INTRODUCTION TO CONVEX OPTIMIZATION

by

Neda Tanoumand

Submitted to the Graduate School of Engineering and Natural Sciences

in partial fulfillment of the requirements for the degree of

Master of Science

Sabancı University

Spring, 2019

Introduction to Convex Optimization

APPROVED BY:

Assoc. Prof. Dr. Nihat Gökhan Göğüş
(Thesis Supervisor)

Asst. Prof. Dr. Turgay Bayraktar
(Jury Member)

Asst. Prof. Dr. Sibel Şahin
(Jury Member)

DATE OF APPROVAL: 17/07/ 2019

INTRODUCTION TO CONVEX OPTIMIZATION

Neda Tanoumand

Master of Science in Mathematics

Thesis Supervisor: Nihat Gökhan Göğüş

Abstract

In this thesis, we touched upon the concept of convexity which is one of the essential topics in optimization. There exist many real world problems that mathematically modelling these problems and trying to solve them are the focus point of many researchers. Many algorithms are proposed for solving such problems. Almost all proposed methods are very efficient when the modelled problems are convex. Therefore, convexity plays an important role in solving those problems. There are many techniques that researchers use to convert a non-convex model to a convex one. Also, most of the algorithms that are suggested for solving non-convex problems try to utilize the notions of convexity in their procedures. In this work, we begin with important definitions and topics regarding convex sets and function. Next, we will introduce optimization problems in general, then, we will discuss convex optimization problems and give important definitions in relation with the topic. Furthermore, we will touch upon Linear Programming which is one of the most famous and useful cases of Convex Optimization problems. Finally, we will discuss the Generalized Inequalities and their application in vector optimization problems.

Keywords: Convexity, Convex Sets, Convex Functions, Convex Optimization, Linear Programming, Vector Optimization.

DIŞBÜKEY OPTİMİZASYONUNA GİRİŞ

Neda Tanoumand

Matematik Yüksek Lisansı

Tez Danışmanı: Nihat Gökhan Göğüş

Özet

Bu tez çalışmasında, optimizasyondaki en temel konulardan biri olan dışbükeylik kavramına değinilmiştir. Gerçek dünya problemlerinin matematiksel olarak modellenmesi ve çözümü birçok araştırmacının odak noktası olmuştur. Bu tür problemleri çözmek için birçok algoritma önerilmiştir. Modellenen problemler dışbükey olduğunda hemen hemen tüm önerilen yöntemler çok etkilidir. Bu nedenle, dışbükeylik bu problemleri çözmede önemli bir rol oynamaktadır. Araştırmacıların dışbükey olmayan bir modeli dışbükey bir modele dönüştürmek için kullandıkları birçok teknik vardır. Ayrıca, dışbükey olmayan problemleri çözmek için önerilen algoritmaların çoğu, dışbükeylik kavramlarını prosedürlerinde kullanmaya çalışmaktadır. Bu çalışmaya dışbükey kümeler ve fonksiyonlarla ilgili önemli tanımlar ve konularla başlanacaktır. Daha sonra genel olarak optimizasyon problemleri tanıtılıp, dışbükey optimizasyon problemleri tartışılacak ve konuyla ilgili önemli tanımlamalar yapılacaktır. Ayrıca, en ünlü ve faydalı dışbükey optimizasyon problemlerinden biri olan Doğrusal Programlamaya da değinilecektir. Son olarak, Genelleştirilmiş Eşitsizlikleri ve bunların vektör optimizasyon problemlerindeki uygulamaları tartışılacaktır.

Anahtar Kelimeler: Dışbükeylik, Dışbükey Kümeler, Dışbükey Fonksiyonlar, Dışbükey Optimizasyon, Doğrusal Programlama, Vektör Optimizasyon.

To my best friend...

Acknowledgments

I would like to express my sincere gratitude to my thesis supervisor Assoc. Prof. Nihat Gökhan Göğüş for his help and encouragement during the course of my master's thesis. I want to sincerely thank Farzin Asghari Arpatappeh for his endless support and endurance.

I gratefully acknowledge Sabancı University for the scholarships received and for becoming part of my family for last four years.

Contents

List of Figures

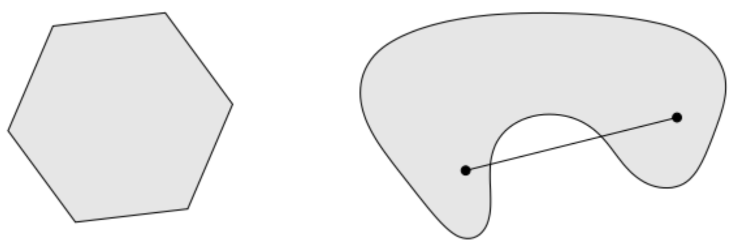
1 Preliminaries (convex sets and convex functions)

In this chapter we begin by presenting basic definitions regarding convex sets and convex functions. Next, some important examples of convex sets which are used frequently in optimization area are provided. Additionally, we investigate some operations that preserve the convexity of sets and functions. Finally, we introduce the concept of cones and generalized inequalities and provide the definition of convex functions with respect to the generalized inequalities.

1.1 Convex sets and related definitions

Definition 1.1. A set $C \subseteq \mathbb{R}^n$ is *convex*, if the line segment joining two distinct points in the set lies completely in C. Mathematically speaking, the set C is convex if for $x_1, x_2 \in C$ and any $0 \leq \theta \leq 1$, we have $\theta x_1 + (1-\theta)x_2 \in C$.

Figure 1.1: Examples of convex and non-convex sets



In Figure 1.1 we can see simple examples of a convex set on the left and a non-convex set on the right. The line joining each two points in the convex sets lies completely in the set; however, as we can see in the right figure, in non-convex sets a line segment joining arbitrary two points of the set does not completely lie in the set.

Definition 1.2. A *convex combination* of set of points $x_1, ..., x_k \in \mathbb{R}^n$ is a point $\theta_1 x_1 + ... + \theta_k x_k$ where $\theta_1 + ... + \theta_k = 1$ and $\theta_i \geq 0$ for $i = 1, ..., k$.

Theorem 1.1. *A set $C \subseteq \mathbb{R}^n$ is convex if and only if it contains all convex combinations of its points.*

Proof. We use induction in k. Suppose (P_k) is the statement: "$x_1, ..., x_k \in C, \theta_1, ..., \theta_k \geq 0$, $\sum_{j=1}^{k} \theta_j = 1$ implies $\sum_{j=1}^{k} \theta_j x_j \in C$".

Suppose C is convex. This means (P_2) is true. Suppose (P_k) is true. We want to prove (P_{k+1}).

Let $x_1, ..., x_k, x_{k+1} \in C$, $\theta_1, ..., \theta_k, \theta_{k+1} \geq 0$, and $\sum_{j=1}^{k+1} \theta_j = 1$. We have to show that $\sum_{j=1}^{k+1} \theta_j x_j \in C$.

Let $\beta_k = \theta_1 + ... + \theta_k$. If $\beta = 0$ then $\theta_j = 0$ for all $j = 1, ..., k$, $\theta_{k+1} = 1$, the statement trivially holds. We may assume that $\beta_k > 0$. Since $\frac{\theta_1}{\beta_k} + ... + \frac{\theta_k}{\beta_k} = 1$, by induction hypothesis (P_k), $\sum_{j=1}^{k} \frac{\theta_j}{\beta_k} x_j \in C$. Notice that $\beta_k + \theta_{k+1} = 1$. Since C is convex

$$\beta_k [\frac{\theta_1}{\beta_k} x_1 + ... + \frac{\theta_k}{\beta_k} x_k] + \theta_{k+1} x_{k+1} \in C.$$

The proof of the converse part follows immediately from the definition of convex sets. □

Theorem 1.2. *Let $\{C_\alpha\}_{\alpha \in I}$ be a family of convex sets in $\mathbb{R}^n$. Then, $C = \cap_{\alpha \in I} C_\alpha$ is convex.*

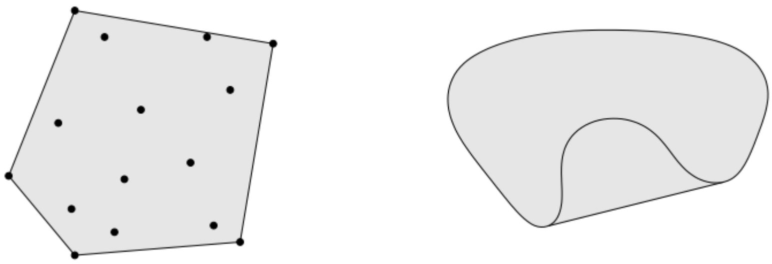
Proof. Let $x_1, x_2 \in C, \theta_1, \theta_2 \geq 0$ such that $\theta_1 + \theta_2 = 1$. Then for each $\alpha \in I$, $x_1, x_2 \in C_\alpha$, hence, $\theta_1 x_1 + \theta_2 x_2 \in C_\alpha$ for each $\alpha \in I$. That is, $\theta_1 x_1 + \theta_2 x_2 \in C$. □

Definition 1.3. Let $S \subseteq \mathbb{R}^n$ be any set, the set of all convex combinations of the points in S is called *convex hull* of S and denoted by **conv** S:

$$\textbf{conv}\, S = \{\theta_1 x_1 + ... + \theta_k x_k |\ x_i \in S,\ \theta_i \geq 0,\ i = 1, ..., k,\ \theta_1 + ... + \theta_k = 1\}.$$

In Figure 1.2, the right figure illustrates the convex hull of a kidney shaped non-convex set and the left one shows the convex hull of set of distinct points which create a pentagon.

Figure 1.2: Examples of convex hulls of non-convex sets



Corollary 1.1. ***conv**S is the smallest convex set that contains S. Equivalently, **conv**S is the intersection of all convex sets containing S.*

Proof. Clearly, $S \subset$ **conv**S. Let C be an arbitrary convex set so that $S \subset C$. By Theorem 1.1, any convex combination of the points in S is contained in C, that is **conv**$S \subset C$.

To prove the second statement, let's define

$$F = \{C : C \text{ is convex, } S \subset C\}.$$

Let $\tilde{C} = \cap_{C \in F} C$. By Theorem 1.2, $\tilde{C}$ is convex. Since $S \subset \tilde{C}$. By the first statement, **conv**$S \subset \tilde{C}$. Also, **conv**$S \in F$. Therefore, **conv**$S = \tilde{C}$. □

Cones

Definition 1.4. A set $C \subseteq \mathbb{R}^n$ is called a *cone*, if $\theta x \in C$ for every $x \in C$ and $\theta \in \mathbb{R}^+$.

Definition 1.5. A set C is called a *convex cone*, if it is a convex set and it satisfies the properties of a cone. (*i.e.* for all $x_1, x_2 \in C$ and $\theta_1, \theta_2 \in \mathbb{R}^+$, $\theta_1 x_1 + \theta_2 x_2 \in C$).

Definition 1.6. Let $x_1, ..., x_k \in \mathbb{R}^n$ and $\theta_1, ..., \theta_k \in \mathbb{R}^+$. The point $\theta_1 x_1 + ... + \theta_k x_k$ is called *conic combination* of x_i's.

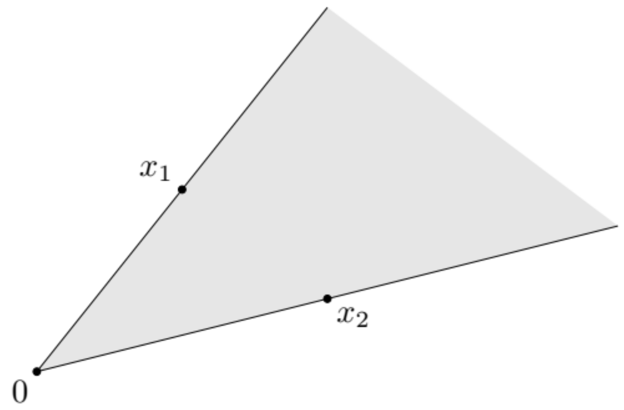
Definition 1.7. Let $C \subseteq \mathbb{R}^n$ be a set. The set of all conic combination of the points in C is called *conic hull* of the set C.

$$\{\theta_1 x_1 + ... + \theta_k x_k | \; x_i \in C, \;\; \theta_i \geq 0, \;\; i = 1, ..., k\}.$$

Note that conic hull of set C is the smallest convex cone that contains C.

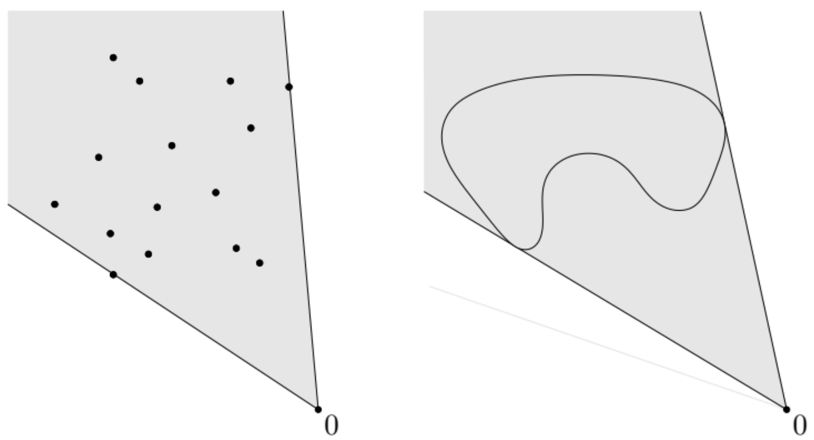
The Figure 1.3 illustrates the smallest cone created by $x_1, x_2 \in \mathbb{R}^2$. In other words, the cone contains all points of the form $\lambda_1 x_1 + \lambda_2 x_2$, where $\lambda_1, \lambda_2 \geq 0$. The point

Figure 1.3: The smallest cone created by two points $x_1, x_2 \in \mathbb{R}^2$



$\lambda_1 = \lambda_2 = 0$ corresponds to the apex of the cone which is 0. This cone is the subset of all cones that contain x_1 and x_2.

Figure 1.4: Conic hulls of two non-convex sets



In Figure 1.4, we can see the conic hull of a kidney shaped non-convex set and conic hull of the set of points.

1.1.1 Some important examples of Convex sets

In this section we will discuss some examples of convex sets that are frequently used in the optimization problems.

- **Hyperplanes and halfspaces**

  Hyperplane is the solution set of a group of linear equations. The set can be shown as following:

$$\{x \mid a^T x = b\}$$

where $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$. This set geometrically can be interpreted as the set of points with the same inner product to a given vector a.

Theorem 1.3. *Hyperplanes are convex sets.*

Proof. Let $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$ define the following hyperplane:

$$H = \{x \mid a^T x = b\}$$

.

We want to show that H contains the convex combination of its points. Suppose $x_1, x_2 \in H$ (*i.e.* $a^T x_1 = b$ and $a^T x_2 = b$) and $0 \leq \lambda \leq 1$.

$$a^T(\lambda x_1 + (1-\lambda)x_2) = \lambda a^T x_1 + (1-\lambda)a^T x_2 = \lambda b + (1-\lambda)b = b$$

This means that $\lambda x_1 + (1-\lambda)x_2 \in H$ and consequently the set H is convex. $\square$

Any hyperplane divides the space (*i.e.* $\mathbb{R}^n$) into two *halfspaces* that can be shown as following:

$$\{x \mid a^T x \leq b\} \quad \text{and} \quad \{x \mid a^T x \geq b\}$$

where $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$. The halfspace can be interpreted as the solution set of linear inequalities.

Theorem 1.4. *Halfspaces are convex sets.*

Proof. The proof is analogous to the proof given for convexity of hyperplanes. Let $K = \{x \mid a^T x \leq b\}$ be the halfspace defined by $a \in \mathbb{R}^n$, $a \neq 0$, and $b \in \mathbb{R}$. We want to show that K contains the convex combination of its points. Suppose $x_1, x_2 \in K$ (*i.e.* $a^T x_1 \leq b$ and $a^T x_2 \leq b$) and $0 \leq \lambda \leq 1$.

$$a^T(\lambda x_1 + (1-\lambda)x_2) = \lambda a^T x_1 + (1-\lambda)a^T x_2 \leq \lambda b + (1-\lambda)b = b$$

$$a^T(\lambda x_1 + (1-\lambda)x_2) \le b$$

This means that $\lambda x_1 + (1-\lambda)x_2 \in K$ and consequently the halfspace K is convex. It can be easily shown that the halfspace $\{x \mid\ a^T x \ge b\}$ is also a convex set. $\square$

Figure 1.5: Hyperplane defined by $a^T x = b$

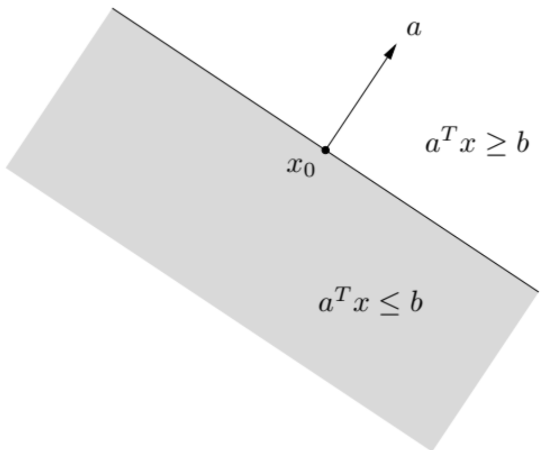


Figure 1.5 illustrates the hyperplane defined by $a^T x = b$ in $\mathbb{R}^2$. The hyperplane creates two halfspaces as $a^T x \ge b$ and $a^T x \le b$.

We will use the convexity of hyperplanes and halfspaces in the upcoming chapters.

- **Euclidean balls**

We can define a *Euclidean ball* in $\mathbb{R}^n$ using the definition of Euclidean norm. Let $u \in \mathbb{R}^n$ be a vector, Euclidean norm is denoted by $\|.\|_2$ and can be defined as $\|u\|_2 = (u^T u)^{1/2}$. Using this definition, Euclidean ball with center $x_c \in \mathbb{R}^n$ and radius $r \in \mathbb{R}^+$ is as following:

$$B(x_c, r) = \{x \mid \|x - x_c\|_2 \le r\} = \{x \mid (x - x_c)^T (x - x_c) \le r^2\}$$

A Euclidean ball can be interpreted as the set of points that their distance to x_c is less than or equal to $r \ge 0$.

Theorem 1.5. *A Euclidean ball is a convex set.*

Proof. Let's denote the Euclidean ball with center $x_c \in \mathbb{R}^n$ and radius $r \in \mathbb{R}^+$ by $B(x_c, r)$. Suppose $x_1, x_2 \in B(x_c, r)$ (*i.e.* $\|x_1 - x_c\|_2 \leq r$ and $\|x_2 - x_c\|_2 \leq r$) and $0 \leq \lambda \leq 1$. We want to explore whether the point $\lambda x_1 + (1 - \lambda)x_2$ belongs to the ball $B(x_c, r)$ or not.

$$\begin{aligned}
\|\lambda x_1 + (1 - \lambda)x_2 - x_c\|_2 &= \|\lambda(x_1 - x_c) + (1 - \lambda)(x_2 - x_c)\|_2 \\
&\leq \lambda\|x_1 - x_c\|_2 + (1 - \lambda)\|x_2 - x_c\|_2 \\
&\leq r.
\end{aligned}$$

□

Note that in the above proof we used the triangular inequality for Euclidean norm.

- **Polyhedra**

  A *polyhedra* is the intersection of finite number of halfspaces and hyperplanes. In other words, it is the solution set of linear equalities and inequalities which can be illustrated as following:

$$P = \{x \mid a_i^T x \leq b_i, \;\; i = 1, ..., m, \;\; c_j^T x = d_j, \;\; j = 1, ..., p\}.$$

  A polyhedra can be shown in the compact form as following:

$$P = \{x \mid Ax \preceq b, \;\; Cx = d\}$$

  where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{p \times n}$, and $d \in \mathbb{R}^p$. We denote *vector inequality* or *componentwise inequality* with $\preceq$ symbol.

  Theorem 1.6. *A polyhedra is a convex set.*

  Proof. Prove of this theorem follows immediately from Theorem 1.2 as a polyhedra is an intersection of finite number of convex sets. □

  Definition 1.8. A bounded polyhedra is called *polytope*.

- **The positive semi-definite cone**

Definition 1.9. Let $A \in \mathbb{R}^{n\times n}$ be a symmetric matrix. The matrix A is called positive semi-definite if for any $x \in \mathbb{R}^n$ we have $x^T Ax \geq 0$. It is called positive definite if for any $x \in \mathbb{R}^n$ we have $x^T Ax > 0$. We denote the set of $n\times n$ symmetric matrices by S^n and the set of symmetric $n \times n$ positive semi-definite matrices by S^n_+:

$$S^n = \{X \in \mathbb{R}^{n\times n} \mid X = X^T\},$$

$$S^n_+ = \{X \in S^n \mid X \succeq 0\}$$

Theorem 1.7. *The set of symmetric positive semi-definite matrices, S^n_+, is a convex cone.*

Proof. Let's first show that S^n_+ is a cone. Suppose $A \in S^n_+$ and $\theta \in \mathbb{R}^+$. From the definition of positive semi-definiteness we have for any $x \in \mathbb{R}^n$ we have $x^T Ax \geq 0$. Then $x^T \theta Ax \geq 0$ which implies that $\theta A \in S^n_+$. This proves that S^n_+ is a cone.

Now let's investigate the convexity of S^n_+. Suppose $A, B \in S^n_+$ and $\theta_1, \theta_2 \in \mathbb{R}^+$, then for any $x \in \mathbb{R}^n$:

$$x^T(\theta_1 A + \theta_2 B)x = \theta_1 x^T Ax + \theta_2 x^T Bx \geq 0$$

which implies that $\theta_1 A + \theta_2 B$ is a positive semi-definite matrix (*i.e.* $\theta_1 A + \theta_2 B \in S^n_+$). □

1.1.2 Operations that preserve convexity of sets

In this section we will discuss two operations that does not change the convexity of sets.

- **Intersection**

  Theorem 1.8. *Let $A, B \subseteq \mathbb{R}^n$ be two convex sets. Then $A \cap B$ is a convex set.*

  Proof. A more general case of this theorem is stated in Theorem 1.2. □

- **Affine functions**

Definition 1.10. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a function. We say that L is a *linear function*, if

- for any vector $x, y \in \mathbb{R}^n$, $L(x + y) = L(x) + L(y)$, and
- for any $x \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, $L(\alpha x) = \alpha L(x)$.

Definition 1.11. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. We say that the function f is *affine*, if there exist a linear function $L : \mathbb{R}^n \to \mathbb{R}^m$ and a vector $b \in \mathbb{R}^m$ such that f can be written as $f(x) = L(x) + b$.

Theorem 1.9. *Let $S \subseteq \mathbb{R}^n$ be a convex set and $f : S \to \mathbb{R}^m$ be an affine function. The image of S under f is a convex set.*

Proof. Let's define the image of a set S under the affine function f as following:

$$f(S) = \{f(x) \mid x \in S\}$$

Since, the function f is an affine function, there exist a linear function $L : S \to \mathbb{R}^m$ and a vector $b \in \mathbb{R}^m$ such that $f(x) = L(x) + b$.

Suppose $y_1, y_2 \in f(S)$ are two points in the image set and $0 \leq \lambda \leq 1$. We explore whether the convex combination of two points are in the image set or not (*i.e.* $\lambda y_1 + (1 - \lambda)y_2 \in f(S)$).

Since, $y_1, y_2 \in f(S)$, there exist $x_1, x_2 \in S$ such that $f(x_1) = y_1$ and $f(x_2) = y_2$. Then,

$$\begin{aligned}
\lambda y_1 + (1 - \lambda)y_2 &= \lambda f(x_1) + (1 - \lambda)f(x_2) \\
&= \lambda(L(x_1) + b) + (1 - \lambda)(L(x_2) + b) \\
&= (\lambda L(x_1) + (1 - \lambda)L(x_2)) + b \\
&= L(\lambda x_1 + (1 - \lambda)x_2) + b \\
&= f(\lambda x_1 + (1 - \lambda)x_2)
\end{aligned}$$

Since, S is a convex set and $x_1, x_2 \in S$, then for any $0 \leq \lambda \leq 1$ the point $\lambda x_1 +$

$(1-\lambda)x_2 \in S$. Therefore, $f(\lambda x_1 + (1-\lambda)x_2) \in f(S)$ and $\lambda y_1 + (1-\lambda)y_2 \in f(S)$ which proves the convexity of the image set. $\square$

Remark. *Let $S \subseteq \mathbb{R}^n$ be a convex set. Then the images of S under translation, scaling, and projection are convex sets.*

Theorem 1.10. *Let $A, B \subseteq \mathbb{R}^n$ be two convex sets. Then, the set $A + B = \{a + b \mid a \in A \ , \ b \in B\}$ is convex.*

Proof. Let $y_1, y_2 \in A + B$. Then, there exists $a_1, a_2 \in A$ and $b_1, b_2 \in B$ such that $y_1 = a_1 + b_1$ and $y_2 = a_2 + b_2$. Let $\lambda \in [0, 1]$, we want to show that $\lambda y_1 + (1-\lambda)y_2 \in A + B$.

$$\begin{aligned} \lambda y_1 + (1-\lambda)y_2 &= \lambda(a_1 + b_1) + (1-\lambda)(a_2 + b_2) \\ &= [\lambda a_1 + (1-\lambda)a_2] + [\lambda b_1 + (1-\lambda)b_2] \end{aligned}$$

Since A and B are convex sets, hence, $\lambda a_1 + (1-\lambda)a_2 \in A$ and $\lambda b1 + (1-\lambda)b_2 \in B$. This proves that $\lambda y_1 + (1-\lambda)y_2 \in A + B$.

$\square$

Theorem 1.11. *Let $A, B \subseteq \mathbb{R}^n$ be two convex sets. Then, the set $A \times B = \{(a, b) \mid a \in A \ , \ b \in B\}$ is convex.*

Proof. Let $y_1, y_2 \in A \times B$. Then, there exists $a_1, a_2 \in A$ and $b_1, b_2 \in B$ such that $y_1 = (a_1, b_1)$ and $y_2 = (a_2, b_2)$. Let $\lambda \in [0, 1]$, we want to show that $\lambda y_1 + (1-\lambda)y_2 \in A \times B$.

$$\begin{aligned} \lambda y_1 + (1-\lambda)y_2 &= \lambda(a_1, b_1) + (1-\lambda)(a_2, b_2) \\ &= (\lambda a_1 + (1-\lambda)a_2, \lambda b_1 + (1-\lambda)b_2) \end{aligned}$$

Since A and B are convex sets, hence, $\lambda a_1 + (1-\lambda)a_2 \in A$ and $\lambda b1 + (1-\lambda)b_2 \in B$. This proves that $\lambda y_1 + (1-\lambda)y_2 \in A \times B$.

$\square$

1.1.3 Proper Cones and Generalized Inequalities

Definition 1.12. Let $K \subseteq \mathbb{R}^n$ be a cone. We call K a *proper cone* if:

- K is a convex cone.
- K is a closed cone.
- The interior if K is nonempty. In other words, K is *solid*.
- No line is contained in K (*i.e.* if $x \in K$ and $-x \in K$ then $x = 0$). In other words, K is *pointed*.

Definition 1.13. $\leq$ is a partial ordering on a set S if $\forall x, y \in S$,

1. $x \leq x$,
2. $x \leq y, \;\; y \leq x$ implies $x = y$,
3. $x \leq y, \;\; y \leq z$ implies $x \leq z$.

A proper cone $K \subset \mathbb{R}^n$ can be utilized to define a partial ordering in $\mathbb{R}^n$. *Generalized inequality* can be defined as following:

$$x \preceq_K y \iff y - x \in K$$

Additionally, a strict partial ordering can be defined as following:

$$x \prec_K y \iff y - x \in \textbf{int}\, K$$

where $x, y \in \mathbb{R}^n$.

1.1.4 Properties of Generalized inequalities

Generalized inequality $\preceq_K$ satisfies many of the properties of standard ordering in $\mathbb{R}$:

- if $x \preceq_K y$ and $u \preceq_K v$, then $x+u \preceq_K y+v$ (*i.e.* generalized inequality is preserved under addition),
- if $x \preceq_K y$ and $y \preceq_K z$, then $x \preceq_K z$ (*i.e.* generalized inequality is *transitive*),

- if $x \preceq_K y$ and $\alpha \in \mathbb{R}^+$, then $\alpha x \preceq_K \alpha y$ (*i.e.* generalized inequality is preserved under nonnegative scaling),
- $x \preceq_K x$ (*i.e.* generalized inequality is *reflexive*),
- if $x \preceq_K y$ and $y \preceq_K x$, then $x = y$ (*i.e.* generalized inequality is *antisymmetric*),
- if $x_i \preceq_K y_i$ for $i = 1, 2, ..., x_i \to x$, and $y_i \to y$ as $i \to \infty$, then $x \preceq_K y$ (*i.e.* generalized inequality is preserved under limit).

Also, note that there are some properties for strict partial ordering,$\prec_K$, which are similar to that of strict standard ordering in $\mathbb{R}$.

1.1.5 Minimum and minimal elements

Generalized inequalities and standard ordering in $\mathbb{R}$ share some properties which are mentioned in the previous section. However, there is an important property that holds for standard ordering, but it does not hold for generalized inequalities. In $\mathbb{R}$ all points are *comparable* that is if $x, y \in \mathbb{R}$ then either $x \leq y$ or $y \leq x$ holds. This property is not true for the case when we use generalized inequalities. This means that there are some points that are not comparable. To elucidate the concept let us investigate an example. Let's consider $\mathbb{R}^2_+$ as a proper cone, $x = [2 \;\; 5]$, and $y = [5 \;\; 1]$. Then, neither $x \preceq_{\mathbb{R}^2_+} y$ holds nor $y \preceq_{\mathbb{R}^2_+} x$. This means that two points x and y are not comparable with respect to generalized inequality $\preceq_{\mathbb{R}^2_+}$.

Since generalized inequalities affect the comparability of points, the concepts of *minimum* elements and *minimal* elements are more complex.

Definition 1.14. Let $S \subseteq \mathbb{R}^n$ be a set and $K \subset \mathbb{R}^n$ be a proper cone. The point $x \in S$ is the *minimum element* of S with respect to generalized inequality $\preceq_K$, if for every $y \in S$ we have $x \preceq_K y$.

Definition 1.15. Minimum element (alternative definition) Let $S \subseteq \mathbb{R}^n$ be a set and $K \subset \mathbb{R}^n$ be a proper cone. Let us $x + K = \{x + y \;:\; y \in K\}$ denote the set of all points which are comparable to x and are greater than or equal to x with respect to the generalized inequality $\preceq_K$. The point x is the minimum element of the set S if

$$S \subseteq x + K$$

Figure 1.6: A set which has a minimum element

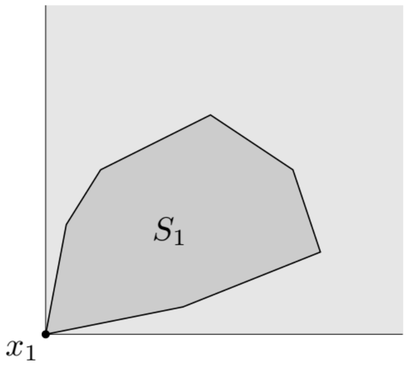


Figure 1.6 illustrates a set S_1 that has a minimum element x_1 in $\mathbb{R}^2$ with respect to $\mathbb{R}^2_+$. As it is shown in the figure, the set S_1 is the subset of $x_1 + K$ which is shaded in the figure.

Theorem 1.12. *Minimum element (if it exists) is unique.*

Proof. Let K be a proper cone, and S be a convex set. Suppose, $x_1, x_2 \in S$ be two minimum elements of S with respect to a generalized inequality $\preceq_K$. Then,

$$x_1 \preceq_K x_2, \ \ and \ \ x_2 \preceq_K x_1,$$

hence, $x_1 = x_2$. □

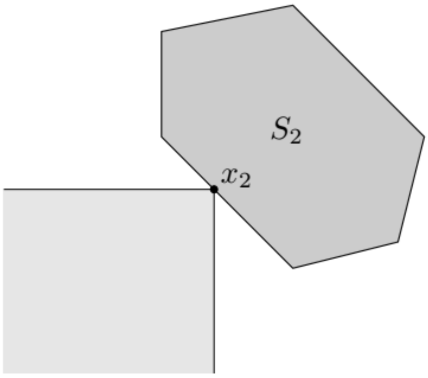
Definition 1.16. Let $S \subseteq \mathbb{R}^n$ be a set and $K \subset \mathbb{R}^n$ be a proper cone. The point $x \in S$ is the *minimal element* of S with respect to generalized inequality $\preceq_K$, if for $y \in S$, $y \preceq_K x$ only if $y = x$. In other words, the point $y \in S$ is either incomparable with x or $x \preceq_K y$.

Definition 1.17. Minimal element (alternative definition) Let $S \subseteq \mathbb{R}^n$ be a set and $K \subset \mathbb{R}^n$ be a proper cone. Let $x - K = \{x - y \ : \ y \in K\}$ denote the set of all points which are comparable to x and are less than or equal to x with respect to the generalized inequality $\preceq_K$. The point x is a minimal element of the set S if

$$(x - K) \cap S = \{x\}$$

The Figure 1.7 shows that the only point in the intersection of S_2 and $x_2 - K$ is x_2. Therefore, the point x_2 is minimal point of S_2. Clearly, a minimal element is not unique. As it can be seen from the figure, all points on the edge of S_2 that contains x_2 are minimal

Figure 1.7: The point x_2 is the minimal point of the set S_2



elements. The set S_2 does not contain any element that satisfies the definition of minimum element. Hence, it does not contain any minimum elements.

Note that the concepts of *maximum* element and *maximal* element can be defined in the similar way.

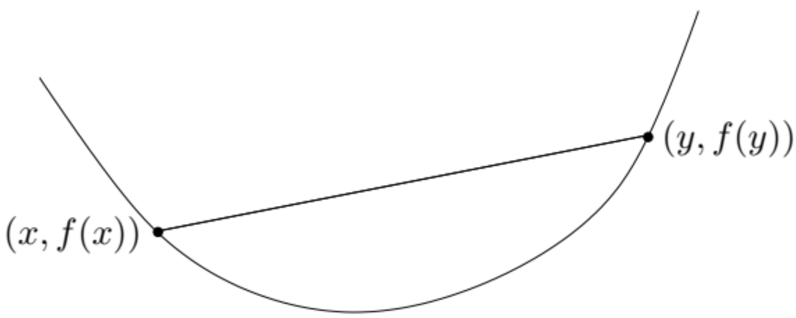
1.2 Convex Functions and Related Definitions

Definition 1.18. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and $\mathbf{dom} f$ denote the domain of it. The function f is convex if its domain is a convex set and for every $x, y \in \mathbf{dom} f$, and $0 \leq \theta \leq 1$:

$$f(\theta x + (1-\theta)y) \leq \theta f(x) + (1-\theta)f(y). \tag{1}$$

The geometric interpretation of inequality (1) is that the line segment joining two points $(x, f(x))$ and $(y, f(y))$ lies above the graph of the function f.

Figure 1.8: Graph of a convex function



The Figure 1.8 pictures a convex function in $\mathbb{R}^2$. As it is illustrated, the line segment joining any two points on the graph lies above it.

Definition 1.19. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and **dom**f denote the domain of it. The function f is concave if its domain is a convex set and $-f$ is a convex function.

Theorem 1.13. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function and* **dom**f *denote its domain. The function f is convex if and only if for all $x \in$* ***dom**f* *and for all v the function $g(t) = f(x + tv)$ is a convex function on its domain (i.e. $\{t \mid x + tv \in$* ***dom**f*$\}$).*

Proof. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function and $t_1, t_2 \in$ **dom**g. Then, for every v and $0 \leq \theta \leq 1$ we have the following relations:

$$\begin{aligned}
g(\theta t_1 + (1-\theta)t_2) &= f(x + (\theta t_1 + (1-\theta)t_2)v) \\
&= f((\theta + 1 - \theta)x + (\theta t_1 + (1-\theta)t_2)v) \\
&= f(\theta(x + t_1 v) + (1-\theta)(x + t_2 v)) \\
&\leq \theta f(x + t_1 v) + (1-\theta)f(x + t_2 v)) \\
&= \theta g(t_1) + (1-\theta)g(t_2))
\end{aligned}$$

which proves the convexity of $g(t)$. To prove the converse, suppose $g : \mathbb{R} \to \mathbb{R}$ be a convex function on its domain. Since **dom** g is a convex set, one can conclude that **dom**f is also a convex set. Let $x, r, s \in$ **dom**f, since **dom**f is convex, there exist v and $t_1, t_2 \in$ **dom** g such that $r = x + t_1 v$ and $s = x + t_2 v$. Then, for every $0 \leq \theta \leq 1$ we have the following relations:

$$\begin{aligned}
f(\theta r + (1-\theta)s) &= f(\theta(x + t_1 v) + (1-\theta)(x + t_2 v)) \\
&= g(\theta t_1 + (1-\theta)t_2) \\
&\leq \theta g(t_1) + (1-\theta)g(t_2)) \\
&\leq \theta f(x + t_1 v) + (1-\theta)f(x + t_2 v)) \\
&= \theta f(r) + (1-\theta)f(s))
\end{aligned}$$

which proves the convexity of $f(x)$.

□

Note that the above theorem implies that a convex function is convex on all lines that intersects with its domain. Since there is an if and only if condition in the theorem, the reverse of the previous statement is also true.

Definition 1.20. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then, the *gradient* of function f is denoted by $\nabla f(x)$ and defined as following:

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

Theorem 1.14. ***First-order condition*** *Let f be a differentiable function over its domain. The function f is convex if and only if for every $x, y \in dom f$ we have*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) \tag{2}$$

Proof. Let's first consider a convex function $f : \mathbb{R} \to \mathbb{R}$. The theorem can be expressed as f is convex if and only if for all x and y in dom$f$

$$f(y) \geq f(x) + f(x)'(y - x) \tag{3}$$

Consider two points $x, y \in$ dom$f$, then since the domain of f is a convex function $(1 - t)x + ty \in$ dom$f$ for all values of $0 < t \leq 1$. Based on the convexity of f we have:

$$f((1 - t)x + ty) \leq (1 - t)f(x) + tf(y)$$

.

Let's divide both sides with by t and rearrange the inequality, then we have

$$f(y) \geq f(x) + \frac{f((1 - t)x + ty) - f(x)}{t}$$

we can replace $(1 - t)x + ty$ by $x + t(y - x)$ and obtain

$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t}$$

now, let's take the limit as $t \to 0$ then we will obtain

$$f(y) \geq f(x) + (y - x)f'(x)$$

For proving the converse, let f be a function that satisfies inequality (3) for all $x, y \in \text{dom} f$. Let's take a point $z = \theta x + (1 - \theta)y$ for $0 \leq \theta \leq 1$ and $x \neq y$. Now we apply the inequality (3) for x and z

$$f(x) \geq f(z) + f'(z)(x - z)$$

multiplying the above inequality by θ and replacing $x - z$ by $(1 - \theta)(x - y)$ we have

$$\theta f(x) \geq \theta f(z) + \theta(1 - \theta)f'(z)(x - y) \tag{4}$$

Now let's apply the inequality (3) for y and z

$$f(y) \geq f(z) + f'(z)(y - z)$$

multiplying the above inequality by $(1 - \theta)$ and replacing $y - z$ by $\theta(y - x)$ we have

$$(1 - \theta)f(y) \geq (1 - \theta)f(z) + \theta(1 - \theta)f'(z)(y - x) \tag{5}$$

summing two inequalities (4) and (5) we obtain

$$\theta f(x) + (1 - \theta)f(y) \geq f(z)$$

which proves the convexity of f.

Now using the previous part, we want to prove the theorem for a function $f : \mathbb{R}^n \to \mathbb{R}$. Let's define a function $g(t) = f(ty + (1 - t)x)$ for $x, y \in \mathbb{R}^n$ and $0 \leq t \leq 1$ where $g'(t) = \nabla f(ty + (1 - t)x)^T(y - x)$.

First let's consider that f is a convex function. The function g is a composition of a convex function with an affine function, so, it is also convex. Since $g : \mathbb{R} \to \mathbb{R}$ we can use the above mentioned results and obtain $g(1) \geq g(0) + g'(0)$ or equivalently

$$f(y) \geq f(x) + \nabla f(x)^T(y - x).$$

To prove the converse of the theorem, let's assume that the inequality (2) holds for every $x, y \in \text{dom} f$. Also let $ty + (1-t)x \in \text{dom} f$ and $\tilde{t}y + (1-\tilde{t})x \in \text{dom} f$ for $0 \leq t, \tilde{t} \leq 1$. Using the inequality we have

$$f(ty + (1-t)x) \geq f(\tilde{t}y + (1-\tilde{t})x) + \nabla f(\tilde{t}y + (1-\tilde{t})x)^T (y-x)(t-\tilde{t})$$

this inequality corresponds to $g(t) \geq g(\tilde{t}) + g'(\tilde{t})(t-\tilde{t})$ which shows the convexity of g and consequently the convexity of f.

□

The inequality (2) states that tangent line at any point $x \in \text{dom} f$ is always a *global underestimator* for a convex function f. Also, the inequality illustrates that a point x is a global minimizer of a convex function f, if $\nabla f(x) = 0$, since, for all $y \in \text{dom} f$ $f(y) \geq f(x)$.

Definition 1.21. Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then, the *hessian* of function f is an $n \times n$ matrix that is denoted by $\nabla^2 f(x)$ and defined as following:

$$\nabla^2 f(x) = \left[\frac{\partial^2 f}{\partial x_j \partial x_k}(x) \right]_{j,k=1,\dots,n}.$$

Theorem 1.15. ***Second-order conditions*** *Let $S \subseteq \mathbb{R}^n$ be an open convex set and $f : S \to \mathbb{R}$ be a twice differentiable function at each point of its domain (i.e. Hessian or second order derivative exists at every point). Then, the function f is convex if and only if for all $x \in S$ Hessian is positive semi-definite ($\nabla^2 f(x) \succeq 0$).*

Proof. Let's first consider $n = 1$ and $S \subseteq \mathbb{R}$. Also, let $f : S \to \mathbb{R}$ is a convex function and $x, y \in S$ with $y > x$. Using Theorem 1.14 we have the following inequalities:

$$f(y) \geq f(x) + f^{'}(x)(y-x) \tag{6}$$

$$f(x) \geq f(y) + f^{'}(y)(x-y) \tag{7}$$

The above inequalities yields the following one:

$$f^{'}(x)(y-x) \leq f(y) - f(x) \leq f^{'}(y)(y-x)$$

which means that

$$f^{'}(x)(y-x) \leq f^{'}(y)(y-x)$$

If we subtract left term from the right one and dividing the resulting term by $(y-x)^2$ we will obtain the following inequality:

$$\frac{f^{'}(y)-f^{'}(x)}{y-x} \geq 0$$

.

Now if we take the limit $y \rightarrow x$, we will have $f^{''}(x) \geq 0$.

To prove the converse direction, suppose $f^{''}(z) \geq 0$ for all $z \in \mathbf{dom} f$. Let $x, y \in \mathbf{dom} f$ and $x < y$. Then we have the following inequality:

$$\int_x^y f^{''}(z)(y-z)dz \geq 0$$

If we solve the above integral using integration by part, we will have the following results:

$$\begin{aligned}\int_x^y f^{''}(z)(y-z)dz &= (f^{'}(z)(y-z))|_{z=x}^{z=y} + \int_x^y f^{'}(z)dz \\ &= -f^{'}(x)(y-x) + f(y) - f(x),\end{aligned}$$

The above results implies that $f(y) \geq f(x) + f^{'}(x)(y-x)$ which based on the Theorem 1.14 shows that f is a convex function.

To prove the general case where $n > 1$, we use Theorem 1.13. Based on the theorem, f is a convex function on its domain if and only if the function $g(t) = f(x_0 + tv)$ is convex on its domain. Based on our proof for $n = 1$, one can conclude that $g(t)$ is convex if and only if $g^{''}(t) \geq 0$ for all $v \in \mathbb{R}^n$,$t \in \mathbf{dom}\ g$, and $x_0 \in \mathbf{dom} f$ which means that

$$g^{''}(t) = v^T \nabla^2 f(x_0 + tv) v \geq 0$$

The inequality illustrates that Hessian is positive semi-definite ($\nabla^2 f(x) \succeq 0$ for all $x \in$**dom**f). Therefore, it is necessary and sufficient condition for a convex function f to have a positive semi-definite Hessian in every point of its domain.

□

Sublevel sets

Definition 1.22. Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a function and C_α be a set that is defined as

$$C_\alpha = \{x \in \mathrm{dom} f \mid f(x) \leq \alpha\}.$$

The set C_α is called α-*sublevel set* of function f.

Theorem 1.16. *Let $\alpha \in \mathbb{R}$ and $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Then, for all values of α, sublevel sets of f are convex.*

Proof. Let's fix $\alpha \in \mathbb{R}$ and define C_α as sublevel set of f. Suppose, $0 \leq \theta \leq 1$ and $x, y \in C_\alpha$ (*i.e.* $f(x) \leq \alpha$ and $f(y) \leq \alpha$). Then, using the convexity of f, we have following relations:

$$\begin{aligned} f(\theta x + (1-\theta)y) &\leq \theta f(x) + (1-\theta)f(y) \\ &\leq \theta\alpha + (1-\theta)\alpha = \alpha \end{aligned}$$

which implies that $\theta x + (1-\theta)y \in C_\alpha$ and C_α is a convex set.

□

1.2.1 Operations that preserve the convexity of a function

In this section we will discuss a number of operations which do not affect the convexity of functions. To get the complete list of such operations and more details interested readers are referred to [1].

- **Composition with an affine function**

  Theorem 1.17. *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a convex function. Let $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$ so that we can define a new function $g(x)$ as follows:*

  $$g(x) = f(Ax + b),$$

where $g : \mathbb{R}^m \to \mathbb{R}$ with $\textbf{dom}\,g = \{x \mid Ax + b \in \textbf{dom}\,f\}$. *If f is a convex function, then g is a convex function.*

Proof. Let's first examine whether domain of g is a convex set or not. The set is the intersection of hyperplane defined by $A \in \mathbb{R}^{n\times m}$ and $b \in \mathbb{R}^n$ with domain of f. Hyperplanes are convex sets. Domain of f is also a convex set, since, the function f is a convex function. Additionally, intersection preserves the convexity of sets. Therefore, one can conclude that the domain of g is a convex set.

Now, let $\lambda \in [0,1]$ and $x_1, x_2 \in \textbf{dom}\,g$ so that

$$g(x_1) = f(Ax_1 + b), \quad g(x_2) = f(Ax_2 + b)$$

Then, using the convexity of f we have the following results:

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= f(A(\lambda x_1 + (1-\lambda)x_2) + b) \\ &= f(\lambda(Ax_1 + b) + (1-\lambda)(Ax_2 + b)) \\ &\le \lambda f(Ax_1 + b) + (1-\lambda) f(Ax_2 + b) \\ &= \lambda g(x_1) + (1-\lambda) g(x_2), \end{aligned}$$

which shows the convexity of $g(x)$. □

- **Pointwise maximum and supremum**

Theorem 1.18. *Let $f_1, f_2 : \mathbb{R}^n \to \mathbb{R}$ be to convex functions with domain* $\textbf{dom}\,f_1$ *and* $\textbf{dom}\,f_2$*, respectively. Let f be a pointwise maximum function of f_1 and f_2 as follows:*

$$f(x) = \max\{f_1(x), f_2(x)\},$$

with domain, $\textbf{dom}\,f = \textbf{dom}\,f_1 \cap \textbf{dom}\,f_2$. *The function f is a convex function.*

Proof. Since, $\textbf{dom}\,f_1$ and $\textbf{dom}\,f_2$ are convex sets and intersection preserves convexity of sets, $\textbf{dom}\,f$ is a convex set. Let $0 \le \lambda \le 1$ and $x, y \in \textbf{dom}\,f$. Also, without loss of generality, suppose that $f(\lambda x+(1-\lambda)y) = f_1(\lambda x+(1-\lambda)y)$. Then, using the convexity of f_1 we have the followings:

$$\begin{aligned} f(\lambda x+(1-\lambda)y) &= f_1(\lambda x+(1-\lambda)y) \\ &\le \lambda f_1(x) + (1-\lambda) f_1(y) \\ &\le \lambda \max\{f_1(x), f_2(x)\} + (1-\lambda)\max\{f_1(y), f_2(y)\} \\ &= \lambda f(x) + (1-\lambda) f(y), \end{aligned}$$

The above inequalities together with the convexity of **dom**f prove the convexity of the function f.

$\square$

Note that the above theorem can be extended as follows. If the functions $f_1, ..., f_m$ are convex function. Their pointwise maximum function f

$$f(x) = \max\{f_1(x), ..., f_m(x)\},$$

is a convex function. The prove follows the similar direction as in the previous theorem.

The pointwise maximum can be extended to the pointwise supremum of an infinite set of convex functions. In the next theorem we will present the extended case.

Theorem 1.19. *Let A be any set with infinite number of elements and $f : \mathbb{R}^n \times A \to \mathbb{R}$ be a function so that $f(x, y)$ is a convex function in x for every $y \in A$. Suppose $g : \mathbb{R}^n \to \mathbb{R}$ is a function as follows:*

$$g(x) = \sup_{y\in A} f(x, y)$$

*where **dom**g = $\{x \mid (x, y) \in$ dom$f$ for all $y \in A,\ \sup_{y\in A} f(x, y) < \infty\}$. The function $g(x)$ which is a pointwise supremum over an infinite set of convex functions is a convex function.*

Proof. From the definition of **dom**g one can conclude that it is a convex set. Let $x_1, x_2 \in$**dom**g and $0 \leq \lambda \leq 1$, then we have

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= \sup_{y \in A} f(\lambda x_1 + (1-\lambda)x_2, y) \\ &\leq \sup_{y \in A} \lambda f(x_1, y) + (1-\lambda) f(x_2, y) \\ &\leq \lambda \sup_{y \in A} f(x_1, y) + (1-\lambda) \sup_{y \in A} f(x_2, y) \\ &= \lambda g(x_1) + (1-\lambda) g(x_2) \end{aligned}$$

which proves the convexity of the function $g(x)$. □

- **Minimization**

Theorem 1.20. *Suppose $C \subseteq \mathbb{R}^m$ be a convex nonempty set and $f : \mathbb{R}^{n \times m} \to \mathbb{R}$ be a convex function. Let $g(x) = \inf_{y \in C} f(x, y)$ be a function that $g(x) > -\infty$ for all x. Then the function g is a convex function in its domain (i.e.* ***dom****g $=$ $\{x \mid (x, y) \in$* ***dom****f for some $y \in C\}$).*

Proof. Let $x_1, x_2 \in$ **dom**g and $\epsilon > 0$. There exist $y_1, y_2 \in C$ such that $f(x_i, y_i) \leq g(x_i) + \epsilon$ for $i = 1, 2$. Suppose $0 \leq \lambda \leq 1$; then we have:

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= \inf_{y \in C} f(\lambda x_1 + (1-\lambda)x_2, y) \\ &\leq f(\lambda x_1 + (1-\lambda)x_2, \lambda y_1 + (1-\lambda)y_2) \\ &\leq \lambda f(x_1, y_1) + (1-\lambda) f(x_2, y_2) \\ &\leq \lambda g(x_1) + (1-\lambda) g(x_2) + \epsilon. \end{aligned}$$

Since the above inequalities are true for any $\epsilon > 0$, we have the following results:

$$g(\lambda x_1 + (1-\lambda)x_2) \leq \lambda g(x_1) + (1-\lambda) g(x_2).$$

□

- **Composition**

  Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function that is the composition of two functions $h : \mathbb{R}^k \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^k$ where $f(x) = h(g(x)) : \mathbb{R}^n \to \mathbb{R}$ and $\textbf{dom}f = \{x \in \textbf{dom}g \mid g(x) \in \textbf{dom}h\}$. In this section we will discuss the conditions on h and g that guarantee the convexity or concavity of f.

  For the simplicity we assume $n = 1$. For more general case, $n > 1$, one can use Theorem 1.13 and its implication. The theorem states that it is enough to show the convexity of f on an arbitrary line which intersects with $\textbf{dom}f$.

  We start with considering $k = 1$ (*i.e.* $h : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$). Let's assume that h and g are twice differentiable functions. With those conditions, the function $f(x) = h(g(x))$ is convex if $f^{''}(x) \geq 0$ for all $x \in \mathbb{R}$. Now taking the second derivative from f will yield the following equation:

  $$f^{''}(x) = h^{''}(g(x))g^{'}(x)^2 + h^{'}(g(x))g^{''}(x). \tag{8}$$

  The conditions that make $f^{''} \geq 0$ ($f^{''} \leq 0$) are the ones that make each term of the equation (8) non-negative (non-positive). From the equation (8) the following results can be obtained:

  - if h is a convex and non-decreasing function (*i.e.* $h^{''} \geq 0$ and $h^{'} \geq 0$, respectively) and g is a convex function (*i.e.* $g^{''} \geq 0$), then f is a convex function (*i.e.* $f^{''} \geq 0$),
  - if h is a convex and non-increasing function (*i.e.* $h^{''} \geq 0$ and $h^{'} \leq 0$, respectively) and g is a concave function (*i.e.* $g^{''} \leq 0$), then f is a convex function (*i.e.* $f^{''} \geq 0$),
  - if h is a concave and non-decreasing function (*i.e.* $h^{''} \leq 0$ and $h^{'} \geq 0$, respectively) and g is a concave function (*i.e.* $g^{''} \leq 0$), then f is a concave function (*i.e.* $f^{''} \leq 0$),
  - if h is a concave and non-increasing function (*i.e.* $h^{''} \leq 0$ and $h^{'} \leq 0$, respectively) and g is a convex function (*i.e.* $g^{''} \geq 0$), then f is a concave function (*i.e.* $f^{''} \leq 0$).

Now let's consider more general case, $k > 1$. Suppose $h : \mathbb{R}^k \to \mathbb{R}$, $g : \mathbb{R}^n \to \mathbb{R}^k$, and $g_i : \mathbb{R}^n \to \mathbb{R}$ for $i = 1, ..., k$ be twice differentiable functions. Then, the function $f : \mathbb{R}^n \to \mathbb{R}$ can be illustrated as following:

$$f(x) = h(g(x)) = h(g_1(x), ..., g_k(x)).$$

Again in this case, without loss of generality we consider $n = 1$ and **dom**$g = \mathbb{R}$ and **dom**$h = \mathbb{R}^k$. To investigate the conditions that guarantee the convexity/concavity of f, let's take its second derivative:

$$f^{''}(x) = g^{'}(x)^T \nabla^2 h(g(x)) g^{'}(x) \ + \ \nabla h(g(x))^T g^{''}(x), \tag{9}$$

From equation (9), one can obtain the following conditions to guarantee the convexity or concavity of f:

- if h is a convex and non-decreasing in each argument (*i.e.* $\nabla^2 h(g(x)) \geq 0$ and $\nabla h(g_i(x)) \geq 0$ for all $i = 1, ..., k$, respectively), and g_i's are convex functions (*i.e.* $g_i^{''} \geq 0$, and $g^{''} \geq 0$), then f is a convex function (*i.e.* $f^{''} \geq 0$),
- if h is a convex and non-increasing in each argument (*i.e.* $\nabla^2 h(g(x)) \geq 0$ and $\nabla h(g_i(x)) \leq 0$ for all $i = 1, ..., k$, respectively), and g_i's are concave functions (*i.e.* $g_i^{''} \leq 0$, and $g^{''} \leq 0$), then f is a convex function (*i.e.* $f^{''} \geq 0$),
- if h is a concave and non-decreasing in each argument (*i.e.* $\nabla^2 h(g(x)) \leq 0$ and $\nabla h(g_i(x)) \geq 0$ for all $i = 1, ..., k$, respectively), and g_i's are concave functions (*i.e.* $g_i^{''} \leq 0$, and $g^{''} \leq 0$), then f is a concave function (*i.e.* $f^{''} \leq 0$).

Note that we discussed the conditions that are valid when both h and g are twice differentiable functions. For exploring the conditions in more general setting where h and g might not be differentiable, the extended-value functions must be defined. This concept is not included in this work. Interested readers can be referred to [1].

1.2.2 Convexity with respect to generalized inequality

Definition 1.23. Let $K \subseteq \mathbb{R}^m$ be a proper cone and $\preceq_K$ be the associated generalized inequality. The function $f : \mathbb{R}^n \to \mathbb{R}^m$ is convex with respect to the generalized inequality or it is called *K-convex* if for all $x, y \in \mathbf{dom} f$, and $0 \leq \lambda \leq 1$, we have

$$f(\lambda x + (1-\lambda)y) \preceq_K \lambda f(x) + (1-\lambda) f(y). \tag{10}$$

The function is called *strictly K-convex* if for all $x \neq y \in \mathbf{dom} f$ and $0 \leq \lambda \leq 1$ we have:

$$f(\lambda x + (1-\lambda)y) \prec_K \lambda f(x) + (1-\lambda) f(y). \tag{11}$$

The generalized inequalities and the related concepts are mostly utilized in vector optimization that we will discuss later.

2 Optimization problems

2.1 Basic Terminology

Let $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$, $i = 0, ..., m$ and $h_j : \mathbb{R}^n \rightarrow \mathbb{R}$, $j = 1, ..., p$ be functions and let's define the set D as following:

$$D = \bigcap_{i=0}^{m} dom f_i(x) \cap \bigcap_{i=1}^{p} dom h_j(x)$$

where $dom f_i(x)$ and $dom h_j(x)$ are domains of the mentioned functions. If we define the set C as

$$C = \{x \in D | f_i(x) \leq 0, i = 1, ..., m, h_j(x) = 0, j = 1, ..., p\}$$

then the optimization problem is the problem of finding $x \in C$ which minimizes the function $f_0(z)$. An optimization problem can be formulated as following:

$$\begin{aligned} \text{minimize} \quad & f_0(z) \\ \text{subject to} \quad & f_i(z) \leq 0 \quad i = 1, ..., m \\ & h_j(z) = 0 \quad j = 1, ..., p \end{aligned} \tag{12}$$

where z is the *optimization variable* and $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ is the *objective function* of the problem. The inequalities $f_i(x) \leq 0$ are called *inequality constraints* and the equalities $h_j(x) = 0$ are *equality constraints*. If an optimization problem does not have any constraints then the problem (12) is called *unconstrained*.

Any point $x \in C$ is called a *feasible point*, so the set C can be considered as the collection of feasible points and it is called the *feasible set*. The problem (12) is *feasible* if the set C contains at least one point and it is *infeasible* if the set C is empty. The inequality constraint $f_i(x) \leq 0$ is called *active* at a feasible point x if $f_i(x) = 0$ and it is called *inactive* if $f_i(x) < 0$.

When the problem (12) has a solution at a point $x \in C$, the value of the objective function at x is called the *optimal value*. Let p^* denote the optimal value. By definition

$$p^* = inf\{f_0(x)|x \in C\}$$

where p^* can take the values $\pm\infty$. When the problem is infeasible, then by definition we set $p^* = \infty$. The problem (12) is *unbounded below* if there are some feasible points $x_k \in C$ such that $f_0(x_k) \to -\infty$ as $k \to \infty$. In this case $p^* = -\infty$.

Optimal and locally optimal points

If the problem (12) is feasible and there exists $x^* \in C$ such that $f_0(x^*) = p^*$ then x^* is an *optimal point* for problem (12).

A feasible point $x \in C$ is called *locally optimal* if there exists $R > 0$ such that x solves the following optimization problem with variable z:

$$\begin{array}{lll} \text{minimize} & f_0(z) & \\ \text{subject to} & f_i(z) \leq 0 & i = 1, ..., m \\ & h_j(z) = 0 & j = 1, ..., p \\ & \|z - x\|_2 \leq R & \end{array}$$

which means that the feasible point x is the minimizer of f_0 over the neighbourhood points. To distinguish between local optimal and optimal points the term "global optimal" sometimes is used.

Maximization problems

The maximization problem

$$\begin{array}{lll} \text{maximize} & f_0(x) & \\ \text{subject to} & f_i(x) \leq 0 & i = 1, ..., m \\ & h_j(x) = 0 & j = 1, ..., p \end{array} \tag{13}$$

is equivalent to the problem of minimizing $-f_0(x)$ over the same feasible region. Therefore, for solving maximization problems we can convert them to an equivalent minimization problem and solve them.

2.2 Equivalent Problems

Two optimization problems are *equivalent* if from solution of one of them we can find the solution of the other one easily. In this part we will discuss some techniques for obtaining equivalent problems.

- **Scaling**

  We can obtain the equivalent of the standard optimization problem (12) by scaling the objective function and inequality constraints by a positive scalar and equality constraints by a non-zero scalar. Let $\alpha_i > 0$ for $i = 0, ..., m$ and $\beta_j \neq 0$ for $j = 1, ..., p$ be scalars. The equivalent problem can be written as following:

$$\begin{array}{ll} \text{minimize} & \tilde{f}(x) = \alpha_0 f_0(x) \\ \text{subject to} & \tilde{f}_i(x) = \alpha_i f_i(x) \leq 0 \quad i = 1, ..., m \\ & \tilde{h}_j(x) = \beta_j h_j(x) = 0 \quad j = 1, ..., p \end{array} \tag{14}$$

  Since the scaling of inequality constraints are made by positive scalars and scaling of equality constraints are made by non-zero scalars the feasible region of problem (14) and problem (12) are the same. Also, the optimal solution for one of the problems is also an optimal point for another one since the scaling of the objective function is made by a positive scalar. Note that although the feasible region and the optimal solution of two problems are the same, two problems are not the same since their objective functions and constraints are different. Two problems are the same if $\alpha_i = 1$ for $i = 0, ..., m$ and $\beta_j = 1$ for $j = 1, ..., p$, otherwise they are equivalent.

- **Change of variables**

  Another form of obtaining an equivalent problem is to substitute the original decision variable with a new one. For this purpose, let $\phi : \mathbb{R}^n \to \mathbb{R}^n$ be a one-to-one function such that range of ϕ be the subset of the domain of the optimization problem (*i.e.* $Range(\phi) \subseteq D$). Then the new decision variable can be defined as z such that $x = \phi(z)$. The equivalent problem with new decision variable z can be formulated as following:

$$
\begin{array}{lll}
\text{minimize} & \tilde{f}_0(z) & \\
\text{subject to} & \tilde{f}_i(z) \leq 0 & i = 1, ..., m \\
& \tilde{h}_j(z) = 0 & j = 1, ..., p
\end{array} \tag{15}
$$

where $\tilde{f}_i(z) = f_i(\phi(z))$ for $i = 0, ..., m$ and $\tilde{h}_j(z) = h_j(\phi(z))$ for $j = 1, ..., p$. If x is a solution for problem (12) then $z = \phi^{-1}(x)$ is a solution for the problem (15). Also, if z is the solution for problem (15) then $x = \phi(z)$ is the solution for problem (12). In this case, two problems are equivalent with *change of variable*.

- **Transformations of objective and constraint functions**

  Consider the following problem:

$$
\begin{array}{lll}
\text{minimize} & \tilde{f}_0(x) & \\
\text{subject to} & \tilde{f}_i(x) \leq 0 & i = 1, ..., m \\
& \tilde{h}_j(x) = 0 & j = 1, ..., p
\end{array} \tag{16}
$$

  where the functions $\tilde{f}_i$ and $\tilde{h}_j$ are defined as composition of functions. Let $\psi_i : \mathbb{R} \rightarrow \mathbb{R}$ be a function such that $\tilde{f}_i = \psi_i(f_i(x))$ for $i = 0, ..., m$ and $\tilde{h}_j = \psi_{m+j}(h_j(x))$ for $j = 1, ..., p$. The problems (16) and (12) are equivalent if the functions ψ_i satisfy the following conditions:

  - ψ_0 is a monotone increasing function,
  - $\psi_i(u) \leq 0$ for $i = 1, ..., m$ if and only if $u \leq 0$,
  - $\psi_i(u) = 0$ for $i = m + 1, ..., p$ if and only if $u \leq 0$

  Consequently, the feasible region and the optimal set of the problem (16) is the same as the feasible and optimal set of problem (12). Note that the scaling method, discussed above, is a special case of obtaining equivalent problem by transforming objective and constraint functions where all ψ_is are linear.

- **Slack variables**

  One common way to obtain an equivalent problem is to use slack variables to change inequality constraints into equality ones. We introduce new variables $s_1, ..., s_m \geq$

0 so that $f_i(x) + s_i = 0$ for every $i = 1, ..., m$, and $x \in C$. Using this fact we can obtain an equivalent problem as following:

$$\begin{array}{lll} \text{minimize} & f_0(x) & \\ \text{subject to} & s_i \geq 0, & i = 1, ..., m \\ & f_i(x) + s_i = 0 & i = 1, ..., m \\ & h_j(x) = 0 & j = 1, ..., p \end{array} \tag{17}$$

where the variables $x \in \mathbb{R}^n$ and $s \in \mathbb{R}^m$. The variables s_i that is used to replace inequality constraints with equality and non-negativity constraints are *slack variables*. The problem (17) has $n + m$ decision variables, m inequality (non-negativity) constraints, and $m + p$ equality constraints.

Note that if the feasible (optimal) solution of problem (17) is (x, s) then x is a feasible (optimal) solution for the original problem (12). The converse is also true. If x is a feasible (optimal) solution for problem (12), then the solution (x, s) where $s_i = -f_i(x)$ for $i = 1, ..., m$ is feasible (optimal) for the problem (17).

- **Eliminating equality constraints**

Recall the equality constraints of an optimization problem (12):

$$h_j(x) = 0, \quad j = 1, ..., p,$$

For eliminating equality constraints we need a function $\phi : \mathbb{R}^k \to \mathbb{R}^n$ such that x satisfies the above equality if and only if there exists $z \in \mathbb{R}^k$ such that $x = \phi(z)$. In other words, the solution set of the equality constraints can be parametrized by variable $z \in \mathbb{R}^k$. Then the equivalent optimization problem can be formulated as following:

$$\begin{array}{ll} \text{minimize} & \tilde{f}_0(z) = f_0(\phi(z)) \\ \text{subject to} & \tilde{f}_i(z) = f_i(\phi(z)) \leq 0 \quad i = 1, ..., m \end{array} \tag{18}$$

where $z \in \mathbb{R}^k$ is the decision variable. The equivalent problem has no equality and m inequality constraints. Note that if x is the optimal solution for the problem (12), then any z that satisfies $x = \phi(z)$ is the optimal solution for the equivalent problem. Since x is a feasible solution for problem (12), there exists at least one such z. Converse of this statement is also true. If z is an optimal solution for the equivalent problem, then $x = \phi(z)$ is an optimal solution for the problem (12).

- **Eliminating linear equality constraints**

  Let's suppose that the equality constraints are linear of the form $Ax = b$ and the solution set of them are not empty which otherwise the problem is infeasible. Let's $F \in \mathbb{R}^{n\times k}$ be any full rank matrix such that $R(F) = N(A)$ and x_0 be any solution of the equality constraints. Then the general solutions of $Ax = b$ can be written as $Fz + x_0$ where $z \in \mathbb{R}^k$. Then we can substitute the general solution form in the problem (12) and eliminate the equality constraints. The equivalent problem can be formulated as following:

$$\begin{array}{ll} \text{minimize} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0 \quad i = 1, ..., m \end{array} \tag{19}$$

  The equivalent problem with $z \in \mathbb{R}^k$ as decision variable has m inequality and zero equality constraints. Additionally, the new problem has rank of A fewer decision variables, since from *rank-nullity* theorem we have $k = n - rank(A)$.

- **Introducing equality constraints**

  In this part we will discuss the method which is the converse of the above mentioned technique. In the problems that the objective and inequality constraint functions are in the form of composition of the functions with affine functions this method can be implemented to tackle the complexity of the problem. Consider the following problem:

$$\begin{array}{lll} \text{minimize} & f_0(A_0x + b_0) & \\ \text{subject to} & f_i(A_ix + b_i) \leq 0 & i = 1, ..., m \\ & h_j(x) = 0, & j = 1, ..., p, \end{array} \tag{20}$$

  where $x \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{k_i\times n}$, $f_i : \mathbb{R}^{k_i} \to \mathbb{R}$, and $A_ix + b_i$ are affine functions. In order to obtain the equivalent of the above problem new decision variable $y_i \in \mathbb{R}^{k_i}$ such that $y_i = A_ix + b_i$ for $i = 0, ..., m$ can be introduced. The equivalent problem can be formulated as following:

$$\begin{array}{lll} \text{minimize} & f_0(y_0) & \\ \text{subject to} & f_i(y_i) \leq 0 & i = 1, ..., m \\ & y_i = A_i x + b_i, & i = 0, ..., m \\ & h_j(x) = 0, & j = 1, ..., p, \end{array} \tag{21}$$

where $y_0 \in \mathbb{R}^{k_0}, \ \ldots, y_m \in \mathbb{R}^{k_m}$. Therefore, the equivalent problem has $k_0 + ... + k_m$ new decision variables and new equality constraints in addition to the decision variables and constraints of the original problem (20).

- **Optimizing over some variables**

It is always true that in problems with more than one decision variables, we can minimize the problem in terms of one variable and then minimize the resulted problem in terms of other ones. Therefore, the fact

$$\inf_{x,y} f(x, y) = \inf_{x} \tilde{f}(x)$$

where $\tilde{f}(x) = \inf_y f(x, y)$ can be used to obtain an equivalent problem is some specific problems.

Suppose that $x \in \mathbb{R}^n$ can be partitioned into $x_1 \in \mathbb{R}^{n_1}$ and $x_2 \in \mathbb{R}^{n_2}$ so that $x = (x_1, x_2)$ and $n_1 + n_2 = n$. Consider an optimization problem that the constraint functions depends on only x_1 or x_2:

$$\begin{array}{lll} \text{minimize} & f_0(x_1, x_2) & \\ \text{subject to} & f_i(x_1) \leq 0 & i = 1, ..., m_1 \\ & \tilde{f}_i(x_2) \leq 0, & i = 1, ..., m_2. \end{array} \tag{22}$$

In order to obtain an equivalent problem, the problem first is minimized in terms of one of the decision variables. Here we first implement the optimization on x_2. For this purpose the new objective function $\tilde{f}_0$ is defined as following:

$$\tilde{f}_0(x_1) = \inf\{f_0(x_1, z) \mid \tilde{f}_i(z) \leq 0, \ i = 1, ..., m_2\}$$

Using the new objective function, the equivalent problem of (22) can be formulated as following:

$$\begin{aligned} &\text{minimize} \quad \tilde{f}_0(x_1) \\ &\text{subject to} \quad f_i(x_1) \leq 0 \quad i = 1, ..., m_1. \end{aligned} \tag{23}$$

- **Epigraph problem form**

Let $X \subseteq \mathbb{R}^n$ be the domain of a function $f : X \to [-\infty, +\infty]$. Then, the *epigraph* of the function f is a set that is subset of $\mathbb{R}^{n+1}$ and defined as following:

$$epi(f) = \{(x, t) \mid x \in X, \ t \in \mathbb{R}, \ f(x) \leq t\}.$$

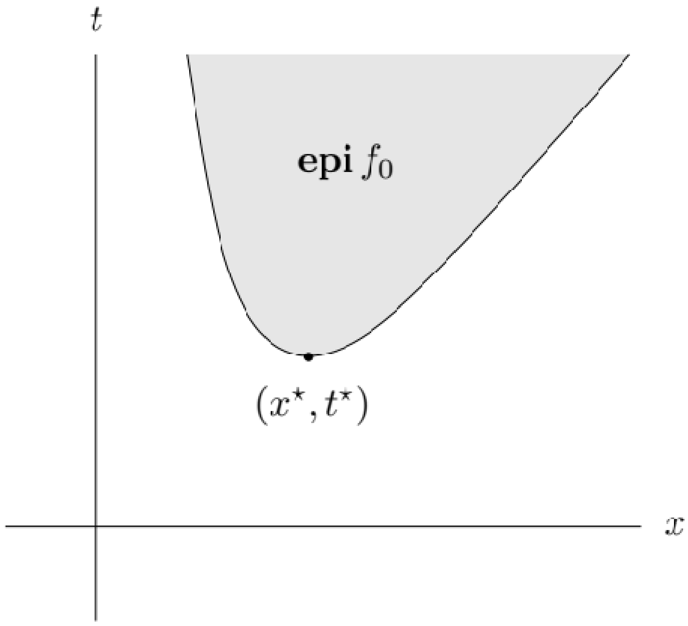
The *epigraph form* of an optimization problem (12) is as following:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \qquad i = 1, ..., m \\ & h_j(x) = 0, \qquad j = 1, ..., p, \end{aligned} \tag{24}$$

where $t \in \mathbb{R}$ and $x \in \mathbb{R}^n$ are the decision variables. Geometrically, the problem (24) can be described as minimizing the decision variable t over the epigraph of the function f_0 such that the constraints on x are satisfied. As an example, let us consider an optimization problem that minimizes the function that is illustrated in Figure 2.1. The problem is to minimize $f_0(x)$ over its domain. The epigraph form problem is to find the lowest point in the epigraph of $f_0(x)$. Therefore, the point (x^*, t^*) is the optimal point.

Note that the optimization problems (12) and (24) are equivalent. We note that (x^*, t^*) is the optimal solution of the problem (24) if and only if x^* is the optimal solution of the original problem (12) where $t^* = f_0(x^*)$.

Figure 2.1: Geometric interpretation of epigraph form of an optimization problem



3 Convex Optimization

3.1 Convex optimization problems

A convex optimization problem can be formulated as following:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0 \quad i = 1, .., m \\ & a_j^T x = b_j \quad j = 1, .., p \end{aligned} \tag{25}$$

where $f_0, .., f_m$ are convex functions. Convex optimization problem requires three additional conditions in comparison with problem (12):

- the objective function is convex,
- the inequality constraint functions must be convex,
- the equality constraint functions $h_j(x) = a_j^T x - b_j$ must be affine.

The domain of problem (25) which is $D = \bigcap_{i=0}^{m} dom f_i$ is a convex set. Also, the inequality constraints are sublevel sets of convex function (*i.e.* $\{x | f_i(x) \leq 0\}$). Since by Theorem 1.16 the sublevel set of convex functions are convex and intersection of convex sets yields a convex set, we can conclude that the set created by inequality constraints is convex. Furthermore, since the equality constraints are affine the set created by intersection of them is the intersection of hyperplanes $\{x | a_j^T x = b_j\}$ which is also a convex set. Finally, the feasible set of problem (25) is the intersection of set D, sublevel sets, and hyperplanes which therefore is a convex set. We can conclude that a convex optimization problem is a problem of minimizing a convex function over a convex set.

Concave maximization problem

The problem below is called *convex optimization problem* if the objective function is a concave function, inequality constraint functions are convex, and equality constraint functions are affine.

$$\begin{aligned} &\text{maximize} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0 \quad i = 1, .., m \\ & && a_j^T x = b_j \quad j = 1, .., p \end{aligned} \tag{26}$$

This problem can be solved by the minimizing $-f_0(x)$ (which is a convex function) over the same feasible region.

3.2 Local and global optima

Theorem 3.1. *Any feasible point that is local optimal for a convex optimization problem is also global optimal.*

Proof. To prove this property, let us assume that x is a locally optimal point for problem (25) that is for some $R > 0$ we have

$$f_0(x) = \inf\{f_0(z) | z \text{ is feasible}, \|z - x\|_2 \leq R\}$$

Let us on the contrary suppose that x is not globally optimal. This means that there exists a feasible y such that $\|y - x\|_2 > R$ and $f_0(y) < f_0(x)$. Let t be the point between x and y such that

$$t = (1 - \theta)x + \theta y$$

for $\theta = \frac{R}{2\|y-x\|_2}$. Then using this value for θ, we have $\|t - x\|_2 = R/2 < R$ which means that t is in the neighborhood of x and by our assumption $f_0(x) \leq f_0(t)$. However, using the convexity of the function f_0:

$$f_0(t) \leq (1 - \theta)f_0(x) + \theta f_0(y) < (1 - \theta)f_0(x) + \theta f_0(x) = f_0(x)$$

which contradicts with our assumption. Therefore, x is globally optimal and $f_0(x) \leq f_0(y)$ for all feasible y.

□

3.3 An optimality criterion for differentiable convex function

In this section, we will discuss optimality condition for convex function and prove this property by using Theorem 1.14 which states First-order condition for being a convex function.

Theorem 3.2. ***optimality condition*** *Let X be the feasible set of convex optimization problem (25), a point x is optimal if and only if $x \in X$ and*

$$\nabla f_0(x)^T(y - x) \geq 0 \textit{ for all } y \in X \tag{27}$$

Proof. Let's first suppose that x is an optimal point and the condition (27) does not hold for x that is there exist some $y \in X$ such that

$$\nabla f_0(x)^T(y - x) < 0.$$

We want to show that this condition contradicts with the optimality of x in its neighbourhood. Let us consider $z(t) = ty + (1 - t)x$ for $t \in [0, 1]$. Since the feasible set is a convex set $z(t)$ is feasible. We want to show that for small value of t the point x is not optimal. To show this we look at the derivative of f_0 in $z(t)$

$$\frac{d}{dt} f_0(z(t))|_{t=0} = \nabla f_0(x)^T(y - x) < 0$$

This means that f_0 is decreasing at $z(t)$ for small values of t. Therefore $f_0(z(t)) < f_0(x)$ which contradicts with our assumption that x is an optimal point.

Conversely, let $x \in X$ which satisfies (27), and since f_0 is a convex function then, based on Theorem 1.14, it satisfies the first-order condition of convex functions for $y \in$ domf_0 (*i.e.* $f(y) \geq f(x) + \nabla f(x)^T(y - x)$). Therefore, $f_0(y) \geq f_0(x)$ for $y \in$ domf_0 which proves the optimality of x.

□

Example 3.1. ***Unconstrained problems*** *We want to obtain the optimality criteria for an unconstrained optimization problem using condition (27). In this case the feasibility condition simply reduces to $x \in$ domf_0. Let x be an optimal point for our problem. Then we have $\nabla f_0(x)^T(y - x) \geq 0$ for all feasible y.*

The function f_0 is a differentiable function; therefore, its domain is an open set. This means that for every feasible x there exists $r > 0$ such that $B(x, r) = \{z \mid \|x - z\| \leq r\} \subseteq$ domf_0 (i.e. an open ball containing points which are feasible for the problem). If we define a point $y = x - t\nabla f_0(x)$, for small and positive amounts of t we have $y \in B$ which implies such y is a feasible point.

Thus, the optimality condition can be modified as following:

$$\nabla f_0(x)^T(y - x) = -t\|\nabla f_0(x)\|_2^2 \geq 0$$

the above condition implies that

$$\nabla f_0(x) = 0. \tag{28}$$

Therefore, the optimality condition (27) reduces to (28).

Every solution to the equality (28) is a minimizer of f_0. On the other hand, if the equation does not have any solution, then there does not exist any optimal solutions.

Example 3.2. ***Problems with equality constraints*** *Consider an optimization problem with only equality constraints*

$$\begin{array}{ll} \textit{minimize} & f_0(x) \\ \textit{subject to} & Ax = b \end{array}$$

where $A \in \mathbb{R}^{p\times n}$. Let's assume that the problem is feasible (i.e. the feasible set is non-empty). For a feasible point x the optimality condition is as following

$$\nabla f_0(x)^T(y - x) \geq 0$$

for all y satisfying $Ay = b$. The feasible region in this problem is affine, so, there exists $\nu \in N(A)$ so that y can be written as $y = x + \nu$. The optimality condition can be modified as following

$$\nabla f_0(x)^T\nu \geq 0 \textit{ for all } \nu \in N(A)$$

For $\nu \in N(A)$ we have $-\nu \in N(A)$, so, $-\nabla f_0(x)^T\nu \geq 0$ which implies that $\nabla f_0(x)^T\nu = 0$ for all $\nu \in N(A)$. This means that $\nabla f_0(x) \perp N(A)$ and using the fact *that $N(A)^{\perp} = R(A^T)$, the optimality condition can be reduced to $\nabla f_0(x) \in R(A^T)$. Which means that there exists $u \in R^p$ so that*

$$\nabla f_0(x) + A^T u = 0 \tag{29}$$

To show the fact that $N(A)^{\perp} = R(A^T)$, let $x \in N(A)$ which means that $Ax = 0$. This equation means that $x \perp \{$row space of $A\}$ and consequently $x \perp \{$column space (range) of $A^T\}$. Since this is true for all $x \in N(A)$, we have $N(A) \perp R(A^T)$.

Conversely, let $y \in R(A^T)$ (i.e. there exists x such that $y = A^T x$) and $z \in N(A)$ (i.e. $Az = 0$). Then, we have

$$y^T z = (A^T x)^T z = (x^T A) z = x^T (Az) = 0.$$

Since this is true for every $y \in R(A^T)$, we have $R(A^T) \perp N(A)$. Consequently, $R(A^T) = N(A)^{\perp}$.

Example 3.3. ***Minimization over the nonnegative orthant*** *Consider the following optimization problem which is minimization of a convex function over nonnegative orthant that is the only constraints are nonnegativity constraints:*

$$\begin{aligned} &\textit{minimize} \quad f_0(x) \\ &\textit{subject to} \quad x \succeq 0 \end{aligned} \tag{30}$$

Recall the optimality condition (27) which will be modified for problem (30) as following:

$$x \succeq 0, \quad \nabla f_0(x)^T (y - x) \geq 0 \;\; \textit{for all} \;\; y \succeq 0. \tag{31}$$

Since the condition $\nabla f_0(x)^T y - \nabla f_0(x)^T x \geq 0$ must be true for all $y \geq 0$, some conditions on $\nabla f_0(x)$ and $-\nabla f_0(x)^T x$ are required to be specified.

First we need the first term, $\nabla f_0(x)^T y$, to be bounded below (i.e. nonnegative) for $y \geq 0$. So, $\nabla f_0(x)^T \succeq 0$. Then, we need $-\nabla f_0(x)^T x \geq 0$, where $x \geq 0$ and $\nabla f_0(x)^T \succeq 0$. This only happens when $\nabla f_0(x)^T x = 0$, that is $\sum_{i=1}^{n} (\nabla f_0(x))_i x_i = 0$. Each term in the summation is the product of nonnegative numbers; therefore, for the summation be equal to zero, each term must be equal to zero ($(\nabla f_0(x))_i x_i = 0$ for $i = 1, ..., n$).

Finally the optimality condition for an optimization of a convex function over nonnegative orthant can be illustrated as following:

$$x \succeq 0, \quad \nabla f_0(x) \succeq 0, \quad x_i(\nabla f_0(x))_i = 0, \quad i = 1, ..., n. \tag{32}$$

The last condition means that the set of indices corresponding to non-zero elements in two vectors, x and $\nabla f_0(x)$ have no intersections. This property is called complementarity, the set are indices are complement to each other.

3.4 Equivalent convex problems

In this section we investigate the transformations that preserve convexity. We want to know whether the transformations that yield equivalent optimization problems preserve the convexity property of the original problem or not.

- **Eliminating equality constraints**

  Recall the procedure of eliminating equality constraint form an optimization problem in section 2.2. For a convex optimization problem the equality constraints are affine and of the form $Ax = b$. Let x_0 be a solution for this system of equations and F be a matrix that its range is equal to the nullity of A. Then, the equivalent optimization problem which contains no equality constraints is formulated as following:

  $$\begin{array}{ll} \text{minimize} & f_0(Fz + x_0) \\ \text{subject to} & f_i(Fz + x_0) \leq 0, \quad i = 1, ..., m. \end{array}$$

  Note that in the problem (3.4) the objective function and the inequality constraint functions are composition of a convex function and an affine function, Such a composition does not destroy the convexity of the problem. Therefore, eliminating the equality constraints in a convex optimization problem preserves convexity of the problem.

- **Introducing equality constraints**

  In section 2.2 we discussed the instruction to introduce new equality constraints to the problem when the objective and constraint functions are the composition of a convex function with an affine function. The resulted equivalent problem formulated as 21. If the new introduced constraints are linear, then, the convexity of the problem is not affected. In other words, introducing equality constraints to the problem preserves the convexity as long as the new constraints are linear.

- **Slack variables**

  Introducing slack variables changes the inequality constraints into equality ones ($f_i(x) \leq 0$ becomes $f_i(x) + s_i = 0$). Since in the convex optimization problem the equality constraints must be affine, the convexity of the problem is preserved as long as $f_i(x)$ is a linear function. Therefore, adding slack variables to the linear inequality constraints does not affect the convexity of the problem.

- **Epigraph form**

  Recall the epigraph form of a convex optimization problem with variables (x, t),

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_0(x) - t \leq 0 \\ & f_i(x) \leq 0, \qquad i = 1, ..., m \\ & h_j(x) = 0, \qquad j = 1, ..., p, \end{array}$$

  Note that the objective function of the epigraph form problem is linear in terms of t, the inequality constraint function $f_0(x) - t$ is convex function of (x, t). Therefore, the epigraph form of a convex optimization problem is also convex.

- **Minimizing over some variables**

  In section 2.2 we followed an instruction to obtain an equivalent problem by minimization over a decision variable. In problem (22), if the objective function is jointly convex in (x_1, x_2), and inequality constraint functions are convex, then equivalent problem (23) is a convex optimization problem. This claim is valid since based on Theorem 1.20 minimization of a convex function over a variable preserves convexity.

4 Linear Programming and Applications

4.1 Linear optimization problems

The problem (25) is called *linear program* (LP) when the objective function and the constraints are linear (or affine) functions. We can formulate a general linear program as following:

$$\begin{aligned} \text{minimize} \quad & c^T x + d \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b \end{aligned} \tag{33}$$

where $G \in \mathbb{R}^{m \times n}$ and $A \in \mathbb{R}^{p \times n}$. Note that the inequality sign means componentwise and the decision variables are unrestricted in sign which means that they can be any number. Since the feasible region of an LP is intersection of half-spaces and hyperplanes, it is a polyhedron. Therefore, an LP can be interpreted geometrically as the minimization of an linear cost function over a polyhedron.

Inequality and Standard form of an LP

An LP with only inequalities is called *inequality form LP* and is formulated as following:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Gx \preceq h \end{aligned} \tag{34}$$

An LP in *standard form* is the minimization of an affine function over the intersection of non-negative orthant $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x \succeq 0\}$ and a feasible hyperplane $\{x : Ax = b\}$. An LP in *standard form* can be shown as following:

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & Ax = b \\
& x \succeq 0
\end{aligned} \tag{35}
$$

where the only inequality is non-negativity of decision variables which means that every component of decision variable must be non-negative.

These two forms of LP are very common and are used in designing the algorithm for solving LPs.

Converting LPs to standard form

To convert an LP (33) to standard form we can define slack variables s and add to the inequality constraint in order to convert it to equality constraint. Then, we can define two non-negative new decision variables x^+ and x^- such that $x = x^+ - x^-$ in order to get rid of unrestricted variables and have non-negativity constraint. After introducing new decision variables and substituting them in the LP (33) we obtain the following formulation:

$$
\begin{aligned}
\text{minimize} \quad & c^T x^+ - c^T x^- + d \\
\text{subject to} \quad & Gx^+ - Gx^- + s = h \\
& Ax^+ - Ax^- = b \\
& x^+,\ x^-,\ s \succeq 0
\end{aligned} \tag{36}
$$

the above formulation is called an LP in standard form.

Equivalence of an general LP and standard form LP

Now let us analyze the equivalence of both formulations (33) and (36). Let x be a feasible solution for problem (33). Then let's define s, x^+ and x^- as following:

$$x_i^+ = max\{0, x_i\} \text{ for all } i = 1, .., n$$

$$x_i^- = max\{0, -x_i\} \text{ for all } i = 1, .., n$$

and let $s = h - Gx$. These new variables are non-negative and feasible for the problem (36). Also the objective function of problem (36) can be calculated using the solution of problem (33) as following:

$$c^T x^+ - c^T x^- + d = c^T(x^+ - x^-) + d = c^T x + d$$

Conversely, let s, x^+ and x^- be feasible solution for the problem (36), then let x be defined as $x = x^+ - x^-$, then x will be a feasible solution for problem (33) with the following objective function value:

$$c^T x + d = c^T x^+ - c^T x^- + d = c^T(x^+ - x^-) + d$$

From these two observations, we can conclude that the optimal objective value of the both problems are equal. Also, from feasible solution of one of them we can obtain a feasible solution for another one. Therefore, both problems (33) and (36) are equivalent and every LP in general form can be converted to an LP in standard form.

Note that the concept of linear programming and its applications in operations research is very extensive. For more information on this topic interested readers can be referred to [2].

4.2 Applications of Linear Programming

In this section we will discuss some important applications of linear programming in Operations Research.

- Diet problem

Diet problem is one of the earliest optimization problems that were suggested and modelled by linear programming. The problem initially was suggested by George Stigler as the problem of deciding the amount of foods that should be take by a normal person. The amount of intake must satisfy the recommended dietary allowance for some nutrients with a minimum cost [3].

In order to model this problem for a single person and for a single day, let F and N denote the set of foods and nutrients, respectively. Let $Fmin_i$ and $Fmax_i$ denote the minimum and maximum number of food $i \in F$ that a person can eat in a day. Also, let $Nmin_j$ and $Nmax_j$ denote the minimum and maximum amount of nutrient $j \in N$ that a person is daily allowed to take. Let a_{ij} denote the amount of nutrient $j \in N$ in food $i \in F$ which has a cost of c_i.

The problem is to choose the best combination of foods that minimizes the total cost and satisfies nutrient constraints and the constraints regarding number of foods. We define decision variable x_i as the number of food $i \in F$ to be consumed. The problem can be mathematically modelled as following:

$$\begin{aligned}
\text{minimize} \sum_{i \in F} c_i x_i & \\
\text{subject to} \sum_{i \in F} a_{ij} x_i \geq Nmin_j & \qquad \forall j \in N \\
\sum_{i \in F} a_{ij} x_i \leq Nmax_j, & \qquad \forall j \in N \\
x_i \geq Fmin_i, & \qquad \forall i \in F \\
x_i \leq Fmax_i, & \qquad \forall i \in F
\end{aligned}$$

This simple problem has been the focus point of researchers for many years. They modified this problem and introduced much challenging problems. For example in one of the latest studies, the author suggests a different objective function for Diet problem. In the study, the objective function is considered as minimization of amount of Glycemic Load (GL) in the foods that a person will take. They also proposed that the GL amounts of foods can only be estimated and cannot be measured accurately. Therefore, the amount of GL in foods are uncertain and the problem can be considered as stochastic optimization [4].

- Assignment problem

In Assignment Problem, we have n jobs to be done by machines. There are m machines that each have different capacities. The problem is to assign each job to a machine by considering the capacity of the machines. Optimization problem is to do this assignment with minimum cost. For mathematical formulation, let I and J denote the set of machines and jobs, respectively. Let u_i denote the capacity of machine $i \in I$. The assignment of job $j \in J$ to machine $i \in I$ consumes d_{ij} units of machine i's capacity and costs c_{ij}. In order to formulate this optimization problem, we define x_{ij} as the decision variable of the problem. If job $j \in J$ is assigned to a machine $i \in I$ then variable x_{ij} takes value of 1 and otherwise it is zero. The Assignment problem can be model as following:

$$\text{minimize} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{37}$$

$$\text{subject to} \sum_{i=1}^{m} x_{ij} = 1, \qquad \forall j = 1, ..., n, \tag{38}$$

$$\sum_{j=1}^{n} d_{ij} x_{ij} \leq u_i, \qquad \forall i = 1, ..., m, \tag{39}$$

$$x_{ij} \in \{0, 1\}, \qquad \forall i = 1, ..., m, \ \forall j = 1, ..., n \tag{40}$$

The objective function in above model is to minimize the total cost of assignment. Constraint (38) ensures that each job is assigned to exactly one machine and constraint (39) ensures that the total capacity consumption of the jobs that are assigned to specific machine is less than the capacity of that machine. Finally, in the model we have the constraints regarding possible values for x_{ij} [5].

This problem can be used in different settings. In [6], the authors considered the assignment problem under uncertainties where the capacities of the machines are random. The authors also proposed an efficient solution methodology for solving new model.

This problems is a well-known optimization problem that have many application areas. The assignment problem can be modified to be applied to assignment of air-planes to flight legs by considering the appropriate scheduling of the flight legs [7].

Assignment problem is also very useful in scheduling and assignment of operating rooms in hospitals. In such problems, there are limited number of operating rooms and specialists that must share the room during different time slots. Assigning each specialist to an operating room by considering the availibility of the rooms is a challenging problem [8]. Another version of this problem is to assign patients to operating rooms by considering the availability of the rooms and urgency state of patients [9] [10]. For more information about application of assignment problems in operating rooms scheduling an interested reader can be referred to [11].

- Location-Allocation problem

In Location-Allocation problems, we have a set of destinations which have fixed and known locations and demand for a specific material. The problem is to determine the number, location, and size (capacity) of resource points that can serve the destinations in an optimal way. The objective is to minimize the cost of shipment between resource points and destinations. Common application areas are determination of warehouses, distribution centers, and production facilities [12]. Besides, the most important application of this problems is determination of emergency response points in the cases of disaster. This problem is applicable for determining location of emergency response facilities and allocation of medical resources to those facilities after earthquake and other disasters. There are many researchers who work on post disaster allocation problems by considering various constraints and different settings [13] [14] [15].

The Location-Allocation problem is a kind of challenging and complicated optimization problem. The mathematical modelling of this problem is too complicated to be discussed in our work. Interested readers can be referred to [16] for mathematical modelling and detailed information regarding this problem.

- Transportation problem

Let's assume that there are m origins and n destination points. Each origin i has a_i amount of a commodity and each destination j has d_j amount of demand for that commodity. The Transportation Problem is the problem of deciding how much commodity must be shipped from each origin to destination points to meet the demands with minimum cost. Let I and J be the set of origins and destinations, respectively. Let c_{ij} be the transportation cost between origin $i \in I$ and destination $j \in J$. The decision variable x_{ij} is to decide how much commodity must be shipped from origin $i \in I$ to destination $j \in J$. The Transportation problem can be modelled as following:

$$\text{minimize} \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij} \tag{41}$$

$$\text{subject to} \sum_{i=1}^{m} x_{ij} = d_j, \qquad \forall j \in J \tag{42}$$

$$\sum_{j=1}^{n} x_{ij} = a_i, \qquad \forall i \in I \tag{43}$$

$$x_{ij} \geq 0, \qquad \forall i \in I,\ \forall j \in J \tag{44}$$

In the model the objective function is to minimize the total transportation cost. Constraint (42) ensures that the demand in each destination point will be met and constraint

(43) ensures that the total commodity shipped from each origin is equal to the amount of the commodity that exist there. In some cases, researchers prefer to use $\geq$ in constraint (42) and $\leq$ in constraint (43) instead of equality signs.

Transportation problem can be adapted to many settings. In some cases the researchers consider the problem as multi-objective problem. Minimization of distance, cost, and time can be considered as objective functions. In multi-objective case, the authors considers the combination of different objective functions and try to find the best solutions with respect to those objectives [17] [18]. There are some studies in which the authors consider the problem in uncertain setting [19]. In [20], the transportation cost, demands and the amount of a commodity in each origin are considered as uncertain parameters.

5 Generalized case and vector optimization

5.1 Generalized inequality constraints

Convex optimization problem with generalized inequality constraints can be formulated as following:

$$\begin{aligned} &\text{minimize} && f_0(x) \\ &\text{subject to} && f_i(x) \preceq_{K_i} 0, \quad i = 1, .., m \\ & && Ax = b \end{aligned} \tag{45}$$

where $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function, $K_i \subseteq \mathbb{R}^{k_i}$ are proper cones. and $f_i : \mathbb{R}^n \rightarrow \mathbb{R}^{k_i}$ are $K_i - convex$ constraint functions. Convex optimization problem with generalized inequality constraints is the general case of the problem (25) where the inequality constraint functions are vector valued and the inequalities are general inequalities.

Many general properties of convex optimization problems are also true for convex optimization problems with general inequality constraints. The following properties are some of those:

- All sets including the set of feasible points, the set of optimal points, and all sublevel sets are convex.
- Any local optimal point for problem (45) is a global optimal point.
- The optimality condition given in Theorem 27 holds for convex optimization problem with general inequality constraints without any change.

5.2 Conic form problems

Conic form problems(or *cone programs*) are simple type convex optimization problems with generalized inequality constraints. This kind of problems have a linear objective function and affine inequality and equality constraint functions. Cone programs can be formulated as following:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Fx + g \preceq_K 0 \\ & Ax = b \end{aligned} \tag{46}$$

Since the inequality constraint function is affine, it is a K-convex function. Therefore, the set of feasible points is a convex set. This form of problems are generalization of linear problems that the componentwise inequality is substituted by a generalized linear inequality. If K is replaced by a nonnegative orthant, then the cone programs are converted to linear programs. The conic form problem in standard form can be formulated as following:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & x \succeq_K 0 \\ & Ax = b \end{aligned} \tag{47}$$

5.3 Vector optimization

In section (5.1) we investigated the types of problems where inequality constraint functions can take vector values. In this section we analyze one kind of problems in which the objective function can take vector values. In such problems the comparison of two points are different from aforementioned problem types. Since the objective function values are vectors, we need a cone to define the a generalized inequality for comparison of those values. Therefore, the minimization of objective function should be considered with respect to a proper cone. The *vector optimization problem* can be formulated as following:

$$\begin{aligned} \text{minimize (with respect to K)} \quad & f_0(x) \\ \text{subject to} \quad & f_i(x) \leq 0, \quad i = 1, .., m \\ & h_j(x) = 0 \quad j = 1, .., p. \end{aligned} \tag{48}$$

where $x \in \mathbb{R}^n$ is the decision variable and $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}^q$ is the vector valued objective function. Here the proper cone $K \subseteq \mathbb{R}^q$ is used for comparison of the objective functions. Note that the only difference between standard optimization problem (12) and vector optimization problem (48) is that the objective function in the former takes scalar values and the comparison is the regular comparison between scalar values; whereas in later one it takes values in $\mathbb{R}^q$ and for comparison purpose a proper cone K is specified.

Let us assume that x and y are two feasible point for the problem (48) and $f_0(x)$ and $f_0(y)$ are respective objective function values. We say that "x is better than y" when $f_0(x) \preceq_K f_0(y)$, that is two objective values are compared with respect to a general inequality specified by the proper cone K. Note that the two objective values need not to be comparable with each other. There might be a case that neither $f_0(x) \preceq_K f_0(y)$ nor $f_0(y) \preceq_K f_0(x)$. In this case, we say "they are not comparable". This happens when $f_0(x)$ is better than $f_0(y)$ in one coordinate and worse in another one. Note that this cannot happen in the scalar case.

The problem (48) is called *convex vector optimization problem* when the objective function is a K-convex function, inequality constraint functions are convex, and equality constraint functions are affine.

5.4 Optimal points and values

In this section we first consider the case in which all objective values are comparable. Let us D be the domain of an optimization problem and let O be the set of *achievable objective values* which is defined as following:

$$O = \{f_0(x) | \exists x \in D, f_i(x) \leq 0, i = 1, .., m, h_j(x) = 0, j = 1, .., p\}$$

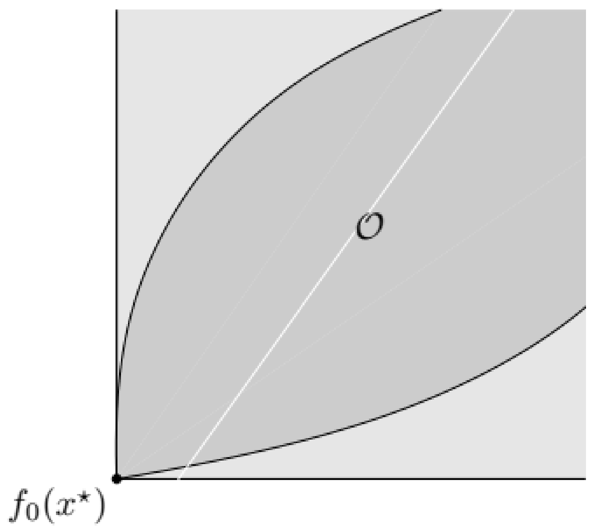
where $O \subseteq \mathbb{R}^q$. In problem (48), if there exists a feasible point x that $f_0(x) \leq_K f_0(y)$ for all feasible point y, then x is an *optimal* point for the problem and $f_0(x)$ is called the *optimal value*. This means that $f_0(x)$ is comparable to all achievable objective values and it is better than or equal to all of them.

Optimality condition for a vector optimization problem is as following. Let $f_0(x) + K$ be the set of all values that are worse than or equal to $f_0(x)$. Then a point x^* is optimal for problem (48) if and only if

$$O \subseteq f_0(x^*) + K \tag{49}$$

This condition means that every achievable value for the problem is worse than or equal to the optimal value. Note that in the special cases of the vector optimization problem there exists an optimal point and corresponding optimal value; however, this is not the case for the most vector optimization problems. Figure 5.1 illustrates an example in $\mathbb{R}^2$. The set O is the set of all achievable values and $f_0(x^*)$ is the optimal value of the example. If we consider $K = \mathbb{R}^2_+$ as a cone, then $f_0(x^*) + K$ can be illustrated as the lightly shaded region in the figure. As it can be observed from the figure, $O \subseteq f_0(x^*)+K$.

Figure 5.1: An example of a problem in $\mathbb{R}^2$ which has an optimal point and optimal values



5.5 Pareto optimal points and values

In the most cases of vector optimization problem, the set of achievable objective values does not have a minimum value *i.e.* the problem does not have an optimal point and optimal value. In this case, we investigate the *minimal* points of the set. If x is a feasible point for a vector optimization problem and $f_0(x)$ is the minimal element of the set of achievable objective values, then the point x is called *Pareto optimal* (or *efficient*) and $f_0(x)$ is called *Pareto optimal value*.

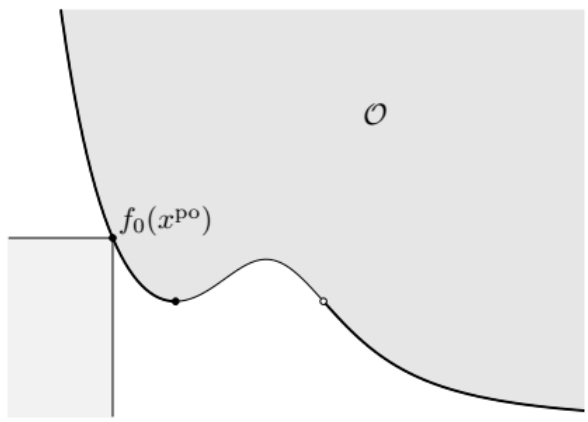
The feasible point x is Pareto optimal if $f_0(y) \leq_K f_0(x)$ implies $f_0(y) = f_0(x)$ for any feasible point y. This means that any achievable value either is worse than $f_0(x)$ or it is not comparable to it. In other words, if x is a Pareto optimal point, then any feasible point y that has the objective value better than or equal to x, has the objective value equal to $f_0(x)$.

The condition for a feasible point x to be a Pareto optimal point is as following. Let $f_0(x) - K$ be the set of all values that are better than or equal to $f_0(x)$. Then the feasible point x is a Pareto optimal point for problem (48) if and only if

$$O \cap (f_0(x) - K) = \{f_0(x)\} \tag{50}$$

The condition means that $f_0(x)$ is the only achievable value that is better than or equal to itself. Figure 5.2 illustrates an example of a problem in $\mathbb{R}^2$ with cone $K = \mathbb{R}^2_+$ which does not have any optimal points and values but has many Pareto optimal points and values. The set O is the set of all achievable values and all points in the boundary of O which are on the darkened curve are Pareto optimal values. As an illustration, the point $f_0(x^{po})$ is a Pareto optimal value and x_{po} is a Pareto optimal point. Note that the cone which is lightly shaded in the figure is $f_0(x^{po}) - K$. Therefore, the condition (50) is satisfied for the point x^{po}.

Figure 5.2: An example of a problem in $\mathbb{R}^2$ which has many Pareto optimal points and values



Note that a vector optimization problem might have many Pareto optimal points and values.

Theorem 5.1. *Let P denotes the set of Pareto optimal values and O be the set of achievable objective values for a vector optimization problem. Then, P lies on the boundary of the set O, i.e. $P \subseteq O \cap$ **bd** O.*

Proof. The proof comes from the definition of Pareto optimal point and value. We know that for a feasible point x to be a Pareto optimal point

$$O \cap (f_0(x) - K) = \{f_0(x)\}$$

Now let's consider $p \in P$ be a Pareto optimal value corresponding to a Pareto optimal point x (*i.e.* $f_0(x) = p$). Let us on the contrary assume that $p \in$ **int** O. In this case, the condition stated will not be satisfied. That is $O \cap (f_0(x) - K) \neq \{f_0(x)\}$. Therefore, for p to be a Pareto optimal value, it must lie on the boundary of the achievable objective values O. $\square$

References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex optimization*. Cambridge university press, 2004.

[2] Mokhtar S Bazaraa, John J Jarvis, and Hanif D Sherali. *Linear programming and network flows*. John Wiley and Sons, 2011.

[3] George J. Stigler. The cost of subsistence. *Journal of Farm Economics*, 27(2):303–314, 1945.

[4] Esra Bas. A robust optimization approach to diet problem with overall glycemic load as objective function. *Applied Mathematical Modelling*, 38(19):4926 – 4940, 2014.

[5] Ruslan Sadykov, François Vanderbeck, Artur Pessoa, and Eduardo Uchoa. Column generation based heuristic for the generalized assignment problem. *XLVII Simpósio Brasileiro de Pesquisa Operacional, Porto de Galinhas, Brazil*, 2015.

[6] Yelin Fu, Jianshan Sun, K. K. Lai, and John W. K. Leung. A robust optimization solution to bottleneck generalized assignment problem under uncertainty. *Annals of Operations Research*, 233(1):123–133, Oct 2015.

[7] Oumaima Khaled, Michel Minoux, Vincent Mousseau, Stéphane Michel, and Xavier Ceugniet. A compact optimization model for the tail assignment problem. *European Journal of Operational Research*, 264(2):548 – 557, 2018.

[8] Roberto Aringhieri, Paolo Landa, Patrick Soriano, Elena Tànfani, and Angela Testi. A two level metaheuristic for the operating room scheduling and assignment problem. *Computers and Operations Research*, 54:21 – 34, 2015.

[9] Bernardetta Addis, Giuliana Carello, Andrea Grosso, and Elena Tànfani. Operating room scheduling and rescheduling: a rolling horizon approach. *Flexible Services and Manufacturing Journal*, 28(1):206–232, Jun 2016.

[10] Sara Ceschia and Andrea Schaerf. Dynamic patient admission scheduling with operating room constraints, flexible horizons, and patient delays. *Journal of Scheduling*, 19(4):377–389, Aug 2016.

[11] Michael Samudra, Carla Van Riet, Erik Demeulemeester, Brecht Cardoen, Nancy Vansteenkiste, and Frank E. Rademakers. Scheduling operating rooms: achievements, challenges and pitfalls. *Journal of Scheduling*, 19(5):493–525, Oct 2016.

[12] Leon Cooper. Location-allocation problems. *Operations Research*, 11(3):331–343, 1963.

[13] F Fiedrich, F Gehbauer, and U Rickers. Optimized resource allocation for emergency response after earthquake disasters. *Safety Science*, 35(1):41 – 57, 2000.

[14] Christophe Duhamel, Andréa Cynthia Santos, Daniel Brasil, Eric Châtelet, and Babiga Birregah. Connecting a population dynamic model with a multi-period location-allocation problem for post-disaster relief operations. *Annals of Operations Research*, 247(2):693–713, Dec 2016.

[15] S. Basu, S. Roy, S. Bandyopadhyay, and S. Das Bit. A utility driven post disaster emergency resource allocation system using dtn. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pages 1–13, 2018.

[16] Riccardo Manzini and Elisa Gebennini. Optimization models for the dynamic facility location and allocation problem. *International Journal of Production Research*, 46(8):2061–2086, 2008.

[17] Heinz Isermann. The enumeration of all efficient solutions for a linear multiple-objective transportation problem. *Naval Research Logistics Quarterly*, 26(1):123–139, 1979.

[18] S.K. Das, A. Goswami, and S.S. Alam. Multiobjective transportation problem with interval cost, source and destination parameters. *European Journal of Operational Research*, 117(1):100 – 112, 1999.

[19] Gurupada Maity and Sankar Kumar Roy. Multiobjective transportation problem using fuzzy decision variable through multi-choice programming. In *Intelligent Transportation and Planning: Breakthroughs in Research and Practice*, pages 866–882. IGI Global, 2018.

[20] Hasan Dalman. Uncertain programming model for multi-item solid transportation problem. *International Journal of Machine Learning and Cybernetics*, 9(4):559–567, Apr 2018.