} \tag{2.12}$$

for some scalar λ. The scalar λ is called **eigenvalue** of $\mathbf{A}$ and the vector $\mathbf{x}$ is said to be an **eigenvector corresponding to λ**

Theorem 2.1 (Characteristic Equation)**.** *If $\mathbf{A}$ is an $n\times n$ matrix, then λ is an eigenvalue of $\mathbf{A}$ iff*

$$\det(\lambda I - A) = 0 \tag{2.13}$$

Corollary 2.1.1.

$$\sum \lambda_A = tr(\mathbf{A}) \tag{2.14}$$

Corollary 2.1.2.

$$\prod \lambda_A = \det(\mathbf{A}) \tag{2.15}$$

Notice: Gaussian elimination changes the eigenvalues.

2.8 Decompositions

Chapter 3

Review of Real Analysis

3.1 Sequences and Series

Chapter 4

Review of Topology

4.1 Open Sets and Closed Sets

Definition 4.1.1 (Metric space)**.** A **metric space** is a set X where we have a notion of distance. That is, if $x, y \in X$, then $d(x, y)$ is the distance between x and y. The particular distance function must satisfy the following conditions:

- $d(x, y) > 0, \forall x, y \in X$
- $d(x, y) = 0 \iff x = y$
- $d(x, y) = d(y, x)$
- $d(x, z) \leq d(x, y) + d(y, z)$

Definition 4.1.2 (Ball)**.** Let X be a metric space. A **ball** B of radius r around a point $x \in X$ is

$$B = \{y \in X | d(x, y) < r\} \tag{4.1}$$

Definition 4.1.3 (Open set)**.** A subset $O \subseteq X$ is **open** if $\forall x \in O, \exists r, B = \{x \in X | d(x, y) < r\} \subseteq O$

Theorem 4.1. *The union of any collection if open sets is open.*

Proof. Sets $S_1, S_2, ..., S_n$ are open sets, let $S = \cup_{i=1}^n S_i$, then $\forall i, S_i \subseteq S$. $\forall x \in S, \exists i, x \in S_i$. Given that S_i is an open set, then for x, $\exists r$ that $B = \{x \in S_i | d(x, y) < r\} \subseteq S_i \subseteq S$, therefore S is an open set. □

Theorem 4.2. *The intersection of any finite number of open sets is open.*

Proof. Sets $S_1, S_2, ..., S_n$ are open sets, let $S = \cap_{i=1}^n S_i$, then $\forall i, S \subseteq S_i$. $\forall x \in S, x \in S_i$. For any i, we can define an r_i, such that $B_i = \{x \in S_i | d(x, y) < r_i\} \subseteq S_i$. Let $r = \min_i\{r_i\}$. Noticed that $\forall i, B' = \{x \in S_i | d(x, y) < r\} \subseteq B_i \subseteq S_i$. Therefore S is an open set. □

Remark. The intersection of infinite number of open sets is not necessarily open.

Here we find an example that the intersection of infinite number of open sets can be closed.

Example. Let $A_n \in \mathbb{R}$ and $B_n \in \mathbb{R}$ be two infinite series, with the following properties.

- $\forall n, A_n < a, \lim A_n = a$
- $\forall n, B_n > b, \lim B_n = b$
- $a < b$

Then we define infinite number of sets S_i, the ith set is defined as

$$S_i = (A_i, B_i) \subset \mathbb{R} \tag{4.2}$$

Then

$$S = \cap_{i=1}^{\infty} S_i = [a, b] \subset \mathbb{R} \tag{4.3}$$

and S is a closed set.

Definition 4.1.4 (Limit point)**.** A point z is a **limit point** for a set A if every open set U that $z \in U$ intersects A in a point other than z.

Notice: z is not necessarily in A.

Definition 4.1.5 (Closed set)**.** A set C is **closed** iff it contains all of its limit points.

Theorem 4.3. $S \in \mathbb{R}^n$ *is closed* $\iff \forall \{x_k\}_{k=1}^{\infty} \in S, \lim_{k \to \infty} \{x_k\}_{k=1}^{\infty} \in S$

Theorem 4.4. *Every intersection of closed sets is closed.*

Theorem 4.5. *Every finite union of closed sets is closed.*

Remark. The union of infinite number of closed sets is not necessarily closed.

Theorem 4.6. *A set C is a closed set if $X \setminus C$ is open*

Proof. Let S be an open set, $x \notin S$, for any open set S_i that $x \in S_i$, we can find a correspond $r_i > 0$, such that $B_i = \{x \in S_i | d(x, y) < r_i\}$. Take $r = \min_{\forall i}\{r_i\}$, set $B = \{x \notin S | d(x, y) < r\} \neq \emptyset$. Which means for any $x \notin S$, we can find at least one point $x' \in B$ that for all open set S_i, $x' \in S_i$, which makes x a limit point of the complement of the open set. Notice that x is arbitrary, then the collection of x, i.e., the complement of S is a closed set. □

Remark. The empty set is open and closed, the whole space X is open and closed.

Part II

Nonlinear Programming

Chapter 5

Convex Analysis

5.1 Convex Sets

Definition 5.1.1. A set $S \in \mathbb{R}^n$ is said to be convex if $\forall x_1, x_2 \in S, \lambda \in (0,1) \Rightarrow \lambda x_1 + (1-\lambda)x_2 \in S$

The following are some familes of convex sets.

Example. Empty set is by convention considered as convex.

Example. Polyhedrons are convex sets.

Example. Let $P = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x}^\top \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ where $\mathbf{A} \in \mathbb{S}_+^{n\times n}$ and $\mathbf{b} \in \mathbb{R}_+$. The set P is a convex subset of $\mathbb{R}^n$.

Example. Let $\|.\|$ be any norm in $\mathbb{R}^n$. Then, the unit ball $B = \{\mathbf{x} \in \mathbb{R}^n | \|\mathbf{x}\| \leq b, b > 0\}$ is convex.

Let S_1, S_2 be convex set, then:

- $S_1 \cap S_2$ is convex set
- $S_1 \oplus S_2$ (Minkowski addition) is convex set, where
$$S_1 \oplus S_2 = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \mathbf{x_1} + \mathbf{x_2}, \mathbf{x_1} \in S_1, \mathbf{x_2} \in S_2\} \tag{5.1}$$
- $S_1 \ominus S_2$ is convex set, where
$$S_1 \oplus S_2 = \{\mathbf{x} \in \mathbb{R}^n | \mathbf{x} = \mathbf{x_1} - \mathbf{x_2}, \mathbf{x_1} \in S_1, \mathbf{x_2} \in S_2\} \tag{5.2}$$
- $f(S_1)$ is convex iff $f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}, \mathbf{A} \in \mathbb{R}^{m\times n}, \mathbf{b} \in \mathbb{R}^m$

Theorem 5.1 (Carathéodory's Theorem). *Let $S \subseteq \mathbb{R}^n$. Then $\forall \mathbf{x} \in conv(S)$, there exists $\mathbf{x}^1, \mathbf{x}^2, ..., \mathbf{x}^p \in S$, where $p \leq n+1$ such that $\mathbf{x} \in conv\{\mathbf{x}^1, \mathbf{x}^2, ...\mathbf{x}^p\}$.*

Notice: This theorem means, any point $\mathbf{x} \in \mathbb{R}^n$ in a convex hull of S, i.e., $conv(S)$, can be included in a convex subset $S' \subseteq conv(S)$ that has $n+1$ extreme points.

Theorem 5.2. *Let S be a convex set with nonempty interior. Let $\mathbf{x}_1 \in cl(S)$ and $\mathbf{x}_2 \in int(S)$, then $\mathbf{y} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2 \in int(S), \forall \lambda \in (0,1)$*

5.2 Convex Functions

Definition 5.2.1. Let $C \in \mathbb{R}^n$ be a convex set. A function $f : C \to \mathbf{R}$ is (resp. strictly) convex if

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \tag{5.3}$$
$$\forall \mathbf{x}_1, \mathbf{x}_2 \in C, \forall \lambda \in (0,1) \tag{5.4}$$

(resp.)

$$f(\lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2) < \lambda f(\mathbf{x}_1) + (1-\lambda)f(\mathbf{x}_2) \tag{5.5}$$
$$\forall \mathbf{x_1} \neq \mathbf{x}_2 \in C, \forall \lambda \in (0,1) \tag{5.6}$$

Notice: When calling a function convex, we imply that its domain is convex.

Example. Given any norm $\|.\|$ on $\mathbb{R}^n$, the function $f(x) = \|x\|$ is convex over $\mathbb{R}^n$.

Definition 5.2.2. Let S be a nonempty convex subset of $\mathbb{R}^n$, $f : S \to \mathbb{R}$ is (resp. strictly) **concave** if $-f(x)$ is (resp. strictly) convex.

Notice: A function may be neither convex nor concave.

Theorem 5.3. *Consider $f : \mathbb{R}^n \to \mathbb{R}$. $\forall \bar{\mathbf{x}} \in \mathbb{R}^n$ and a nonzero direction $\mathbf{d} \in \mathbb{R}^n$. Define $F_{\bar{\mathbf{x}},d}(\lambda) = f(\bar{\mathbf{x}} + \lambda\mathbf{d}$. Then f is (resp. strictly) convex iff $F_{\bar{\mathbf{x}},d}(\lambda)$ is (resp. strictly) convex for all $\bar{\mathbf{x}} \in \mathbb{R}^n, \forall \mathbf{d} \in \mathbb{R}^n \setminus \{0\}$.*

Definition 5.2.3 (Level-set)**.** Given a function $f : \mathbb{R}^n \to \mathbb{R}$ and a scalar $\alpha \in \mathbb{R}$, we refer to the set $S_\alpha = \{\mathbf{x} \in S | f(\mathbf{x}) \leq \alpha\} \subseteq \mathbb{R}^n$ as the **α-level-set** of f.

Lemma 5.4. *Let S be a nonempty convex set in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}^n$ be a convex function, then the **α-level-set** of f is a convex set for each value of $\alpha \in \mathbb{R}$.*

Notice: The converse is not necessarily true.

Definition 5.2.4 (Epigraphs, Hypographs)**.** Let $S \in \mathbb{R}^n$ be such that $S \neq \emptyset$. The **epigraph** of f, denoted by $epi(f)$ is

$$epi(f) = \{(\mathbf{x}, y) \in S | \mathbf{x} \in S, y \in \mathbb{R}, y \geq f(x)\} \in \mathbb{R}^{n+1} \tag{5.7}$$

The **hypograph** of f, denoted by $hypo(f)$ is

$$hypo(f) = \{(\mathbf{x}, y) \in S | \mathbf{x} \in S, y \in \mathbb{R}, y \leq f(x)\} \in \mathbb{R}^{n+1} \tag{5.8}$$

Theorem 5.5. *Let S be a nonempty convex subset in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$. Then f is convex iff $epi(f)$ is convex.*

Theorem 5.6. *Let S be a nonempty convex subset in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a convex function on S. Then f is continuous in $int(S)$.*

5.3 Subgradients and Subdifferentials

Definition 5.3.1 (Subgradient)**.** Let S be a nonempty convex set in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a convex function, then ξ is a **subgradient** of f at $\bar{\mathbf{x}}$ if

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^\top (\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S \tag{5.9}$$

Definition 5.3.2 (Subdifferential)**.** The set of all subgradients of f at $\bar{\mathbf{x}}$ is called **subdifferential** of f at $\bar{\mathbf{x}}$, denoted as $\partial f(\bar{\mathbf{x}})$

Theorem 5.7. *Let S be a nonempty convex set in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a convex function. Then for $\bar{\mathbf{x}} \in int(S)$, there exists a vector ξ such that*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^\top (\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S \tag{5.10}$$

In particular, the hyperplane

$$\mathcal{H} = \{(\mathbf{x}, y) | y = f(\bar{\mathbf{x}}) + \xi^\top (\mathbf{x} - \bar{\mathbf{x}})\} \tag{5.11}$$

is a supporting plane of $epi(f)$ at $(\bar{\mathbf{x}}, f(\bar{\mathbf{x}}))$

Theorem 5.8. *Let S be a nonempty convex set in $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be a convex function. Suppose that for each $\bar{\mathbf{x}} \in S$, there exists ξ such that*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \xi^\top (\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S \tag{5.12}$$

Then f is convex on $int(S)$

Notice: Not all convex functions are continuous, it has to be continuous in its interior, but it may not be continuous at the boundary.

5.4 Differentiable Functions

Definition 5.4.1 (Differentiable Functions)**.** Let S be a nonempty subset of $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$.Then f is said to be **differentiable** at $\bar{\mathbf{x}} \in int(S)$ if there exists a vector $\nabla f(\bar{\mathbf{x}})$ and a function $\alpha : \mathbb{R}^n \to \mathbb{R}$ such that

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})^\top (\mathbf{x} - \bar{\mathbf{x}}) + \alpha(\bar{\mathbf{x}}, \mathbf{x} - \bar{\mathbf{x}}) \|\mathbf{x} - \bar{\mathbf{x}}\| \tag{5.13}$$

for all $\mathbf{x} \in S$ where $\lim_{\mathbf{x} - \bar{\mathbf{x}}} \alpha(\bar{\mathbf{x}}, \mathbf{x} - \bar{\mathbf{x}}) = 0$

Remark. If function f is differentiable, then $\nabla f(\bar{\mathbf{x}}) = (\frac{\partial f(\bar{\mathbf{x}})}{\partial \mathbf{x}_1}, \frac{\partial f(\bar{\mathbf{x}})}{\partial \mathbf{x}_2}, \cdots, \frac{\partial f(\bar{\mathbf{x}})}{\partial \mathbf{x}_n})$, and the gradient is unique.

Lemma 5.9. *Let $S \neq \emptyset$ be a convex set of $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be convex. If f is differentiable at $\bar{\mathbf{x}} \in int(S)$, then the subdifferential of f at $\bar{\mathbf{x}}$ is the singleton, $\{\nabla f(\bar{\mathbf{x}})\}$*

Theorem 5.10. *Let S be a nonempty subset of $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be differentiable on S. Then f is (resp. strictly) convex on S iff $\forall \bar{\mathbf{x}} \in S$*

$$f(\mathbf{x}) \geq f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \in S \tag{5.14}$$

(resp.)

$$f(\mathbf{x}) > f(\bar{\mathbf{x}}) + \nabla f(\bar{\mathbf{x}})(\mathbf{x} - \bar{\mathbf{x}}), \forall \mathbf{x} \neq \bar{\mathbf{x}} \in S \tag{5.15}$$

Theorem 5.11 (Mean-value Theorem)**.** *Let S be a nonempty subset of $\mathbb{R}^n$. Let $f : S \to \mathbb{R}$ be differentiable on S. Then for all $\mathbf{x}_1, \mathbf{x}_2 \in S$, there exists $\lambda \in (0, 1)$ such that*

$$f(\mathbf{x}_2) = f(\mathbf{x}_2) + \nabla f(\hat{\mathbf{x}})(\mathbf{x}_2 - \mathbf{x}_1) \tag{5.16}$$

where

$$\hat{\mathbf{x}} = \lambda \mathbf{x}_1 + (1 - \lambda) \mathbf{x}_2 \tag{5.17}$$

Chapter 6

KKT Optimality Conditions

Chapter 7

Lagrangian Duality

Chapter 8

Unconstrained Optimization

Chapter 9

Penalty and Barrier Functions