GEOMETRIC APPROACH TO LINEAR OPTIMIZATION AND PRIMAL DUAL THEORY

M.Tech Thesis

Submitted in partial fulfillment for the award of the Degree of Master of Technology in Computer Science and Engineering

by

ANIL KUMAR S. (Roll No: M130426CS)

under the guidance of

Dr. K. MURALIKRISHNAN

तमसो मा ज्योतिर्गमय

Department of Computer Science and Engineering

NATIONAL INISTITUTE OF TECHNOLOGY CALICUT

NIT Campus P.O., Calicut - 673601, India

2015

CERTIFICATE

This is to certify that the thesis entitled GEOMETRIC APPROACH TO LINEAR OPTIMIZATION AND PRIMAL DUAL THEORY is a bona fide record of the major project done by ANIL KUMAR S. (Roll No: M130426CS) under my supervision and guidance, in partial fulfillment of the requirements for the award of Degree of Master of Technology in Computer Science and Engineering from National Institute of Technology Calicut in the year 2015.

Project Guide

Head of the Department

Place: NIT-Calcut

Date : 15-07-2015

Declaration

I declare that this written submission represents my ideas in my own words and where others' ideas or words have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be cause for disciplinary action by the Institute.

Anil Kumar S.
M130426CS

Place: NIT-Calcut
Date : 15-07-2015

Acknowledgements

Every successful project is the outcome of hardwork and strong support and guidance given by a number of people. I sincerely appreciate the inspiration, support and guidance of all those people who have been instrumental in making this project a success.

I express my gratitude to *Dr.G.R.C. Reddy, Director, NIT-Calicut* for making the resources available at right time without which this project would not have been a success. I take this opportunity to express a deep sense of gratitude to *Dr.K.A.Abdul Nazeer, Head of the Department of Computer Science and Engineering* and *Dr.Priya Chandran* for the cordial support, valuable information and guidance, which helped me in completing this task through various stages. I also take this opportunity to express my profound gratitude and deep regards to my guide *Dr.K. Muralikrishnan, Associate Professor* for his exemplary guidance, monitoring and constant encouragement throughout the course of this thesis. The blessing, help and guidance given by him from time to time shall carry me a long way in the journey of life on which I am about to embark. I am obliged to staff members of the Cloud Computing Lab and Computational Intelligence Lab, for their support and cooperation during the period of my assignment. Last but not the least, I thank almighty, my mother, brother, sister and friends for their constant encouragement without which this assignment would not be possible.

Anil Kumar S.

Place: NIT-Calcut
Date : 15-07-2015

Abstract

Linear Optimization is essentially the process of optimizing a linear objective function subjected to a finite number of linear constraints. Such optimization problems are usually expressed as linear programming problems. The computational strategies for linear programming problems are well explained in the literature. From a geomeric perspective, we see that linear programming problems have well defined structural geometry. This thesis explores the structural foundations of linear programming by investigating the structural geometry of polyhedral sets and derives the core results in linear programming like Caratheodory characterization theorem and the Fundamental theorem of linear programing by means of these structural concepts with elementary linear algebra and real analysis. It also presents the primal dual theory and proves the duality theorems with the aid of certain algebraic results formulated on the platform provided by the fundametal theorem of linear programming. The approach taken in this thesis in deriving various results give more emphasis to the underlying geometry. Hence it seems that the thesis gives a better visual intuition of the notions and results in linear programming.

Contents

List of Figures

List of Tables and Algorithms

Tables

Algorithms

List of Symbols and Abbreviations

$\mathbb{R}^{m\times n}$	Vector space of $m \times n$ matrices of real numbers.
$\mathbb{R}^n$	Vector space of $n \times 1$ column vectors of real numbers.
$\mathbf{0}_V$	Zero vector of vector space V.
$\mathbf{0}_n$	Zero vector of $\mathbb{R}^n$.
$\mathbf{1}_n$	The $n \times 1$ vector whose all entries are 1.
$\mathbf{e}_i$	A column vector whose i^{th} entry is 1 and remaining entries are 0.
$Cone(S)$	Conic hull of set S.
$Conv(S)$	Convex hull of set S.
$R_C(S)$	Recession cone of set S.
$F(P)$	Feasible region of linear program P.
$Img(P)$	Image of linear program P.
$OPT(P)$	Finite optimum of linear program P.
$\overline{H}_{\leq}$, $\overline{H}_{\geq}$	The closed half spaces of hyperplane H.
$H_{<}$, $H_{>}$	The open half spaces of hyperplane H.
$R_N(F(P))$	Normalised recession direction set of feasible region $F(P)$.
$\gamma_{\mathbf{xy}}$	Duality gap between the linear programs with respect to the respective feasible solutions $\mathbf{x}$ and $\mathbf{y}$.
$deg(\mathbf{x})$	Degeneracy of the basic feasible solution $\mathbf{x}$.
$N_\delta(\mathbf{x})$	The δ-Neighbourhood of the vector $\mathbf{x}$.
$\overline{S}$	Closure of the set S.
$B_r(\mathbf{x})$	Open ball of radius r centered at $\mathbf{x}$.
$\overline{B}_r(\mathbf{x})$	Closed ball of radius r centered at $\mathbf{x}$.
$Span(S)$	Span of the vectors in the set S.
$dim(V)$	Dimension of the vector space V.
$\langle \mathbf{x}, \mathbf{y} \rangle$	Inner product of the vectors $\mathbf{x}$ and $\mathbf{y}$.
$\lVert \mathbf{x} \rVert$	Norm of vector $\mathbf{x}$ in general and denotes the l_2 norm of $\mathbf{x}$ if $\mathbf{x}$ is a vector in the Eucledian space.
$\lVert \mathbf{x} \rVert_p$	The l_p norm of the vector $\mathbf{x}$.
$\lVert \mathbf{x} \rVert_\infty$	The ∞ norm of the vector $\mathbf{x}$.
$d(\mathbf{x}, \mathbf{y})$	Distance between vectors $\mathbf{x}$ and $\mathbf{y}$.

Chapter 1

Introduction

1.1 History of linear programing

The linear programming method was first developed by *Leonid Kantorovich* in 1937. The motivation for the development of the area is to plan the total expenditure and returns so as to reduce the costs to the army and increase the losses incurred by the enemy in the Second World War. This method was kept secret until *George B. Dantzig* published the *Simplex method* in 1947 [1].

Dantzig's original example was aimed at finding the best assignment of 70 people to 70 jobs. The computing power required to test all the permutations to select the best one was quiet large as the number of possible configurations was beyond a practically computable limit. However, by expressing the problem as a linear program and then applying simplex algorithm , it took only a reasonable time to find the optimal assignment. Immediately after Dantzig published the linear programming methodology, another mathematician namely *John von Neumann* introduced the concept of duality and established the famous *Min-Max Theorem* in Game Theory. This invention also led to the further development of primal-dual theory and its use in designing approximation algorithms for NP-Complete problems.

Even though the linear programming Method became popular in 1950's several researchers developed the idea in the past. Among those, the contributions of *Jean-Baptiste Joseph Fourier* and *de la Valle Poussin* are of great relevance. Each published papers describing the linear programming theory in the year 1823 and 1911 respectively.

The linear programming problem was first shown to be solvable in polynomial time by *Leonid Khachiyan* in 1979. Later on an Indian mathematician *Narendra Karmarkar* laid a milestone in the field of linear programing with the advent of a famous polynomial time algorithm known as *interior point method* for solving linear programming Problems and published the same in the year 1984. Now this area has grown to a full fledged field of Mathematics with strong theoretical as well as practical support.

1.2 Scope of linear programming

The field of linear programming has wide scope in the field of Science and Technology. The application of this area ranges from solving simple numerical problems to computable problems with extremely high computational complexity. Linear programming is one of the classical fields of optimization. Several problems in Operations Research can be represented as linear programming problems. The ideas in linear programming forms the backbone for the development of Primal-Dual based Approximation Algorithm Design for classical NP-Complete Problems. Linear programming is widely used in the field of Economics and Company Management where the primary focus is on maximizing the profit with least possible consumption of the resources. The field of linear programming based combinatorial optimization finds its applications in various areas of VLSI design like floorplanning so as to produce highly efficient integrated chips with least space and power consumption.

1.3 Objectives of the Project

The primary focus of the project is to explore the geometric foundations of linear programming and derive the core results in linear programming like Caratheodory characterisation theorem [1] and the fundamental theorem of linear programming by means of these geometric characteristics. This not only makes the understanding of the subject simpler, but also gives the geometric visualisation of various results in linear programming and thereby strengthen the conceptual knowledge of the subject. This makes the thesis to cover the entire theory of the graphical method of solving linear programs. The project also focuses on presenting the essentials of primal dual theory with a specific attention to prove the strong duality theorem from the first principles with the aid of elementary linear algebra and real analysis.

1.4 Motivation for the Work

The basic facts which motivate this work are the following.

1. An in-depth study of linear optimization is difficult and requires strong foundation in mathematical areas like linear algebra and real analysis. Most of the linear programming related works found in the literature gives emphasis to the computational aspects of linear optimization. Therefore presenting the fundamentals of linear optimization from a geometric perspective with elementary linear algebra and real analysis seems to be very useful to a novice person in the field of Engineering Optimization.

2. An extensive study of approximation algorithms requires strong foundation in primal dual theory which include the weak duality theorem, the strong duality theorem etc [2] [3]. Most of the classical approaches of proving results like strong duality theorem requires very strong foundation in multivariable calculus or advanced results like Farkas Lemma, Theorem of Alternatives etc. Hence it seems that a material which derives the duality theorems based on the geometric characterisation of linear programming reduces the efforts to be taken by the reader to have a clear understanding of the primal dual theory.

1.5 Prerequiste for the Reader

While presenting various notions and establishing results, it seems that the fundamental notions in Topology like closed sets, bounded sets, compact sets etc and basic knowledge in linear algebra are inevitable. Therefore the reader is expected to have a basic knowledge in *Real Analysis* and *Linear Algebra*.

1.6 Overview of the Thesis

The thesis is organised into six chapters and three appendices. The second chapter deals with the preliminaries of linear programming including the general, canonical and standard orms of linear programs, their equivalence and the concepts regarding feasibility and unboundedness of linear programs. The chapter also contains algorithms for converting one form to other. The third chaper introduces the underlying geometric concepts in various linear programming results. This chapter primarly includes the discussion of convex sets and cones in the Eucledian space with a view point of illustrating the theorems and results essential for deriving linear programing results. The fourth chapter is dedicated to discuss the geometry of linear programming.

This chapter starts with the discussion of vertices, extreme points and basic feasible solutions and their equivalence. This chapter outlines the formulation of geometric notions such as recession directions, extreme directions in linear programing context. It also contains a detailed discussion on Caratheodory Theorem and the Fundamental Theorem of linear programing along with their proof and algorithms introduced wherever necessary and hence explores the foundation for the graphical method of solving linear programs. The fifth chapter discuss the primal dual theory in detail. This chapter starts with the definitions of primal and dual linear programs and proceed to the discussions on the weak Duality theorm, complementary slackness conditions and terminates with the stong Duality theorem giving the proofs of each. The sixth chapter concludes the thesis with a summary of the core ideas illustated in the thesis and giving the recomendations for future work.

As pointed out earlier, the thesis also consists of three appendices. Appendix A discuss the technical results which are essential for proving the strong duality theorem. Both appendices B and C essentially outlines various notions and results which are essential in the preceeding chapters without proofs. These results are reffered to as *Facts* throughout the thesis and are freely used wherever required. Appendix B introduces the essential linear algebra required to derive the results. The proof of various Facts introduced in Appendix B can be seen in any standard text book on linear algebra such as the one by *Kenneth Hoffman* and *Ray Kunze*. Appendix C is dedicated to the discussion of Real Analysis concepts which are essential to meet the objectives of the thesis. The proofs of various Facts in this appendix are available in textbooks on Modern Topology such as the one by *G.F. Simmons*.

Chapter 2

Geometric Notions in Linear Programing

2.1 Introduction

In this chapter, we present the geometric notions and results in the Eucledian space which are essential for deriving various linear programming results. The results from Linear Algebra which are used in this chapter are discussed briefly in Appendix A without proof. A detailed discussion of these results and their proof can be seen in any standard textbook on Linear Algebra. The results from Real Analysis which support the concepts in this chapter are outlined in Appendix B. An in-depth discussion on these topics are available standard textbooks on Modern Toplogy such as the book *Introduction to Topology and Modern analysis* by *George F. Simmons* [4].

2.2 Cones and Convex Sets

In this section, we define the notions of convex sets and cones and explores the geometric characteristics of these sets [5] [6].

Definition 2.2.1.
Let V be a vector space and $S \subseteq V$. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ in S. Then

(a) any expression of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ where $\lambda_i \in \mathbb{R}$ for each $i \in \{1, 2, 3, \cdots, k\}$ and $\sum_{i=1}^{k} \lambda_i = 1$ is called an *affine combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$.

(b) any expression of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ where $\lambda_i \geq 0, 1 \leq i \leq k$ is called a *conic combination* or *positive combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$. In particular, any expression of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ where $\lambda_i > 0, 1 \leq i \leq k$ is called a *strict conic combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$.

(c) any expression of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ where $\lambda_i \geq 0, 1 \leq i \leq k$ and $\sum_{i=1}^{k} \lambda_i = 1$ is called a *convex combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$. In particular, any expression of the form $\sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ where $\lambda_i > 0, 1 \leq i \leq k$ and $\sum_{i=1}^{k} \lambda_i = 1$ is called a *strict convex combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$.

Definition 2.2.2.
Let V be a vector space and $S \subseteq V$. Then

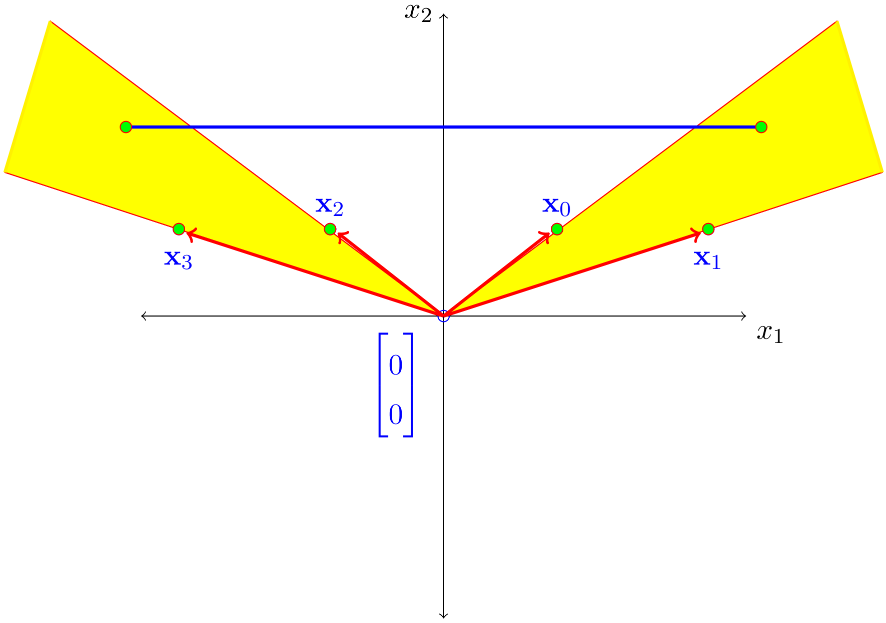


Figure 2.1: The Shaded Region is a cone, but not a convex cone.

(a) S is said to be an *affine set* if S is closed under affine combinations.

(b) S is said to be a *convex set* if S is closed under convex combinations.

(c) S is said to be a *cone* if for any $\mathbf{x} \in S, \lambda > 0$ we must have $\lambda \mathbf{x} \in S$.

(d) S is said to be a *convex cone* if S is a cone as well as a convex set.

Remark 2.2.1.

Let V be a vector space and cone $S \subseteq V$. Since S is a cone, we see that corresponding to any $\mathbf{x} \in S$, we must have $0\mathbf{x} \in S$. This implies that the zero vector $\mathbf{0}_V$ of V is an element of S.

Remark 2.2.2.

By the definiton of convex sets, it is clear that any convex combination of two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in a convex set $S \subset \mathbb{R}^n$ is also an element of S. That is $\forall\, \mathbf{x}_1, \mathbf{x}_2 \in S,$ we must have $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in S, \forall \lambda_1, \lambda_2 \geq 0$ such that $\lambda_1 + \lambda_2 = 1$. This means that the line segment joining the points $\mathbf{x}_1$ and $\mathbf{x}_2$ must fully lie in S.

Now we are going to prove that intersection of convex sets is also convex.

Lemma 2.2.1.

Convex sets are closed under intersection.

Proof.

Let $S_1, S_2, S_3, \cdots, S_k$ be k convex subsets of a vector space V. We are going to prove that $\bigcap_{i=1}^{k} S_i$ is convex.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_t$ be the elements of $\bigcap_{i=1}^{k} S_i$. Let $\mathbf{y}$ be any convex combination of the elements $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_t$. Therefore there exist positive real numbers $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_t$ such that $\mathbf{y} = \sum_{j=1}^{t} \lambda_j \mathbf{x}_j$ and $\sum_{j=1}^{t} \lambda_j = 1$

It is easy to see that each of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_t$ is an element of all of $S_1, S_2, S_3, \cdots, S_k$. Since each of $S_1, S_2, S_3, \cdots, S_k$ is convex, we see that $\mathbf{y}$ which is a convex combination of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_t$ is an element of each of $S_1, S_2, S_3, \cdots, S_k$. This further implies that $\mathbf{y}$ is an element of $\bigcap_{i=1}^{k} S_i$. Since the choice of elements $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_t$ within $\bigcap_{i=1}^{k} S_i$ is arbitrary,

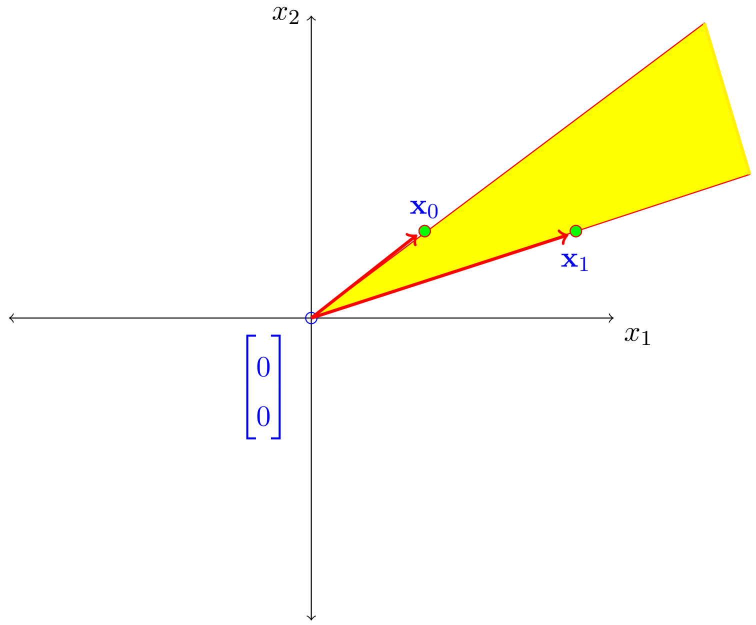


Figure 2.2: For any two points $\mathbf{x}_1, \mathbf{x}_2$ in $\mathbb{R}^n$ $Cone(\{\mathbf{x}_1, \mathbf{x}_2\})$ is the region bounded by the line segments originating from the origin to $\mathbf{x}_1$ and $\mathbf{x}_2$ within the orthant(s) containing $\mathbf{x}_1$ and $\mathbf{x}_2$. It is easy to se that $Cone(\{\mathbf{x}_1, \mathbf{x}_2\})$ is a convex cone.

we conclude that $\bigcap_{i=1}^{k} S_i$ is closed under convex combinations. Hence $\bigcap_{i=1}^{k} S_i$ is convex. Hence the Lemma. □

Lemma 2.2.2.

Let V be a vector space $S \subseteq V$. S is a ***convex cone*** *if and only if S is closed under conic combinations.*

Proof.

If part:-

Let S be closed under conic combinations. This implies that for any vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$, all conic combinations of the form $\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 + \lambda_3\mathbf{x}_3 + \cdots + \lambda_k\mathbf{x}_k$ where $\lambda_i \geq 0, 1 \leq i \leq k$ are also elements of S. This implies that all convex combinations of the form $\gamma_1\mathbf{x}_1 + \gamma_2\mathbf{x}_2 + \gamma_3\mathbf{x}_3 + \cdots + \gamma_k\mathbf{x}_k$ where $\gamma_i \geq 0, 1 \leq i \leq k$ and $\sum\limits_{i=1}^{k} \gamma_i = 1$ are also elements of S. Therefore S is a convex set. Moreover for any $\mathbf{x} \in S$, it must be the case that all conic combinations of the form $\lambda\mathbf{x} + 0.\mathbf{x} = \lambda\mathbf{x}, \lambda \geq 0$ are also elements of S. Hence we see that S is a cone. Thus we conclude that S is a convex cone.

Only if Part:-

Given that S is a convex cone. Therefore by definition, S must contain all convex combinations of the elements of S. Let T be the set of all conic combinations of the elements of S. We prove this part by proving that each of S and T are subsets of each other.

The fact that $S \subseteq T$ is trivial. To prove that $T \subseteq S$, let us take an arbitrary element $\mathbf{x}$ in T. Therefore $\mathbf{x}$ can be expressed as a conic combination of the form $\mathbf{x} = \sum\limits_{1=1}^{m} \lambda_i\mathbf{y}_i$ for some $m \in \mathbb{N}, \{\mathbf{y_i} \mid 1 \leq i \leq m\} \subset S$ and $\lambda_i \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. Clearly $\frac{\lambda_i}{\sum\limits_{j=1}^{m} \lambda_j} \in [0, 1]$ for each $i \in \{1, 2, 3, \cdots, m\}$ and $\sum\limits_{i=1}^{m} \left[\frac{\lambda_i}{\sum\limits_{j=1}^{m} \lambda_j}\right] = 1$, we see that $\sum\limits_{i=1}^{m} \left[\frac{\lambda_i}{\sum\limits_{j=1}^{m} \lambda_j}\right] \mathbf{y}_i$ is a convex combination of the elements of S and hence it must be a member of S. This further implies that $\mathbf{x} = \sum\limits_{i=1}^{m} \lambda_i \mathbf{y}_i = \sum\limits_{j=1}^{m} \lambda_j \left(\sum\limits_{i=1}^{m} \left[\frac{\lambda_i}{\sum\limits_{j=1}^{m} \lambda_j} \right] \mathbf{y}_i \right) \in S$ since S is a cone.

Since the choice of $\mathbf{x}$ is arbitrary in T, we claim that $\forall \mathbf{x} \in T$, it must be the case that $\mathbf{x} \in S$. Thus we claim that $S \subseteq T$ and $T \subseteq S$ and hence $S = T$. Hence the Lemma. □

Remark 2.2.3.

By the definition of convex cone, it is clear that any conic combination of two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in a convex cone $S \subset \mathbb{R}^n$ is also an element of S. That is $\lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 \in S \ \forall \lambda_1, \lambda_2 \geq 0$. This means that the entire area bounded by the two lines - one through $\mathbf{0}_n$ and $\mathbf{x}_1$ and the other through $\mathbf{0}_n$ and $\mathbf{x}_2$ within the orthant(s) containing $\mathbf{x}_1$ and $\mathbf{x}_2$ is also included in S (See Figure 2.2).

Definition 2.2.3.

Let V be a vector space and $S \subseteq V$. Then

(a) the set of all elements of V obtained by the convex combinations of the elements of S is called the *convex hull* of S and is denoted by $conv(S)$.

$$Conv(S) = \{\mathbf{x} \mid \mathbf{x} \in V \ \text{ and } \ \textstyle\sum_{\mathbf{x}_i \in S} \lambda_i \mathbf{x}_i \text{ where each } \lambda_i \geq 0 \text{ and } \sum \lambda_i = 1\}$$

(b) the set of all elements of V obtained by the conic combinations of the elements of S is called the *conic hull* of S and is denoted by $Cone(S)$.

$$Cone(S) = \{\mathbf{x} \mid \mathbf{x} \in V \ \text{ and } \ \textstyle\sum_{\mathbf{x}_i \in S} \lambda_i \mathbf{x}_i \text{ where each } \lambda_i \geq 0\}$$

Remark 2.2.4.

We follow the convention that $Cone(\phi) = Conv(\phi) = \phi$.

Remark 2.2.5.

Let V be a vector space and S be a non-empty subset of V. Then $Conv(S) \subseteq Cone(S)$. This follows from the fact that every convex combination of the elements of S is a conic combination also.

Lemma 2.2.3.

Let V be a vector space and S be a non-empty subset of V. Then $Cone(S)$ is the smallest convex cone containing S.

Proof.

The fact that $S \subseteq Cone(S)$ is trivial. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be any elements of $Cone(S)$. Then by the definition of $Cone(S)$, $\mathbf{x}_i = \sum\limits_{j=1}^{k_i} \alpha_{ij} \mathbf{y}_{ij}$ such that corresponding to each $i \in \{1, 2, 3, \cdots, k\}$, we have the relations $\alpha_{ij} \geq 0, \mathbf{y}_{ij} \in S, 1 \leq j \leq k_i$ for some $k_i \in \mathbb{N}$

Let $\mathbf{x} = \sum\limits_{i=1}^{m} \lambda_i \mathbf{x}_i$ where $\lambda_i \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$.

Now we have

$$\mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i = \sum_{i=1}^{m} \lambda_i \left\{ \sum_{j=1}^{k_i} \alpha_{i_j} \mathbf{y}_{i_j} \right\} = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \lambda_i \alpha_{i_j} \mathbf{y}_{i_j}$$

Clearly $\lambda_i \alpha_{i_j} \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. Thus $\mathbf{x}$ is a conic combination of several elements of S and hence belongs to $Cone(S)$. Since the choice of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ is arbitrary within $Cone(S)$, we further infer that any conic combination of the elements of $Conv(S)$ is also an element of $Cone(S)$. Hence we conclude that $Cone(S)$ is a convex cone by Lemma 2.2.2.

Let T be convex cone in V such that $S \subseteq T$ and $T \subset Cone(S)$. This implies that there must have some element $\mathbf{u}$ in $Cone(S) \setminus T$ such that $\mathbf{u} = \sum_{i=1}^{k} \gamma_i \mathbf{z}_i$ where each $\mathbf{z}_i \in S$ and $\gamma_i \geq 0$. Since $S \subseteq T$, this further implies that each of the elements $\mathbf{z}_i$ belongs to T. Therefore we infer that $\mathbf{u}$ is a conic combination of the elements of T but not a member of T which contradicts the fact that T is a cone. Hence we conclude that there does not have a convex cone T in V such that $S \subseteq T$ and $T \subset Cone(S)$.

Hence the Lemma. □

Lemma 2.2.4.

Let V be a vector space and $S \subseteq V$. Then $Conv(S)$ is the smallest convex set containing S.

Proof.

The fact that $S \subseteq Conv(S)$ is trivial. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be any elements of $Conv(S)$. Then by the definition of $Conv(S)$, $\mathbf{x}_i = \sum_{j=1}^{k_i} \alpha_{ij} \mathbf{y}_{ij}$ such that corresponding to each $i \in \{1, 2, 3, \cdots, k\}$, we have the relations $\alpha_{ij} \geq 0, \mathbf{y}_{ij} \in S, 1 \leq j \leq k_i$ for some $k_i \in \mathbb{N}$ such that $\sum_{j=1}^{k_i} \alpha_{i_j} = 1$.

Let $\mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i$ where $\lambda_i \geq 0, \sum_{i=1}^{m} \lambda_i = 1, 1 \leq i \leq m$.

Now we have

$$\mathbf{x} = \sum_{i=1}^{m} \lambda_i \mathbf{x}_i = \sum_{i=1}^{m} \lambda_i \left\{ \sum_{j=1}^{k_i} \alpha_{i_j} \mathbf{y}_{i_j} \right\} = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \lambda_i \alpha_{i_j} \mathbf{y}_{i_j}$$

Clearly $\lambda_i \alpha_{i_j} \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$.

Moreover

$$\sum_{i=1}^{m} \sum_{j=1}^{k_i} \lambda_i \alpha_{i_j} = \sum_{i=1}^{m} \lambda_i \left\{ \sum_{j=1}^{k_i} \alpha_{i_j} \right\} = 1$$

Thus $\mathbf{x}$ is a convex combination of several elements of S and hence belongs to $Conv(S)$. Since the choice of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ is arbitrary within $Conv(S)$, we further infer that any convex combination of the elements of $Conv(S)$ is also an element of $Conv(S)$. Hence we conclude that $Conv(S)$ is a convex set.

Let T be convex subset of V such that $S \subseteq T$ and $T \subset Conv(S)$. This implies that there must have some element $\mathbf{u}$ in $Conv(S) \setminus T$ such that $\mathbf{u} = \sum_{i=1}^{k} \gamma_i \mathbf{z}_i$ where each $\mathbf{z}_i \in S$ and $\gamma_i \geq 0$ such that $\sum_{i=1}^{m} \gamma_i = 1$. Since $S \subseteq T$, this further implies that each of the elements $\mathbf{z}_i$ belongs to T. Therefore we infer that $\mathbf{u}$ is a convex combination of the elements of T but not a member of T which contradicts the fact that T is a convex set. Hence we conclude that there does not have a convex subset T of V such that $S \subseteq T$ and $T \subset Conv(S)$.

Hence the Lemma. □

Lemma 2.2.5.

Let V be a vector space and S be a non-empty subset of V. Then $Cone(Conv(S)) = Cone(S)$.

Figure 2.3: The figure on the RHS is the convex hull determined by the points in the set on the LHS

Proof.
We prove the Lemma by showing that each of $Cone(Conv(S))$ and $Cone(S)$ is a subset of the other.

Since $S \subseteq Conv(S)$, it is easy to see that $Cone(S) \subseteq Cone(Conv(S))$. Let $\mathbf{y}$ be any element of $Cone(Conv(S))$. This means that $\mathbf{y}$ is a conic combination of the elements of $Conv(S)$. Since $Conv(S) \subseteq Cone(S)$, we further infer that $\mathbf{y}$ is a conic combination of the elements of $Cone(S)$ and is obviously an element of $Cone(S)$. Since $\mathbf{y}$ can be any element of $Cone(Conv(S))$, we conclude that $Cone(Conv(S)) \subseteq Cone(S)$.

Hence the Lemma. □

Lemma 2.2.6.
For any finite subset S of $\mathbb{R}^n$, $Conv(S)$ is a bounded set .

Proof.
Let $S = \{\mathbf{x}_i \mid 1 \leq i \leq k\}$ where each $\mathbf{x}_i \in \mathbb{R}^n$ and $k \in \mathbb{N}$. Let $\delta = \max(\{\|y\| \mid y \in S\})$. Let $\mathbf{x}$ be any element of $Conv(S)$. Therefore we must have

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \text{ such that } \lambda_i \geq 0, 1 \leq i \leq k \text{ and } \sum_{i=1}^{k} \lambda_i = 1$$

Now we have

$$\|\mathbf{x}\| \quad = \quad \left\| \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \right\| = \sum_{i=1}^{k} \lambda_i \|\mathbf{x}_i\| \leq \sum_{i=1}^{k} \lambda_i \delta$$

which imples that $\|\mathbf{x}\| \leq \delta$ as $\sum_{i=1}^{k} \lambda_i = 1$. Since $\mathbf{x}$ can be chosen any vector in $Conv(S)$, we infer that $\forall \mathbf{x} \in Conv(S)$ it must be the case that $\|\mathbf{x}\| \leq \delta$. Hence $Conv(S) \subset B_{\delta+1}(\mathbf{0}_n)$. Hence the Lemma. □

Now we introduce the notions of recession directions in convex sets.

Definition 2.2.4.
Given $\mathbf{x}, \mathbf{d} \in \mathbb{R}^n$. Then the set $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\}$ is called a *half-line* anchored at $\mathbf{x}$ and direction $\mathbf{d}$.

Definition 2.2.5.
Given a convex set $C \subseteq \mathbb{R}^n$. Then any non-zero vector $\mathbf{d} \in \mathbb{R}^n$ is said to be a *recession direction* of C if $\forall \mathbf{x} \in C$, the half-line $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$.

Lemma 2.2.7.
Let $C \subseteq \mathbb{R}^n$ be a closed convex set . Then the following statements are equivalent.

(1) The vector $\mathbf{d} \in \mathbb{R}^n$ is a recession direction of C.

(2) $\forall \mathbf{x} \in C, \{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$

(3) $\exists \mathbf{x} \in C, \{\mathbf{x} + \alpha \mathbf{d} \mid \alpha \geq 0\} \subseteq C$

Proof.
(1) $\Rightarrow$ *(2)* by the definition of recession directions,
(2) $\Rightarrow$ *(3)* is trivial.
What remains to prove the equivalence is *(3)* $\Rightarrow$ *(1)*.
Let there be some $\mathbf{x} \in C$ such that the half-line $\{\mathbf{x}+\alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$. Let $\mathbf{x}'$ be any vector in C. Now we need to consider two cases.

Case 1: $\mathbf{x}'$ *lies on the line* $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \in \mathbb{R}\}$.
In this case, we need to consider two cases.

Case 1.1: $\mathbf{x}'$ *lies on the half-line* $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\}$.
In this case, there exists $\alpha_0 \geq 0$ such that $\mathbf{x}' = \mathbf{x} + \alpha_0\mathbf{d}$. Therefore

$$\{\mathbf{x}' + \alpha\mathbf{d} \mid \alpha \geq 0\} = \{\mathbf{x} + (\alpha + \alpha_0)\mathbf{d} \mid \alpha \geq 0\} \subseteq C$$

Case 1.2: $\mathbf{x}'$ *lies on the half-line* $\{\mathbf{x} - \alpha\mathbf{d} \mid \alpha \geq 0\}$.
In this case, there exists $\alpha_1 \geq 0$ such that $\mathbf{x}' = \mathbf{x} - \alpha_1\mathbf{d}$. Therefore $\mathbf{x} = \mathbf{x}' + \alpha_1\mathbf{d}$. Since C is convex and the points $\mathbf{x}', \mathbf{x}' + \alpha_1\mathbf{d}$, the line segment $\{\mathbf{x}' + \alpha\mathbf{d} \mid 0 \leq \alpha \leq \alpha_1\} \subseteq C$. Moreover $\{\mathbf{x}' + \alpha\mathbf{d} \mid \alpha \geq \alpha_1\} = \{\mathbf{x} + (\alpha - \alpha_1)\mathbf{d} \mid \alpha \geq \alpha_1\} = \{\mathbf{x}' + \alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$. These two results together imply that the half-line $\{\mathbf{x}' + \alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$.

Case 2: $\mathbf{x}'$ *does not lie on the line* $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \in \mathbb{R}\}$.
Consider the sequence $\{\mathbf{y}_k\}_{k\in\mathbb{N}}$ defined by $\mathbf{y}_k = \mathbf{x} + k\mathbf{d}, \forall k \in \mathbb{N}$. Obviously $\{\mathbf{y}_k\}_{k\in\mathbb{N}}$ is a sequence in C. It is not hard to see that $\{\|\mathbf{y}_k\|\} \to \infty$ (Refer to Fact C.1.2 in Appendix C). This further implies that $\{\|\mathbf{y}_k - \mathbf{x}\|\} \to \infty$. Hence there exists a subsequence $\{\mathbf{u}_j\}_{j\in\mathbb{N}}$ of the sequence $\{\mathbf{y}_k - \mathbf{x}'\}_{k\in\mathbb{N}}$ defined by $\mathbf{u}_j = \mathbf{y}_{k_j} - \mathbf{x}', \forall j \in \mathbb{N}$ such that the sequence $\{\|\mathbf{u}_j\|\}_{j\in\mathbb{N}} = \{\left\|\mathbf{y}_{k_j} - \mathbf{x}'\right\|\}_{j\in\mathbb{N}}$ is strictly increasing and $\{\|\mathbf{u}_j\|\} = \{\left\|\mathbf{y}_{k_j} - \mathbf{x}'\right\|\} \to \infty$.

Let α be any real number. Since the sequence $\{\mathbf{u}_j\}_{j\in\mathbb{N}}$ is strictly increasing and $\{\|\mathbf{u}_j\|\} \to \infty$., there exists a least integer t such that $\alpha\|\mathbf{d}\| < \|\mathbf{u}_{t+j}\|, \forall j \in \mathbb{N}$. Let the sequence $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ be defined by $\mathbf{z}_j = \mathbf{x}' + \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\mathbf{u}_{t+j}, \forall j \in \mathbb{N}$.
Now

$$\mathbf{z}_j = \mathbf{x}' + \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\mathbf{u}_{t+j} = \mathbf{x}' + \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\left(\mathbf{y}_{k_j} - \mathbf{x}'\right) = \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\mathbf{y}_{k_j} + \mathbf{x}'\left(1 - \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\right)$$

This implies that each term $\mathbf{z}_j$ is a convex combination of the two vectors $\mathbf{y}_{k_j}$ and $\mathbf{x}'$ in C as $\alpha\|\mathbf{d}\| < \|\mathbf{u}_{t+j}\|, \forall j \in \mathbb{N}$. Hence $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ is a sequence over C.

[*Note*:- $\|\mathbf{z}_j - \mathbf{x}'\| = \alpha\|\mathbf{d}\|, \forall j \in \mathbb{N}$. Thus each $\mathbf{z}_j$ lies on the boundary of a closed ball of radius $\alpha\|\mathbf{d}\|$ centered at $\mathbf{x}'$ (Refer to Definition C.1.11 in Appendix C) as shown in figure 2.4
We have

$$\begin{aligned}
\mathbf{z}_j &= \mathbf{x}' + \frac{\alpha\|\mathbf{d}\|}{\|\mathbf{u}_{t+j}\|}\mathbf{u}_{t+j} \\
&= \mathbf{x}' + \frac{\alpha\|\mathbf{d}\|}{\left\|\mathbf{y}_{k_j} - \mathbf{x}'\right\|}\left(\mathbf{y}_{k_j} - \mathbf{x}'\right) \\
&= \mathbf{x}' + \frac{1}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|}(\mathbf{x} - \mathbf{x}') + \frac{k_{t+j}\|\mathbf{d}\|}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|}\alpha\mathbf{d}
\end{aligned}$$

Now

$$\begin{aligned}
\lim_{j\to\infty} \frac{1}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|} &= \lim_{k_{t+j}\to\infty} \frac{1}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|} = 0 \\
\lim_{j\to\infty} \frac{k_{t+j}\|\mathbf{d}\|}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|} &= \lim_{k_{t+j}\to\infty} \frac{k_{t+j}\|\mathbf{d}\|}{\|\mathbf{x} - \mathbf{x}' + k_{t+j}\mathbf{d}\|} = 1 \quad \text{(By L'Hospital's rule)}
\end{aligned}$$

Therefore

$$\begin{aligned}\lim_{j\to\infty}\mathbf{z}_j &= \lim_{j\to\infty}\mathbf{x}' + \lim_{j\to\infty}\frac{1}{\|\mathbf{x}-\mathbf{x}'+k_{t+j}\mathbf{d}\|}(\mathbf{x}-\mathbf{x}') + \lim_{j\to\infty}\frac{k_{t+j}\|\mathbf{d}\|}{\|\mathbf{x}-\mathbf{x}'+k_{t+j}\mathbf{d}\|}\alpha\mathbf{d}\\ &= \lim_{j\to\infty}\mathbf{x}' + \lim_{k_{t+j}\to\infty}\frac{1}{\|\mathbf{x}-\mathbf{x}'+k_{t+j}\mathbf{d}\|}(\mathbf{x}-\mathbf{x}') + \lim_{k_{t+j}\to\infty}\frac{k_{t+j}\|\mathbf{d}\|}{\|\mathbf{x}-\mathbf{x}'+k_{t+j}\mathbf{d}\|}\alpha\mathbf{d}\\ &= \mathbf{x}'+\alpha\mathbf{d}\end{aligned}$$

Hence we see that the sequence $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ converges to $\mathbf{x}'+\alpha\mathbf{d}$ (See Figure 2.4).

Since $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ is a sequence over C and C is closed, it follows that $\mathbf{x}'+\alpha\mathbf{d}\in C$ (Refer to Fact C.1.4 in Appendix C). Since α can be chosen as any arbitrary non-negative real number, we conclude that the set $\{\mathbf{x}'+\alpha\mathbf{d}\,|\,\alpha\geq 0\}\subseteq C$. Thus we see that in both cases, the set $\{\mathbf{x}'+\alpha\mathbf{d}\,|\,\alpha\geq 0\}\subseteq C$. Since $\mathbf{x}'$ can be chosen as arbitrarily within C, we conclude that $\forall\mathbf{x}'\in C,\alpha\geq 0$ the vector $\mathbf{x}'+\alpha\mathbf{d}\in C$. Therefore $\mathbf{d}$ is a recession direction of C. Thus *(3)* $\Rightarrow$ *(1)*. Hence the Lemma. □

Lemma 2.2.8.
Let $C\subseteq\mathbb{R}^n$ be a closed convex set. C is unbounded if and only if it contains a half-line .

Proof.
if part:-
Let C consist of some half-line $\{\mathbf{x}+\alpha\mathbf{d}\mid\alpha\geq 0\}$ for some $\mathbf{x}\in C$. Clearly the sequence $\{\mathbf{x}+k\mathbf{d}\}_{k\in\mathbb{N}}$ is a sequence over C and $\{\|\mathbf{x}+k\mathbf{d}\|\}\to\infty$. Therefore C must be unbounded.
only-if part:-
Let C be an unbounded set. Let $\mathbf{x}_0$ be any vector in C. Since C is unbounded, $C\setminus\{\mathbf{x}_0\}$ is also unbounded. Therefore there exists a sequence $\{\mathbf{x}_k\}_{k\in\mathbb{N}}$ in $C\setminus\{\mathbf{x}_0\}$ such that the sequence $\{\|\mathbf{x}_k\|\}_{k\in\mathbb{N}}$ is strictly increasing sequence in $\mathbb{R}$ and $\|\mathbf{x}_k\|\to\infty$ (Refer to Fact C.1.8 in Appendix C). Furthermore there exists a sequence $\{\mathbf{x}_k-\mathbf{x}_0\}_{k\in\mathbb{N}}$ in $\mathbb{R}^n$ such that the sequence $\{\|\mathbf{x}_k-\mathbf{x}_0\|\}_{k\in\mathbb{N}}$ is strictly increasing and $\|\mathbf{x}_k-\mathbf{x}_0\|\to\infty$ (Refer to Fact C.1.1 in Appendix C).

Consider the sequence $\left\{\frac{\mathbf{x}_k-\mathbf{x}_0}{\|\mathbf{x}_k-\mathbf{x}_0\|}\right\}_{k\in\mathbb{N}}$. Each vector in this sequence has unit l_2 norm (Refer to Example B.1.3 in Appendix B) in $\mathbb{R}^n$ and therefore this sequence is a bounded sequence. Therefore by Bolzano-Weistrass Theorem (Refer to Fact C.1.9 in Appendix C), there exists a converging subsequence $\left\{\frac{\mathbf{x}_{k_j}-\mathbf{x}_0}{\|\mathbf{x}_{k_j}-\mathbf{x}_0\|}\right\}_{j\in\mathbb{N}}$ of the sequence $\left\{\frac{\mathbf{x}_k-\mathbf{x}_0}{\|\mathbf{x}_k-\mathbf{x}_0\|}\right\}_{k\in\mathbb{N}}$ such that $\frac{\mathbf{x}_{k_j}-\mathbf{x}_0}{\|\mathbf{x}_{k_j}-\mathbf{x}_0\|}\to\mathbf{d}$ for some $\mathbf{d}\in\mathbb{R}^n$. It is obvious that $\mathbf{d}\neq\mathbf{0}_n$.

Let α be any non-negative real number. Since $\{\|\mathbf{x}_k-\mathbf{x}_0\|\}_{k\in\mathbb{N}}$ is strictly increasing and $\|\mathbf{x}_k-\mathbf{x}_0\|\to\infty$, the subsequence $\{\|\mathbf{x}_{k_j}-\mathbf{x}_0\|\}_{j\in\mathbb{N}}$ is also strictly increasing and $\|\mathbf{x}_{k_j}-\mathbf{x}_0\|\to\infty$ (Refer to Fact C.1.3 in Appendix C). This ensures the existence of a least positive integer t such that $\alpha<\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|,\forall j\in\mathbb{N}$.

Now consider the sequence $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ defined by $\mathbf{z}_j=\mathbf{x}_0+\alpha\frac{\mathbf{x}_{k_{t+j}}-\mathbf{x}_0}{\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|},\forall j\in\mathbb{N}$.

We have

$$\forall j\in\mathbb{N},\mathbf{z}_j=\mathbf{x}_0+\frac{\mathbf{x}_{k_{t+j}}-\mathbf{x}_0}{\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|}=\frac{\alpha}{\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|}\mathbf{x}_{k_{t+j}}+\left(1-\frac{\alpha}{\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|}\right)\mathbf{x}_0$$

Hence we see that each $\mathbf{z}_j$ is a convex combination of the vectors $\mathbf{x}_{k_{t+j}}$ and $\mathbf{x}_0$ both belonging to C as $\alpha<\|\mathbf{x}_{k_{t+j}}-\mathbf{x}_0\|,\forall j\in\mathbb{N}$. Thus it follows that $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ is a sequence over C.

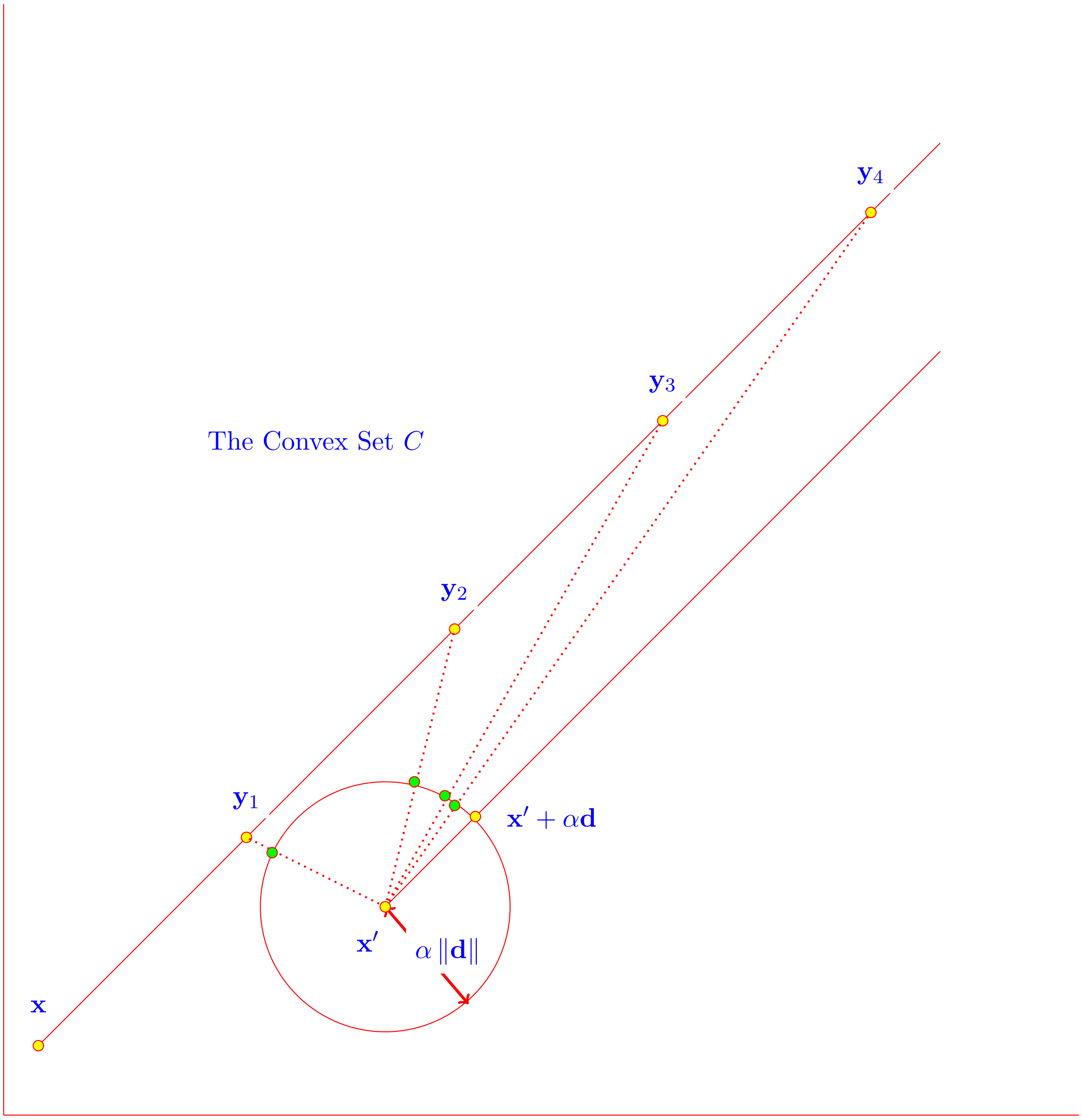


Figure 2.4: Illustration of Case 2 of Lemma 2.2.7. The points marked on the circle from left to right are $\mathbf{z_1}, \mathbf{z_2}, \mathbf{z_3}, \mathbf{z_4}$ and $\mathbf{x}' + \alpha\mathbf{d}$ respectively.

We also have

$$\begin{aligned}
\lim_{j\to\infty} \mathbf{z}_j &= \lim_{j\to\infty} \mathbf{x}_0 + \lim_{j\to\infty} \alpha \frac{\mathbf{x}_{k_{t+j}} - \mathbf{x}_0}{\left\|\mathbf{x}_{k_{t+j}} - \mathbf{x}_0\right\|} \\
&= \lim_{j\to\infty} \mathbf{x}_0 + \lim_{j\to\infty} \frac{\mathbf{x}_{k_{t+j}} - \mathbf{x}_0}{\left\|\mathbf{x}_{k_{t+j}} - \mathbf{x}_0\right\|} \\
&= \mathbf{x}_0 + \alpha\mathbf{d}
\end{aligned}$$

Hence we see that the sequence $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ converges to $\mathbf{x}_0 + \alpha\mathbf{d}$(See Figure 2.5). Since $\{\mathbf{z}_j\}_{j\in\mathbb{N}}$ is a sequence over C and C is closed, we claim that $\mathbf{x}_0 + \alpha\mathbf{d} \in C$. Since α can be chosen as any arbitrary non-negative real number, we further claim that the set $\{\mathbf{x}_0 + \alpha\mathbf{d} \,|\, \alpha \geq 0\} \subseteq C$. This implies that C contains a half-line. Hence the Lemma. □

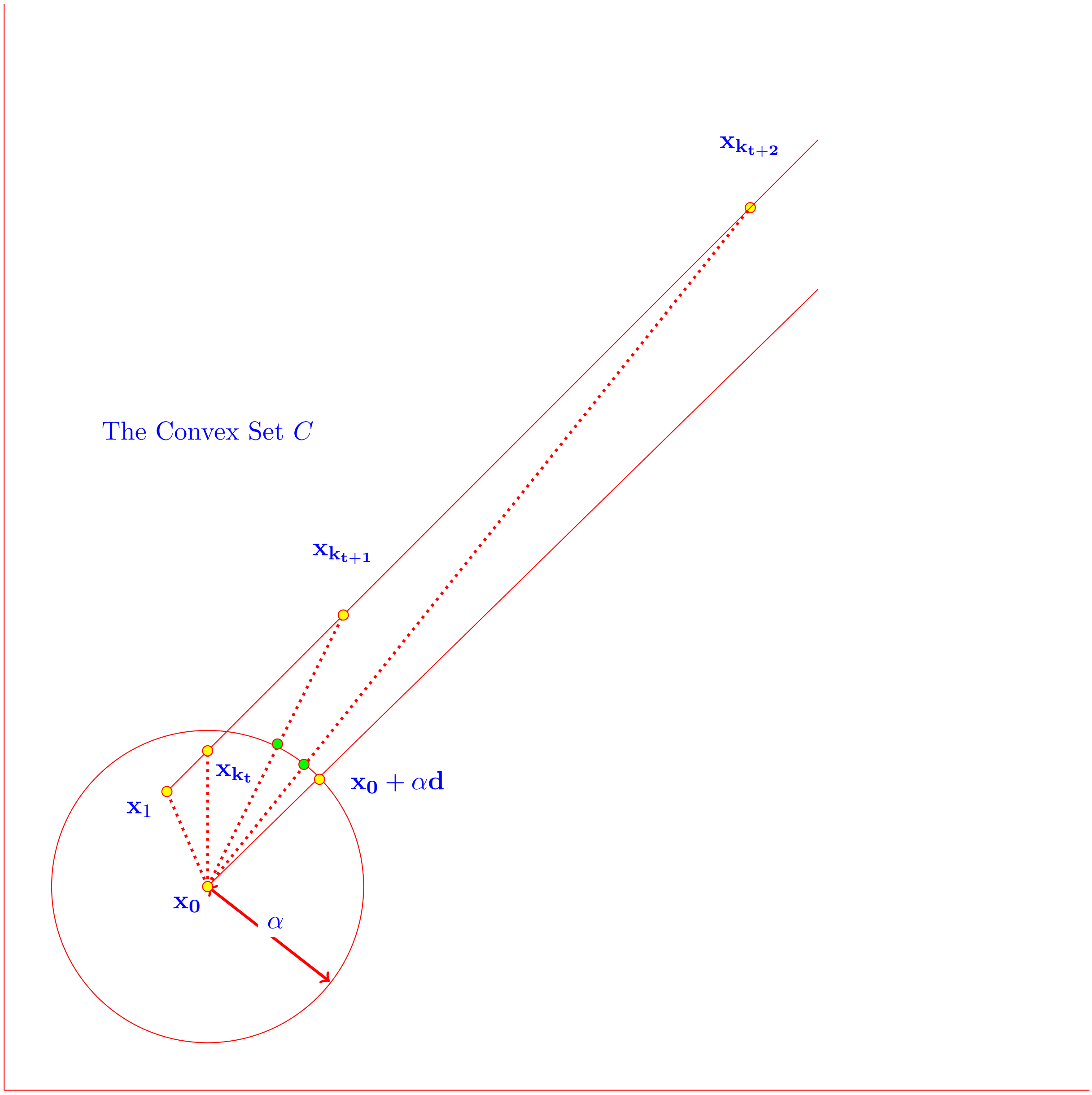


Figure 2.5: Illustration of only-if part of Lemma 2.2.8. The points marked on the circle from left to right are $\mathbf{z_1}, \mathbf{z_2}$ and $\mathbf{x_0} + \alpha\mathbf{d}$ respectively.

Corollary 2.2.1.

A non-empty closed convex set in $\mathbb{R}^n$ is unbounded if and only if it has at least one recession direction.

Proof.

Let C be any non-empty closed convex set in $\mathbb{R}^n$. By Lemma 2.2.8, C is unbounded if and only if it contains some half-line $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\}$. This further implies that C is unbounded if and only if C has a recession direction by Lemma 2.2.7. Hence the Lemma. □

Corollary 2.2.2.

A non-empty closed convex set in $\mathbb{R}^n$ is bounded if and only if it has no recession directions.

Lemma 2.2.9.

The set of all recession directions of a convex set in $\mathbb{R}^n$ together with $\mathbf{0}_n$ forms a convex cone .

Proof.
Let $C \in \mathbb{R}^n$ be a convex set. Let S be the set of all recession directions of C. To show that $S \cup \{\mathbf{0}_n\}$ is a convex cone, it is enough to show that $S \cup \{\mathbf{0}_n\}$ contains all the conic combinations of the elements of S by Lemma 2.2.2.

Let $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_k$ be any elements of $S \cup \{\mathbf{0}_n\}$. Let $\mathbf{d} = \sum_{i=1}^{k} \lambda_i \mathbf{d}_i, \lambda_i \geq 0, 1 \leq i \leq k$.

Claim:- d is a recession direction of C.

We prove the claim by applying mathematical induction on k.

Basis:- $k = 1$. Since $\mathbf{d}_1$ is a recession direction of C, we see that for each $\mathbf{x}$ in C, the set $\{\mathbf{x} + \alpha \mathbf{d}_1 \mid \alpha \geq 0\} \subseteq C$. Hence we see that the set $\{\mathbf{x} + \alpha\lambda_1 \mathbf{d}_1 \mid \alpha \geq 0\} \subseteq C$. This implies that $d = \lambda_1 \mathbf{d}_1$ is a recession direction of C. Hence Basis proved.

Inductive Hypothesis:-

Assume that $\mathbf{d} = \sum_{i=1}^{k} \lambda_i \mathbf{d}_i, \lambda_i \geq 0$ is a recession direction of C for $1 \leq k \leq t$ for some $t \in \mathbb{N}$.

Induction:-$k = t + 1$.

Now $\mathbf{d} = \sum_{i=1}^{t+1} \lambda_i \mathbf{d}_i = \sum_{i=1}^{t} \lambda_i \mathbf{d}_i + \lambda_{t+1}\mathbf{d}_{t+1} = \mathbf{d}' + \lambda_{t+1}\mathbf{d}_{t+1}$ where $\mathbf{d}' = \sum_{i=1}^{t} \lambda_i \mathbf{d}_i$.

By Inductive Hypothesis, $\mathbf{d}'$ is a recession direction of C. Therefore we see that for each $\mathbf{x}$ in C and each $\beta \geq 0$, it must be the case that $\mathbf{x} + \beta\mathbf{d}' \in C$.

Since $\mathbf{d}_{t+1}$ is a recession direction of C, we see that for each $\mathbf{y}$ in C and each $\gamma \geq 0$, it must be the case that $\mathbf{y} + \gamma\mathbf{d}_{t+1} \in C$. This futher implies that for each $\mathbf{y}$ in C and each $\gamma \geq 0$, the vector $\mathbf{y} + \gamma\lambda_{t+1}\mathbf{d}_{t+1} \in C$.

Combining the two inferences, we see that for each $\mathbf{x}$ in C and each $\alpha \geq 0$, it must be the case that $\mathbf{x} + \alpha(\mathbf{d}' + \lambda_{t+1}\mathbf{d}_{t+1}) = \mathbf{x} + \alpha\mathbf{d}' + \alpha\lambda_{t+1}\mathbf{d}_{t+1} \in C$. Hence it follows that $\mathbf{d}' + \lambda_{t+1}\mathbf{d}_{t+1}$ is a recession direction of C.

Thus we conclude that $\forall k \in \mathbb{N}, \mathbf{d} = \sum_{i=1}^{k} \lambda_i \mathbf{d}_i, \lambda_i \geq 0, 1 \leq i \leq k$ is a recession direction of C.

Since $\mathbf{d}$ is a recession direction of C, it must be a member of $S \cup \{\mathbf{0}_n\}$. Thus we see that any conic combination of $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_k$ is a member of $S \cup \{\mathbf{0}_n\}$. Since the vectors $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_k$ can be arbitrarily within C, we further infer that $S \cup \{\mathbf{0}_n\}$ contains all the conic combinations of the elements of its own. Hence $S \cup \{\mathbf{0}_n\}$ is a convex cone. Hence the Lemma. □

Remark 2.2.6.

The convex cone formed by the recession directions of an unbounded convex set C in $\mathbb{R}^n$ together with $\mathbf{0}_n$ is called *recession cone* of C and is denoted by $Rc(C)$.

Lemma 2.2.10.

Given convex set $C \subseteq \mathbb{R}^n$ and its recession direction $\mathbf{d} \in \mathbb{R}^n$. Then there exists another vector $\mathbf{d}' \in \mathbb{R}^n$ such that

(a) $\langle \mathbf{1_n}, \mathbf{d}' \rangle = 1$

(b) $\mathbf{d}'$ *is also a recession direction of C.*

Proof.
Letting $\mathbf{d}' = \frac{1}{\langle \mathbf{1_n}, \mathbf{d} \rangle}\mathbf{d}$, the result follows directly. □

2.3 Hyperplanes and Polyhedral Sets

In this section, we first introduce the notions of hyperplanes, polyhedral sets and polytopes. Following this discussion, we bring about the geometric notions - vertices and extreme points

Definition 2.3.1.
A *hyperplane* H in $\mathbb{R}^n$ is the set $H = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle = \delta\}$ for some $\mathbf{a} \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$.

Definition 2.3.2.
A set $H_s \subseteq \mathbb{R}^n$ is said to be a *closed half space* if there exists some $\mathbf{a} \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ such that $H_s = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle (\leq, \geq) \delta\}$.

Definition 2.3.3.
A set $H_s \subseteq \mathbb{R}^n$ is said to be an *open half space* if there exists some $\mathbf{a} \in \mathbb{R}^n$ and $\delta \in \mathbb{R}$ such that $H_s = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle (<, >) \delta\}$.

Remark 2.3.1.
Intutively, a hyperplane $H = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle = \delta, \mathbf{a} \in \mathbb{R}^n, \delta \in \mathbb{R}\}$ divides $\mathbb{R}^n$ into closed half spaces $\overline{H_{\geq}} = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle \geq \delta\}$ and $\overline{H_{\leq}} = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle \leq \delta\}$. It is easy to see that $\overline{H_{\geq}} = \overline{H'_{\leq}}$ and $\overline{H_{\leq}} = \overline{H'_{\geq}}$ where $H' = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle -\mathbf{a}, \mathbf{x} \rangle = -\delta\}$

Similarly hyperplane H divides $\mathbb{R}^n$ into disjoint open half spaces $H_{>} = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle > \delta\}$ and $H_{<} = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle < \delta\}$. It easily follows that $H_{>} = H'_{<}$ and $H_{<} = H'_{>}$ where $H' = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle -\mathbf{a}, \mathbf{x} \rangle = -\delta\}$

Definition 2.3.4.
Given the non empty subset $S \subseteq \mathbb{R}^n$. The Hyperplane $H = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}, \mathbf{x} \rangle = \delta$ for some $\mathbf{a} \in \mathbb{R}^n$ and $\delta \in \mathbb{R}\}$ is said to be a *supporting hyperplane* for S if the following two conditions are satisfied.

(i) $\exists \mathbf{y} \in S$ such that $\langle \mathbf{a}, \mathbf{y} \rangle = \delta$.

(ii) $S \subseteq \overline{H_{\geq}}$ or $S \subseteq \overline{H_{\leq}}$.

It is equivalent to say that H supports S if either $\delta = \min\left\{\langle \mathbf{a}, \mathbf{x} \rangle \;\middle|\; \mathbf{x} \in S\right\}$ or $\delta = \max\left\{\langle \mathbf{a}, \mathbf{x} \rangle \;\middle|\; \mathbf{x} \in S\right\}$

Definition 2.3.5.
The intersection of a finite number of closed half spaces in $\mathbb{R}^n$ is called a *polyhedral set* or simply *polyhedron* . In particular a bounded polyhedral set in $\mathbb{R}^n$ is called a *polytope* .

Remark 2.3.2.
Since the intersection of a finite number of closed sets is closed, polyhedral sets are closed.

Lemma 2.3.1.
Given $P \subseteq \mathbb{R}^n$. P is a polyhedral set if and only if there exists $m \in \mathbb{N}, \mathbf{A} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$ such that $P = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$.

Proof.
if part:-

Let m be any positive integer. Let $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_m \end{bmatrix}^T$ where $\mathbf{a}_i \in \mathbb{R}^n, 1 \leq i \leq m$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & b_3 \cdots, b_m \end{bmatrix}^T$.

Now

$$\begin{aligned} \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{A}\mathbf{x} \leq \mathbf{b}\} &= \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}_i, \mathbf{x} \rangle \leq b_i \forall i \text{ such that } 1 \leq i \leq m\} \\ &= \bigcap_{i=1}^{m} \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}_i, \mathbf{x} \rangle \leq b_i\} \end{aligned}$$

Thus we see that the set $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ is a polyhedral set, each of the sets $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}_i, \mathbf{x} \rangle \leq b_i\}$ are closed half-spaces in $\mathbb{R}^n$.

only if part:-

Let P be a polyhedral set in $\mathbb{R}^n$. Therefore there exists positive integer m and hyperplanes $H_i = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{a}_i, \mathbf{x} \rangle = b_i\}$ such that $\mathbf{a}_i \in \mathbb{R}^n$ and $b_i \in \mathbb{R}$ for each $i \in \{1, 2, 3, \cdots, m\}$ and $P = \bigcap_{i=1}^{m} \overline{H_{i\leq}}$. Therefore P can be expressed as $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{Ax} \leq \mathbf{b}\}$ where $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \cdots & \mathbf{a}_m \end{bmatrix}^T$ and $\mathbf{b} = \begin{bmatrix} b_1 & b_2 & b_3 & \cdots & b_m \end{bmatrix}^T$.

Hence the Lemma. □

Definition 2.3.6.

Given convex set $C \subseteq \mathbb{R}^n$ and $\mathbf{x} \in P$. Then $\mathbf{x}$ is said to be a *vertex* or *corner point* of C if C has a supporting hyperplane H in $\mathbb{R}^n$ such that $C \cap H = \{\mathbf{x}\}$. In other words $\mathbf{x}$ is said to be a *vertex* of C if there exists some hyperplane H in $\mathbb{R}^n$ such that $C \cap H = \{\mathbf{x}\}$ and $C \subseteq \overline{H_{\leq}}$.

Lemma 2.3.2.

Given convex set $C \subseteq \mathbb{R}^n$ and $\mathbf{x} \in C$. Then $\mathbf{x}$ is a vertex of C if and only if there exists some vector $\mathbf{c} \in \mathbb{R}^n$ such that $\mathbf{x}$ is the unique point in C such that $\langle \mathbf{c}, \mathbf{x} \rangle > \langle \mathbf{c}, \mathbf{y} \rangle$ for each $\mathbf{y} \in C \setminus \{\mathbf{x}\}$.

Proof.

if part:-

Let $\mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{c}, \mathbf{x} \rangle > \langle \mathbf{c}, \mathbf{y} \rangle$ for each $\mathbf{y} \in C \setminus \{\mathbf{x}\}$.

Consider the hyperplane $H = \{\mathbf{z} \mid \mathbf{z} \in \mathbb{R}^n \text{ such that } \langle \mathbf{c}, \mathbf{z} \rangle = \langle \mathbf{c}, \mathbf{x} \rangle\}$. Clearly $C \cap H = \{\mathbf{x}\}$ and $C \setminus \{\mathbf{x}\} \subseteq H_<$. This implies that $C \cap H = \{\mathbf{x}\}$ and $C \subseteq \overline{H_{\leq}}$. Thus $\mathbf{x}$ is a vertex of C.

only if part:-

Let $\mathbf{x}$ be a vertex of C. Therefore there exists hyperplane $H = \{\mathbf{z} \mid \mathbf{z} \in \mathbb{R}^n \text{ such that } \langle \mathbf{c}, \mathbf{z} \rangle = \delta\}$ where $\mathbf{c} \in \mathbb{R}^n, \delta \in \mathbb{R}$ and $C \cap H = \{\mathbf{x}\}$. This further implies that $C \cap H = \{\mathbf{x}\}$ and $C \subseteq \overline{H_{\leq}}$. That is $\mathbf{x} \in H$ and $C \setminus \{\mathbf{x}\} \subseteq H_<$. Hence we see that $\langle \mathbf{c}, \mathbf{x} \rangle = \delta$ and $\forall \mathbf{y} \in P \setminus \{\mathbf{x}\}, \langle \mathbf{c}, \mathbf{y} \rangle < \delta$. In other words, $\langle \mathbf{c}, \mathbf{x} \rangle > \langle \mathbf{c}, \mathbf{y} \rangle$ for each $\mathbf{y} \in C \setminus \{\mathbf{x}\}$. Hence the Lemma. □

Definition 2.3.7.

Given convex set $C \subseteq \mathbb{R}^n$ and $\mathbf{x} \in C$. Then $\mathbf{x}$ is said to be an *extreme point* of C if there exist no distinct points $\mathbf{y}, \mathbf{z}$ in C such that $\mathbf{x} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{z}, \lambda \in (0, 1)$. In other words, $\mathbf{x}$ is an extreme point of C if $\mathbf{x}$ cannot be expressed as a strict convex combination of two other points in C. Geometrically this means that $\mathbf{x}$ does not lie on a line joining any two vectors in C.

Lemma 2.3.3.

Given convex set $C \subseteq \mathbb{R}^n$ and $\mathbf{y} \in C$. If $\mathbf{y}$ is a vertex of C, then $\mathbf{y}$ will be an extreme point of C.

Proof.

Let $\mathbf{y}$ be a vertex of C. Therefore by Lemma 2.3.2, there exists vector $\mathbf{c} \in \mathbb{R}^n$ such that for each $\mathbf{z}$ in $C \setminus \{\mathbf{y}\}$, we have the relation $\langle \mathbf{c}, \mathbf{y} \rangle > \langle \mathbf{c}, \mathbf{z} \rangle$. Assume that $\mathbf{y}$ is not an extreme point of C. Therefore there exist vectors $\mathbf{u}$ and $\mathbf{v}$ in C such that $\mathbf{y} = \lambda \mathbf{u} + (1 - \lambda)\mathbf{v}$ for some $\lambda \in (0, 1)$.

Now

$$\begin{aligned} \langle \mathbf{c}, \mathbf{y} \rangle &= \langle \mathbf{c}, \lambda \rangle \mathbf{u} + (1 - \lambda)\mathbf{v} \rangle = \lambda \langle \mathbf{c}, \mathbf{u} \rangle + (1 - \lambda) \langle \mathbf{c}, \mathbf{v} \rangle < \lambda \langle \mathbf{c}, \mathbf{y} \rangle + (1 - \lambda) \langle \mathbf{c}, \mathbf{y} \rangle \\ & \quad (\because \mathbf{u}, \mathbf{v} \in C \Rightarrow \langle \mathbf{c}, \mathbf{u} \rangle < \langle \mathbf{c}, \mathbf{y} \rangle, \langle \mathbf{c}, \mathbf{v} \rangle < \langle \mathbf{c}, \mathbf{y} \rangle) \end{aligned}$$

which lead to the contradiction that $\langle \mathbf{c}, \mathbf{y} \rangle < \langle \mathbf{c}, \mathbf{y} \rangle$. Hence we claim that $\mathbf{y}$ is an extreme point of C. □

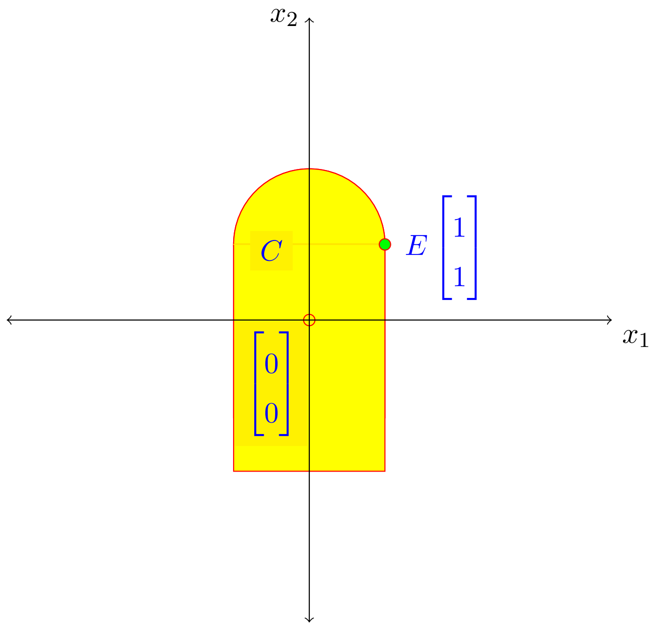


Figure 2.6: Illustration of Remark 2.3.3. The point E is an extreme point of C but not a vertex.

Remark 2.3.3.

The converse of Lemma 2.3.3 is not true. To see thts, consider the convex set $C \subset \mathbb{R}^2$ given by $C = \left\{ \mathbf{x} \middle| \mathbf{x} \in \mathbb{R}^2 \text{ and } \left\| \mathbf{x} - \begin{bmatrix}0\\1\end{bmatrix} \right\| \leq 1 \right\} \cup \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^2 \text{ and } \|\mathbf{x}\|_\infty = 1\}$ (See Figure 2.6). It is easy to see that the vector $\begin{bmatrix}1\\1\end{bmatrix} \in C$ and is an extreme point of C. In $\mathbb{R}^2$, C has only one supporting hyperplane $H = \left\{ \mathbf{x} \middle| \mathbf{x} \in \mathbb{R}^2 \text{ and } \left\langle \begin{bmatrix}1\\0\end{bmatrix}, \mathbf{x} \right\rangle = 1 \right\} = \left\{ \begin{bmatrix}1\\y\end{bmatrix} \middle| y \in \mathbb{R} \right\}$ which passes through $\begin{bmatrix}1\\1\end{bmatrix}$, but $H \cap C = \left\{ \begin{bmatrix}1\\y\end{bmatrix} \middle| \mid y \mid \leq 1 \right\} \neq \begin{bmatrix}1\\1\end{bmatrix}$. Therefore $\begin{bmatrix}1\\1\end{bmatrix}$ is not a vertex of C.

We close the chapter by giving two observations which will be used exclusively in the subsequent chapters.

(a) The presence of a half-line within a polyhedral set in $\mathbb{R}^n$ is a necessary and sufficient condition for the polyhedral set to be unbounded.

(b) A polytope in $\mathbb{R}^n$ does not have any recession directions and hence no extreme directions.

2.4 Summary

In this chapter, we discussed the basic geometric notions and results which we use in the subsequent chapters for deriving various results in linear programming. We had started the chapter with the notions of cones, convex sets, convex hull, recession directions and extreme directions in $\mathbb{R}^n$. We proved that the presence of a single half-line within a closed convex set confirms that the set is unbounded which further implies that a bounded convex set in the eucledian space does not have any recession directions. Following this, we came across the concept of polyhedral sets and polytopes which are of great importance in the subsequent chapters. In this section, the notions of vertices and extreme points of polyhedral sets were formally defined and established the necessary and sufficient condition for a point in a polyhedral set to be a vertex by showing the existence of a linear function on $\mathbb{R}^n$ which is uniquely optimized at the point under consideration. In the next chapter, we introduce the mathematical model of linear programming with supporting notions and results.

Chapter 3

Linear Programming Fundamentals

3.1 Introduction

In this chapter, we present the fundamentals of linear programming. In section 3.2, we give the informal definition of linear programming and its features. Section 3.3 starts with the formal definition of a linear program as a 5-tupe along with the general form of a linear program and proceed to the notions of feasibility, unboundedness and equivalence in linear programming. This section also discusses the canonical and the standared forms of linear programs and the algorithms for converting one form to another form. We close this chapter by establishing the equivalence of the three forms of linear programs.

3.2 What is Linear Programming?

According to George B. Dantzig, *Linear programming can be viewed as part of a great revolutionary development which has given mankind the ability to state general goals and to lay out a path of detailed decisions to take in order to best achieve its goals when faced with practical situations of great complexity* [1]

Linear programming, sometimes known as linear optimization, is *the problem of maximizing or minimizing a linear function over a convex polyhedron specified by a set of linear constraints. In other words, linear programming is the optimization of an outcome based on some set of constraints using a linear mathematical model.*

Linear programming can also be defined in a more simple way as *the problem of optimising a linear function of a given set of decision variables(called objective function) subjected to one or more linear constraints involving these variables.* From this definition, the following inferences can easily be made.

- Linear programming is an optimization technique.
- The number of decision variables is finite.
- Both the objective function and the constraints are linear functions.
- There should be a single objective function.

Following are the essential assumptions which must be made while formulating a given problem as a linear programming problem. [7]

- *Proportionality assumption:-* The value of the objective function and the left hand side of the constraints are directly proportional to the value of the decision variables.
- *Independence and additivity assumption:-* The value of each decision variable is completely independent of the values of other decision variables and therefore the contributions of

each decision variable to the objective function and various constraints may be added to give the total value of the objective function and the total value of the left hand side of each constraint.

- *Divisibility assumption:-* The decision variables are assumed to be infinitely divisible which means that the decision variables may be assigned fractional values. In other words decision variables are assumed to be continuous decision variables.
- *Certainity assumption:-* The coefficients of various decision variables and constraints must be known completely without any uncertainity and they should never undergo any changes.
- *Finiteness Assumption:-* The number of decision variables and constraints are assumed to be finite without which the optimal solution can't be determined.

3.3 Mathematical model of LP

In this section, we formally define the term *linear program.*

3.3.1 Formal Definition and General Form of linear programs

Definition 3.3.1.
A linear programming problem or simply linear program is a 5-tuple $P = (V, obj, C, N, T)$ where

V is the set of $n > 0$ *decision variables* each of which is assumed to take real values.
obj is called the *objective function* which is a linear function of the decision variables.

$$obj : \mathbb{R}^n \to \mathbb{R} \text{ of the form } \sum_{j=1}^{n} c_j x_j \text{ where } c_j \in \mathbb{R}, 1 \le j \le n, \text{ to be optimized}$$

C is the finite set of $m \ge 0$ linear constraints of the form

$$\sum_{j=1}^{n} a_{ij} x_j \; op_i \; b_i \text{ where } b_i \in \mathbb{R}, a_{ij} \in \mathbb{R}, op_i \in \{\le, \ge, =\} \text{ and } 1 \le i \le m, 1 \le j \le n$$

N is the set of constraints of the form $x_j \; op_j \; 0$ where $op_j \in \{\ge, \le, \gtreqless\}, 1 \le j \le n$
T is the type of optimization of the objective function.$i.e, T \in \{$*Maximize, Minimize*$\}$

That is P has the general form

$$\begin{aligned} \text{Minimize / Maximize } & \sum_{j=1}^{n} c_j x_j \text{ subject to} \\ & \sum_{j=1}^{n} a_{ij} x_j (\le, \ge, =) b_i \\ & x_j (\ge, \le, \gtreqless) 0 \quad 1 \le i \le m, 1 \le j \le n \end{aligned} \tag{3.1}$$

Remark 3.3.1.
If T is *Maximize*, then P is called a *maximization linear program.* Similarly if T is *Minimize*, then P is called a *minimization linear program* . A constraint of the form $x_j \ge 0$ is called a *non-negativity constraint* of P. Similarly a constraint of the form $x_j \le 0$ is called a *non-positivity constraint* of P. A decision variable x_j of P that can take any real value is called a *free variable*. All constraints in C having the form $\alpha = \beta$ are called *equality constraints* of P . All other constraints are called *inequality constraints* of P.

3.3.2 Feasibility, Unboundedness and Equivalence in LP

In this section, we bring about the concepts of feasible region, feasible linear programs, bounded linear programs and equivalence of linear programs [3].

Definition 3.3.2.

Given linear program $P = (V, obj, C, N, T)$. Then any $\mathbf{x} \in \mathbb{R}^n$ is said to be a *feasible solution* of P if $\mathbf{x}$ satisfies all the constraints in $C \cup N$. The set of all feasible solutions of P is called the *feasible solution set* or *feasible region* of P and is denoted by $F(P)$.

$F(P) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \mathbf{x} \text{ satisfies the constraints in } C \cup N\}$

Definition 3.3.3.

The *image of a linear program* P is denoted by $Image(P)$ and is defined as

$$Image(P) = \left\{ obj(\mathbf{x}) \,\middle|\, \mathbf{x} \in F(P) \right\}.$$

Definition 3.3.4.

A linear program $P = (V, obj, C, N, T)$ is said to have *finite optimum* if there exists some $\mathbf{x} \in F(P)$ such that

$obj(\mathbf{x}) \leq obj(\mathbf{y}), \forall \mathbf{x} \in F(P)$ if $T =$ *Minimize*

$obj(\mathbf{x}) \geq obj(\mathbf{y}), \forall \mathbf{x} \in F(P)$ if $T =$ *Maximize*

In such case, $\mathbf{x}$ is called the *optimal solution* of P and $obj(\mathbf{x})$ is the *finite optimum* of P. We denote the optimum of P by the notation $OPT(P)$ and is given by

$$OPT(P) = \begin{cases} \max\left(Img(P)\right) & \text{if } T\text{=}Maximize \\ \min\left(Img(P)\right) & \text{if } T\text{=}Minimize \end{cases}$$

Definition 3.3.5.

A linear program P is said to be an *infeasible linear program* if the feasible solution set $F(P)$ is empty. Otherwise P is called a *feasible linear program.*

Definition 3.3.6.

A linear program $P = (V, obj, C, N, T)$ is said to be a *unbounded linear program* if for every $\alpha \in \mathbb{R}$, there exists $\mathbf{x} \in F(P)$ such that

$obj(\mathbf{x}) < \alpha$ if $T =$ *Minimize*

$obj(\mathbf{x}) > \alpha$ if $T =$ *Maximize*

In other words, P is said to be unbounded if $\sup(Img(P)) = \infty$ or $\inf(Img(P)) = -\infty$ according to whether T is *Maximize* or *Minimize.*

Remark 3.3.2.

The fact that a linear program has finte optimum does not imply that the corresponding feasible region is a bounded set. This can easily be seen from the following linear program.

$$\begin{aligned} \text{Minimize } x_1 + x_2 & \text{ subject to} \\ x_1 + x_2 &\geq 0 \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned} \tag{3.2}$$

The feasible region of this linear program is the first quadrant of $\mathbb{R}^2$ which is an unbounded set. But the linear program has optimum 0 at $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Example 3.3.1.

Consider the following linear programs.

$$
\begin{aligned}
\text{Maximize } & x_1 + x_2 \text{ subject to} \\
& x_1 + 2x_2 \leq 6 \\
& 2x_1 + x_2 \leq 6 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
\tag{3.3}
$$

$$
\begin{aligned}
\text{Maximize } & x_1 + x_2 \text{ subject to} \\
& x_1 + 2x_2 \leq -6 \\
& 2x_1 + x_2 \leq -6 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
\tag{3.4}
$$

$$
\begin{aligned}
\text{Maximize } & x_1 + x_2 \text{ subject to} \\
& x_1 + 2x_2 \geq 6 \\
& 2x_1 + x_2 \geq 6 \\
& x_1 \geq 0 \\
& x_2 \geq 0
\end{aligned}
\tag{3.5}
$$

The linear program(3.6) is feasible as $\mathbf{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ is a feasible solution of the program whereas the linear program (3.4) is infeasible. It is easy to see that the linear program(3.5) is unbounded.

Remark 3.3.3.

By Definitions 3.3.4, 3.3.5 and 3.3.6, we see that a linear program has finite optimum if and only if the linear program is neither infeasible nor unbounded. Thus it follows that any linear program P satisfies exactly one of the following statements.

1. P is infeasible.
2. P is unbounded.
3. P has finite optimum.

Definition 3.3.7.

Given linear programs $P_1 = (V_1, obj_1, C_1, N_1, T_1)$ and $P_2 = (V_2, obj_2, C_2, N_2, T_2)$ where $T_1 = T_2$. We say that P_1 and P_2 are equivalent if for any $\mathbf{x} \in F(P_1)$ there exists some $\mathbf{y} \in F(P_2)$ such that $obj_1(\mathbf{x}) = obj_2(\mathbf{y})$ and vice versa. In such case, we see that $Img(P_1) = Img(P_2)$.

Given linear programs $P_1 = (V_1, obj_1, C_1, N_1, \mathit{Maximize})$ and $P_2 = (V_2, obj_2, C_2, N_2, \mathit{Minimize})$. We say that P_1 and P_2 are equivalent if for any $\mathbf{x} \in F(P_1)$ there exists some $\mathbf{y} \in F(P_2)$ such that $obj_1(\mathbf{x}) = -obj_2(\mathbf{y})$ and vice versa.

Remark 3.3.4.

Given linear program $P = (V, obj, C, N, \mathit{Minimize})$. Then we can form another linear program $P' = (V, -obj, C, N, \mathit{Maximize})$. The fact that P and P' are equivalent follows from Defintion 3.3.7.

3.3.3 Canonical Form of a linear program

In this section, we introduce the canonical form of a linear program and an algorithm to convert a linear program in general form to canonical form [3].

Definition 3.3.8.
Given linear program $P = (V, obj, C, N, T)$. Then P is said to be in *canonial form* if it satisfies the following conditions.

1. T=*Maximize*.
2. Each constraint in C is of the form $\sum_{j=1}^{n} h_j x_j \leq t$
3. Each decision variable must be non-negative. That is for each decision variable $x_j \in V$, there exists the constraint $x_j \geq 0$ in N.

Remark 3.3.5.
Given linear program $P = (V, obj, C, N, T)$ in general form. Using the following steps, we can form another linear program P' in canonical form such that P and P' are equivalent.

1. If $T =$ *Minimize*, then obtain the linear program $P_1 = (V_1, obj_1, C_1, N_1, Maximize)$ where

$$V_1 = V,\ obj_1 = -obj,\ C_1 = C,\ N_1 = N$$

   If $T =$ *Maximize*, then set $P_1 = P$.

2. If P_1 has free decision variables or non-positive decision variables, then obtain the linear program P_2 given by $P_2 = (V_2, obj_2, C_2, N_2, Maximize)$ where
   (a) V_2 consists of all decision variables of P_1 except the free variables and non-positive variables of P_1 and new decision variables ${x_j}^+$ and ${x_j}^-$ for each free decision variable x_j and x'_k for each non-positive decision variable x_k.
   (b) obj_2 is obtained by replacing each free variable x_j in obj_1 by the expression ${x_j}^+ - {x_j}^-$ and each non-positive decision variable x_k by the expression $-x'_k$.
   (c) C_2 is the set of all constraints obtained by replacing each free decision variable x_j and each non-positive variable x_k in the constraints belonging to C_1 by the expressions ${x_j}^+ - {x_j}^-$ and $-x'_k$ respectively.
   (d) N_2 consists of all constraints of N_1 except the non-positivity constraints and new constraints ${x_j}^+ \geq 0, {x_j}^- \geq 0$ and $x'_k \geq 0$ corresponding to each free decision variable x_j and each non-positive decision variable x_k respectively.

   If P_1 has no free variables, then set $P_2 = P_1$.

3. Obtain the linear program $P' = (V', obj', C', N', Maximize)$ where V', obj' and N' are same as V_2, obj_2 and N_2 respectively.
   The set C' consists of
   (a) all constraints in C_2 of the form $\alpha \leq \beta$.
   (b) constraint $-\alpha \leq -\beta$ corresponding to each constraint in C_2 that is in the form $\alpha \geq \beta$
   (c) constraints $\gamma \leq \delta$ and $-\gamma \leq -\delta$ corresponding to each constraint in C_2 that is in the form $\gamma = \delta$.

   It is easy to see that P' is in Canonical form.

Algorithm *toCanonicalForm*(P)
(∗ The algorithm converts the linear program in general form to canonical form. ∗)
Input: The linear program $P = (V, obj, C, N, T)$.
Output: The linear program $P' = (V', obj', C', N', Maximize)$ which is in canonical form and equivalent to P

$C' := \phi, N' = \phi$;
$S_1 := \{x_j \mid x_j$ is a free variable of $P\}$;
$S_2 := \{{x_j}^+, {x_j}^- \mid x_j$ is a free decision variable of $P\}$;
$S_3 := \{{x_k}' \mid x_k$ is a non-positive decision variable of $P\}$;
$V' := (V \setminus S_1) \cup S_2 \cup S_3$;
$obj_1 := obj$ with x_j replaced by ${x_j}^+ - {x_j}^-$ for each free decision variable x_j of P;
$obj_2 := obj_1$ with x_k replaced by $-{x_k}'$ for each non-positive decision variable x_k of P;
if $T = Minimize$
  then $obj' := -obj_2$
  else $obj' := obj_2$;
for each constraint $\alpha \leq \beta$ in C
  do $\alpha_1 := \alpha$ with x_j replaced by ${x_j}^+ - {x_j}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-{x_k}'$ for each non-positive decision variable x_k of P;
    $C' := C' \cup \{\alpha_2 \leq \beta\}$;
for each constraint $\alpha \geq \beta$ in C
  do $\alpha_1 := \alpha$ with x_j replaced by ${x_j}^+ - {x_j}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-{x_k}'$ for each non-positive decision variable x_k of P;
    $C' := C' \cup \{-\alpha_2 \leq -\beta\}$;
for each constraint $\alpha = \beta$ in C
  do $\alpha_1 := \alpha$ with x_j replaced by ${x_j}^+ - {x_j}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-{x_k}'$ for each non-positive decision variable x_k of P;
    $C' := C' \cup \{\alpha_2 \leq \beta, -\alpha_2 \leq -\beta\}$;
$N' := \{v \geq 0 \mid v \in V'\}$;
$P' := (V', obj', C', N', Maximize)$;
return P';

Algorithm 3.1: Algorithm that returns the linear program in canonical form which is equivalent to the input linear program in general form.

Referring to these steps, we may write the Algorithm *toCanonicalForm* (See Algorithm 3.1) that converts a given linear program in P in general form to an equivalent linear program P' in canonical form .

Example 3.3.2.
The linear programs LP 3.6 and LP 3.4 in Example 3.3.1 are in canonical form.

Remark 3.3.6.
Given linear program $P = (V, obj, C, N, T)$ where

$$\begin{aligned} V &= \{x_1, x_2, x_3, \cdots, x_n\} \\ obj &= \sum_{j=1}^{n} c_j x_j \\ C &= \left\{ \sum_{j=1}^{n} a_{ij} x_j \geq b_i \mid 1 \leq i \leq m \right\} \\ N &= \{x_j \geq 0 \mid 1 \leq j \leq n\} \\ T &= Maximize \end{aligned}$$

It is easy to see that P is in canonical form. P can be represented in matrix form as

$$\begin{aligned} \text{Maximize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A} = [a_{ij}]_{m \times n} \in \mathbb{R}^{m \times n} \\ & \mathbf{x} = [x_j]_{n \times 1} \in \mathbb{R}^n \\ & \mathbf{b} = [b_i]_{m \times 1} \in \mathbb{R}^m \\ & \mathbf{c} = [c_j]_{n \times 1} \in \mathbb{R}^n \end{aligned} \tag{3.6}$$

The feasible region of P is given by $F(P) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{Ax} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}_n\}$. By Lemma 2.3.1, we see that $F(P)$ is a polyhedral set in the positive orthant of $\mathbb{R}^n$. Since every polyhedral set is a convex set, $F(P)$ is convex too. Note that $F(P)$ is closed since polyhedral sets are closed by definition.

3.3.4 Standard Form of a linear program

In this section, we introduce the standard form of a linear program and an algorithm to convert a linear program in general form to stanadrd form [3]. Before this, we intoduce the notions of slack vector and surplus vector.

Definition 3.3.9.
Let P be the linear program

$$\begin{aligned} \text{Maximize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{Ax} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A} \in \mathbb{R}^{m \times n}, \ \mathbf{b} \in \mathbb{R}^m, \ \mathbf{c} \in \mathbb{R}^n \end{aligned} \tag{3.7}$$

Let $\mathbf{x}$ be any feasible solution of P. Then the vector $\mathbf{b} - \mathbf{Ax}$ is called the *slack vector* of P associated with $\mathbf{x}$ and is denoted by $\mathbf{s_x}$.

Definition 3.3.10.
Let P be the linear program

$$\begin{aligned} \text{Minimize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{Ax} \geq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A} \in \mathbb{R}^{m \times n}, \ \mathbf{b} \in \mathbb{R}^m, \ \mathbf{c} \in \mathbb{R}^n \end{aligned} \tag{3.8}$$

Let $\mathbf{x}$ be any feasible solution of P. Then the vector $\mathbf{Ax} - \mathbf{b}$ is called the *surplus vector* of P associated with $\mathbf{x}$ and is denoted by $\mathbf{t_x}$.

Now we give the definition of standard form.

Definition 3.3.11.
Given linear program Problem $P = (V, obj, C, N, T)$. Then P is said to be in *standard form* if it satisfies the following conditions.

1. $T = Maximize$.
2. Each constraint in C is an equality constraint. That is each constraint in C is of the form $\sum\limits_{j=1}^{n} h_j x_j = t$.

3. Each decision variable must be non-negative. That is for each decision variable $x_j \in V$, there exists the constraint $x_j \geq 0$ in N.

Remark 3.3.7.
Given a linear program $P = (V, obj, C, N, T)$ in the general form. Using the following steps, we can form another linear program $P' = (V', obj', C', N', Maximize)$ in standard form such that P and P' are equivalent .

1. Follow the first two steps of converting P to the Canonical form as given in Remark 3.3.6. This results in the linear program $P_2 = (V_2, obj_2, C_2, N_2, Maximize)$.
2. Obtain the linear program $P' = (V', obj', C', N', Maximize)$ where
   (a) V' consists of all decision variables of P_2 and new decision variables s_i, t_j corresponding to each constraint in C_2 having the form $\alpha_i \leq \beta_i$ and $\gamma_j \geq \delta_j$ respectively. Here each s_i is called a *slack variable* and each t_j is called a *surplus variable*.
   (b) obj' is same as obj_2.
   (c) C' consists of
      (i) all equality constraints in C_2.
      (ii) $\alpha_i + s_i = \beta_i$ corresponding to each constraint in C_2 that is in the form $\alpha_i \leq \beta_i$.
      (iii) $\gamma_j + t_j = \delta_j$ corresponding to each constraint in C_2 that is in the form $\gamma_j \geq \delta_j$.
   (d) N' consists of constraints of the form $v \geq 0$ corresponding to each decision variable v in V'.

   It is to see that P' is in standard form.

Referring to these steps, we may write the Algorithm *toStandardForm* (See Algorithm 3.2) that converts a given linear program in P in general form to an equivalent linear program P' in standard form.

Remark 3.3.8.
Let P_1 be the linear program

$$\begin{aligned} \text{Minimize } & \langle \mathbf{c}_1, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}_1\mathbf{x} \geq \mathbf{b}_1 \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A}_1 \in \mathbb{R}^{m\times n},\ \mathbf{b}_1 \in \mathbb{R}^m,\ \mathbf{c}_1 \in \mathbb{R}^n \end{aligned} \tag{3.9}$$

and P_2 be the linear program

$$\begin{aligned} \text{Maximise } & \langle \mathbf{c}_2, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}_2\mathbf{x} \leq \mathbf{b}_2 \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A}_2 \in \mathbb{R}^{m\times n},\ \mathbf{b}_2 \in \mathbb{R}^m,\ \mathbf{c}_2 \in \mathbb{R}^n \end{aligned} \tag{3.10}$$

Let $P_1{}'$ be the linear program

$$\begin{aligned} \text{Minimize } & \langle \mathbf{c}_1, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}_1\mathbf{x} - \mathbf{t}_\mathbf{x} = \mathbf{b}_1 \\ & \mathbf{x} \geq \mathbf{0}_n \\ & \mathbf{t}_\mathbf{x} \geq \mathbf{0}_m \end{aligned} \tag{3.11}$$

and $P_2{}'$ be the linear program

$$\begin{aligned} \text{Maximise } & \langle \mathbf{c}_2, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}_2\mathbf{x} + \mathbf{s}_\mathbf{x} = \mathbf{b}_2 \\ & \mathbf{x} \geq \mathbf{0}_n \\ & \mathbf{s}_\mathbf{x} \geq \mathbf{0}_m \end{aligned} \tag{3.12}$$

It is easy to see that $P_1{}'$ is in standard form and is equivalent to P_1. Similarly $P_2{}'$ is in standard form and is equivalent to P_2.

Algorithm *toStandardForm*(P)
(∗ The algorithm converts the linear program in general form to standard form. ∗)
Input: The linear program $P = (V, obj, C, N, T)$.
Output: The linear program $P' = (V', obj', C', N', Maximize)$ which is in canonical form and equivalent to P

$C' := \phi, N' = \phi$;
$S_1 := \{x_j \mid x_j$ is a free decision variable of $P\}\}$;
$S_2 := \{x_j{}^+, x_j{}^- \mid x_j$ is a free decision variable of $P\}\}$;
$S_3 := \{x_k{}' \mid x_k$ is a non-positive decision variable of $P\}$;
$V' := (V \setminus S_1) \cup S_2 \cup S_3$;
$obj_1 := obj$ with x_j replaced by $x_j{}^+ - x_j{}^-$ for each free decision variable x_j of P;
$obj_2 := obj_1$ with x_k replaced by $-x_k{}'$ for each non-positive decision variable x_k of P;
if $T = Minimize$
  then $obj' := -obj_2$
  else $obj' := obj_2$;
for each constraint $\alpha_i \leq \beta_i$ in C
  do $\alpha_1 := \alpha_i$ with x_j replaced by $x_j{}^+ - x_j{}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-x_k{}'$ for each non-positive decision variable x_k of P;
    $V' := V' \cup \{s_i\}$;
    $C' := C' \cup \{\alpha_2 + s_i = \beta_i\}$;
for each constraint $\alpha_j \geq \beta_j$ in C
  do $\alpha_1 := \alpha_j$ with x_j replaced by $x_j{}^+ - x_j{}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-x_k{}'$ for each non-positive decision variable x_k of P;
    $V' := V' \cup \{t_j\}$;
    $C' := C' \cup \{\alpha_2 - t_j = \beta_j\}$;
for each constraint $\alpha = \beta$ in C
  do $\alpha_1 := \alpha$ with x_j replaced by $x_j{}^+ - x_j{}^-$ for each free decision variable x_j of P;
    $\alpha_2 := \alpha_1$ with x_k replaced by $-x_k{}'$ for each non-positive decision variable x_k of P;
    $C' := C' \cup \{\alpha_2 = \beta\}$;
$N' := \{v \geq 0 \mid v \in V'\}$;
$P' := (V', obj', C', N', Maximize)$;
return P';

Algorithm 3.2: Algorithm that returns the linear program in standard form which is equivalent to the input linear program in general form.

3.3.5 Equivalence of General, Canonical and Standard Forms

From the discussions done so far, we see that corresponding to any linear program in general form, there exist linear programs in canonical form and standard form and may be obtained by calling Algorithms *toCanonicalForm* and *toStandardForm* respectively. Each linear program in the canonical form or standard form is in the general form and therefore an equivalent linear program in the third form may obtained by calling Algorithm *toCanonicalForm* or Algorithm *toStandardForm* accordingly. The following lemma summarises the discussions done so far.

Lemma 3.3.1.
The general, canonical and the standard forms of linear programs are equivalent.

3.4 Summary

In this chapter, we introduced the fundamental concepts in linear programming. We have started the chapter with the general form of a linear program which follows the discussion on notions of feasible region, feasible linear programs, unbounded linear programs and equivalent linear programs with illustrative examples. Then we defined the canonical and standard forms of LP and established the equivalence of various forms with algorithms for converting from one form to another form. In the next chapter, we shall discuss the geometry of linear programming which provides the platform for the graphical method of solving linear programs.

Chapter 4

Geometry of Linear Programming

4.1 Introduction

This chapter is intended to give a geometric interpretation for linear programming problems and a graphical method to solve linear programs. In section 2.2, we introduce the notion of basic feasible solution and establish its equivalence with vertices and extreme points. Section 2.3 is dedicated to the discussion of recession directions and extreme directions of a polyhedral set. In section 2.4, we prove and illustrate the General representation theorem [7] for linear programming due to *Constantin Carathodory* and the fundamental theorem of linear prgramming. We complete this chapter by giving a brute force algorithm for solving linear programs with a brief graphical illustration of the algorithm in section 2.5.

Throughout this chapter we assume that the linear program under consideration is in the canonical form given by

$$\begin{aligned} &\text{Maximize } \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ &\mathbf{A}\mathbf{x} \leq \mathbf{b} \\ &\mathbf{x} \geq \mathbf{0}_n \end{aligned} \tag{4.1}$$

where $\mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{c} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m$.

The feasible region of the linear program can be expressed as $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n, \mathbf{G}\mathbf{x} \leq \mathbf{h}\}$ where

$$\begin{aligned} \mathbf{G} &= \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots & \mathbf{g}_{m+n} \end{bmatrix}^T = \begin{bmatrix} \mathbf{A} \\ -\mathbf{I_n} \end{bmatrix} \in \mathbb{R}^{(m+n) \times n} \\ \mathbf{h} &= \begin{bmatrix} h_1 & h_2 & h_3 & \cdots & h_{m+n} \end{bmatrix}^T = \begin{bmatrix} \mathbf{b} \\ \mathbf{0}_n \end{bmatrix} \in \mathbb{R}^{m+n} \end{aligned}$$

4.2 Basic Feasible Solutions

In this section, we introduce the concept of basic feasible solutions of a linear program and establish that the basic feasible solution of a linear program, vertices and extreme points of the polyhedral set representing the feasible region of the linear program are equivalent.

Definition 4.2.1.

Given any feasible linear program P in canonical form and $S \subseteq \{1, 2, 3, \cdots, m+n\}$. Then the set of constraints $\{\langle \mathbf{g}_i, \mathbf{x} \rangle \leq h_i \mid i \in S\}$ of P is said to be *linearly independent* if $\{\mathbf{g}_i \mid i \in S\}$ forms a linearly independent set of vectors in $\mathbb{R}^n$.

Definition 4.2.2.

Given any feasible linear program P in canonical form and $\mathbf{x} \in \mathbb{R}^n$. The number of linearly independent constraints of P which are tight at $\mathbf{x}$ is called the *rank* of $\mathbf{x}$ with respect to P and is denoted by $rank_P(\mathbf{x})$.

Remark 4.2.1.
Since the maximum number of linearly indpendent vectors in any set of vectors in $\mathbb{R}^n$ is n, we see that the maximum number of linearly independent constraints of P is n. This implies that for any linear program P in canonical form, the rank of all elements of $F(P)$ with respect to P is at most n.

The following Lemma characterizes the interior points in $F(P)$.

Lemma 4.2.1.
Given linear Program P in the canonical form and $\mathbf{y} \in F(P)$. $\mathbf{y}$ is an interior point of $F(P)$ if and only if $rank_P(\mathbf{y}) = 0$.

Proof.
Let the constraints of P be expressed as

$$\langle \mathbf{g}_i, \mathbf{x} \rangle \leq h_i, \;\; \mathbf{g}_i \in \mathbb{R}^n, \; b_i \in \mathbb{R}, \; 1 \leq i \leq m+n$$

if part:-
Let $rank_P(\mathbf{x}) = 0$. Clearly $\mathbf{x} > \mathbf{0}_n$. Since $rank_P(\mathbf{x}) = 0$, we see that none of the constraints of P is tight at $\mathbf{x}$ and hence $\langle \mathbf{g}_i, \mathbf{x} \rangle < h_i$ for each $i \in \{1, 2, 3, \cdots, m+n\}$. This implies that corresponding to each $i \in \{1, 2, 3, \cdots, m+n\}$, there exists $\delta_i > 0$ such that $\langle \mathbf{g}_i, \mathbf{x} \rangle = h_i - \delta_i$. This further implies that corresponding to each i in $\{1, 2, 3, \cdots, m+n\}$, there exists some vector $\mathbf{d}_i = \begin{bmatrix} \epsilon_i{}^1 & \epsilon_i{}^2 & \epsilon_i{}^3 & \cdots & \epsilon_i{}^n \end{bmatrix}^T > \mathbf{0}_n$ such that $\langle \mathbf{g}_i, \mathbf{d}_i \rangle \leq \delta_i$ and hence $\langle \mathbf{g}_i, \mathbf{x} \pm \mathbf{d}_i \rangle \leq h_i$.
Let $\epsilon = \min\{\epsilon_i{}^j \mid 1 \leq i \leq k, 1 \leq j \leq n\}$ and $\mathbf{d} = \epsilon \mathbf{1}_n$. Then we see that $\langle \mathbf{g}_i, \mathbf{x} \pm \mathbf{d} \rangle \leq h_i$ for each $i \in \{1, 2, 3, \cdots, m+n\}$. We futher see that for each vector $\mathbf{d}'$ in $\mathbb{R}^n$ such that $\mathbf{0}_n \leq \mathbf{d}' \leq \mathbf{d}$, it must be the case that $\langle \mathbf{g}_i, \mathbf{x} \pm \mathbf{d}' \rangle \leq h_i$ for each $i \in \{1, 2, 3, \cdots, m+n\}$. In other words, there exists $\epsilon > 0$ such that corresponding to each vector $\mathbf{d}'$ in $\mathbb{R}^n$ with $\|\mathbf{d}'\|_\infty \leq \epsilon$, we must have $\langle \mathbf{g}_i, \mathbf{x} \pm \mathbf{d}' \rangle \leq h_i$ for each $i \in \{1, 2, 3, \cdots, m+n\}$. Thus it follows that there exists $\epsilon > 0$ such that all vectors $\mathbf{z}$ in $\mathbb{R}^n$ with $\|\mathbf{z} - \mathbf{x}\|_\infty \leq \epsilon$ are members of $F(P)$. Since the norms in $\mathbb{R}^n$ are equivalent (Refer to Fact in Appendix), we further see that there exists some $\delta > 0$ such that all vectors $\mathbf{z}$ in $\mathbb{R}^n$ with $\|\mathbf{z} - \mathbf{x}\| \leq \delta$ are members of $F(P)$. This confirms that there exists some $\delta > 0$ such that $N_\delta(\mathbf{x}) \subseteq F(P)$. Hence we conclude that $\mathbf{x}$ is an interior point of $F(P)$.
only if part:-
Let $\mathbf{y}$ be an interior point of $F(P)$. Therefore there exists $\delta > 0$ such that $N_\delta(\mathbf{y}) \subseteq F(P)$. This ensures that coresponding to each $\mathbf{d} \in \mathbb{R}^n$, there exists $\epsilon > 0$ such that $\mathbf{y} \pm \epsilon\mathbf{d} \in F(P)$. In other words, corresponding to each $\mathbf{d} \in \mathbb{R}^n$, there exists $\epsilon > 0$ such that $\langle \mathbf{g}_i, \mathbf{y} \pm \epsilon\mathbf{d} \rangle = \langle \mathbf{g}_i, \mathbf{y} \rangle \pm \epsilon \langle \mathbf{g}_i, \mathbf{d} \rangle \leq h_i$ for each $i \in \{1, 2, 3, \cdots, m+n\}$. Since $\epsilon > 0$ and $\mathbf{g}_i \neq \mathbf{0}_n$, we further infer that $\langle \mathbf{g}_i, \mathbf{x} \rangle < h_i$ for each i in $\{1, 2, 3, \cdots, m+n\}$. Hence we see that $rank_P(\mathbf{y}) = 0$
Hence the Lemma. □

Corollary 4.2.1.
Given linear Program P in the canonical form and $\mathbf{y} \in \mathbb{R}^n$. $rank_P(\mathbf{y}) = 0$ if and only if for each $\mathbf{d} \in \mathbb{R}^n$, there exists $\delta > 0$ such that $\mathbf{x} \pm \delta\mathbf{d} \in F(P)$.

Definition 4.2.3.
Given linear Program P in the canonical form and $\mathbf{y} \in \mathbb{R}^n$. Then $\mathbf{y}$ is said to be a *basic feasible solution(bfs)* [7] of P if

(i) $rank_P(\mathbf{y}) = n$.

(ii) $\mathbf{y} \in F(P)$.

Theorem 4.2.1.

Given a feasible linear Program P in canonical form and $\mathbf{y} \in F(P)$. Then the following statements are equivalent.

(a) $\mathbf{y}$ *is a vertex of* $F(P)$.

(b) $\mathbf{y}$ *is an extreme point of* $F(P)$.

(c) $\mathbf{y}$ *is a basic feasible solution of* P.

Proof.

Let the constraints of P be expressed as

$$\langle \mathbf{g}_i, \mathbf{x} \rangle \leq h_i,\ \mathbf{g}_i \in \mathbb{R}^n,\ h_i \in \mathbb{R},\ 1 \leq i \leq m+n$$

$(a) \Rightarrow (b)$ by Lemma 2.3.3.

(b) $\Rightarrow$ *(c)*

Let $\mathbf{y}$ be an extreme point of $F(P)$. Assume that $\mathbf{y}$ is not a basic feasible solution of P. Therefore $rank_P(\mathbf{y}) < n$ since every extreme point of $F(P)$ is a member of $F(P)$. Now we shall consider the following two cases.

Case 1: $rank_P(\mathbf{y}) = 0$

Let $\mathbf{d}$ be any non-zero vector in $\mathbb{R}^n$. Therefore by Corollary 4.2.1, there exists $\delta > 0$ such that $\mathbf{x} \pm \delta\mathbf{d} \in F(P)$. This gives a representation for $\mathbf{y}$ as a convex combination of two distinct points $\mathbf{y} + \delta\mathbf{d}$ and $\mathbf{y} - \delta\mathbf{d}$ in $F(P)$ given by $\mathbf{y} = \frac{1}{2}(\mathbf{y} + \delta\mathbf{d}) + \frac{1}{2}(\mathbf{y} - \delta\mathbf{d})$ which leads to the contradiction that $\mathbf{y}$ is not an extreme point of $F(P)$.

Case 2: $1 \leq rank_P(\mathbf{y}) < n$

Let $r = rank_P(\mathbf{y})$. Without loss of generality, we may assume that the first $t \geq r$ constraints of P are are tight at $\mathbf{y}$ and the first r among these costraints are linearly independent. That is

$$\langle \mathbf{g}_j, \mathbf{y} \rangle = h_j, 1 \leq j \leq t \text{ and } \langle \mathbf{g}_j, \mathbf{y} \rangle < h_j, t+1 \leq j \leq m+n$$

Claim. *There exists a non-zero vector* $\mathbf{d} \in \mathbf{R}^n$ *and* $\delta > 0$ *such that* $\mathbf{y} \pm \delta\mathbf{d} \in F(P)$.

Since $r < n$, the homogeneous system of linear equations $\begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots & \mathbf{g}_t \end{bmatrix}^T \mathbf{z} = \mathbf{O_r}$ has infinitely many solutions. Let $\mathbf{d}$ be any non zero solution of this homogeneous system . That is $\langle \mathbf{g}_j, \mathbf{d} \rangle = 0$ for each j in $\{1, 2, 3, \cdots, t\}$.

Let j be any index in $\{1, 2, 3, \cdots, m+n\}$.

If $j \leq t$, we see that for each $\epsilon \geq 0$

$$\langle \mathbf{g}_j, \mathbf{y} \pm \epsilon\mathbf{d} \rangle = \langle \mathbf{g}_j, \mathbf{y} \rangle \pm \epsilon\langle \mathbf{g}_j, \mathbf{d} \rangle = h_j \pm 0 = h_j$$

If $t+1 \leq j \leq m+n$, we see that $\langle \mathbf{g}_j, \mathbf{y} \rangle < h_j$. This means that there exists some positive real number α_j such that $\langle \mathbf{g}_j, \mathbf{y} \rangle = h_j - \alpha_j$. Thus it follows that there exists some $\delta_j > 0$ such that $\langle \mathbf{g}_j, \delta_j\mathbf{d} \rangle \leq \alpha_j$ and hence $\langle \mathbf{g}_j, \mathbf{y} \pm \delta_j\mathbf{d} \rangle \leq h_j$.

Now by letting $\delta = \min\{\delta_j \mid t+1 \leq j \leq m+n\}$, we see that $\langle \mathbf{g}_j, \mathbf{y} \pm \delta\mathbf{d} \rangle \leq h_j$ for each $j \in \{1, 2, 3, \cdots, m+n\}$. Thus it follows that there exists a non-zero vector $\mathbf{d} \in \mathbf{R}^n$ and $\delta > 0$ such that $\mathbf{y} + \delta\mathbf{d} \in F(P)$ and $\mathbf{y} - \delta\mathbf{d} \in F(P)$.

Now $\mathbf{y} = \frac{1}{2}(\mathbf{y} + \delta\mathbf{d}) + \frac{1}{2}(\mathbf{y} - \delta\mathbf{d})$ which is a strict convex combination of two points $\mathbf{y} + \delta\mathbf{d}$ and $\mathbf{y} - \delta\mathbf{d}$ in $F(P)$ which leads to the contradiction that $\mathbf{y}$ is not an extreme point of $F(P)$.

Hence in either case we get results contradicting to the fact that $\mathbf{y}$ is not an extreme point of $F(P)$. Hence we claim that $\mathbf{y}$ is a basic feasible solution of P.

(c) $\Rightarrow$ *(a)*:-

Let $\mathbf{y}$ be a basic feasible solution of P. Therefore $rank_P(\mathbf{y}) = n$ and hence n linearly independent constraints of P are tight at $\mathbf{y}$. Without loss of generality, we may assume that the first $t \geq n$ constraints of P are are tight at $\mathbf{y}$ and the first n among these costraints are linearly independent. That is

$$\langle \mathbf{g}_j, \mathbf{y} \rangle = h_j, 1 \leq j \leq t \text{ and } \langle \mathbf{g}_j, \mathbf{y} \rangle < h_j, t+1 \leq j \leq m+n$$

Clearly the rank of the matrix $\begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots, \mathbf{g}_n \end{bmatrix} \mathbf{z} = \mathbf{0}_n$ is n and hence the system of linear equations $\begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots, \mathbf{g}_n \end{bmatrix} \mathbf{z} = \mathbf{h}$ has a unique solution which is $\mathbf{y}$. This implies that $\mathbf{y}$ is the unique point in $\mathbb{R}^n$ such that $\langle \mathbf{g}_i, \mathbf{y} \rangle = h_i$ for each $i \in \{1, 2, 3, \cdots, n\}$. This further implies that for each vector $\mathbf{v}$ in $F(P) \setminus \{\mathbf{y}\}$, there exists $k \in \mathbb{N}$ such that $1 \leq k \leq n$ and $\langle \mathbf{g}_k, \mathbf{v} \rangle < h_k$.

Let $\mathbf{c} \in \mathbb{R}^n$ defined by $\mathbf{c} = \sum_{j=1}^{n} \mathbf{g}_j$. Now $\langle \mathbf{c}, \mathbf{y} \rangle = \sum_{j=1}^{n} \langle \mathbf{g}_j, \mathbf{y} \rangle = \sum_{j=1}^{n} h_j$. For each vector $\mathbf{v}$ in $F(P) \setminus \{\mathbf{y}\}$, we have the relation $\langle \mathbf{c}, \mathbf{v} \rangle = \sum_{j=1}^{n} \langle \mathbf{g}_j, \mathbf{v} \rangle < \sum_{j=1}^{n} h_j$.

Hence we see that there exists vector $\mathbf{c} \in \mathbb{R}^n$ such that $\langle \mathbf{c}, \mathbf{y} \rangle > \langle \mathbf{c}, \mathbf{v} \rangle$ for each vector $\mathbf{v}$ in $F(P) \setminus \{\mathbf{y}\}$. Therefore by Lemma 2.3.2, $\mathbf{y}$ is a vertex of $F(P)$.

Hence the Theorem. □

Example 4.2.1.

Consider the linear program

$$\begin{aligned} &\text{Maximize } x_1 + x_2 \text{ subject to} \\ &2x_1 + x_2 \leq 6 \\ &x_1 + 2x_2 \leq 6 \\ &x_1 + x_2 \leq 4 \\ &x_1 \geq 0 \\ &x_2 \geq 0 \end{aligned} \tag{4.2}$$

This linear program is in the canonical form and its feasible region is shown in Figure 4.1. We see that The basic feasible solutions of P are $\begin{bmatrix} 0 \\ 3 \end{bmatrix}$, $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$, $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$. It is easy to see that these vectors are the vertices and the extreme points of $F(P)$.

Remark 4.2.2.

Referring to Theorem 4.2.1, we may write the Algorithm *getExtremePoints* (See Algorithm 4.1) that returns the extreme points of a linear program in canonical form .

Remark 4.2.3.

In theorem 4.2.1, we proved the equivalence of the notions of vertices and extreme points for polyhedral sets in the first orthant of $\mathbb{R}^n$. This equivalence can be generalised to any polyhedral set in $\mathbb{R}^n$. However, this equivalence does not hold for an arbitrary convex set in $\mathbb{R}^n$ as shown in Remark 2.3.3.

Now we are going to prove an important result that states that any constraint which is tight at any two feasible solutions of a linear program in canonical form is tight all strict convex combinations of the two feasible solutions and vice versa.

Lemma 4.2.2.

Let P be any feasible linear program in canonical form and two vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Let C be any constraint of P. Then the following statements are equivalent.

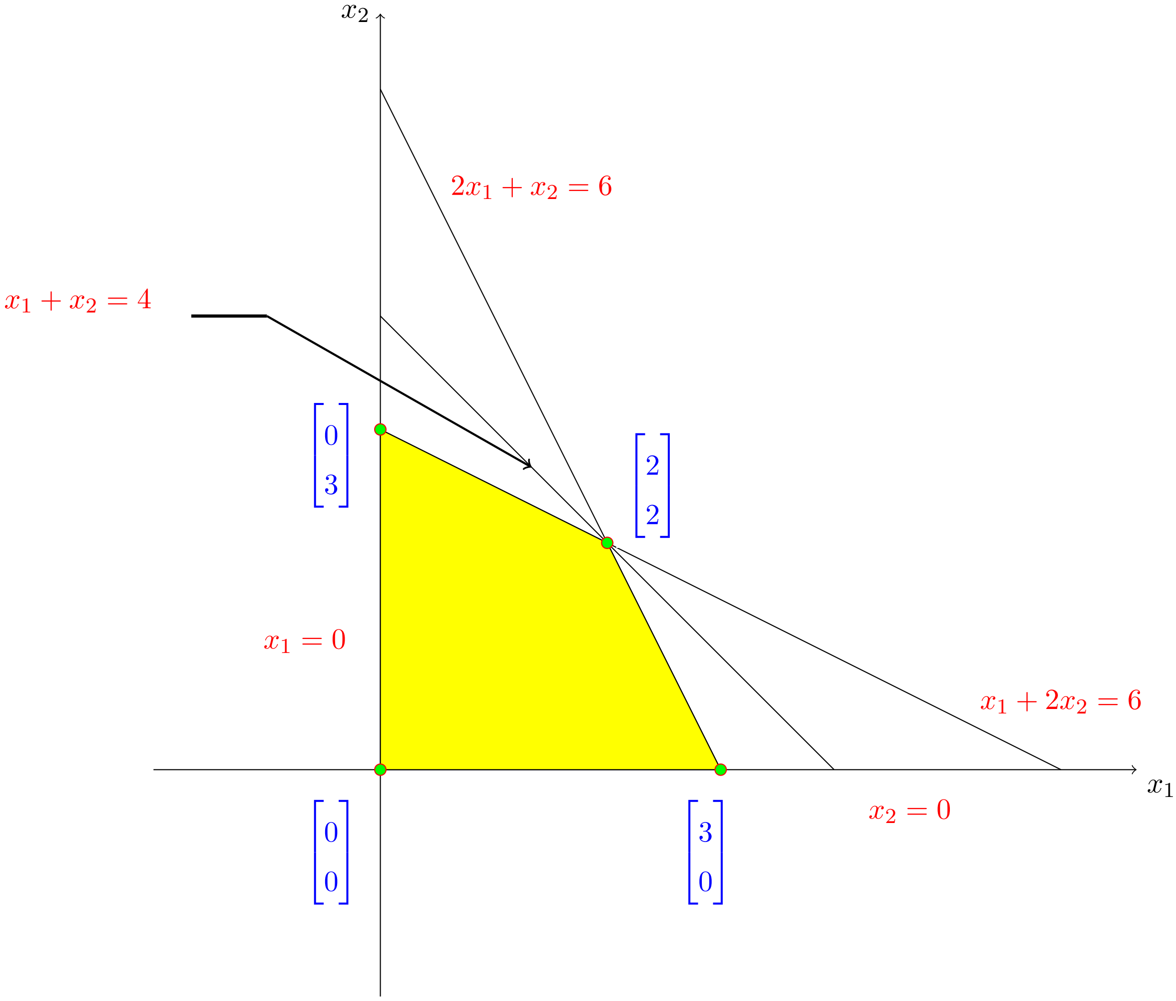


Figure 4.1: The shaded region is the feasible region of LP 4.2.

Algorithm *getExtremePoints*$(F(P))$
(∗ The algorithm returns the extreme points of a linear program P in canonical form. ∗)
Input: $F(P) = \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{g}_i, \mathbf{x} \rangle \leq h_i, i \in \{1, 2, 3 \cdots, m+n\}\}$
Output: The extreme points of P
1. $E_P := \phi$;
2. **for** each permutation I of $\{1, 2, 3 \cdots, m+n\}$ of size n
3. $B = \left[\mathbf{g}_i\right]_{i \in I}$;
4. **if** $rank(B) := n$
5. **then** Solve $\{\langle \mathbf{g}_i, \mathbf{y} \rangle = h_i \mid i \in I\}$ using Gauss Elimination method.
6. $E_P := E_P \cup \{\mathbf{y}\}$;
7. **return** E_P;

Algorithm 4.1: Algorithm that returns the extreme points of a linear program in canonical form.

(1) C is tight at both $\mathbf{u}$ *and* $\mathbf{v}$.

(2) C is tight at every strict convex combination of $\mathbf{u}$ *and* $\mathbf{v}$.

(3) C is tight at some strict convex combination of $\mathbf{u}$ *and* $\mathbf{v}$.

Proof.
Let the constraints of P be expressed as

$$\langle \mathbf{g}_i, \mathbf{x} \rangle \leq h_i \; \mathbf{g}_i \in \mathbb{R}^n, \; h_i \in \mathbb{R}, \; 1 \leq i \leq m+n$$

Let C be the constraint $\langle \mathbf{a}_k, \mathbf{x} \rangle \leq b_k$ where $k \in \{1, 2, 3, \cdots, m+n\}$.
(1) ⇒ (2):-
Let C be tight at $\mathbf{u}$ and $\mathbf{v}$. That is $\langle \mathbf{g}_k, \mathbf{u} \rangle = \langle \mathbf{g}_k, \mathbf{v} \rangle = h_k$.
Let $\mathbf{z}$ be any strict convex combination of $\mathbf{u}$ and $\mathbf{v}$. Therefore there exists some $\lambda \in (0, 1)$ such that $\mathbf{z} = \lambda \mathbf{u} + (1 - \lambda)\mathbf{v}$.
Now

$$\langle \mathbf{g}_k, \mathbf{z} \rangle \;\; = \;\; \langle \mathbf{g}_k, \lambda \mathbf{u} + (1-\lambda)\mathbf{v} \rangle = \lambda \langle \mathbf{g}_k, \mathbf{u} \rangle + (1-\lambda) \langle \mathbf{g}_k, \mathbf{v} \rangle = h_k$$

Thus we see that C is tight at $\mathbf{z}$. Since $\mathbf{z}$ can be any strict convex combination of $\mathbf{u}$ and $\mathbf{v}$, it follows that C is tight at all strict convex combinations of $\mathbf{u}$ and $\mathbf{v}$.
(2) ⇒ (3) is trivial.
(3) ⇒ (1):-
Let $\mathbf{y}$ be some strict convex combination of $\mathbf{u}$ and $\mathbf{v}$. Therefore there exists some $\alpha \in (0, 1)$ such that $\mathbf{y} = \alpha \mathbf{u} + (1 - \alpha)\mathbf{v}$. Let C be tight at $\mathbf{y}$. Assume that C is not tight at at least one of $\mathbf{u}$ and $\mathbf{v}$. Now

$$\langle \mathbf{g}_k, \mathbf{y} \rangle \;\; = \;\; \langle \mathbf{g}_k, \alpha \mathbf{u} + (1-\alpha)\mathbf{v} \rangle = \alpha \langle \mathbf{g}_k, \mathbf{u} \rangle + (1-\alpha) \langle \mathbf{g}_k, \mathbf{v} \rangle < \alpha h_k + (1-\alpha) h_k$$

which implies that $\langle \mathbf{g}_k, \mathbf{y} \rangle < h_k$ and leads to the contradiction that C is not tight at $\mathbf{y}$. Hence we see that the constraint C is tight at both $\mathbf{u}$ and $\mathbf{v}$.
Hence the Lemma □

Now we are going to extend Lemma 4.2.2 to affine combinations. However, the extension is possible only in one direction as illustrated below.

Lemma 4.2.3.
Let P be any feasible linear program in canonical form and two vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Then any constraints of P that is tight at both $\mathbf{u}$ and $\mathbf{v}$ is tight at any feasible affine combinations of $\mathbf{u}$ and $\mathbf{v}$.

Proof.
Let k be any index in $\{1, 2, 3, \cdots, m+n\}$ such that the constraint $\langle \mathbf{g}_k, \mathbf{x} \rangle \leq b_k$ is tight at both $\mathbf{u}$ and $\mathbf{v}$. That is $\langle \mathbf{g}_k, \mathbf{u} \rangle = \langle \mathbf{g}_k, \mathbf{v} \rangle = h_k$.
Let $\mathbf{z}$ be any feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$. Therefore there exists real number β such that $\mathbf{z} = \beta \mathbf{u} + (1 - \beta)\mathbf{v}$. Now

$$\langle \mathbf{g}_k, \mathbf{z} \rangle \;\; = \;\; \langle \mathbf{g}_k, \beta \mathbf{u} + (1-\beta)\mathbf{v} \rangle = \beta \langle \mathbf{g}_k, \mathbf{u} \rangle + (1-\beta) \langle \mathbf{g}_k, \mathbf{v} \rangle = h_k$$

Thus we see that the constraint $\langle \mathbf{g}_k, \mathbf{x} \rangle \leq h_k$ is tight at $\mathbf{z}$. Since $\mathbf{z}$ can be chosen as any feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$ and the choice of k is arbitrary in the range $[1, m+n]$, any constraint of P that is tight at both $\mathbf{u}$ and $\mathbf{v}$ will also be tight at any feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$. Hence the Lemma. □

It can be shown that the converse of Lemma 4.2.3 is not true in general. We illustrate this by means of an example given in the following remark.

Remark 4.2.4.
Consider the linear program P given by

$$\begin{aligned} \text{Maximize } & x_1 + x_2 \text{ subject to} \\ & x_1 + 2x_2 \leq 6 \\ & 2x_1 + x_2 \leq 6 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned} \tag{4.3}$$

The feasible region of P shown in Figure 4.2. The constraint $2x_1 + x_2 \leq 6$ is tight at the feasible points $\begin{bmatrix} 2.3 \\ 1.6 \end{bmatrix}$ (point $\mathbf{u}$ in Figure 4.2) and $\begin{bmatrix} 5.4 \\ 0.6 \end{bmatrix}$ (point $\mathbf{v}$ in Figure 4.2). It is easy to see that this constraint is tight at the point $\begin{bmatrix} 2.5 \\ 1 \end{bmatrix}$ (point $\mathbf{z}$ in Figure 4.2) which is a strict convex combination of $\mathbf{u}$ and $\mathbf{v}$ and at the point $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ (point $\mathbf{y}$ in Figure 4.2) which is a feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$. It can also be seen that the constraint $x_1 + 2x_2 \leq 6$ is also tight at $\mathbf{y}$ and is not tight at any of $\mathbf{u}, \mathbf{v}$ or $\mathbf{z}$.

The following corollary easily follows from Lemma 4.2.2, Lemma 4.2.3 and Remark 4.2.4.

Corollary 4.2.2.
Let P be any feasible linear program in canonical form and two vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Then the rank of any feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$ with respect to P is at least the rank of any strict convex combination of $\mathbf{u}$ and $\mathbf{v}$ with respect to P.

Now we discuss a lemma which essentially states that it is impossible to move infinitely along both the forward and backward directions the feasible region of a linear program while retaining feasiblity in both the directions.

Lemma 4.2.4.
Let P be any feasible linear program in canonical form and $\mathbf{z}$ be any strict convex combination of two vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Let S_1 and S_2 be two subsets of real numbers defined by $S_1 = \{\alpha \mid \alpha \geq 0 \text{ and } \mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\}$ and $S_2 = \{\alpha \mid \alpha \geq 0 \text{ and } \mathbf{z} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\}$. Let $\alpha_1 = \sup(S_1)$ and $\alpha_2 = \sup(S_2)$. Then

(1) $\alpha_1 > 0$ *and* $\alpha_2 > 0$.

(2) $\min\{\alpha_1, \alpha_2\} < \infty$.

(3) If $\alpha_1 < \infty$, then

  (a) $\alpha_1 = \max(\{\alpha \mid \alpha \geq 0 \text{ and } \mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\})$.
  (b) $\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v}) \in F(P)$.
  (c) $rank_P(\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v})) > rank_P(\mathbf{z})$.

(4) If $\alpha_2 < \infty$, then

  (a) $\alpha_2 = \max(\{\alpha \mid \alpha \geq 0 \text{ and } \mathbf{z} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\})$.
  (b) $\mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u}) \in F(P)$.
  (c) $rank_P(\mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u})) > rank_P(\mathbf{z})$.

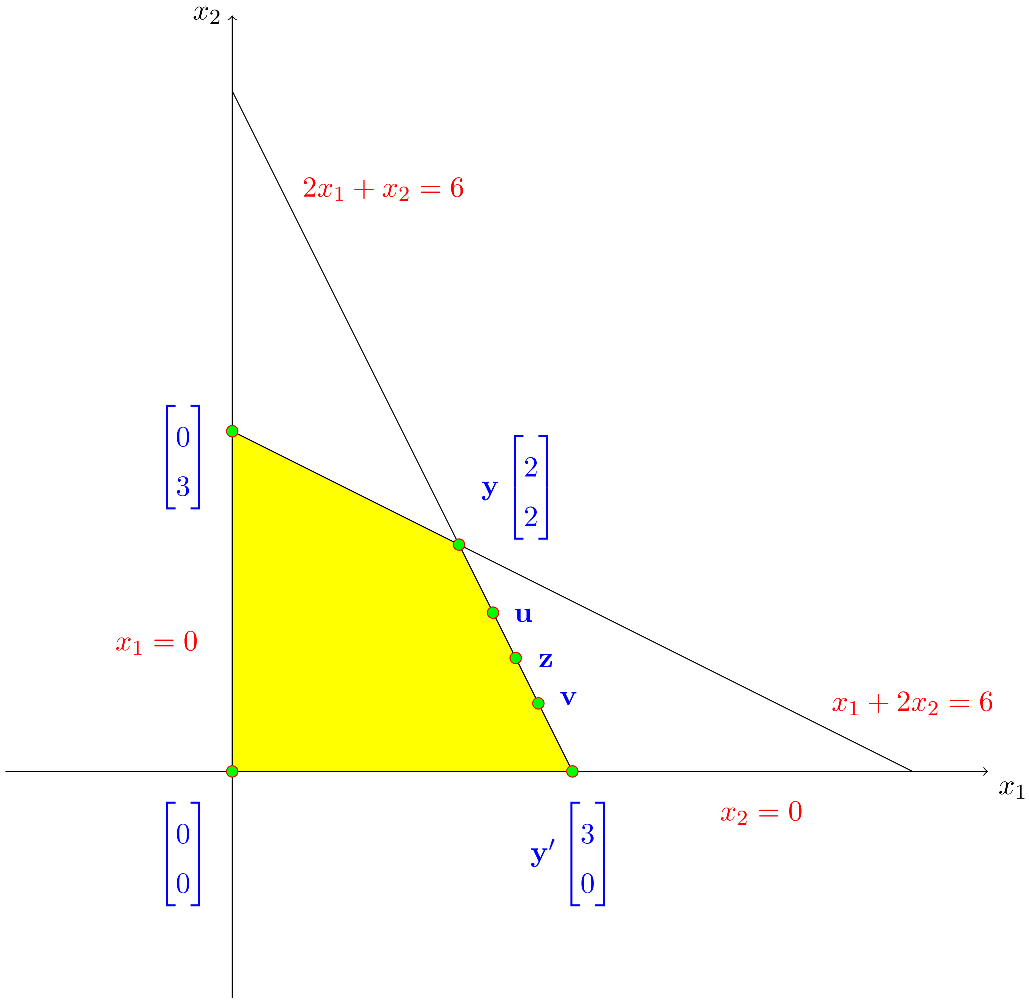


Figure 4.2: The shaded region is the feasible region of LP 4.3. It is easy to see that the constraint $2x_1 + x_2 \leq 6$ is tight at both $\mathbf{u}$ and $\mathbf{v}$. This constraint is also tight at $\mathbf{z}$ which is a strict convex combination of $\mathbf{u}$ and $\mathbf{v}$ and at $\mathbf{y}$ which is a feasible affine combination of $\mathbf{u}$ and $\mathbf{v}$. It is also seen that the constraint $x_1 + 2x_2 \leq 6$ is also tight at $\mathbf{y}$ and is not tight at any of $\mathbf{u}, \mathbf{v}$ or $\mathbf{z}$.

Proof.

(1):- Since $\mathbf{z}$ is a strict convex combination of $\mathbf{u}$ and $\mathbf{v}$, there exists some $\lambda \in (0, 1)$ such that $\mathbf{z} = \lambda\mathbf{u} + (1 - \lambda)\mathbf{v}$. Hence $\mathbf{u} = \mathbf{z} + (1 - \lambda)(\mathbf{u} - \mathbf{v})$ and $\mathbf{v} = \mathbf{z} + \lambda(\mathbf{v} - \mathbf{u})$. This implies that there exists $\alpha_1, \alpha_2 > 0$ such that $\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v}) \in F(P)$ and $\mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u}) \in F(P)$. Thus we can conclude that $\alpha_1 > 0$ and $\alpha_2 > 0$.

(2):- It is easy to see that as α increases, one among $\mathbf{z}+\alpha(\mathbf{u}-\mathbf{v})$ and $\mathbf{z}+\alpha(\mathbf{v}-\mathbf{u})$ violates the non-negativity constraints of P. This implies that there exists some $\alpha_0 < \infty$ such that at least one of $\mathbf{z}+\alpha(\mathbf{u}-\mathbf{v})$ and $\mathbf{z}+\alpha(\mathbf{v}-\mathbf{u})$ will not belong to $F(P)$ for each $\alpha > \alpha_0$. Therefore we see that at least one of S_1 and S_2 is bounded. Hence we conclude that at least one of S_1 and S_2 has a finite supremum (Refer to Fact C.1.7 in Appendix C). Hence $\min\{\alpha_1, \alpha_2\} < \infty$ (See Figure 4.3).

(3):- If $\alpha_1 < \infty$, then we see that S_1 is bounded. It is not hard to see that S_1 is closed. Thus we see that S_1 is compact by Heine Borel Theorem (Refer to Fact C.2.1 in Appendix C). Hence S_1 has a finite maximum (Refer to Fact C.2.2 in Appendix C). Thus it follows that S_1 has a finite maximum α_1 such that $\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v}) \in F(P)$.

Let $T_1 = \{\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \mid \alpha \in S_1\}$ and $\mathbf{y}_1 = \mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v})$. We have

$$\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) = \lambda\mathbf{u} + (1 - \lambda)\mathbf{v} + \alpha(\mathbf{u} - \mathbf{v}) = (\lambda + \alpha)\mathbf{u} + (1 - \lambda - \alpha)\mathbf{v}$$

Thus we see that $\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v})$ is an affine combination of $\mathbf{u}$ and $\mathbf{v}$ for all real values of α. Thus it follows that $\mathbf{y}_1$ is an affine combination of $\mathbf{u}$ and $\mathbf{v}$. Now the application of Lemma 4.2.2 and Lemma 4.2.3 implies that all constraints of P which are tight at $\mathbf{z}$ are tight at all vectors in S_1 which includes $\mathbf{y}_1$ also . Moreover the fact that $\alpha_1 < \infty$ implies that for all values of $\alpha > \alpha_1$, it must be the case that $\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \notin F(P)$. This further implies that in addition to the constraints which are tight at $\mathbf{z}$, there is at least one more constraint of P becomes tight at $\mathbf{y}_1$.

Now we are going to prove that the constraint of P which is exclusively tight at $\mathbf{y}_1$ is not linearly dependent on the other constraints which are tight at all points in S_1 as follows. Let $r = rank_P(\mathbf{z})$. Without loss of generality, we may take the first t constraints of P are tight at $\mathbf{z}$ for some t satisfying $t \geq r$. Therefore these constraints are tight at all vectors in T_1 by Lemma 4.2.3. Let k be an index such that $t + 1 \leq k \leq m + n$ and the constraint $\langle \mathbf{g}_k, \mathbf{x}\rangle \leq h_k$ is not tight at any of the points of T_1 except $\mathbf{y}_1$. Assume that this constraint is linearly dependent on the first t constraints which are tight at $\mathbf{y}_1$. Therefore there exists real numbers $\beta_1, \beta_2, \beta_3, \cdots, \beta_t$ not all of which are zero such that $\mathbf{g}_k = \sum_{i=1}^{t} \beta_i \mathbf{g}_i$.

Now we see that for each vector $\mathbf{x}$ in T_1, it must be the case that

$$\langle \mathbf{g}_k, \mathbf{x}\rangle = \left\langle \sum_{i=1}^{t} \beta_i \mathbf{g}_i, \mathbf{x} \right\rangle = \sum_{i=1}^{t} \beta_i \langle \mathbf{g}_i, \mathbf{x}\rangle = \sum_{i=1}^{t} \beta_i h_i$$

Since $\mathbf{z} \in T_1$ and the constraint $\langle \mathbf{g}_k, \mathbf{x}\rangle \leq h_k$ is tight at $\mathbf{z}_1$, we conclude that $\sum_{i=1}^{t} \beta_i h_i = h_k$. Thus we see that $\langle \mathbf{g}_k, \mathbf{x}\rangle = h_k$ for each vector $\mathbf{x} \in T_1$ which contradicts the fact that the constraint $\langle \mathbf{g}_k, \mathbf{x}\rangle \leq h_k$ is tight at none of the points of T_1 other than $\mathbf{y}_1$. Thus we conclude that the constraint of P which is exclusively tight at $\mathbf{y}_1$ is not linearly dependent on the other constraints which are tight at all points in T_1.

Thus we see that the number of tight linearly independent constraints at $\mathbf{y}_1$ is more than that at $\mathbf{z}$. Hence $rank_P(\mathbf{y}_1) > rank_P(\mathbf{z})$.

(4) can proved in the same way as *(3)* was proved.

Hence the Lemma. □

Corollary 4.2.3.

Let P be a linear program in canonical form such that $F(P)$ is bounded. Let $\mathbf{z}$ be any strict convex combination of two given vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Let $S_1, S_2 \subseteq \mathbb{R}$ defined by $S_1 = \{\alpha \mid \alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\}$ and $S_2 = \{\alpha \mid \alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\}$. Let $\alpha_1 = \sup(S_1)$ and $\alpha_2 = \sup(S_2)$. Then

1. $0 < \alpha_1 < \infty$ *and* $0 < \alpha_2 < \infty$.
2. $\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v}) \in F(P)$ *and* $\mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u}) \in F(P)$.
3. $rank_P(\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v})) > rank_P(\mathbf{z})$ *and* $rank_P(\mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{v})) > rank_P(\mathbf{z})$

Corollary 4.2.4.

Let P be a linear program in canonical form such that $F(P)$ is unbounded. Let $\mathbf{z}$ be any strict convex combination of two given vectors $\mathbf{u}$ and $\mathbf{v}$ in $F(P)$. Let $S_1, S_2 \subseteq \mathbb{R}$ defined by $S_1 = \{\alpha \mid \alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\}$ and $S_2 = \{\alpha \mid \alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\}$. Let $\alpha_1 = \sup(S_1)$ and $\alpha_2 = \sup(S_2)$. Then

1. *at most one of α_1 and α_2 is infinity.*

2. *if* $\alpha_1 = \infty$*, then* $\mathbf{u} - \mathbf{v}$ *is a recession direction of* $F(P)$*.*
3. *if* $\alpha_2 = \infty$*, then* $\mathbf{v} - \mathbf{u}$ *is a recession direction of* $F(P)$*.*

Example 4.2.2.

1. In Figure 4.2, we see that the feasible region $F(P)$ is bounded. The vectors $\mathbf{y}$ and $\mathbf{y}'$ marked in the figure are affine combinations of $\mathbf{u}$ and $\mathbf{v}$ and the vectors $\mathbf{z}$ is a strict convex combination of $\mathbf{u}$ and $\mathbf{v}$. We see that $rank_P(\mathbf{z}) = 1$ and $rank_P(\mathbf{y}) = rank_P(\mathbf{y}') = 2$.
2. In Figure 4.3, we see that the feasible region $F(P)$ is unbounded. Here we see that $\alpha_1 = \infty$ and $\mathbf{u} - \mathbf{v}$ is a recession direction of $F(P)$.

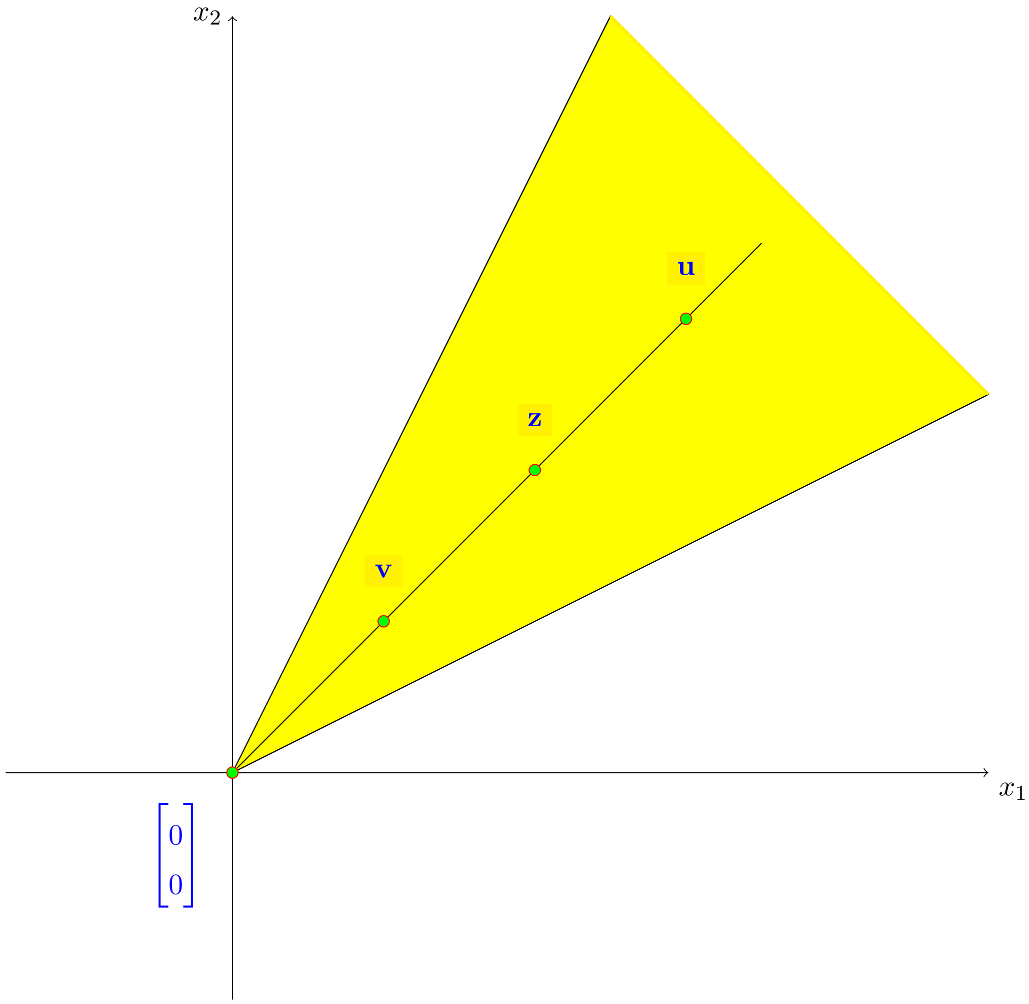


Figure 4.3: Ilustration of observation(2) in Example 4.2.2.

Theorem 4.2.2.

For any feasible linear Program P *in canonical form,* $F(P)$ *has finite non-zero number of extreme points [7].*

Proof.

It follows from Theorem 4.2.1 that the set of extreme points of $F(P)$ and the set of basic feasible solutions of P are the same and hence the number of extreme points of $F(P)$ is same as the number of basic feasible solutions of P which is at most $\binom{m+n}{n}$. Thus we see that $F(P)$ has finite number of extreme points.

The no trivial part is to prove that $F(P)$ has at least one extreme point. Let $\mathbf{z} \in F(P)$. Clearly $rank_P(\mathbf{z}) \leq n$. If $rank_P(\mathbf{z}) = n$, then $\mathbf{z}$ will be an extreme point of $F(P)$ and the proof is finished. Therefore we shall consider the case where $rank_P(\mathbf{z}) < n$. Therefore $\mathbf{z}$ is not a basic feasible solution of $F(P)$ and hence by Theorem 4.2.1, it is not an extreme point of $F(P)$. Hence it is clear that there exist two points $\mathbf{u}$ and $\mathbf{v}$ such that $\mathbf{z}$ is a strict convex combination of $\mathbf{u}$ and $\mathbf{v}$. That is $\mathbf{z} = \lambda\mathbf{u} + (1-\lambda)\mathbf{v}$ for some $\lambda \in (0,1)$.

Let $S = \{\alpha \mid \alpha \geq 0, \mathbf{z} + \alpha(\mathbf{u}-\mathbf{v}) \in F(P) \text{ and } \mathbf{z} + \alpha(\mathbf{v}-\mathbf{u}) \in F(P)\}$. Let $\alpha_1 = \max(S)$. By Lemma 4.2.4, we see that $0 < \alpha_1 < \infty$ and rank of at least one of the two feasible solutions $\mathbf{y} + \alpha_1(\mathbf{u}-\mathbf{v})$ and $\mathbf{y} - \alpha_1(\mathbf{u}-\mathbf{v})$ must be strictly greater than $rank_P(\mathbf{y})$. Without loss of generality we may assume that $rank_P(\mathbf{y} + \alpha_1(\mathbf{u}-\mathbf{v})) > rank_P(\mathbf{y})$. Now applying the strategy mentioned above on $\mathbf{y} + \alpha_1(\mathbf{u}-\mathbf{v})$, we will obtain another feasible solution whose rank with respect to P is strictly greater than $rank_P(\mathbf{y} + \alpha_1(\mathbf{u}-\mathbf{v}))$. Continuing like this, we will get a point in $F(P)$ whose rank with respect to P is n which will be a basic feasible solution of P and hence an extreme point of $F(P)$ by Theorem 4.2.1. Thus we see that $F(P)$ must have at least one extreme point. □

Now we are going to prove that a bounded feasible region can be represented as the convex hull of the set of extreme points of the feasible region using the ideas developed so far in this section.

Theorem 4.2.3.
Let P a linear program such that $F(P)$ is bounded. Then $F(P)$ is the convex hull of the set of extreme points of $F(P)$.

Proof.
Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$. By Lemma 4.2.2, $1 \leq k < \infty$. We need to show that $F(P) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. We prove the Theorem by showing that $F(P)$ and $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$ are subsets of each other.

Claim. $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \subseteq F(P)$:-

Let $\mathbf{x}$ be an arbitrary element of $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. Therefore $\mathbf{x}$ can be expressed as $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ such that $\lambda_i \in [0,1]$ for each $i \in \{1,2,3,\cdots,k\}$ and $\sum_{i=1}^{k} \lambda_i = 1$.

Now

$$\mathbf{G}\mathbf{x} = \mathbf{G}\left(\sum_{i=1}^{k} \lambda_i \mathbf{x}_i\right) = \sum_{i=1}^{k} \lambda_i \mathbf{G}\mathbf{x}_i \leq \sum_{i=1}^{k} \lambda_i \mathbf{h}$$

which implies that $\mathbf{G}\mathbf{x} \leq \mathbf{h}$ since $\sum_{i=1}^{k} \lambda_i = 1$. Hence $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \subseteq F(P)$.

Claim. $F(P) \subseteq Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$:-

In this case, we prove that whenever $\mathbf{z} \in F(P)$, $\mathbf{z} \in Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. We prove this result by applying reverse induction on $rank_P(\mathbf{z})$.

Basis :

If $rank_P(\mathbf{z}) = n$, then $\mathbf{z}$ must be an extreme point of $F(P)$. Hence $\mathbf{z} \in \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}$ and the result follows immediately.

Inductive Hypothesis:-

Assume that each $\mathbf{z} \in F(P)$ satisfying $n \geq rank_P(\mathbf{x}) \geq r$ is contained in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$ for some positive integer r.

Induction:-

Let $rank_P(\mathbf{z}) = r - 1$. Clearly $\mathbf{y}$ is not an extreme point of $F(P)$. Hence there exist vectors $\mathbf{u}, \mathbf{v}$ in $F(P)$ such that $\mathbf{z} = \lambda\mathbf{u} + (1-\lambda)\mathbf{v}$ for some $\lambda \in (0,1)$. Let $\alpha_1 = \sup\{\alpha \mid$

$\alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\}$ and $\alpha_2 = \sup\{\alpha \mid \alpha \geq 0$ and $\mathbf{z} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\}$. By Lemma 4.2.4 and Corollary 4.2.3, we see that $0 < \alpha_1 < \infty, 0 < \alpha_2 < \infty$ and ranks of both the feasible solutions $\mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v}), \mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u})$ with respect to P are strictly greater than $rank_P(\mathbf{z})$. That is there exist two feasible solutions $\mathbf{y}_1 = \mathbf{z} + \alpha_1(\mathbf{u} - \mathbf{v})$ and $\mathbf{y}_2 = \mathbf{z} + \alpha_2(\mathbf{v} - \mathbf{u})$ such that $rank_P(\mathbf{y}_1) \geq r$ and $rank_P(\mathbf{y}_2) \geq r$. By inductive Hypthesis $\mathbf{y}_1$ and $\mathbf{y}_2$ are contained in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$.

It is easy to see that $\mathbf{z} = \frac{\alpha_2}{\alpha_1+\alpha_2}\mathbf{y}_1 + \frac{\alpha_1}{\alpha_1+\alpha_2}\mathbf{y}_2$ which is a convex combination of two vectors in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. Hence $\mathbf{z}$ is also contained in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$.

Thus by induction we see that each $\mathbf{z} \in F(P)$ is contained in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. Since any element of $F(P)$ may be chosen as $\mathbf{z}$, we further conclude that all vectors in $F(P)$ are contained in $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. Hence $F(P) \subseteq Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$.

Thus it follows that $F(P) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$ by the two claims. Hence the Theorem. □

Theorem 4.2.4.

Given a linear program P in canonical form such that $F(P)$ is a polytope. Then P has an optimal extreme point solution.

Proof.

Since the objective function of P is a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}$, we see that the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ is continuous (Refer to Fact B.2.1 in Appendix B). Furthermore $F(P)$ is a polytope and hence $F(P)$ is closed and bounded. Thus we see that $F(P)$ is a compact subset of $\mathbb{R}^n$ by Heine Borel Theorem . Hence $Image(P)$ is a compact subset of $\mathbb{R}$ and must have a finite maximum (Refer to Fact C.2.4 and Fact C.2.2 in Appendix C). Thus it follows that P has an optimal solution.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$. By Theorem 4.2.2, we see that $1 \leq k < \infty$. Since $F(P)$ is a polytope, we see that $F(P) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. Let t be the index satisfying the conditions $1 \leq t \leq k$ and $\langle \mathbf{c}, \mathbf{x}_t \rangle = \max(\{\langle \mathbf{c}, \mathbf{x}_i \rangle \mid 1 \leq i \leq k\})$.

Let $\mathbf{x}^*$ be the optimal solution of P. Therefore there exist non-negative real numbers $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_k$ such that $\mathbf{x}^* = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i$ and $\sum_{i=1}^{k} \lambda_i = 1$. Now we have

$$\langle \mathbf{c}, \mathbf{x}^* \rangle = \left\langle \mathbf{c}, \sum_{i=1}^{k} \lambda_i \mathbf{x}_i \right\rangle = \sum_{i=1}^{k} \lambda_i \langle \mathbf{c}, \mathbf{x}_i \rangle \leq \langle \mathbf{c}, \mathbf{x}_t \rangle \sum_{i=1}^{k} \lambda_i$$

Hence $\mathbf{x}_t$ must be optimal. Hence the Theorem. □

The following corollary is an immediate consequence of Theorem 4.2.4.

Corollary 4.2.5.

Let P be a feasible linear program in canonical form. If P is unbounded, then $F(P)$ is unbounded.

To summarise what discussed so far, we see that for any linear program having a bounded feasible region, there will be an extreme point optimal solution. In the forthcoming sections, we try to extend this result to linear programs with unbounded feasible region.

4.3 Recession Directions and Extreme Directions of Feasible Region

In this section we introduce the notion of extreme directions and its characterisation using extreme points [7].

Lemma 4.3.1.
Given a linear Program P in canonical form and $\mathbf{d} \in \mathbb{R}^n$. $\mathbf{d}$ is a recession direction of $F(P)$ if and only if $\mathbf{Gd} \leq \mathbf{O_{m+n}}, \mathbf{d} \neq \mathbf{0}_n$.

Proof.
if part:-
Let $\mathbf{d}$ be a recession direction of $F(P)$. This implies that for every $\mathbf{x}$ in $F(P)$, the half-line $\{x + \alpha d \mid \alpha \geq 0\} \subseteq F(P)$ which requires $\mathbf{G}(\mathbf{x} + \alpha\mathbf{d}) = \mathbf{Gx} + \alpha\mathbf{Gd} \leq \mathbf{h}$ for all $\alpha \geq 0$. This is posible only if $\mathbf{Gd} \leq \mathbf{O_{m+n}}$. Note that $\mathbf{d} \neq \mathbf{O_n}$ by the definition of recession direction.
only- if part:-
Let $\mathbf{Gd} \leq \mathbf{O_{m+n}}$ and $\mathbf{d} \neq \mathbf{0}_n$. Therefore we see that for every $\mathbf{x}$ in $F(P)$, it is the case that $\mathbf{G}(\mathbf{x} + \alpha\mathbf{d}) = \mathbf{Gx} + \alpha\mathbf{Gd} \leq \mathbf{h}$ for each $\alpha \geq 0$. This further implies that the half-line $\{\mathbf{x} + \alpha\mathbf{d} \mid \alpha \geq 0\} \subseteq C$. Therefore $\mathbf{d}$ is a recession direction of $F(P)$.
Hence the Lemma. □

Corollary 4.3.1.
Given linear Program P in canonical form. The recession cone of the feasible region $F(P)$ is given by $R_C(F(P)) = \{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n$ and $\mathbf{Gd} \leq \mathbf{0}_{m+n}\}$.

Remark 4.3.1.
Since $\mathbf{G} = \begin{bmatrix} \mathbf{A} \\ -\mathbf{I}_n \end{bmatrix}$, the constraint $\mathbf{Gd} \leq \mathbf{0}_{m+n}$ implies that $\mathbf{d} \geq \mathbf{0}_n$.

It is easy to see that if $\mathbf{d}$ is a recession direction, so will be $\alpha\mathbf{d}$ for any $\alpha > 0$. To have a unique representative of all such collinear recession directions, we bring about the notion of *Normalised Recession Direction.*

Definition 4.3.1.
Given linear Program P in canonical form such that $F(P)$ is unbounded. A recession direction of $F(P)$ is said to be *normalised recession direction* if $\langle \mathbf{1}_n, \mathbf{d} \rangle = 1$. The set of all normalised recession directions of $F(P)$ is called the *normalised recession direction set* of $F(P)$ is denoted by $R_N(F(P))$. That is

$$R_N(F(P)) = \{\mathbf{d} \mid \mathbf{d} \in R_C(F(P)) \text{ and } \langle \mathbf{1}_n, \mathbf{d} \rangle = 1\}$$

Remark 4.3.2.
For any linear program P in canonical form having an unbounded feasible region, it is easy to see that $R_N(F(P)) \subseteq R_C(F(P))$. Moreover every recession direction of $F(P)$ lies in the first orthant of $\mathbb{R}^n$ by Remark 4.3.1. This implies that the l_1 norm of any recession direction in $R_N(F(P))$ is unity.

Lemma 4.3.2.
Given linear Program P in canonical form such that $F(P)$ is unbounded. $R_N(F(P))$ is a convex compact subset of $R_C(F(P))$.

Proof.
$R_N(F(P))$ is a subset of $R_C(F(P))$ by definition.
We see that

$$\begin{aligned} R_N(F(P)) &= \{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n, \mathbf{Gd} \leq \mathbf{0}_{m+n} \text{ and } \langle \mathbf{1}_n, \mathbf{d} \rangle = 1\} \\ &= \{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n, \mathbf{Gd} \leq \mathbf{0}_{m+n}\} \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{1}_n, \mathbf{x} \rangle = 1\} \end{aligned}$$

Therefore $R_N(F(P))$ is a polyhedral set and is thus closed.
By Remark 4.3.2, each vector $\mathbf{d}$ in $R_N(F(P))$ has unit l_1 norm. Thefore there exists some finite positive real number t such that the l_2 norm of each vector $\mathbf{d}$ in $R_N(F(P))$ is at most t

by equivalence of norms in $\mathbb{R}^n$. This further implies that $R_N(F(P)) \subset B_{t+1}(\mathbf{0}_n)$. Thus we see that $R_N(F(P))$ is bounded. Since $R_N(F(P))$ is closed and bounded, it is compact by Heine Borel Theorem .

It is easy to see that the polyhedral set $\{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n, \mathbf{G}\mathbf{d} \leq \mathbf{0}_{m+n}\}$ is convex and the set $\{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{1}_n, \mathbf{x} \rangle = 1\}$ is convex. Since the intersection of two convex sets are convex, it follows that $R_N F(P)$ is also convex by Lemma 2.2.1. □

Lemma 4.3.3.

Given linear Program P in canonical form such that $F(P)$ is unbounded. Then $R_C(F(P)) = Cone\,(R_N(F(P)))$.

Proof.

We prove the Lemma by proving that each of $Cone\,(R_N(F(P)))$ and $R_C(F(P))$ is contained in the other.

By definition, $R_N(F(P))$ is a subset of $R_C(F(P))$ and therefore the vectors in $R_N(F(P))$ are recession directions of $F(P)$. Therefore any conic combination of the vectors in $R_N(F(P))$ is also a recession direction of $F(P)$ by Lemma 2.2.2 and Lemma 2.2.9 and hence a member of $R_C(F(P))$. This further implies that each element of $Cone\,(R_N(F(P)))$ is also a member of $R_C(F(P))$ and hence $Cone\,(R_N(F(P))) \subseteq R_C(F(P))$.

Let $\mathbf{d}$ be any recession direction in $R_C(F(P))$. By Lemma 2.2.8, there exists recession direction $\mathbf{d}' = \frac{1}{\langle \mathbf{1}_n, \mathbf{d} \rangle}\mathbf{d}$ of $F(P)$. Clearly $\mathbf{d}'$ is a member of $R_N(F(P))$ since $\langle \mathbf{1}_n, \mathbf{d}' \rangle = 1$. This further impies that $\mathbf{d} = \langle \mathbf{1}_n, \mathbf{d} \rangle \mathbf{d}'$ is a member of $Cone\,(R_N(F(P)))$. Since the choice of $\mathbf{d}$ is arbitrary within $R_C(F(P))$, we conclude that every recession direction in $R_C(F(P))$ is also a member of $Cone\,(R_N(F(P)))$ and hence $R_C(F(P)) \subseteq Cone\,(R_N(F(P)))$.

Hence the Lemma. □

Now we introduce the notions of extreme directions and normalised extreme directions.

Definition 4.3.2.

Given linear Program P in canonical form such that $F(P)$ is unbounded. Then any recession direction $\mathbf{d}$ of $F(P)$ is said to be an *extreme direction* of C if and only if there do not exist two normalised recession directions $\mathbf{d}_1$ and $\mathbf{d}_2$ of $F(P)$ such that $\mathbf{d} = \lambda_1 \mathbf{d}_1 + \lambda_2 \mathbf{d}_2$ for some $\lambda_1, \lambda_2 > 0$. That is an extreme direction of $F(P)$ cannot be expressed as a strict conic combination of two normalised recession directions of $F(P)$.

Definition 4.3.3.

Given linear Program P in canonical form such that $F(P)$ is unbounded. An extreme direction of $F(P)$ is said to be *normalised extreme direction* if $\langle \mathbf{1}_n, \mathbf{d} \rangle = 1$.

The following lemma characterises the normalised extreme directions of an unbounded feasible region as the extreme points of the normalised recession direction set.

Lemma 4.3.4.

Let P be a linear program in canonical form such that $F(P)$ is unbounded. Let $\mathbf{d} \in R_N(F(P))$. Then $\mathbf{d}$ is a normalised extreme direction of $F(P)$ if and only if $\mathbf{d}$ is an extreme point of $R_N(F(P))$.

Proof.

if Part:

Let $\mathbf{d}$ be an extreme point of $R_N(F(P))$. Assume that $\mathbf{d}$ is not a normalised extreme direction of $F(P)$. Therefore there exist normalised recession directions $\mathbf{d}_1$ and $\mathbf{d}_2$ of $F(P)$ such that $\mathbf{d} = \gamma_1 \mathbf{d}_1 + \gamma_2 \mathbf{d}_2$ where $\gamma_1, \gamma_2 > 0$.

Now

$$1 = \langle \mathbf{1_n}, \mathbf{d} \rangle = \langle \mathbf{1_n}, \gamma_1 \mathbf{d}_1 + \gamma_2 \mathbf{d}_2 \rangle = \gamma_1 \langle \mathbf{1}_n, \mathbf{d}_1 \rangle + \gamma_2 \langle \mathbf{1}_n, \mathbf{d}_2 \rangle = \gamma_1 + \gamma_2$$

$(\because \mathbf{d}_1 \text{ and } \mathbf{d}_2 \text{ are normalised extreme directions and hence } \langle \mathbf{1}_n, \mathbf{d}_1 \rangle = \langle \mathbf{1}_n, \mathbf{d}_2 \rangle = 1)$

Thus we see that $\mathbf{d} = \gamma_1 \mathbf{d}_1{}' + (1-\gamma_1)\mathbf{d}_2{}'$ where $\lambda_1 \in (0,1)$. That is $\mathbf{d}$ can be expressed as a strict convex combination of two normalised recession directions $\mathbf{d}_1, \mathbf{d}_2{}' \in R_N(F(P))$. This contradicts the fact that $\mathbf{d}$ is an extreme point of $R_N(F(P))$. Thus it follows that $\mathbf{d}$ is a normalised extreme direction of $F(P)$.

only if part:-

Let $\mathbf{d}$ be a non-extreme point of $R_N(F(P))$. This implies that there exist normalised recession directions $\mathbf{d}_1$ and $\mathbf{d}_2$ in $R_N(F(P))$ such that $\mathbf{d} = \lambda \mathbf{d}_1 + (1-\lambda)\mathbf{d}_2$ for some $\lambda \in (0,1)$. Thus we see that $\mathbf{d}$ is a strict conic combination of two normalised recession directions $\mathbf{d}_1$ and $\mathbf{d}_2$ of $F(P)$. Hence $\mathbf{d}$ is not a normalised extreme direction of $F(P)$.

Hence the Lemma. □

Example 4.3.1.

Let P be the linear program

$$\begin{aligned} \text{Maximize } & x_1 + x_2 \text{ subject to} \\ & x_1 - 2x_2 \le 2 \\ & -x_1 + x_2 \le 3 \\ & x_1 \ge 0 \\ & x_2 \ge 0 \end{aligned} \tag{4.4}$$

The feasible region $F(P)$ is shown Figure 4.4(a). The Recession cone of $F(P)$ is given by

$$R_C(F(P)) = \{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n, \mathbf{G}\mathbf{d} \le \mathbf{0}_{m+n} \text{ and } \langle \mathbf{1}_n, \mathbf{d}\rangle = 1\} \text{ where } \mathbf{G} = \begin{bmatrix} 1 & -2 \\ -1 & 1 \\ -1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The recession cone and the normalised recession direction set of $F(P)$ is shown in Figure 4.4(b). By Lemma 4.3.4, we see that the vectors marked $\mathbf{d}_1$ and $\mathbf{d}_2$ in Figure 4.4(b) are normalised extreme directions of $F(P)$.

Theorem 4.3.1.

Given linear Program P in canonical form. Then $F(P)$ has finite number of normalised extreme directions.

Proof.

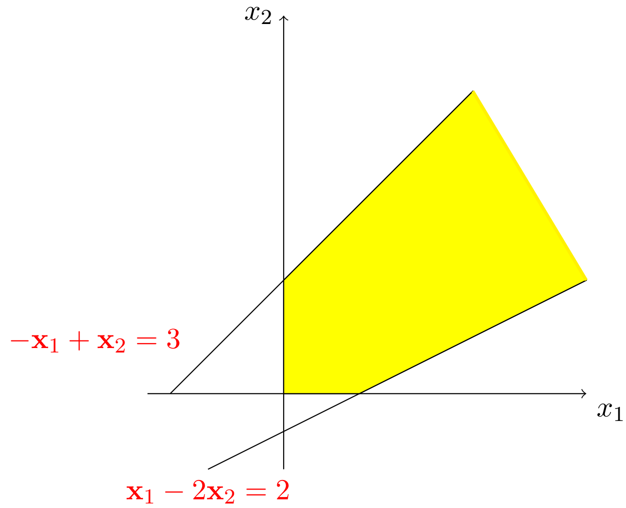
If $F(P)$ is a polytope, there are no extreme directions by corollary 2.2.2. Hence we see that if $F(P)$ is bounded, there are no normalised extreme directions. In case $F(P)$ is unbounded, by Lemma 4.3.4 we see that there are as many normalised extreme directions as the extreme points of $R_N(F(P))$. Since $R_N(F(P)) = \{\mathbf{d} \mid \mathbf{d} \in \mathbb{R}^n, \mathbf{G}\mathbf{d} \le \mathbf{0}_{m+n}, \langle \mathbf{1}_n, \mathbf{d}\rangle \le 1 \text{ and } \langle -\mathbf{1}_n, \mathbf{d}\rangle \le -1\}$, the number of extreme points of $R_N(F(P))$ is at most $\binom{m+n+2}{n}$ by Theorem4.2.2. Thus it follows that the number of normalised extreme directions of $F(P)$ is at most $\binom{m+n+2}{n}$.

Hence the result. □

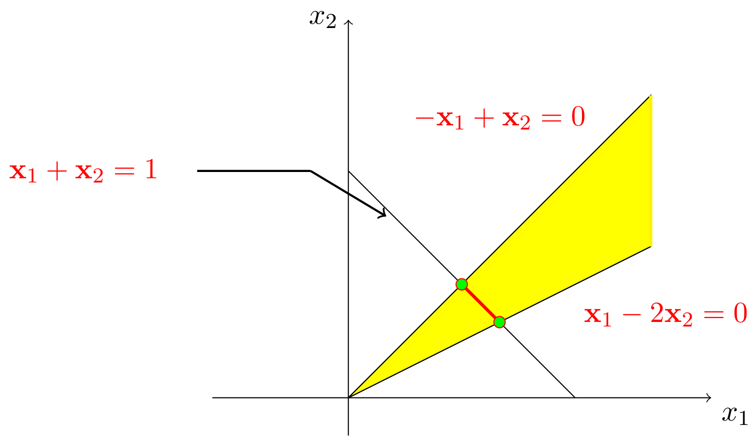
The following Theorem characterises the recession cone of an unbounded feasible region as the conic hull of the set of normalised extreme directions of the feasible region. This has a major role in proving the Caratheodory characterisation theorem.

Theorem 4.3.2.

For any linear program in canonical form with an unbounded feasible region, the recession cone of the feasible region is the conic hull of the set of normalised extreme directions of the feasible region.



(a) The feasible region $F(P)$ (yellow shaded region) in Example 4.3.1.



(b) The recession cone of $F(P)$ in Example 4.3.1 is the yellow shaded region. The normalised recession direction set of $F(P)$ is the thick red line segment and its extreme points are the normalised extreme directions of $F(P)$.

Figure 4.4: Illustration of Lemma 4.3.4. The feasible region, recession cone, normalised recession direction set and the normalised extreme directions of the linear program in Example 4.3.1.

Proof.

Let P be any linear program in canonical form as given by LP 4.1. By Lemma 4.3.2, we see that $R_N(F(P))$ is a polytope and is the feasible region of the linear program P' given by

$$\begin{aligned} \text{Maximize} \quad & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}\mathbf{x} \leq \mathbf{0}_m \\ & \langle \mathbf{1}_n, \mathbf{x} \rangle \leq 1 \\ & \langle -\mathbf{1}_n, \mathbf{x} \rangle \leq -1 \\ & \mathbf{x} \geq \mathbf{0}_n \end{aligned}$$

Therefore we have

$$\begin{aligned} R_C(F(P)) &= Cone\,(R_N(F(P))) && \text{(By Lemma 4.3.4)} \\ &= Cone\,(F(P')) \end{aligned}$$

$$\begin{aligned}
&= Cone\,(Conv\,(\{\mathbf{x} \mid \mathbf{x} \text{ is an extreme point of } F(P')\})) && \text{(By Theorem 4.2.3)}\\
&= Cone\,(\{\mathbf{x} \mid \mathbf{x} \text{ is an extreme point of } R_N(F(P))\}) && \text{(By Lemma 2.2.5)}\\
&= Cone\,(\{\mathbf{x} \mid \mathbf{x} \text{ is a normalised extreme direction of } F(P)\}) && \text{(By Lemma 4.3.4)}
\end{aligned}$$

Hence the result. □

4.4 Fundamental Theorem of LP

In this section we present the fundamental theorem of linear programming which forms the foundation for the graphical method of solving linear programs. In view of proving the fundamental theorem of LP, we prove the Caratheodory characterisation theorem . We start with a lemma that handles most of the technicalities of Caratheodory characterisation theorem.

Lemma 4.4.1.
Let P be a linear program in canonical form such that $F(P)$ is unbounded. Let E_P be the set of extreme points of $F(P)$ and $\gamma \in \mathbb{R}$ such that $\gamma > \max(\{\langle \mathbf{1}_n, \mathbf{x}\rangle \mid \mathbf{x} \in E_P\})$. Let P_γ be a linear program such that $F(P_\gamma) = F(P) \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$ and $\langle \mathbf{1}_n, \mathbf{x}\rangle \le \gamma\}$ and E_{P_γ} be the set of extreme points of $F(P_\gamma)$. Then

(1) $F(P_\gamma)$ is a polytope such that $F(P_\gamma) \subseteq F(P)$ and $E_P \subseteq E_{P_\gamma}$.

(2) For each $\mathbf{w} \in E_{P_\gamma} \setminus E_P$, there exists some $\mathbf{y} \in E_P$ such that $\mathbf{w} = \mathbf{y} + \mathbf{d}$ for some $\mathbf{d} \in R_C(F(P))$. That is, each extreme point of P_γ which is not extreme point of P is the sum of some extreme point of P and some recession direction of P.

Proof.
(1):- The feasible region of P_γ is given by $F(P_\gamma) = F(P) \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$ and $\langle \mathbf{1}_n, \mathbf{x}\rangle < \gamma\}$. Therefore $F(P_\gamma) \subseteq F(P)$. It is obvious that the l_1 norm of each $\mathbf{x}$ in $F(P_\gamma)$ is at most γ which ensures that $F(P_\gamma)$ is bounded. Hence $F(P_\gamma)$ is a polytope.

Since $\gamma > \max(\{\langle \mathbf{1}_n, \mathbf{x}\rangle \mid \mathbf{x} \in E_P\})$ and $F(P_\gamma) = F(P) \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n$ and $\langle \mathbf{1}_n, \mathbf{x}\rangle \le \gamma\}$, we see that each extreme point of $F(P)$ is an extreme point of $F(P_\gamma)$. Hence $E_P \subseteq E_{P_\gamma}$.

(3):- Let $\mathbf{w}$ be any element of $E_{P_\gamma} \setminus E_P$. Since $\mathbf{w}$ is an extreme point of P_γ and not an extreme point of P, we further see that the constraint $\langle \mathbf{1}_n, \mathbf{x}\rangle \le \gamma$ must be tight at $\mathbf{w}$. Thus we see that $rank_P(\mathbf{w}) = n - 1$.

Let $\alpha_1 = \sup\{\alpha \mid \alpha \ge 0$ and $\mathbf{w} + \alpha(\mathbf{u} - \mathbf{v}) \in F(P)\}, \alpha_2 = \sup\{\alpha \mid \alpha \ge 0$ and $\mathbf{w} + \alpha(\mathbf{v} - \mathbf{u}) \in F(P)\}$. By Lemma 4.2.4, we have $\alpha_1, \alpha_2 > 0$. Now we are going to prove that exactly one of $\alpha_1$and α_2 is infinity. Assume that both α_1 and α_2 are finite. Hence we see that rank of each of the feasible points $\mathbf{y} = \mathbf{w} + \alpha_1(\mathbf{u} - \mathbf{v})$ and $\mathbf{z} = \mathbf{w} + \alpha_2(\mathbf{v} - \mathbf{u})$ is n by Lemma 4.2.4. This implies that both $\mathbf{y}$ and $\mathbf{z}$ are extreme points of $F(P)$. Since every extreme point of $F(P)$ is also an extreme point of $F(P_\gamma)$, we see that and $\mathbf{y}$ and $\mathbf{z}$ are extreme points of $F(P_\gamma)$ and hence members of $F(P_\gamma)$.

It is easy to see that $\mathbf{w} = \frac{\alpha_2}{\alpha_1+\alpha_2}\mathbf{y} + \frac{\alpha_1}{\alpha_1+\alpha_2}\mathbf{z}$ which is a strict convex combination of the two elements $\mathbf{y}$ and $\mathbf{z}$ in $F(P_\gamma)$ This imples that $\mathbf{w}$ is not an extreme point of $F(P_\gamma)$ and cannot be an element of E_{P_γ}, which is a contradiction. Thus we conclude that both α_1 and α_2 cannot be finite. Moreover Lemma 4.2.4 implies that at least one of α_1 and α_2 is finite. Thus it follows that exactly one of α_1 and α_2 is finite.

Without loss of generality, we may take $\alpha_1 < \infty$. Therefore the point $\mathbf{y}$ is an extreme point of $F(P)$. Since $\alpha_2 = \infty$, we infer that $\mathbf{v} - \mathbf{u}$ is a recession direction of $F(P)$ and must be present in $Rc(F(P)$. Since $\mathbf{w} - \mathbf{y} = \alpha_1(\mathbf{v} - \mathbf{u})$, we also see that $\mathbf{w} - \mathbf{y} \in Rc(F(P)$. Thus we see that $\mathbf{w} - \mathbf{y}$ is a recession direction of $F(P)$. In other words, $\mathbf{w} = \mathbf{y} + \mathbf{d}$ for some $\mathbf{y} \in E_P$ and $\mathbf{d} \in R_C(F(P))$. Since $\mathbf{w}$ can be any element of $E_{P_\gamma} \setminus E_P$, we conclude that corresponding to each $\mathbf{w} \in E_{P_\gamma} \setminus E_P$, there exists some $\mathbf{y} \in E_P$ such that $\mathbf{w} = \mathbf{y} + \mathbf{d}$ for some $\mathbf{d} \in R_C(F(P))$ □

Example 4.4.1.

Let P be the linear program

$$\begin{aligned} \text{Maximize } & x_1 + x_2 \text{ subject to} \\ & x_1 - 2x_2 \leq 2 \\ & -4x_1 - x_2 \leq -8 \\ & -x_1 + x_2 \leq 3 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned} \tag{4.5}$$

The feasible region $F(P)$ is shown Figure 4.5(a).

Choosing $\gamma = 9$ and adding the new constraint $\mathbf{x}_1 + \mathbf{x}_2 \leq 9$ we get the linear program P_γ given below.

$$\begin{aligned} \text{Maximize } & x_1 + x_2 \text{ subject to} \\ & x_1 - 2x_2 \leq 2 \\ & -4x_1 - x_2 \leq -8 \\ & -x_1 + x_2 \leq 3 \\ & x_1 + x_2 \leq 9 \\ & x_1 \geq 0 \\ & x_2 \geq 0 \end{aligned} \tag{4.6}$$

The feasible region of $F(P_\gamma)$ is a polytope and is shown in Figure 4.5(b) with thick red edges. The extreme points of $F(P_\gamma)$ are $\mathbf{y}_1, \mathbf{y}_2, \mathbf{w}_1$ and $\mathbf{w}_2$. We see that both $\mathbf{w}_1 - \mathbf{y}_1$ and $\mathbf{w}_2 - \mathbf{y}_2$ are recession directions of $F(P)$.

Now we define an important term called minkowski sum of two sets as follows.

Definition 4.4.1.

Let $S, T \subseteq \mathbb{R}^n$. Then the minkowski sum of S and T is denoted by $S \oplus T$ and is defined as $S \oplus T = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in S, \mathbf{y} \in T\}$.

The following lemma establishes the crux of Caratheodory characterisation theorem.

Lemma 4.4.2.

Let P be a linear program in canonical form such that $F(P)$ is unbounded. Then $F(P)$ is the minkowski sum of the set of extreme points of $F(P)$ and the recession cone of $F(P)$.

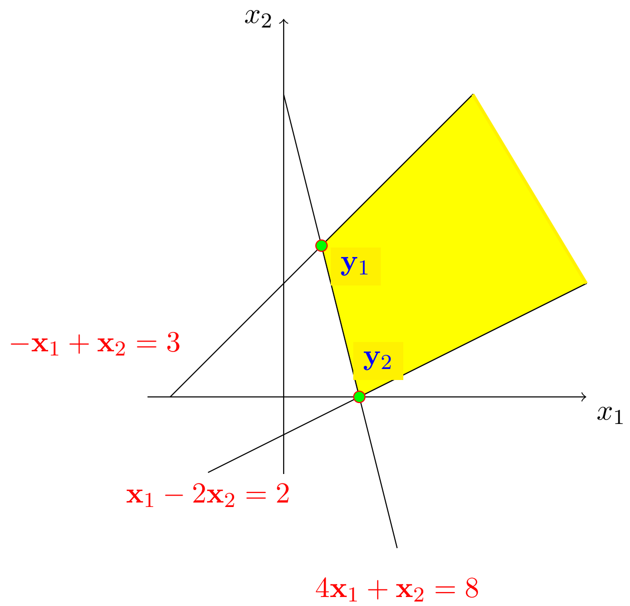
Proof.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$. We are going to prove that $F(P) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$. We prove the lemma by proving that LHS and RHS are subsets of each other.

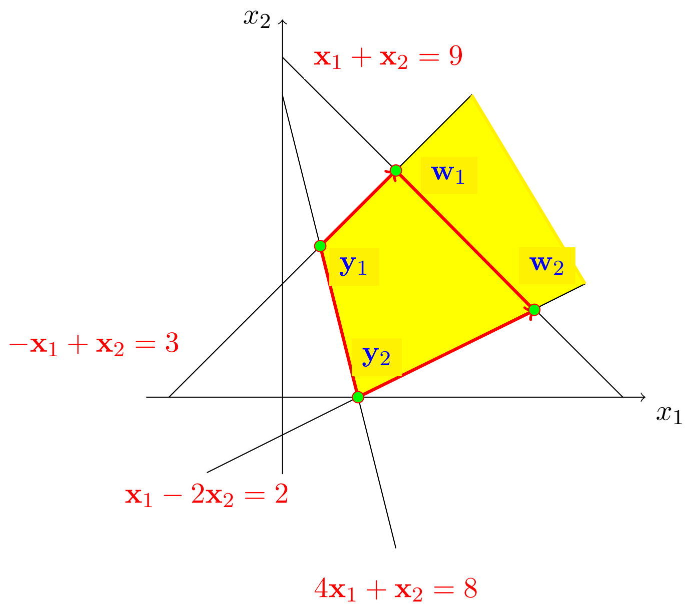
Claim. *$Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P)) \subseteq F(P)$.*

Proof.

Let $\mathbf{x}$ be any element of $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$. Therefore $\mathbf{x}$ can be expressed as $\mathbf{x} = \sum_{i=1}^{k} \beta_i \mathbf{x}_i + \mathbf{d}$ for some recession direction $\mathbf{d}$ of $F(P)$ and non-negative real numbers $\beta_1, \beta_2, \beta_3, \cdots, \beta_k$ such that $\sum_{i=1}^{k} \beta_i = 1$. Since each of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ are extreme points of $F(P)$, we see that $\mathbf{G}\mathbf{x}_i \leq \mathbf{h}$ for each $i \in \{1, 2, 3, \cdots, k\}$. Since $\mathbf{d}$ is a recession direction of $F(P)$, we must have $\mathbf{G}\mathbf{d} \leq \mathbf{0}_{m+n}$ by Lemma 4.3.1.



(a) The feasible region $F(P)$ (yellow shaded region) and the extreme points $\mathbf{y}_1$ and $\mathbf{y}_2$ in Example 4.4.1.



(b) The feasible region $F(P_\gamma)$ in Example 4.4.1. $F(P_\gamma)$ is a subset of $F(P)$ and is the bounded set shown with thick red edges. The extreme points of of $F(P_\gamma)$ are $\mathbf{y}_1, \mathbf{y}_2, \mathbf{w}_1$ and $\mathbf{w}_2$. It is easy to see that $\mathbf{w}_1 - \mathbf{y}_1$ and $\mathbf{w}_2 - \mathbf{y}_2$ are recession directions of $F(P)$.

Figure 4.5: Illustration of Lemma 4.4.1. The feasible regions $F(P)$ and $F(P_\gamma)$ and the respective extreme points in Example 4.4.1.

Now

$$\mathbf{Gx} \;=\; \mathbf{G}\left(\sum_{i=1}^{k}\beta_i\mathbf{x}_i + \mathbf{d}\right) = \sum_{i=1}^{k}\beta_i\mathbf{G}\mathbf{x}_i + \mathbf{Gd} \leq \mathbf{h}$$

Thus we see that $\mathbf{x}$ is an element of $F(P)$. Since the choice of $\mathbf{x}$ within the minkowski sum $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$ is arbitrary, the claim follows. □

Claim. *$Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P)) \subseteq F(P)$.*

Proof.
Let $\mathbf{y}$ be an element of $F(P)$. Let $\gamma \in \mathbb{R}$ such that

$$\gamma > \max\left(\{\langle \mathbf{1}_n, \mathbf{y}\rangle, \langle \mathbf{1}_n, \mathbf{x}_1\rangle, \langle \mathbf{1}_n, \mathbf{x}_2\rangle, \langle \mathbf{1}_n, \mathbf{x}_3\rangle, \cdots, \langle \mathbf{1_n}, \mathbf{x}_k\rangle\}\right)$$

Consider the linear program P_γ such that $F(P_\gamma) = F(P) \cap \{\mathbf{x} \mid \mathbf{x} \in \mathbb{R}^n \text{ and } \langle \mathbf{1}_n, \mathbf{x}\rangle \leq \gamma\}$. By Lemma 4.4.1 each extreme point of $F(P)$ is an extreme point of $F(P_\gamma)$. Therefore we may take the set of extreme points of $F(P_\gamma)$ as $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\} \cup \{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \cdots, \mathbf{w}_l\}$. By Lemma 4.2.2, we see that $1 \leq k + l < \infty$.

By Lemma 4.4.1, we see that $F(P_\gamma)$ is bounded. Hence the feasible region of P_γ can be expressed as $F(P_\gamma) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \cdots, \mathbf{w}_l\})$ by Theorem 4.2.3. It is easy to see that $\mathbf{y} \in F(P_\gamma)$ and hence $\mathbf{y} \in Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \cdots, \mathbf{w}_l\})$.

Since $\mathbf{y} \in Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k, \mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \cdots, \mathbf{w}_l\})$, we see that there exist non-negative real numbers $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_k$ and $\gamma_1, \gamma_2, \gamma_3, \cdots, \gamma_l$ such that $\mathbf{y} = \sum_{i=1}^{k}\lambda_i\mathbf{x}_i + \sum_{i=1}^{l}\gamma_i\mathbf{w}_i$ and $\sum_{i=1}^{k}\lambda_i + \sum_{i=1}^{l}\gamma_i = 1$. Furthermore by Lemma 4.4.1, $\mathbf{w}_i = \mathbf{y}_i + \mathbf{d}_i$ for some $\mathbf{y}_i \in \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}$ and $\mathbf{d}_i \in R_C(F(P))$ for each $i \in \{1, 2, 3, \cdots, l\}$. Hence we have

Now

$$\mathbf{y} \;=\; \sum_{i=1}^{k}\lambda_i\mathbf{x}_i + \sum_{i=1}^{l}\gamma_i(\mathbf{y}_i + \mathbf{d}_i) = \left(\sum_{i=1}^{k}\lambda_i\mathbf{x}_i + \sum_{i=1}^{l}\gamma_i\mathbf{y}_i\right) + \left(\sum_{i=1}^{l}\gamma_i\mathbf{d}_i\right)$$

It follows that $\sum_{i=1}^{k}\lambda_i\mathbf{x}_i + \sum_{i=1}^{l}\gamma_i\mathbf{y}_i$ is a convex combination of the extreme points of $F(P)$ and hence a member of $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\})$. It can also be seen that $\sum_{i=1}^{l}\gamma_i\mathbf{d}_i$ is a conic combination of some recession directions of $F(P)$ and hence a recession direction of $F(P)$ by Lemma 2.2.9. Therefore $\sum_{i=1}^{l}\gamma_i\mathbf{d}_i \in R_C(F(P))$. Thus it follows that $\mathbf{y}$ is an element of $Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$. Since the choice of $\mathbf{y}$ within $F(P)$ is arbitrary, we conclude that $F(P) \subseteq Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$. □

Thus we conclude that the feasible region $F(P) = Conv(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}) \oplus R_C(F(P))$ by the two claims. □

Now we are going to prove an important lemma which specifies the necessary and sufficient conditions for a linear program to be unbounded.

Lemma 4.4.3.
Let P be a feasible linear program in canonical form. Then P is unbounded if and only if there exists some $\mathbf{d} \in R_C(F(P))$ such that $\langle \mathbf{c}, \mathbf{d}\rangle > 0$.

Proof.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$.

if part:-

Let there exists some $\mathbf{d} \in R_C(F(P))$ such that $\langle \mathbf{c}, \mathbf{d} \rangle > 0$. Assume that P has a finite optimum and the corresponding optimal solution is $\mathbf{x}$. By Lemma 4.4.2, we see that there exist non-negative real numbers $\lambda_1, \lambda_2, \lambda_3, \cdots, \lambda_k$ and $\mathbf{d}' \in R_C(F(P))$ such that $\sum_{i=1}^{k} \lambda_i = 1$ and $\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}_i + \mathbf{d}'$.

Let $\mathbf{x}' = \mathbf{x} + \mathbf{d}$. Since $\mathbf{d}'$ is a recession direction of $F(P)$, we see that the conic combination $\mathbf{d}' + \mathbf{d}$ is also a recession direction of $F(P)$. This implies that $\mathbf{x}' \in Conv\left(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}\right) \oplus R_C(F(P))$. Hence by Lemma 4.4.2, we see that $\mathbf{x}' \in F(P)$.

Now

$$\langle \mathbf{c}, \mathbf{x}' \rangle = \langle \mathbf{c}, \mathbf{x} + \mathbf{d} \rangle = \langle \mathbf{c}, \mathbf{x} \rangle + \langle \mathbf{c}, \mathbf{d} \rangle > \langle \mathbf{c}, \mathbf{x} \rangle \quad (\because \langle \mathbf{c}, \mathbf{d} \rangle > 0)$$

This contradicts the optimality of $\mathbf{x}$. Hence we see that P is unbounded. □

only if part:-

Let P be unbounded. By Corollary 4.2.5, we see that $F(P)$ is unbounded. Hence by Lemma 2.2.1, we see that $F(P)$ has recession directions and hence $R_C(F(P))$ is non-empty. Assume that each $\mathbf{d} \in R_C(F(P))$ satisfies the condition $\langle \mathbf{c}, \mathbf{d} \rangle \leq 0$. Let t be the index satisfying the conditions $1 \leq t \leq k$ and $\langle \mathbf{c}, \mathbf{x}_t \rangle = \max\left(\{\langle \mathbf{c}, \mathbf{x}_i \rangle \mid 1 \leq i \leq k\}\right)$.

Let $\mathbf{y}$ be any feasible solution of P. Hence by lemma 4.4.2, there exist non-negative real numbers $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k$ and $\mathbf{d}'' \in R_C(F(P))$ such that $\mathbf{y} = \sum_{i=1}^{k} \alpha_i \mathbf{x}_i + \mathbf{d}''$.

Now

$$\langle \mathbf{c}, \mathbf{y} \rangle = \left\langle \mathbf{c}, \sum_{i=1}^{k} \alpha_i \mathbf{x}_i + \mathbf{d}'' \right\rangle = \sum_{i=1}^{k} \alpha_i \langle \mathbf{c}, \mathbf{x}_i \rangle + \langle \mathbf{c}, \mathbf{d}'' \rangle \leq \langle \mathbf{c}, \mathbf{x}_t \rangle \sum_{i=1}^{k} \alpha_i + \langle \mathbf{c}, \mathbf{d}'' \rangle$$

which implies that $\langle \mathbf{c}, \mathbf{y} \rangle \leq \langle \mathbf{c}, \mathbf{x}_t \rangle$ since $\sum_{i=1}^{k} \alpha_i = 1$ and $\langle \mathbf{c}, \mathbf{d}'' \rangle \leq 0$. Since $\mathbf{y}$ can be any element of $F(P)$, we see that $\mathbf{x}_t$ is an optimal solution of P. This contradicts the fact that P is unbounded. Hence we conclude that P must have some recesion direction $\mathbf{d}$ such that $\langle \mathbf{c}, \mathbf{d} \rangle > 0$. Hence the Lemma.

The following corollary easily follows from Lemma 4.3.1 and Lemma 4.4.3.

Corollary 4.4.1.

Let P be a feasible linear program in canonical form. Then P has a finite optimum if and only if there exists no $\mathbf{d} \in \mathbb{R}^n$ such that $\mathbf{G}\mathbf{d} \leq \mathbf{0}_{m+n}$ and $\langle \mathbf{c}, \mathbf{d} \rangle > 0$.

Theorem 4.4.1. ***(Caratheodory Characterisation Theorem)***

Let P be a linear program in canonical form. Then $F(P)$ is the minkowski sum of the convex hull of the set of extreme points of the feasible region and the conic hull of the set of normalised extreme directions of the feasible region.

Proof.

If $R_C(F(P))$ is empty, then $F(P)$ is bounded and the result follows from Theorem 4.2.3. Otherwise $F(P)$ is unbounded. Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$. By Lemma 4.4.2, we see that $F(P) = Conv\left(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}\right) \oplus R_C(F(P))$. By Theorem 4.3.2, $R_C(F(P)) = Cone\left(\{\mathbf{x} \mid \mathbf{x} \text{ is a normalised extreme direction of } F(P)\}\right)$. Hence it follows that $F(P) = Conv\left(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}\right) \oplus Cone\left(\{\mathbf{x} \mid \mathbf{x} \text{ is a normalised extreme direction of } F(P)\}\right)$. Hence the proof. □

Theorem 4.4.2. ***(Fundamental Theorem of Linear Programming)***
Given feasible linear Program P in canonical form. If P has an optimal solution, then P will also have an extreme point optimal solution.

Proof.
If $F(P)$ is bounded, the theorem follows immediately by Theorem 4.2.4. Therefore we consider the case where $F(P)$ is ubounded. Let $\mathbf{x}^*$ be the optimal solution of P. Therefore $\langle \mathbf{c}, \mathbf{x}^* \rangle \geq \langle \mathbf{c}, \mathbf{x} \rangle$ for each $\mathbf{x} \in F(P) \setminus \{\mathbf{x}^*\}$.

Let $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be the extreme points of $F(P)$ and $\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_l$ be the extreme directions of $F(P)$. By Lemma 4.3.1, we see that $\mathbf{G}\mathbf{d}_j \leq \mathbf{0}_{m+n}$ for each $j \in \{1, 2, 3, \cdots, l\}$. Since P has finite optimum, we see that $\langle \mathbf{c}, \mathbf{d}_j \rangle \leq 0$ or each $j \in \{1, 2, 3, \cdots, l\}$ by Corollary 4.4.1.

By Theorem 4.4.1, we see that $F(P) = Conv\left(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}\right) \oplus Cone\left(\{\mathbf{d}_1, \mathbf{d}_2, \mathbf{d}_3, \cdots, \mathbf{d}_l\}\right)$. Hence there exist non-negative real numbers $\alpha_1, \alpha_2, \alpha_3, \cdots, \alpha_k$ and $\beta_1, \beta_2, \beta_3, \cdots, \beta_l$ such that $\sum_{j=1}^{k} \alpha_j = 1$ and $\mathbf{x}^* = \sum_{j=1}^{k} \alpha_j \mathbf{x}_j + \sum_{j=1}^{l} \beta_j \mathbf{d}_j$.

Let t be the index such that $\langle \mathbf{c}, \mathbf{x}_t \rangle = \max\left(\{\langle \mathbf{c}, \mathbf{x}_i \rangle \mid 1 \leq i \leq k\}\right)$.

Now we have

$$\langle \mathbf{c}, \mathbf{x}^* \rangle = \left\langle \mathbf{c}, \sum_{j=1}^{k} \alpha_j \mathbf{x}_j + \sum_{j=1}^{l} \beta_j \mathbf{d}_j \right\rangle = \sum_{j=1}^{k} \alpha_j \langle \mathbf{c}, \mathbf{x}_j \rangle + \sum_{j=1}^{l} \beta_j \langle \mathbf{c}, \mathbf{d}_j \rangle \leq \langle \mathbf{c}, \mathbf{x}_t \rangle \sum_{j=1}^{k} \alpha_j$$
$$(\because \langle \mathbf{c}, \mathbf{d}_j \rangle \leq 0 \text{ for each } j \in \{1, 2, 3, \cdots, l\}.)$$

which implies that $\langle \mathbf{c}, \mathbf{x}^* \rangle = \langle \mathbf{c}, \mathbf{x}_t \rangle$ as $\sum_{j=1}^{k} \alpha_j = 1$ and $\mathbf{x}^*$ is the optimal solution of P. Thus we see that P has an optimal extreme point solution. Hence the Theorem. □

The following lemma shows that if P is any feasible linear program in canonical form and having a finite optimum, then any other feasible linear program with same objective function and constraint matrix of P will also have finite optimum.

Lemma 4.4.4.
Let P be the feasible linear program

$$\textit{Maximize } \langle \mathbf{c}, \mathbf{x} \rangle \textit{ subject to}$$
$$\mathbf{A}\mathbf{x} \leq \mathbf{b}$$
$$\mathbf{x} \geq \mathbf{0}_n$$

and P' be the linear program

$$\textit{Maximize } \langle \mathbf{c}, \mathbf{x} \rangle \textit{ subject to}$$
$$\mathbf{A}\mathbf{x} \leq \mathbf{b}'$$
$$\mathbf{x} \geq \mathbf{0}_n$$

for some vector $\mathbf{b}' \in \mathbb{R}^m$. If P has a finite optimum and P' is feasible, then P' has an optimal basic feasible solution.

Proof.
Assume that P has a finite optimum and P' is feasible. Since $P$is feasible and has finite optimum, by Corollary 4.4.1, we see that there exists no $\mathbf{d} \in \mathbb{R}^n$ such that $\mathbf{G}\mathbf{d} \leq \mathbf{0}_{m+n}$ and $\langle \mathbf{c}, \mathbf{d} \rangle > 0$. This further implies that P' also has a finite optimum. Hence by Theorem 4.4.2, we see that P' has an optimal extreme point solution. By Theorem 4.2.1, extreme points are equivalent to basic feasible solutions and hence P' has an optimal basic feasible solution. □

Lemma 4.4.5.
Given feasible linear Program P in canonical form. If P has more than one optimal solution , then P has infinite number of optimal solutions.

Proof. Let $\mathbf{x}$ and $\mathbf{y}$ be two optimal solutions of P. Clearly $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{c}, \mathbf{y} \rangle$. Let $\mathbf{z} = \lambda \mathbf{x} + (1-\lambda)\mathbf{y}$ where $\lambda \in [0, 1]$. By convexity of $F(P)$, we see that $\mathbf{z}$ is also a member of $F(P)$.

Now

$$\langle \mathbf{c}, \mathbf{z} \rangle = \langle \mathbf{c}, \lambda \mathbf{x} + (1-\lambda)\mathbf{y} \rangle = \lambda \langle \mathbf{c}, \mathbf{x} \rangle + (1-\lambda)\langle \mathbf{c}, \mathbf{z} \rangle = \langle \mathbf{c}, \mathbf{x} \rangle$$

Hence $\mathbf{z}$ is also an optimal solution of P. Since $\mathbf{x}$ can be chosen as any convex combination of $\mathbf{x}$ and $\mathbf{y}$, we conclude that any convex combination of $\mathbf{x}$ and $\mathbf{y}$ is an optimal solution of P. Hence P has infinite number of optimal solutions. □

4.5 Brute-force algorithm and Graphical method for solving LP

By Theorem 4.4.2, we have seen that the optimal solution of a linear program occurs at some extreme point of the feasible region. Based on this idea, it is possible to write a brute force algorithm *bruteforceLPSolve* to solve a given linear program as follows. The correctness of the algorithm is ensured by the discussions done so far in this chapter.

Algorithm *bruteforceLPSolve*$(F(P), obj)$
(∗ The algorithm returns the optimal solution of a linear program P in the canonical form. ∗)
Input: The feasible region of P and its objective function obj.
Output: The optimal solution of P together with its finite optimum value of the objective function.

S :=*getExtremePoints*$(F(P))$;
$OPT = -\infty$;
for each extreme point $\mathbf{x} \in S$
  do if $obj(\mathbf{x}) \geq OPT$
    then $OPT = obj(\mathbf{x})$;
      $\mathbf{x}^* = \mathbf{x}$;
return $(\mathbf{x}^*, OPT)$;

Algorithm 4.2: A brute force algorithm to solve a linear program in canonical form.

Since the extreme points are the vertices of the feasible region, we can solve a given linear program in canonical form using the following steps.

1. Plot the feasible region of the linear program and identify its vertices.
2. Evaluate the objective function at each of the identified vertices. The vertex which yields the highest value for the objective function will be the optimal solution of the linear program.

This graphical method is illustated in the following example.

Example 4.5.1.
Consider the linear program

$$\begin{aligned} &\text{Maximize } x_1 + x_2 \text{ subject to} \\ &2x_1 + x_2 \leq 6 \\ &x_1 + 2x_2 \leq 6 \\ &2x_1 - x_2 \leq -1 \\ &x_1 \geq 0,\ x_2 \geq 0 \end{aligned} \tag{4.7}$$

This linear program is in the canonical form and its feasible region is shown in Figure 4.6. The vertices of the feasible region and the values of the objective function at each of these vertices are given in Table 4.1. From the Table, it is seen that the linear program has the optimal solution $\begin{bmatrix} 2 & 2 \end{bmatrix}^T$.

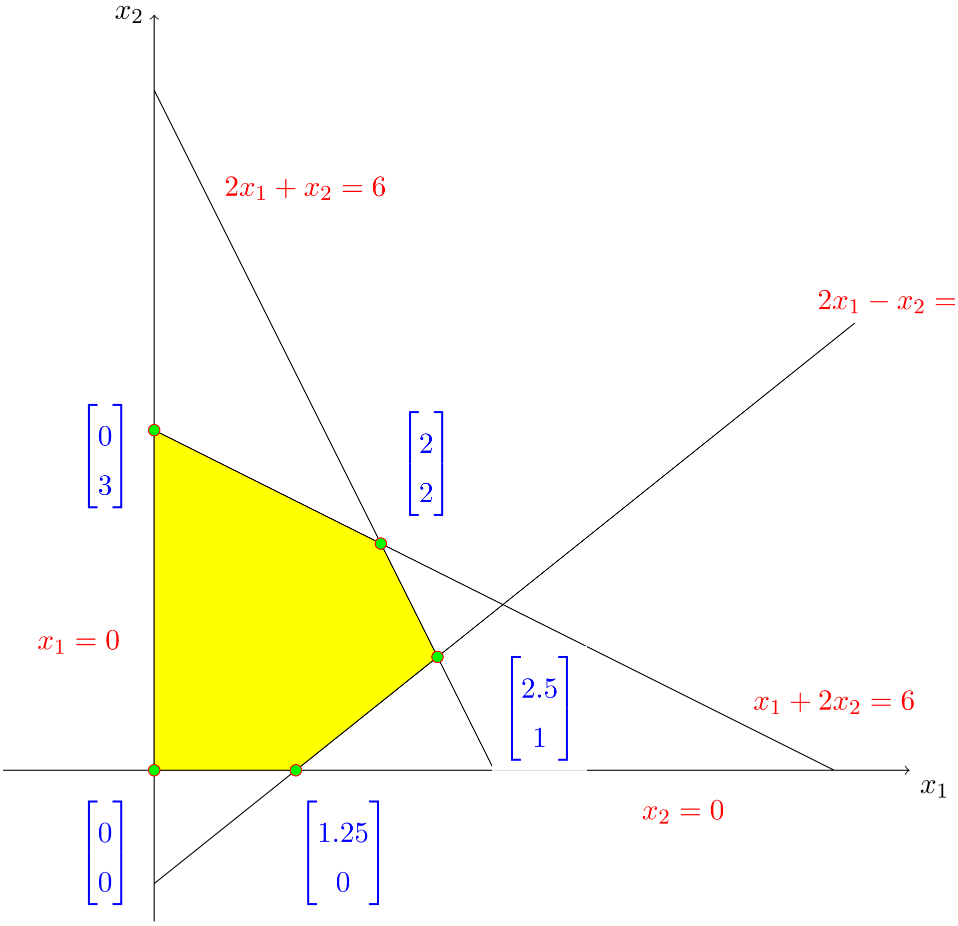


Figure 4.6: The shaded region is the feasible region of LP 4.7.

Sl.No.	Extreme Point	Value of the objective function
1	$\begin{bmatrix} 0 & 0 \end{bmatrix}^T$	0
2	$\begin{bmatrix} 0 & 3 \end{bmatrix}^T$	3
3	$\begin{bmatrix} 2 & 2 \end{bmatrix}^T$	4
4	$\begin{bmatrix} 2.5 & 1 \end{bmatrix}^T$	3.5
5	$\begin{bmatrix} 1.25 & 0 \end{bmatrix}^T$	1.25

Table 4.1: The vertices of the fesible region of LP 4.7 together with the values of the objective function at each of the vertices.

4.6 Summary

In this chapter, we introduced the notion of rank of a feasible solution and characterised the interior points of the feasible regions as feasible solutions with rank 0. We brought about the notion of basic feasible solutions and established the equivalence of the vertices, extreme points and basic feasible solutions and proved that all linear programs have finite non-zero number of extreme points. We defined the notions of normalised recession directions, extreme directions and normalised extreme directions of an unbounded feasible region. We had charcterised the recession cone of an unbounded feasible region as the cone of the set of normalised recession directions. We proved that any linear program with an unbounded feasible region has finite number of normalised extreme directions. We proved the Carathedory characterisation theorem and the fundamental theorem of linear programming and thereby established the foundation for the graphical method of solving linear programs. Finally we presented a brute force algorithm to solve linear programs and illustrated the graphical method for solving linear programs with an example. In the next chapter, we shall discuss the primal dual theory in detail.

Chapter 5

Primal Dual Theory

5.1 Introduction

In this chapter, we introduce the notion of dual of a linear program and explore the relationship beween a linear program and its dual. These results together form the primal dual theory which serves as the basis of the primal-dual schema for designing approximation algorithms for NP-Hard combinatorial optimization problems [2]. The Weak Duality Theorem, Strong Duality Theorem and the Complementary Slackness Conditions are the building blocks of the theory. In the forthcoming sections we shall discuss each of these in detail.

5.2 Dual of a Linear Program

In this section we introduce the concept of dual of a linear program.

Definition 5.2.1.
Given the *maximization linear program P* in canonical form called the *primal program* given by

$$\begin{aligned} \text{Maximize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\ & \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}_n \text{ where} \\ & \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n \end{aligned} \tag{5.1}$$

Then the *minimization linear program D* given by

$$\begin{aligned} \text{Minimize } & \langle \mathbf{b}, \mathbf{y} \rangle \text{ subject to} \\ & \mathbf{A}^T\mathbf{y} \geq \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0}_m \text{ where} \\ & \mathbf{y} \in \mathbb{R}^m \end{aligned} \tag{5.2}$$

is called the *dual program* or simply the *dual* of P.

Remark 5.2.1.
Let the primal program (P) be the maximization program given by

$$\begin{aligned} \text{Maximize } & \sum_{j=1}^{n} c_j x_j \text{ subject to} \\ & \sum_{j=1}^{n} a_{ij} x_j \leq b_i, 1 \leq i \leq m \\ & x_j \geq 0, 1 \leq j \leq n \end{aligned} \tag{5.3}$$

Then the dual program D of P is given by

$$
\begin{aligned}
\text{Minimize } & \sum_{i=1}^{m} b_i y_i \text{ subject to} \\
& \sum_{i=1}^{m} a_{ij} y_i \geq c_j, 1 \leq j \leq n \\
& y_i \geq 0, 1 \leq i \leq m
\end{aligned}
\tag{5.4}
$$

Remark 5.2.2.
Let the primal program (P) be the *minimization program* given by

$$
\begin{aligned}
\text{Minimize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\
& \mathbf{A}\mathbf{x} \geq \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}_n \text{ where} \\
& \mathbf{A} \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m, c \in \mathbb{R}^n
\end{aligned}
$$

Then the corresponding the dual program (D) of P is a *maximization program*

$$
\begin{aligned}
\text{Maximize } & \langle \mathbf{b}, \mathbf{y} \rangle \text{ subject to} \\
& \mathbf{A}^T \mathbf{y} \leq \mathbf{c} \\
& \mathbf{y} \geq \mathbf{0}_m
\end{aligned}
$$

Remark 5.2.3.
It must be noted that there are as many decision variables in the dual program as the number of constraints in the primal program and as many constraints in the dual program as the number of decision variables in the primal Program.

Lemma 5.2.1.
Given primal program P in (5.1) and the dual program D in (5.2). Then the dual of D is P itself.

Proof.
By Remark 5.2.2, the dual of D is the maximization program given by

$$
\begin{aligned}
\text{Maximize } & \langle \mathbf{c}, \mathbf{z} \rangle \text{ subject to} \\
& \mathbf{A}\mathbf{z} \leq \mathbf{b} \\
& \mathbf{z} \geq \mathbf{0}_n
\end{aligned}
\tag{5.5}
$$

Clearly LP (5.5) is same as the primal program P itself. Hence the proof. □

Remark 5.2.4.
In the rest of this section, we assume that P refers to the primal program

$$
\begin{aligned}
\text{Maximize } & \langle \mathbf{c}, \mathbf{x} \rangle \text{ subject to} \\
& \mathbf{A}\mathbf{x} \leq \mathbf{b} \\
& \mathbf{x} \geq \mathbf{0}_n \text{ where} \\
& \mathbf{A} \in \mathbb{R}^{m \times n}, \mathbf{x} \in \mathbb{R}^n, \mathbf{b} \in \mathbb{R}^m, \mathbf{c} \in \mathbb{R}^n
\end{aligned}
\tag{5.6}
$$

and D refers to the dual program

$$\begin{aligned} &\text{Minimize } \langle \mathbf{b}, \mathbf{y} \rangle \text{ subject to} \\ &\mathbf{A}^T\mathbf{y} \geq \mathbf{c} \\ &\mathbf{y} \geq \mathbf{0}_m \end{aligned} \tag{5.7}$$

5.3 Primal Dual Relationships

In this section, we discuss three results stating the relationship between the primal and the dual programs.

5.3.1 Weak Duality Theorem

The weak duality theorem states that any value in the image of the primal program is a lower bound for the image of the dual program. We formally state this Theorem as follows.

Theorem 5.3.1. ***(Weak Duality Theorem)***
Given the primal program P and the corresponding dual program D. Then for any $\mathbf{x} \in F(P)$ and $\mathbf{y} \in F(D)$, it must be the case that $\langle \mathbf{c}, \mathbf{x} \rangle \leq \langle \mathbf{b}, \mathbf{y} \rangle$.

Proof.
Let $\mathbf{x}$ be any feasible solution of P and $\mathbf{y}$ be any feasible solution of D. Therefore we see that $\mathbf{Ax} \leq \mathbf{b}, \mathbf{x} \geq \mathbf{0}_n$ and $\mathbf{A}^T\mathbf{y} \geq \mathbf{c}, \mathbf{y} \geq \mathbf{0}_m$.
Now

$$\mathbf{c}^T\mathbf{x} \leq (\mathbf{A}^T\mathbf{y}^T)\mathbf{x} = \mathbf{y}^T(\mathbf{Ax}) \leq \mathbf{y}^T\mathbf{b} = \mathbf{b}^T\mathbf{y}$$

Hence it follows that $\langle \mathbf{c}, \mathbf{x} \rangle \leq \mathbf{b}, \mathbf{y} \rangle$. Hence the Proof. □

The following corollary is an immediate consequence of weak duality theorem.

Corollary 5.3.1.
Given the primal program P and the corresponding dual program D. If $\mathbf{x}$ and $\mathbf{y}$ are feasible solutions of P and D respectively such that $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{b}, \mathbf{y} \rangle$, then $\mathbf{x}$ is the optimal solution of P and $\mathbf{y}$ is the optimal solution of D.

Now we are going to define a term called *duality gap* which is a measure of how far are a given primal feasible solution and a dual feasible solution from the optimum.

Definition 5.3.1.
Given the primal program P and the corresponding dual program D. Then for any $\mathbf{x} \in F(P)$ and $\mathbf{y} \in F(D)$, the difference $\langle \mathbf{b}, \mathbf{y} \rangle - \langle \mathbf{c}, \mathbf{x} \rangle$ is called the *duality gap between P and D with respect to $\mathbf{x}$ and $\mathbf{y}$* and is denoted by $\gamma_{\mathbf{xy}}$.

Given the primal program P and the corresponding dual program D. Let $\mathbf{x}$ be any feasible solution of the primal program and $\mathbf{y}$ be any feasible solution of the dual program. Then by Theorem 5.3.1, $\gamma_{\mathbf{xy}} \geq 0$. The case where $\gamma_{\mathbf{xy}} = 0$ is of great importance and is the main discussion in the remaining sections.

5.3.2 Complementary Slackness Conditions

Complementary slackness is a relationship between the primal and dual programs which informally suggests that variables in one program are complementary to constraints in the other. Accordingly there are two sets of complementary slackness conditions namely, *primal complementary slackness conditions* and *dual complementary slackness conditions*. In this subsection, we discuss each slackness conditions in detail.

Theorem 5.3.2. ***(Complementary Slackness Conditions)***

Given the primal program P and the corresponding dual program D. Let $\mathbf{x}$ be a primal feasible solution and $\mathbf{y}$ be a dual feasible solution. Then $\gamma_{\mathbf{xy}} = 0$ if and only if both $\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle = 0$ and $\langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0$.

Proof.

if part:-

Let $\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle = 0$ and $\langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0$. Hence we see that

$$\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle + \langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \left\langle \mathbf{A}^T\mathbf{y}, \mathbf{x} \right\rangle - \langle \mathbf{c}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle - \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{b}, \mathbf{y} \rangle - \langle \mathbf{c}, \mathbf{x} \rangle = 0$$
$$(\because \left\langle \mathbf{A}^T\mathbf{y}, \mathbf{x} \right\rangle = \langle \mathbf{y}, \mathbf{A}\mathbf{x} \rangle = \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle)$$

which imples that $\langle \mathbf{b}, \mathbf{y} \rangle - \langle \mathbf{c}, \mathbf{x} \rangle = 0$. That is $\gamma_{\mathbf{xy}} = 0$.

only-if part:-

Let $\gamma_{\mathbf{xy}} = 0$. Since $\mathbf{x}$ is a primal feasible solution and $\mathbf{y}$ is a dual feasible solution, we see that $\mathbf{A}\mathbf{x} \leq \mathbf{b}$, $\mathbf{A}^T\mathbf{y} \geq \mathbf{c}$, $\mathbf{x} \geq \mathbf{0}_n$ and $\mathbf{y} \geq \mathbf{0}_m$. This implies that

$$\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle \geq 0 \text{ and } \langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle \geq 0$$

Now

$$\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle + \langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = \left\langle \mathbf{A}^T\mathbf{y}, \mathbf{x} \right\rangle - \langle \mathbf{c}, \mathbf{x} \rangle + \langle \mathbf{b}, \mathbf{y} \rangle - \langle \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0 \quad (\because \langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{b}, \mathbf{y} \rangle)$$

Thus $\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle = 0$ and $\langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0$. □

Remark 5.3.1.

The conditions $\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle = 0$ and $\langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0$ in Theorem 5.3.2 are called *primal complementary slackness conditions* and *dual complementary slackness conditions* respectively.

Remark 5.3.2.

Let the primal program P be LP 5.3. Then the corresponding dual program D is given by LP 5.4. Let $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_m \end{bmatrix}^T$ be the feasible solutions of P and D respectively such that $\gamma_{\mathbf{xy}} = 0$.

The condition $\left\langle \mathbf{A}^T\mathbf{y} - \mathbf{c}, \mathbf{x} \right\rangle = 0$ is equivalent to $\left(\sum\limits_{i=1}^{m} a_{ij}y_i - c_j \right) x_j = 0$ for each $j \in \{1, 2, 3, \cdots, n\}$. That is $\sum\limits_{i=1}^{m} a_{ij}y_i = c_j$ or $x_j = 0$ for eah $j \in \{1, 2, 3, \cdots, n\}$ which form the primal complementary slackness conditions.

Similarly the condition $\langle \mathbf{b} - \mathbf{A}\mathbf{x}, \mathbf{y} \rangle = 0$ is equivalent to $\left(\sum\limits_{j=1}^{n} a_{ij}x_j - b_i \right) y_i = 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. That is $\sum\limits_{j=1}^{n} a_{ij}x_j = b_i$ or $y_i = 0$ for eah $i \in \{1, 2, 3, \cdots, m\}$ which form the dual complementary slackness conditions..

Remark 5.3.3.

Given the primal program P and the dual program D. Let $\mathbf{x}$ be a primal feasible solution and $\mathbf{y}$ be a dual feasible solution such that $\gamma_{\mathbf{xy}} = 0$. Let $\mathbf{s}_{\mathbf{x}} \in \mathbb{R}^m$ defined by $\mathbf{s}_{\mathbf{x}} = \mathbf{b} - \mathbf{A}\mathbf{x}$ and $\mathbf{t}_{\mathbf{y}} \in \mathbb{R}^n$ defined by $\mathbf{t}_{\mathbf{y}} = \mathbf{A}^T\mathbf{y} - \mathbf{c}$. By Definition 3.3.9 and Definition 3.3.10, we see that $\mathbf{s}_{\mathbf{x}}$ is the primal slack vector associated with $\mathbf{x}$ and $\mathbf{t}_{\mathbf{y}}$ is the dual surplus vector associated with $\mathbf{y}$. Then the primal complementary slackness conditions can be represented as $\langle \mathbf{s}_{\mathbf{x}}, \mathbf{y} \rangle = 0$ and the dual complementary slackness conditions can be represented as $\langle \mathbf{t}_{\mathbf{y}}, \mathbf{x} \rangle = 0$. That is the primal slack vector $\mathbf{s}_{\mathbf{x}}$ and the dual feasible solution $\mathbf{y}$ are orthogonal to each other in $\mathbb{R}^m$ and the dual surplus vector $\mathbf{t}_{\mathbf{y}}$ and the primal feasible solution $\mathbf{x}$ are orthogonal to each other in $\mathbb{R}^n$. Since $\mathbf{s_x}, \mathbf{t_y}, \mathbf{x}$ and $\mathbf{y}$ are non-negative vectors, we infer that corresponding to each $i \in \{1, 2, 3, \cdots, m\}$, at least one among the i^{th} components of $\mathbf{s_x}$ and $\mathbf{y}$ must be 0 and corresponding to each $j \in \{1, 2, 3, \cdots, n\}$, at least one among the j^{th} components of $\mathbf{t_y}$ and $\mathbf{x}$ must be 0.

5.3.3 Strong Duality Theorem

The strong duality theorem is one of the significant milestones in the development of the primal dual theory which essentially states that the finite optima of both the primal and the dual programs are exactly the same [8] [9] [7].

In the following subsections, we discuss the preliminary results for proving the strong duality theorem. Throughout these subsections P_1 refers to the following linear program

$$\begin{aligned} \text{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad & \text{subject to} \\ \mathbf{G}\mathbf{z} &= \mathbf{h} \\ \mathbf{z} &\geq \mathbf{0}_n \quad \text{where} \end{aligned} \tag{5.8}$$

$$\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots & \mathbf{g}_n \end{bmatrix} \in \mathbb{R}^{m \times n} \text{ and } rank(\mathbf{G}) = m, \ \mathbf{d} = \begin{bmatrix} d_1 & d_2 & d_3 & \cdots & d_n \end{bmatrix}^T \in \mathbb{R}^n$$

and $\mathbf{h} = \begin{bmatrix} h_1 & h_2 & h_3 & \cdots & h_m \end{bmatrix}^T \in \mathbb{R}^m$

Remark 5.3.4. Corresponding to any linear program, an equivalent linear program of the type(5.14) can be obtained by converting the given linear program to standard form and then eliminating the redundant constraints if any.

5.3.3.1 Basis Matrix

Definition 5.3.2.

Given linear program P_1. Then any $m \times m$ matrix $\mathbf{B}$ formed out of any m linearly independent column vectors of $\mathbf{G}$ in such a way that the column vectors in $\mathbf{B}$ appear in he same relative order as they are in $\mathbf{G}$ is called a *basis matrix* of P_1. .

Remark 5.3.5. Given linear program P_1. Since the columns of each basis matrix of P_1 are linearly independent, we see that each basis matrix of P_1 is invertible.

Definition 5.3.3.

Given linear program P_1 and a feasible solution $\mathbf{z}$ of P_1. Let $\mathbf{B}$ be a basis matrix of P_1. Let $\mathbf{N}$ be the $m \times (n-m)$ matrix formed out of the column vectors in $\mathbf{G}$ which are not in $\mathbf{B}$ in such a way that the column vectors of $\mathbf{N}$ appear in the same relative order as they are in $\mathbf{G}$. Let $I_\mathbf{B}$ be the m-tuple whose i^{th} entry is the index of the i^{th} column in $\mathbf{B}$ within the matrix $\mathbf{G}$. Let $I_\mathbf{N}$ be the $(n-m)$-tuple whose i^{th} entry is the index of the i^{th} column in $\mathbf{N}$ within the matrix $\mathbf{G}$. Let $\mathbf{z_B} = \begin{bmatrix} z_i \end{bmatrix}_{i \in I_\mathbf{B}}$ and $\mathbf{z_N} = \begin{bmatrix} z_j \end{bmatrix}_{j \in I_\mathbf{N}}$. Then the vectors $\mathbf{z_B}$ and $\mathbf{z_N}$ are called *basis vector* and *non-basis vector* defined by the basis matrix $\mathbf{B}$. The components of $\mathbf{z_B}$ and $\mathbf{z_N}$ are respectively called the *basic components* and the *non-basic components* of $\mathbf{z}$.

Remark 5.3.6.

Let $\mathbf{z}$ be a feasible solution of P_1. In various lemmas which we prove in this section we follow the convention that $\mathbf{B}$ refers to a basis matrix of P_1 and the sets $I_\mathbf{B}$ and $I_\mathbf{N}$, matrices $\mathbf{B}$ and $\mathbf{N}$, vectors $\mathbf{z_B}$ and $\mathbf{z_N}$ are defined as in the Definition 5.3.3 where the vectors $d_\mathbf{B}$ and $d_\mathbf{N}$ are defined as $d_\mathbf{B} = [d_i]_{i \in I_\mathbf{B}}$ and $d_\mathbf{N} = [d_j]_{j \in I_\mathbf{N}}$.

Let $\mathbf{z}$ be a any feasible solution of P_1. Now we are going to find an expression for basic components of a feasible solution of P_1 in terms of the non-basic components.

Since $\mathbf{z}$ is a feasible solution of P_1, we have $\mathbf{Gz} = \mathbf{Bz_B} + \mathbf{Nz_N} = \mathbf{h}$ which implies that

$$\mathbf{z_B} = \mathbf{B}^{-1}(\mathbf{h} - \mathbf{Nz_N}) \tag{5.9}$$

We can also find an expression for the value of the objective function of P_1 at $\mathbf{z}$ as follows.

$$\begin{aligned}
\langle \mathbf{d}, \mathbf{z} \rangle &= \sum_{j=1}^{n} d_j z_j = \sum_{j \in I_\mathbf{B}} d_j z_j + \sum_{j \in I_\mathbf{N}} d_j z_j = \langle \mathbf{d_B}, \mathbf{z_B} \rangle + \sum_{j \in I_\mathbf{N}} d_j x_j \\
&= \langle \mathbf{d_B}, \mathbf{B}^{-1}(\mathbf{h} - \mathbf{Nz_N}) \rangle + \sum_{j \in I_\mathbf{N}} d_j z_j = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{Nz_N} \rangle + \sum_{j \in I_\mathbf{N}} d_j z_j \\
&= \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle - \left\langle \mathbf{d_B}, \sum_{j \in I_\mathbf{N}} \mathbf{B}^{-1}\mathbf{g}_j z_j \right\rangle + \sum_{j \in I_\mathbf{N}} d_j z_j \\
&= \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle - \sum_{j \in I_\mathbf{N}} \left\langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \right\rangle z_j + \sum_{j \in I_\mathbf{N}} d_j z_j
\end{aligned}$$

Thus

$$\langle \mathbf{d}, \mathbf{z} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle + \sum_{j \in I_\mathbf{N}} \left(d_j - \left\langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \right\rangle \right) z_j \tag{5.10}$$

Now we prove a lemma that characterizes the basic feasible solutions of P_1 in terms of the corresponding basis matrices.

Lemma 5.3.1.
Given linear program P_1 and some vector $\mathbf{z} \in \mathbb{R}^n$. Then $\mathbf{z}$ is a basic feasible solution of P_1 if and only if there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$.

Proof.
if part:-
Let there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$. Hence we have $\mathbf{Bz_B} + \mathbf{Nz_N} = \mathbf{h}$ which means that $\mathbf{Gz} = \mathbf{h}$. Since $\mathbf{z_B} \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$, we see that $\mathbf{z} \geq \mathbf{0}_n$. Hence we see that $\mathbf{z}$ is a feasible solution of P_1.
Since $\mathbf{Gz} = \mathbf{h}$ and $rank(\mathbf{G}) = m$, we see that the m constraints of P_1 specified by $\mathbf{G}$ are linearly independent and are tight at $\mathbf{z}$. Since $\mathbf{z_N} = \mathbf{0}_{n-m}$, we also see that $n-m$ non-negativity constraints of P_1 are also tight $\mathbf{z}$. Hence it follows that n linearly independent constraints of P_1 are tight at $\mathbf{z}$. Hence we conclude that $\mathbf{z}$ is a basic feasible solution of P_1.

only if part:-
Let $\mathbf{z}$ be a basic feasible solution of P_1. Therefore $rank_{P_1}(\mathbf{z}) = n$ which means that n linearly independent constraints of P_1 must be tight at $\mathbf{z}$. Since $\mathbf{Gz} = \mathbf{h}$ and $rank(\mathbf{G}) = m$, we see that the m constraints defined by $\mathbf{G}$ are linearly independent and are tight at $\mathbf{z}$. Hence at least $n-m$ non-negativity constraints must be tight at $\mathbf{z}$ in order that n linearly independent constraints are tight at $\mathbf{z}$. Hence it follows that $\mathbf{z}$ consists of at least $n-m$ zero components. Let $\mathbf{B}$ be an $m \times m$ matrix formed out of m column vectors of $\mathbf{G}$ and $\mathbf{N}$ be the $m \times (n-m)$ matrix formed out of the the $n-m$ columns of $\mathbf{G}$ in such a way that $\mathbf{z_N} = \mathbf{0}_{n-m}$ and $\mathbf{Gz} = \mathbf{Bz_B} + \mathbf{Nz_N} = \mathbf{Bz_B} = \mathbf{h}$. We are going to show that $\mathbf{B}$ is a basis matrix of P_1.
Without loss of generality, we may take the first $t \geq n-m$ components of $\mathbf{z}$ are zero. Since $\mathbf{z}$ is a basic feasible solution of P_1, we see that $\mathbf{z}$ is a solution of the system of equations $\mathbf{G}'\mathbf{y} = \mathbf{h}'$ where $\mathbf{G}' = \begin{bmatrix} \mathbf{G}^T & \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 & \cdots & \mathbf{e}_t \end{bmatrix}^T$ such that $rank(\mathbf{G}') = n$ and $\mathbf{h}' = \begin{bmatrix} \mathbf{h} & \mathbf{0}_t \end{bmatrix}^T$. Since $rank(\mathbf{G}) = n$, we see that $\mathbf{z}$ is the only solution of this system. Consider the system of equations $\mathbf{Bx} = \mathbf{h}$. Clearly $\mathbf{z_B}$ is a solution of the system $\mathbf{Bx} = \mathbf{h}$. If this system has another

solution $\mathbf{x}'$, then $\begin{bmatrix} \mathbf{x}' \\ \mathbf{0}_{n-m} \end{bmatrix}$ will be a solution of the system $\mathbf{G}'\mathbf{y} = \mathbf{h}'$. This leads us to the contradiction that the system $\mathbf{G}'\mathbf{y} = \mathbf{h}'$ has more than one solution which in turn contradicts the fact that $rank(\mathbf{G}') = n$. Hence we see that the system of equations $\mathbf{Bx} = \mathbf{h}$ has a unique solution and hence $rank(\mathbf{B}) = m$. Hence we conclude that $\mathbf{B}$ is a basis matrix of P_1.

Hence the Lemma. □

Corollary 5.3.2.

Given linear program P_1 and a basic feasible solution of $\mathbf{z}$ of P_1. Let $\mathbf{B}$ be a basis matrix of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}h \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$. Then $\langle \mathbf{d}, \mathbf{z} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle$.

Proof.

Since $\mathbf{z_N} = \mathbf{0}_{n-m}$, we see that each $z_j = 0$ for each $j \in I_\mathbf{N}$. Substituting this in Equation 5.10, we get $\langle \mathbf{d}, \mathbf{z} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{h} \rangle$. □

The following lemma derives a sufficient condition for a given basic feasible solution of P_1 to be optimal.

Lemma 5.3.2.

Given linear Program P_1 and a basic feasible solution $\mathbf{z}$ of P_1. If $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ for each non-basic component z_j of $\mathbf{z}$, then $\mathbf{z}$ is optimal.

Proof.

Let $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ for each non-basic variable $\mathbf{z}_j$ of $\mathbf{z}$. Let $\mathbf{y}$ be any basic feasible solution of P_1. Let $\mathbf{r} = \mathbf{y} - \mathbf{z}$. Since $\mathbf{Gz} = \mathbf{Gy} = \mathbf{h}$, we see that $\mathbf{Gr} = \mathbf{0}_m$. That is $\mathbf{Br_B} + \mathbf{Nr_N} = \mathbf{Br_B} + \sum_{j \in I_\mathbf{N}} \mathbf{g}_j r_j = \mathbf{0}_m$ which implies that

$$\mathbf{r_B} = -\mathbf{B}^{-1} \sum_{j \in I_\mathbf{N}} \mathbf{g}_j r_j$$

Now

$$\langle \mathbf{d}, \mathbf{r} \rangle = \langle \mathbf{d_B}, \mathbf{r_B} \rangle + \langle \mathbf{d_N}, \mathbf{r_N} \rangle = \langle \mathbf{d_B}, -\mathbf{B}^{-1} \sum_{j \in I_\mathbf{N}} \mathbf{g}_j r_j \rangle + \sum_{j \in I_\mathbf{N}} d_j r_j = \sum_{j \in I_\mathbf{N}} \left(d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g_j} \rangle \right) r_j$$

Corresponding to each $j \in I_\mathbf{N}$, we see that $z_j = 0$. Since $\mathbf{y}$ is a feasible solution of P_1, we see that $\mathbf{y} \geq \mathbf{0}_n$. These two facts shows that corresponding to each $j \in I_\mathbf{N}$, we must have $r_j = y_j - z_j = y_j \geq 0$. This together with the fact that $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ for each $j \in I_\mathbf{N}$ implies that $\langle \mathbf{d}, \mathbf{r} \rangle = \sum_{j \in I_\mathbf{N}} (d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g_j} \rangle) r_j \leq 0$. That is $\langle \mathbf{d}, \mathbf{y} \rangle - \langle \mathbf{d}, \mathbf{z} \rangle \leq 0$. In otherwords $\langle \mathbf{d}, \mathbf{y} \rangle \leq \langle \mathbf{d}, \mathbf{z} \rangle$. Since $\mathbf{y}$ may be chosen as any feasible solution of P_1, we conclude that $\mathbf{z}$ is the optimal solution of P. Hence the proof. □

5.3.3.2 Non-degenerate Basic Feasible Solutions

The basic feasible solutions of P_1 can be classified into non-degenerate basic feasible solutions and degenerate basic feasible solutions according to whether exactly n or more than n constraints are tight at those solutions.

Definition 5.3.4.

Given linear Program P_1 and basic feasible solution $\mathbf{z}$ of P_1. Then $\mathbf{z}$ is said to be *non-degenerate basic feasible solution* if exactly n constraints of P_1 are tight at P_1. If more than n constraints of P_1 are tight at $\mathbf{z}$, then $\mathbf{z}$ is called a *degenerate basic feasible solution*. The number of constraints of P_1 which are tight at $\mathbf{z}$ in excess of n is called *degeneracy* of $\mathbf{z}$ and is denoted by $deg(\mathbf{z})$.

The following lemma characterizes the non-degenerate basic feasible solutions of P_1.

Lemma 5.3.3.

Given linear program P_1 and some vector $\mathbf{z} \in \mathbb{R}^n$. Then $\mathbf{z}$ is a non-degenerate basic feasible solution of P_1 if and only if there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} > \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$.

Proof.

if part:-

Let there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} > \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$. By Lemma 5.3.1, we see that $\mathbf{z}$ is a basic feasible solution of P_1. What remains is to show that $\mathbf{z}$ is non-degenerate.

Since $\mathbf{z}$ is a basic feasible solution of P_1, the m constraints specified by $\mathbf{G}$ are tight at $\mathbf{z}$. Since $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} > \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$, we see that exactly $n - m$ non-negativity constraints of P_1 are also tight at $\mathbf{z}$. Thus we see that exactly n constraints of P_1 are tight at $\mathbf{z}$. Hence $\mathbf{z}$ is a non-degenerate basic feasible solution of P_1.

only if part:-

Let $\mathbf{z}$ be a non-degenerate basic feasible solution of P_1. By Lemma 5.3.1, we see that there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$. What remains is to show that none of the components of $\mathbf{B}^{-1}\mathbf{h}$ is zero.

Since $\mathbf{z}$ is a non-degenerate basic feasible solution of P_1, we see that there cannot have more than n constraints of P_1 tight at $\mathbf{z}$. Since $\mathbf{z}$ is a feasible solution of P_1, we see that the m constraints specified by $\mathbf{G}$ are tight at $\mathbf{z}$. We also see that $n - m$ non-negativity constraints of P_1 are also tight at $\mathbf{z}$ as $\mathbf{z_N} = \mathbf{0}_{n-m}$. Hence it follows that none of the components of $\mathbf{z_B}$ is zero in order that $\mathbf{z}$ is a non-degenerate basic feasible solution of P_1. Hence the Lemma. □

Now we are going to derive the necessary and sufficient conditions for a non-degenerate basic feasible solution of P_1 to be optimal. To derive these conditions, we require the following lemma.

Lemma 5.3.4.

Given linear program P_1 and a non-degenerate basic feasible solution $\mathbf{z}$ of P_1 such that $d_k - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_k \rangle = \epsilon > 0$ for some non-basic component z_k of $\mathbf{z}$. Then there exists some $\delta > 0$ and feasible solution $\mathbf{z}'$ of P_1 such that $\langle \mathbf{d}, \mathbf{z}' \rangle = \langle \mathbf{d}, \mathbf{z} \rangle + \epsilon\delta$.

Proof.

Given that $\mathbf{z}$ is a non-degenerate basic feasible solution of P_1. By Lemma 5.3.3 we see that there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} > \mathbf{0}_m$. Let δ be positive real number and $\mathbf{z}' = \begin{bmatrix} z_1' & z_2' & z_3' & \cdots & z_n' \end{bmatrix}^T$ be the vector defined by

$$z_j' = \begin{cases} z_j & \text{if } j \neq k \\ \delta & \text{if } j = k \end{cases}$$

Let $\mathbf{z_N}' = \begin{bmatrix} z_j' \end{bmatrix}_{j \in I_\mathbf{N}}$. Clearly $\mathbf{z_N}' \geq \mathbf{0}_{n-m}$. In order that $\mathbf{z}'$ is a feasible solution of P_1, it should satisfy the conditions $\mathbf{z_B}' = \mathbf{B}^{-1}\mathbf{h} - \mathbf{B}^{-1}\mathbf{N}\mathbf{z_N}' = \mathbf{B}^{-1}\mathbf{h} - \mathbf{B}^{-1}\mathbf{g_k}\delta$ and $\mathbf{z_B}' \geq \mathbf{0}_m$.

Let $\mathbf{B}^{-1}\mathbf{h} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \cdots & \beta_m \end{bmatrix}^T$ and $\mathbf{B}^{-1}\mathbf{g_k} = \begin{bmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \cdots & \alpha_m \end{bmatrix}^T$. Since $\mathbf{B}^{-1}\mathbf{b} > \mathbf{0}_m$, we see that $\beta_i > 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. Hence we must have

$$\mathbf{z_B}' = \begin{bmatrix} \beta_1 - \delta\alpha_1 & \beta_2 - \delta\alpha_2 & \beta_3 - \delta\alpha_3 & \cdots & \beta_m - \delta\alpha_m \end{bmatrix}^T$$

Now we need to consider two cases.

Case 1: $\alpha_i \leq 0$ *for each* $i \in \{1, 2, 3, \cdots, m\}$.

In this case, we see that for any non-negative value of δ, the vector $\mathbf{z_B}' > \mathbf{0}_m$. This implies that for any non-negative value of δ, the vector $\mathbf{z}'$ will be a feasible solution of P_1.

Case 2:- $\alpha_i > 0$ *for one or more* $i \in \{1, 2, 3, \cdots, m\}$.

In this case, if $\delta = \min\{\frac{\beta_i}{\alpha_i} \mid i \in \{1, 2, 3, \cdot, m\}$ such that $\alpha_i > 0\}$, we see that $\delta > 0$ and $\mathbf{z_B}' \geq \mathbf{0}_m$. This implies that for this choice of of δ, the vector $\mathbf{z}'$ will be a feasible solution of P_1.

Hence in both cases we can find $\delta > 0$ such that the vector $\mathbf{z}'$ will be a feasible solution of P_1.

Now

$$\begin{aligned}
\langle \mathbf{d}, \mathbf{z}' \rangle &= \langle \mathbf{d}, \mathbf{B}^{-1}\mathbf{h} \rangle + \left(d_k - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_k \rangle\right) \delta \quad (\because \text{By Eqn (5.10)}) \\
&= \langle \mathbf{d}, \mathbf{B}^{-1}\mathbf{h} \rangle + \epsilon\delta \\
&= \langle \mathbf{d}, \mathbf{z} \rangle + \epsilon\delta \quad (\because \text{By Corollary 5.3.2})
\end{aligned}$$

Hence the Lemma. □

Lemma 5.3.5.

Given linear program P_1 *and a non-degenerate basic feasible solution* $\mathbf{z}$ *of* P_1. *Then* $\mathbf{z}$ *is an optimal solution of* P_1 *if and only if* $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ *for each non-basic component* z_j *of* $\mathbf{z}$.

Proof.

The *if part* follows directly from Lemma 5.3.2. The *only-if part* can be proved as follows.

Let $\mathbf{z}$ be an optimal solution of P_1. Asssume that there exists some non-basic component z_k such that $d_k - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_k \rangle = \epsilon > 0$. By Lemma 5.3.4, we see that there exists some $\delta > 0$ and feasible solution $\mathbf{z}'$ of P_1 such that $\langle \mathbf{d}, \mathbf{z}' \rangle = \langle \mathbf{d}, \mathbf{z} \rangle + \epsilon\delta$. This means that there exists some feasible solution $\mathbf{z}'$ of P_1 such that $\langle \mathbf{d}, \mathbf{z}' \rangle > \langle \mathbf{d}, \mathbf{z} \rangle$ which contradicts the optimality of $\mathbf{z}$. Hence we conclude that $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ for each non-basic component z_j of $\mathbf{z}$. Hence the Lemma. □

Corollary 5.3.3.

Given linear program P_1 *and a non-degenerate basic feasible solution* $\mathbf{z}$ *of* P_1. *Then* $\mathbf{z}$ *is an optimal solution of* P_1 *if and only if* $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$.

Proof.

By Lemma 5.3.5, we see that $\mathbf{z}$ is a non-degenerate optimal basic feasible solution of P if and only if $d_j - \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{g}_j \rangle \leq 0$ for each non-basic component z_j of $\mathbf{z}$. This implies that $\mathbf{z}$ is a non-degenerate optimal basic feasible solution of P if and only if $\mathbf{d_N}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{N} \leq \mathbf{0}^T_{n-m}$. Since $\mathbf{c_B}^T - \mathbf{c_B}^T\mathbf{B}^{-1}\mathbf{B} = \mathbf{0}^T_m$, we further see that $\mathbf{z}$ is a non-degenerate optimal basic feasible solution of P if and only if $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$. □

We close this subsection with a lemma that has a major role in proving strong duality theorem.

Lemma 5.3.6.

Given primal program P in canonical form (5.6) and the corresponding dual program D in (5.7). Let P' be the linear program in standard form and equivalent to P given by

$$\begin{aligned} &\textit{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad \textit{subject to} \\ &\mathbf{G}\mathbf{z} = \mathbf{b} \\ &\mathbf{z} \geq \mathbf{0}_{m+n} \quad \textit{where} \\ &\mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{m \times (m+n)}, \ \mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{R}^{m+n} \end{aligned}$$

If P' has an optimal non-degenerate basic feasible solution, then there exist feasible solution $\mathbf{x}$ of P and feasible solution $\mathbf{y}$ of D such that $\gamma_{\mathbf{xy}} = 0$.

Proof.

It is not hard to see that P' is a linear program equivalent to the primal program P. Let $\mathbf{z}$ an optimal non-degenerate basic feasible solution of P'. Hence by Corollary‘5.3.2, we see that $\langle \mathbf{d}, \mathbf{z} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{b} \rangle$. Furthermore by corollary 5.3.3, we have the conditions

$$\mathbf{d}^T - \mathbf{d_B}^T \mathbf{B}^{-1} \mathbf{G} \leq \mathbf{0}_{m+n}^T \tag{5.11}$$

Since $\mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix}$, the conditions in (5.11) can be split into two as

$$\mathbf{c}^T - \mathbf{d_B}^T \mathbf{B}^{-1} \mathbf{A} \leq \mathbf{0}_n{}^T \text{ and } \mathbf{0}_m{}^T - \mathbf{d_B}^T \mathbf{B}^{-1} \mathbf{I}_m \leq \mathbf{0}_m{}^T \tag{5.12}$$

Let $\mathbf{y} \in \mathbb{R}^m$ defined by $\mathbf{y} = \left(\mathbf{B}^{-1}\right)^T \mathbf{d_B}$. Hence we have

$$\mathbf{c}^T - \mathbf{y}^T \mathbf{A} \leq \mathbf{0}_n{}^T \text{ and } \mathbf{0}_m{}^T - \mathbf{y}^T \mathbf{I}_m \leq \mathbf{0}_m{}^T$$

That is $\mathbf{A}^T\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}_m$. Hence $\mathbf{y}$ is a feasible solution of the dual program D. Moreover $\langle \mathbf{b}, \mathbf{y} \rangle = \langle \mathbf{b}, {\mathbf{B}^{-1}}^T \mathbf{d_B} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{b} \rangle = \langle \mathbf{d}, \mathbf{z} \rangle$. Since P and P' are equivalent, we see that there exists some feasible solution $\mathbf{x}$ of P such that $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{d}, \mathbf{z} \rangle$. Thus we conclude that there exist some primal feasible solution $\mathbf{x}$ and dual feasible solution $\mathbf{y}$ such that $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{b}, \mathbf{y} \rangle$ and hence $\gamma_{\mathbf{xy}} = 0$. Hence the Lemma. □

Lemma 5.3.6 can be extended to the case where all optimal basic feasible solutions of P' given in the lemma are degenerate. To do this extension, we require some results related to degenerate basic feasible solutions.

5.3.3.3 Handling Degeneracy

The following Lemma characterizes the degenerate basic feasible solutions of P_1.

Lemma 5.3.7.

Given linear program P_1 in the standard form (5.14) and some vector $\mathbf{z} \in \mathbb{R}^n$. Then $\mathbf{z}$ is a degenerate basic feasible solution of P_1 with $deg(\mathbf{z}) = t$ if and only if there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_N} = \mathbf{0}_{n-m}$, $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_B}$ consists of exactly t zero components.

Proof.

if part:-

Let there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_N} = \mathbf{0}_{n-m}$, $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_B}$ consists of t zero components. By Lemma 5.3.1, we see that $\mathbf{z}$ is a basic feasible solution of P_1. What remains is to show that $\mathbf{z}$ is degenerate such that $deg(\mathbf{z}) = t$.

Since $\mathbf{z}$ is a basic feasible solution of P_1, the m constraints specified by $\mathbf{G}$ are tight at $\mathbf{z}$. Since $\mathbf{z_B}$ consists of t zero components and $\mathbf{z_N} = \mathbf{0}_{n-m}$, we see that $n - m + t$ non-negativity constraints of P_1 are also tight at $\mathbf{z}$. Thus we see that exactly $n + t$ constraints of P_1 are tight at $\mathbf{z}$. Hence $\mathbf{z}$ is a degenerate basic feasible solution of P_1. such that $deg(\mathbf{z}) = t$.

only if part:-

Let $\mathbf{z}$ be a degenerate basic feasible solution of P_1 such that $deg(\mathbf{z}) = t$. By Lemma 5.3.1, we see that there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{z_B} = \mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ and $\mathbf{z_N} = \mathbf{0}_{n-m}$. What remains is to show that exactly t components of $\mathbf{z_B}$ are zero.

Since $deg(\mathbf{z}) = t$, we see that there are $n+t$ constraints of P_1 tight at $\mathbf{z}$. Since $\mathbf{z}$ is a feasible solution of P_1, we see that the m constraints specified by $\mathbf{G}$ are tight at $\mathbf{z}$. We also see that $n - m + t$ non-negativity constraints of P_1 are also tight at $\mathbf{z}$. Hence it follows that $\mathbf{z}$ consists of $n - m + t$ zero components out of which $n - m$ components correspond to the non-basic components as $\mathbf{z_N} = \mathbf{0}_{n-m}$. Hence it follows that exactly t components of $\mathbf{z_B}$ are zero. Hence the Lemma. □

The following Lemma has a major role in proving strong duality theorem.

Lemma 5.3.8.

Given linear program P_1 in the standard form (5.14) such that P_1 has an optimal degenerate basic feasible solution. Then there exists a basis matrix $\mathbf{B}$ of P_1 such that $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$ and $\mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$.

Proof.

The proof of this lemma is technical and is moved to Appendix A. □

Remark 5.3.7.

It is important to note that the basis matrix specified in Lemma 5.3.8 need not be the basis matrix corresponding to the degenerate optimal basic feasible solution $\mathbf{z}$ of P_1. However, if $\mathbf{z}$ were a non-degenerate basic feasible solution of P_1, the corresponding basis matrix would satisfy the conditions $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$ and $\mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$ as was shown in Lemma 5.3.3 and Lemma 5.3.5.

Now we are going to extend Lemma 5.3.6 to the case where the linear program P' given in the lemma has a degenerate optimal basic feasible solution.

Lemma 5.3.9.

Given primal program P in canonical form (5.6) and the corresponding dual program D in (5.7). Let P' be the linear program in standard form and equivalent to P given by

$$\textit{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad \textit{subject to}$$

$$\mathbf{G}\mathbf{z} = \mathbf{b}$$

$$\mathbf{z} \geq \mathbf{0}_{m+n} \quad \textit{where } \mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{m \times (m+n)}, \ \mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{R}^{m+n}$$

If P' has an optimal degenerate basic feasible solution, then there exist primal feasible solution $\mathbf{x}$ and dual feasible solution $\mathbf{y}$ such that $\gamma_{\mathbf{xy}} = 0$.

Proof.

It is not hard to see that P' is a linear program equivalent to the primal program P. Let $\mathbf{z}$ be an optimal degenerate basic feasible solution of P'. By Lemma 5.3.8, we see that there exists a basis matrix $\mathbf{B}$ of P' such that $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}^T_{m+n}$ and $\mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}_m$. Since $\mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix}$ and $\mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix}$, the conditions $\mathbf{d}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_{m+n}{}^T$ can be split into two as

$$\mathbf{c}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{A} \leq \mathbf{0}_n{}^T \text{ and } \mathbf{0}_m{}^T - \mathbf{d_B}^T\mathbf{B}^{-1}\mathbf{I}_m \leq \mathbf{0}_m{}^T \tag{5.13}$$

Let $\mathbf{y} \in \mathbb{R}^m$ defined by $\mathbf{y} = (\mathbf{B}^{-1})^T \mathbf{d_B}$. Hence we have

$$\mathbf{c}^T - \mathbf{y}^T\mathbf{A} \leq \mathbf{0}_n{}^T \text{ and } \mathbf{0}_m{}^T - \mathbf{y}^T\mathbf{I}_m \leq \mathbf{0}_m{}^T$$

That is $\mathbf{A}^T\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq \mathbf{0}_m$. Hence $\mathbf{y}$ is a feasible solution of the dual program D. Let $\mathbf{z}'$ be a vector in $\mathbb{R}^{m+n}$ such that $\mathbf{z_B}' = \mathbf{B}^{-1}\mathbf{b} \geq \mathbf{0}_m$ and $\mathbf{z_N}' = \mathbf{0}_n$. By Lemma 5.3.1, we see that $\mathbf{z}'$ is a basic feasible solution of P'. Now by Corollary 5.3.2, we see that $\langle \mathbf{d}, \mathbf{z}' \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{b} \rangle$. Hence $\langle \mathbf{b}, \mathbf{y} \rangle = \langle \mathbf{b}, \mathbf{B}^{-1T}\mathbf{d_B} \rangle = \langle \mathbf{d_B}, \mathbf{B}^{-1}\mathbf{b} \rangle = \langle \mathbf{d}, \mathbf{z}' \rangle$. Since P and P' are equivalent, we see that there exists some feasible solution $\mathbf{x}$ of P such that $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{d}, \mathbf{z}' \rangle$. Thus we conclude that there exist some primal feasible solution $\mathbf{x}$ and dual feasible solution $\mathbf{y}$ such that $\langle \mathbf{c}, \mathbf{x} \rangle = \langle \mathbf{b}, \mathbf{y} \rangle$ and hence $\gamma_{\mathbf{xy}} = 0$. Hence the Lemma. □

Corollary 5.3.4.
Given primal program P in canonical form (5.6) and the corresponding dual program D in (5.7). Let P' be the linear program in standard form and equivalent to P given by

$$\begin{aligned} &\textit{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad \textit{subject to} \\ &\mathbf{Gz} = \mathbf{b} \\ &\mathbf{z} \geq \mathbf{0}_{m+n} \quad \textit{where } \mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{m \times (m+n)}, \ \mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{R}^{m+n} \end{aligned}$$

If P' has an optimal basic feasible solution, then there exist primal feasible solution $\mathbf{x}$ and dual feasible solution $\mathbf{y}$ such that $\gamma_{\mathbf{xy}} = 0$.

Proof.
The result directly follows from Lemma 5.3.6 and Lemma 5.3.9. □

Now we are in a position to prove the strong duality theorem.

5.3.3.4 Proof of Strong Duality Theorem

Theorem 5.3.3. ***(Strong Duality Theorem)***
Given primal program P in canonical form (5.6) and the corresponding dual program D in (5.7). Then

1. *If either P or D has a finite optimum, then so does the other.*
2. *The duality gap beween P and D with respect to the corresponding optimal solutions is zero.*

Proof.
Let P has a finite optimum. Consider the linear program P' which is in standard form and equivalent to P given by

$$\begin{aligned} &\text{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad \text{subject to} \\ &\mathbf{Gz} = \mathbf{b} \\ &\mathbf{z} \geq \mathbf{0}_{m+n} \quad \text{where } \mathbf{G} = \begin{bmatrix} \mathbf{A} & \mathbf{I}_m \end{bmatrix} \in \mathbb{R}^{m \times (m+n)}, \ \mathbf{d} = \begin{bmatrix} \mathbf{c} \\ \mathbf{0}_m \end{bmatrix} \in \mathbb{R}^{m+n} \end{aligned}$$

Since P is equivalent to P' and P has a finite optimum, we see that P' also has a finite optimum. By Theorem 4.4.2 and Theorem 4.2.1, it follows that P' has an optimal basic feasible solution. Hence by Corollary 5.3.4, it follows that there exist primal feasible solution $\mathbf{x}$ and dual feasible solution $\mathbf{y}$ such that $\gamma_{\mathbf{xy}} = 0$. □

Conversely let D has a finite optimum. Consider the linear programs D_1 given by

$$\begin{aligned} D_1\text{:- Maximize } & \langle -\mathbf{b}, \mathbf{y}\rangle \text{ subject to} \\ & -\mathbf{A}^{\mathbf{T}}\mathbf{y} \leq -\mathbf{c} \\ & \mathbf{y} \geq \mathbf{0}_m \end{aligned} \tag{5.14}$$

The dual P_1 of D_1 is given by

$$\begin{aligned} P_1\text{:- Minimize } & \langle -\mathbf{c}, \mathbf{x}\rangle \text{ subject to} \\ & -\mathbf{A}\mathbf{x} \geq -\mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}_n \end{aligned} \tag{5.15}$$

It is not hard to see that P is equivalent to P_1, D is equivalent to D_1, $F(D) = F(D_1)$ and $F(P) = F(P_1)$. Since D has a finite optimum, we see that D_1 also has finite optimumm. Taking D_1 as the primal program and proceeding as above, we see that there exist vectors $\mathbf{x} \in F(P_1)$ and $\mathbf{y} \in F(D_1)$ such that $\langle -\mathbf{c}, \mathbf{x}\rangle = \langle -\mathbf{b}, \mathbf{y}\rangle$. Since $F(D) = F(D_1)$ and $F(P) = F(P_1)$, we see that $\mathbf{x}$ is a primal feasible solution and $\mathbf{y}$ is a dual feasible solution. Since $\langle -\mathbf{c}, \mathbf{x}\rangle = \langle -\mathbf{b}, \mathbf{y}\rangle$, we see that $\langle \mathbf{c}, \mathbf{x}\rangle = \langle \mathbf{b}, \mathbf{y}\rangle$ which implies that the duality gap between P and D with respect to $\mathbf{x}$ and $\mathbf{y}$ is zero. Hence the Theorem.

5.4 Summary

In this chapter, we had presented the fundamentals of primal dual theory. In section 5.2, we introduced the notions of primal and dual linear programs and established that dual of the dual is the primal program itself. In Section 5.3.1, we proved the weak duality theorem and introduced the notion of duality gap. We had seen that the existence of a primal feasible solution and a dual feasible solution with zero duality gap shows that these solutions are optimal to the concerned linear programs. In section 5.3.2, we had discussed the complementary slackness conditions and established that the existence of a primal feasible solution and dual feasible solution with zero duality gap is possible if and only if the primal feasible solution is orthogonal to the dual surplus vector and the dual feasible solution is orthogonal to the primal slack vector. In section 5.3.3, we proceeded to discuss the strong duality theorem. As a prerequisite to prove this theorem, we had introduced the notions of basis matrix, degenerate and non-degenerate basic feasible solutions with reference to a linear program in standard form and established that in case a linear program in standard form that is equivalent to the primal program has an optimal basic feasible solution, there exist a primal feasible solution and a dual feasible solution with zero duality gap. The chapter was completed by giving a proof of strong duality theorem using the results established in section 5.3.3.

Chapter 6

Conclusion

6.1 Thesis Summary

In this thesis, we have presented the structural geometry of linear programming and used it to derive the classical results like Caratheodory characterisation theorem and fundamental theorem of linear programming and algebraic results for deriving the duality theorem in primal dual theory. In chapter 2, we introduced the notions of recession directions and recession cones and established that a polyhedral set is bounded if and only if it has no recession directions. In this chapter, we have brought about the notions of vertices and extreme points in convex sets. In chapter 3, we have formally defined linear programming problem and discussed the notions of feasible linear programs, unbounded linear programs and equivalence of linear programs. We had seen three different forms of linear programs and established that these three forms are equivalent. In chapter 4, we introduced the notion of basic feasible solutions and established that the two geometric notions - vertices and extreme points are equivalent to the algebraic notion - basic feasible solutions. In this chapter, we introduced the notions of extreme directions and normalised recession directions. It was shown that the feasible region of every linear program has finite non-zero number of extreme points and finite number of extreme directions. We had characterised the feasible region of a linear program using the extreme points and the recession cone of the feasible region. We established that the recession cone of an unbounded linear program can be characterised as the conic hull of the normalised extreme directions of the feasile region. Using these structural properties of the feasible region, we proved the Caratheodory characterisation theorem and the Fundamental theorem of linear programming. Following these discussions, we presented a brute force algorithm for linear programming and illustrated the way in which this algorithm can be executed graphically. In chapter 5, we have explained the notion of dual of a linear program and derived the weak duality theorem and the complementary slackness conditions. Our discussion is completed by giving a proof of strong duality theorem based on some algebraic results formulated using the structural properties investigated in the preceeding chapters.

6.2 Pros and Cons of the Work

Unlike the classical methods, the techniques and strategies used in this theis for deriving various results give more emphasis to the underlying geometry. In our opinion, this geometric approach to the subject gives better visual intuition. The classical approach to prove the duality theorem are based on Theorem of Alternatives or methods from multivariable calculus, both being somewhat mathematically sophisticated. In this work, core results like Caratheodory characterisation theorem and Fundamental theorem of linear programming are derived essentially on the basis of the geometric characterisation of polyhedral sets. This enables the reader to form a strong foundation in linear optimization without having complicated mathematical prerequisites. Moreover the algebraic results derived to prove the strong duality theorem in fact form the core results in the simplex method. This makes the reader to have a detailed study of the simplex method with lesser effort.

The difficulty we face in this approach is the technicalities involved in surmounting the notion of degeneracy. The effort to be taken to address this issue requires certain technical results which are given in Appendix A. In fact degenracy does not need any kind of special treatment in other methods of proving duality theorem.

We believe that the thesis may be used as a teaching material to cover a graduate level course on the foundations of linear optimization.

6.3 Recommendations for Future Work

Even though the present work fullfills its objectives, it may be extended in several ways. Some of these extensions are given below.

1. Devise new strategies for deriving the duality theorem without having a special treatment on degenerate basic feasible solutions.

2. Investigate the geometry of primal dual relationship and derive the classical results like Farkas Lemma and Theorem of Alternatives as its consequence.

3. Using the ideas in this thesis, establish the primal dual algorithm and its characteristics.

Appendix A

Proof of Lemma 5.3.8

Before proving Lemma 5.3.8, we prove the following technical lemma [9].

Lemma A.0.1.
Let $\mathbf{r} = \begin{bmatrix} r_0 & r_1 & r_2 & \cdots & r_n \end{bmatrix}^T \in \mathbb{R}^{n+1}$ *such that* $\mathbf{r} \neq \mathbf{0}_{n+1}$. *Let* k *be the least index in* $\{0, 1, 2, \cdots, n\}$ *such that* $r_k \neq 0$. *Let* $\alpha = \max\{|r_j| \big| r_j \neq 0$ *and* $0 \leq j \leq n\}$ *and* $\beta = \min\{|r_j| \big| r_j \neq 0$ *and* $0 \leq j \leq n\}$. *Let* ϵ *be any real number such that* $0 < \epsilon < \frac{\beta}{\alpha+\beta}$. *Then* $\sum_{j=0}^{n} r_j \epsilon^j > 0$ *if* $r_k > 0$ *and* $\sum_{j=0}^{n} r_j \epsilon^j < 0$ *if* $r_k < 0$.

Proof.
We have $\sum_{j=0}^{n} r_j \epsilon^j = r_k \epsilon^k + \sum_{j=k+1}^{n} r_j \epsilon^j = \epsilon^k \left(r_k + \sum_{j=k+1}^{n} r_j \epsilon^{j-k} \right)$. Now we shall consider the following two cases.

Case 1:- $r_k > 0$

In this case, we see that $\beta \leq r_k$. Since $\alpha = \max\{| r_j | \, |r_j \neq 0$ and $0 \leq j \leq n\}$, we see that $-\alpha \leq r_j$ for each $j \in \{0, 1, 2, \cdots, n\}$. Hence we have

$$\begin{aligned}
\sum_{j=0}^{n} r_j \epsilon^j &= \epsilon^k \left(r_k + \sum_{j=k+1}^{n} r_j \epsilon^{j-k} \right) \\
&\geq \epsilon^k \left(\beta - \alpha \sum_{j=1}^{n-k} \epsilon^j \right) \\
&\geq \epsilon^k \left(\beta - \alpha \sum_{j=1}^{\infty} \epsilon^j \right) \\
&= \epsilon^k \left(\beta - \frac{\alpha \epsilon}{1-\epsilon} \right) \\
&= \epsilon^k \left(\frac{\beta - \epsilon(\alpha+\beta)}{1-\epsilon} \right) \\
&> 0 \qquad (\because 0 < \epsilon < \frac{\beta}{\alpha+\beta})
\end{aligned}$$

Case 2:- $r_k < 0$

In this case, we see that $-\beta \leq r_k$. Since $\alpha = \max\{| r_j | \, |r_j \neq 0$ and $0 \leq j \leq n\}$, we see that

$\alpha \geq r_j$ for each $j \in \{0, 1, 2, \cdots, n\}$. Hence we have

$$\begin{aligned}\sum_{j=0}^{n} r_j \epsilon^j &= \epsilon^k \left(r_k + \sum_{j=k+1}^{n} r_j \epsilon^{j-k} \right) \\ &\leq \epsilon^k \left(-\beta + \alpha \sum_{j=1}^{n-k} \epsilon^j \right) \\ &\leq \epsilon^k \left(-\beta + \alpha \sum_{j=1}^{\infty} \epsilon^j \right) \\ &= -\epsilon^k \left(\beta - \frac{\alpha\epsilon}{1-\epsilon} \right) \\ &= -\epsilon^k \left(\frac{\beta - \epsilon(\alpha+\beta)}{1-\epsilon} \right) \\ &< 0 \qquad (\because 0 < \epsilon < \frac{\beta}{\alpha+\beta})\end{aligned}$$

Hence the Lemma. □

Remark A.0.1.

Lemma A.0.1 informally states that the sum $\sum_{j=0}^{n} r_j \epsilon^j$ is positive or negative according to whether the first non zero component of the vector $\mathbf{r}$ is positive or negative.

Now we are going to prove Lemma 5.3.8. In this proof, we construct a new linear program which will always have an optimal non-degenerate basic feasible solution by suitably perturbing the given linear program. The existence of the non-degenerate optimal basic feasible solution of the new linear program can be shown using Lemma A.0.1. While proving this lemma, we follow the notations specified in Remark 5.3.6.

Lemma A.0.2.

Let P be the linear program in the standard form given by

$$\begin{aligned} &\textit{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \quad \textit{subject to} \\ &\mathbf{G}\mathbf{z} = \mathbf{h} \\ &\mathbf{z} \geq \mathbf{0}_n \quad \textit{where} \end{aligned} \tag{A.1}$$

$\mathbf{G} \in \mathbb{R}^{m \times n}$ *such that rank*$(\mathbf{G}) = m$, $\mathbf{d} \in \mathbb{R}^n$ *and* $\mathbf{h} \in \mathbb{R}^m$*. If P has an optimal degenerate basic feasible solution, then there exists a basis matrix* $\mathbf{B}$ *of P such that* $\mathbf{d}^T - \mathbf{d_B}^T \mathbf{B}^{-1} \mathbf{G} \leq \mathbf{0}_n{}^T$ *and* $\mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$.

Proof.

Assume that P has an optimal degenerate basic feasible solution. Let H_P be the set of all basis matrices of P. Since P has a basic feasible solution, we see that H_P is non-empty. Let $\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots & \mathbf{g}_n \end{bmatrix}$. Let $\mathbf{B}$ be any member of H_P and $\mathbf{u}^1, \mathbf{u}^2, \mathbf{u}^3, \cdots, \mathbf{u}^m \in \mathbb{R}^{n+1}$ defined by $\mathbf{u}^i = \begin{bmatrix} {u^i}_0 & {u^i}_1 & {u^i}_2 & \cdots & {u^i}_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{h} \rangle & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_1 \rangle & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_2 \rangle & \cdots & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_n \rangle \end{bmatrix}$ for each $i \in \{1, 2, 3, \cdots, m\}$. It is easy to see that $\mathbf{u}^i$ is the i^{th} row of $\mathbf{B}^{-1}\mathbf{G}$. Since $\mathbf{B}$ is a submatrix of $\mathbf{G}$, we see that $\mathbf{B}\mathbf{B}^{-1} = \mathbf{I}_m$ will be a submatrix of $\mathbf{B}^{-1}\mathbf{G}$ and hence $\mathbf{u}^i$ will have at least one unity entry. Hence it follows that $\mathbf{u}^i \neq \mathbf{0}_{n+1}$ for each $i \in \{1, 2, 3, \cdots, m\}$.

Let $\alpha_i = \max\{|u^i_j| \,\Big|\, u^i_j \neq 0 \text{ and } 0 \leq j \leq n\}$ and $\beta_i = \max\{|u^i_j| \,\Big|\, u^i_j \neq 0 \text{ and } 0 \leq j \leq n\}$ for each $i \in \{1, 2, 3 \cdots, m\}$. Let $\alpha_{\mathbf{B}} = \max\{\alpha_i \mid 1 \leq i \leq m\}$ and $\beta_{\mathbf{B}} = \min\{\beta_i \mid 1 \leq i \leq m\}$.

Let $\alpha = \max\{\alpha_{\mathbf{B}} \mid \mathbf{B} \in H_{\mathbf{B}}\}$ and $\beta = \min\{\beta_{\mathbf{B}} \mid \mathbf{B} \in H_{\mathbf{B}}\}$. Let ϵ be a real number such that $0 < \epsilon < \frac{\beta}{\alpha+\beta}$.

Consider the linear program P' given by

$$\text{Maximize } \langle \mathbf{d}, \mathbf{z} \rangle \text{ subject to}$$
$$\mathbf{G}\mathbf{z} = \mathbf{h} + \sum_{j=1}^{n} \mathbf{g}_j \epsilon^j \tag{A.2}$$
$$\mathbf{z} \geq \mathbf{0}_n$$

Claim.

P' has an optimal non-degenerate basic feasible solution.

Proof.

Let $\mathbf{z}$ be any feasible solution of P. Then we have $\mathbf{G}\mathbf{z} = \mathbf{h}$ and $\mathbf{z} \geq \mathbf{0}_n$. Let $\mathbf{z}' = \mathbf{z} + \begin{bmatrix} \epsilon & \epsilon^2 & \epsilon^3 & \cdots & \epsilon^n \end{bmatrix}^T$. Since $\epsilon > 0$ and $\mathbf{z} \geq \mathbf{0}_n$, it follows that $\mathbf{z}' > \mathbf{0}_n$. Furthermore $\mathbf{G}\mathbf{z}' = \mathbf{G}\mathbf{z} + \sum\limits_{j=1}^{n} \mathbf{g}_j \epsilon^j = \mathbf{h} + \sum\limits_{j=1}^{n} \mathbf{g}_j \epsilon^j$. Thus we see that $\mathbf{z}'$ is a feasible solution of P'. Hence we infer that P' is a feasible linear program. By Lemma 4.4.4, we see that P' has an optimal basic feasible solution. By Lemma 5.3.1, it follows that there exists a basis matrix $\mathbf{B}$ of P' such that $\mathbf{z}'_{\mathbf{B}} = \mathbf{B}^{-1}\left(\mathbf{h} + \sum\limits_{j=1}^{n} \mathbf{g}_j \epsilon^j\right) \geq \mathbf{0}_m$ and $\mathbf{z}'_{\mathbf{N}} = \mathbf{0}_{n-m}$. Since the constraint matrix of both P and P' are the same, we see that every basis matrix of P' is a basis matrix of P and hence $\mathbf{B} \in H_P$.

Let $\mathbf{r}^i = \begin{bmatrix} r^i{}_0 & r^i{}_1 & r^i{}_2 & \cdots & r^i{}_n \end{bmatrix} = \begin{bmatrix} \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{h} \rangle & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_1 \rangle & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_2 \rangle & \cdots & \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_n \rangle \end{bmatrix}$ for each $i \in \{1, 2, 3, \cdots, m\}$. Since $\mathbf{B}$ is a submatrix of $\mathbf{G} = \begin{bmatrix} \mathbf{g}_1 & \mathbf{g}_2 & \mathbf{g}_3 & \cdots & \mathbf{g}_n \end{bmatrix}$, we see that at least one of the entries in each row of the vector $\mathbf{r}^i$ is unity. Hence it follows that $\mathbf{r}^i \neq \mathbf{0}_{n+1}$ for each $i \in \{1, 2, 3, \cdots, m\}$.

By the choice of α and β, it follows that $\alpha = \max\{|r^i{}_j| \,\big|\, r^i{}_j \neq 0 \text{ and } 0 \leq j \leq n\}$ and $\beta = \min\{|r^i{}_j| \,\big|\, r^i{}_j \neq 0 \text{ and } 0 \leq j \leq n\}$. Since $\mathbf{z}'_{\mathbf{B}} = \mathbf{B}^{-1}\left(\mathbf{h} + \sum\limits_{j=1}^{n} \mathbf{g}_j \epsilon^j\right) \geq \mathbf{0}_m$, we see that $\langle \mathbf{e}_i, \mathbf{z}'_{\mathbf{B}} \rangle = \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{h} \rangle + \sum\limits_{j=1}^{n} \langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{g}_j \rangle \epsilon^j = \sum\limits_{j=0}^{n} r^i{}_j \epsilon^j \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. By Lemma A.0.1, the following observations can be made.

1. Since $\langle \mathbf{e}_i, \mathbf{z}'_{\mathbf{B}} \rangle \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$, we see that the first non-zero component of the vector $\mathbf{r}^i$ is not negative and therefore positive for each $i \in \{1, 2, 3, \cdots, m\}$. This implies that $\langle \mathbf{e}_i, \mathbf{z}'_{\mathbf{B}} \rangle > 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. Thus $\mathbf{z}'_{\mathbf{B}} > 0$.

2. Since $\langle \mathbf{e}_i, \mathbf{z}'_{\mathbf{B}} \rangle \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$, we see that the first component of the vector $\mathbf{r}^i$ is not negative for each $i \in \{1, 2, 3, \cdots, m\}$. This implies that $\langle \mathbf{e}_i, \mathbf{B}^{-1}\mathbf{h} \rangle \geq 0$ for each $i \in \{1, 2, 3, \cdots, m\}$. Thus we see that $\mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$.

To summarise, we see that there exists a basis matrix $\mathbf{B}$ of P such that $\mathbf{z}'_{\mathbf{B}} > \mathbf{0}_m$ and $\mathbf{z}'_{\mathbf{N}} = \mathbf{0}_{n-m}$. By Lemma 5.3.3, we see that $\mathbf{z}'$ is a non degenerate basic feasible solution of P'. □

By Corollary 5.3.3, we see that $\mathbf{d}^T - \mathbf{d_B}^T \mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$. Thus we conclude that there exists a basis matrix $\mathbf{B}$ of P such that $\mathbf{d}^T - \mathbf{d_B}^T \mathbf{B}^{-1}\mathbf{G} \leq \mathbf{0}_n{}^T$ and $\mathbf{B}^{-1}\mathbf{h} \geq \mathbf{0}_m$. Hence the Lemma. □

Appendix B

Linear Algebra Fundamentals

In this Appendix, we have outlined the notions and the results in Linear algebra which were used in the preceeding chapters without including proofs [10] [11]. A detailed discussion on these topics may be found in any standard textbook on linear algebra.

B.1 Vector Spaces

Definition B.1.1.
Let $\mathbb{F}$ be a Field whose elements are called *scalars*. Then a *vector space* S over $\mathbb{F}$ is triplet $S = (E, +, \cdot\,)$ where

(a) E is a non-empty set of elements called *vectors*.

(b) $+$ is called the *vector addition* operator which operates on any two vectors in E and produce some vector in E. The vector addition of any two vectors $\mathbf{x}$ and $\mathbf{y}$ in E is denoted by $\mathbf{x} + \mathbf{y}$. Moreover the vector addition operation satisfies the following properties.

   (a) Vector addition is Commutative. That is for any two vectors $\mathbf{x}, \mathbf{y} \in E, \mathbf{x}+\mathbf{y} = \mathbf{y}+\mathbf{x}$.
   (b) Vector addition is associative. That is for any three vectors $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E, \mathbf{x}+(\mathbf{y}+\mathbf{z}) = (\mathbf{x}+\mathbf{y})+\mathbf{z}$.
   (c) There exists a zero vector $\mathbf{0}_S$ in E such that $\mathbf{0}_S + \mathbf{x} = \mathbf{x}$ for any vector $\mathbf{x} \in E$.
   (d) For each vector $\mathbf{x} \in E$, there exists some vector $-\mathbf{x}$ called the additive inverse of $\mathbf{x}$ in E such that $\mathbf{x} + -\mathbf{x} = \mathbf{0}_S$.

(c) $\cdot$ is called *scalar multiplication* operator which takes some scalar from $\mathbb{F}$ and a vector from E and produces some vector in E. The scalar multiplication of any scalar α in $\mathbb{F}$ and vector $\mathbf{x}$ in E is denoted by $\alpha\mathbf{x}$. Moreover the scalar multiplication operation satisfies the following properties.

   (a) Scalar muliplication is distributive over vector addition. That is for any two vectors $\mathbf{x}, \mathbf{y} \in E$ and $\alpha \in \mathbb{F}$, $\alpha(\mathbf{x}+\mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$.
   (b) Scalar muliplication is distributive over scalar addition. That is for any $\alpha, \beta \in \mathbb{F}$ and $\mathbf{x} \in E$, $(\alpha+\beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{y}$.
   (c) Scalar mulitplication is compatible with multiplication of the field elements. That is for $\alpha, \beta \in \mathbb{F}$ and $\mathbf{x} \in E$, $(\alpha\beta)\mathbf{x} = \alpha(\beta\mathbf{x})$.
   (d) For each $\mathbf{x} \in E, 1 \cdot \mathbf{x} = \mathbf{x}$ where 1 is the multiplicative identity of F.

Example B.1.1.
Given $m, n \in \mathbb{N}$. Then the set of all $m \times n$ matrices over the field $\mathbb{R}$ with usual definition of matrix addition and scalar multiplication a vector space. This vector space is usually denoted

by $\mathbb{R}^{m\times n}$. In particular $\mathbb{R}^{m\times 1}, m \geq 2$ is usually abreviated as $\mathbb{R}^m$, which is nothing but the vector spce of $m \times 1$ column vectors over the field of real numbers.

Definition B.1.2.
Let S be a vector space over the field $\mathbb{F}$ and $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be vectors in S. Then a *linear combination* of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ is an expression of the form $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i$ where $\alpha_i \in \mathbb{F}, 1 \leq i \leq k$.

Definition B.1.3.
Let S be a vector space over the field $\mathbb{F}$ and $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be vectors in S. Then the *span* of $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}$ is denoted by $Span\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\}$ and is defined as the set of vectors in S obtained by the linear combinations of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$.

$$\text{Span}\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k\} = \left\{ \sum_{i=1}^{k} \alpha_i \mathbf{x}_i \;\middle|\; \alpha_i \in \mathbb{F}, 1 \leq i \leq k \right\}$$

Definition B.1.4.
Let S be a vector space over the field $\mathbb{F}$ and $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ be vectors in S. Then these vectors are said to be *linearly independent* if $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = \mathbf{0}_S$, then we must have $\alpha_i = 0, \forall i$ such that $1 \leq i \leq k$. Otherwise they are called *linearly dependent* .

Definition B.1.5.
Let S be a vector space over the field $\mathbb{F}$ and $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k$ be vectors in S. Then the set $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k\}$ is said to be a *basis* of S if

(a) $Span\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k\} = S$.

(b) $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k$ are linearly independent.

Fact B.1.1. ***(Unique Representation Theorem)***
Let S be a vector space over the field $\mathbb{F}$ and $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k\}$ be a basis of V. Then any vector in S can be expressed a linear combination of $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3, \cdots, \mathbf{b}_k$ in exactly one way.

Definition B.1.6.
Let S be a vector space over the field $\mathbb{F}$. Then S is said to be *a finite dimensional vector space* if S has a finite basis. Otherwise S is called an *infinite dimensional vector space*.

Fact B.1.2.
All bases of a finite dimensional vector space have the same cardinality.

Definition B.1.7.
Let S be a vector space over the field $\mathbb{F}$. Then the *dimension* of S is denoted by $dim(S)$ and is defined as the cardinality of its bases.

Fact B.1.3.
Let S be a vector space over the field $\mathbb{F}$ with $dim(S) = n$. Then any set of more than n vectors in S is linearly dependent.

Definition B.1.8.
Let $S = (E, +, \cdot\,)$ be a vector space over the field $\mathbb{F}$. Then any vector space $S' = (E', +, \cdot\,)$ where $E' \subseteq E$ is called a *subspace* of S.

Definition B.1.9.
Let S be a vector space over the field $\mathbb{F}$. Then the function $\langle , \rangle : S \times S \to \mathbb{R}$ is said to be an *inner product* or *dot product* on S if it satisfies the following properties.

(a) Symmetry Property:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle, \text{ for each } \mathbf{x}, \mathbf{y} \text{ in } S.$$

(b) Linearity Property:

$$\left\langle \sum_i \alpha_i \mathbf{x}_i, \mathbf{y} \right\rangle = \sum_i \langle \mathbf{x}_i, \mathbf{y} \rangle \text{ for each } \alpha_i \in \mathbb{F} \text{ and each } \mathbf{x}_i, \mathbf{y} \textit{ in } S$$
$$\left\langle \mathbf{x}, \sum_i \alpha_i \mathbf{y}_i \right\rangle = \sum_i \langle \mathbf{x}, \mathbf{y}_i \rangle \text{ foreach } \alpha_i \in \mathbb{F} \text{ and } \mathbf{y}_i, \mathbf{x} \textit{ in } S$$

(c) Positive Definiteness Property: For each $\mathbf{x} \in S \setminus \{\mathbf{0}_S\}, \langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{0}_S, \mathbf{0}_S \rangle = 0$.

Definition B.1.10.
An *inner product space* is a pair $(S, \langle, \rangle)$ where S is a vector space and $\langle, \rangle$ is an inner product on S.

Example B.1.2.
The Eucledian Space $(\mathbb{R}^n, \langle, \rangle)$ is an inner product space where the inner product of any two vectors $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 & \cdots & x_n \end{bmatrix}^T$ and $\mathbf{y} = \begin{bmatrix} y_1 & y_2 & y_3 & \cdots & y_n \end{bmatrix}^T \in \mathbb{R}^n$ is defined as $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$. This inner product is known as the *standard inner product* for $\mathbb{R}^n$.

Definition B.1.11.
Let $I = (S, \langle, \rangle)$ be an inner product space. Then any two vectors $\mathbf{x}$ and $\mathbf{y}$ of I is said to be *orthogonal* if and only if $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

Definition B.1.12.
Let S be a vector space over the field $\mathbb{F}$. Then the function $\|\cdot\| : S \to \mathbb{R}$ is said to be a *norm* of S if it satisfies the following properties.

(a) For each $\mathbf{x}$ in S such that $\mathbf{x} \neq \mathbf{0}_S$, it must be the case that $\|\mathbf{x}\| \geq 0$ and $\|\mathbf{0}_S\| = 0$.

(b) For each $\mathbf{x}$ in S and $\alpha \in \mathbb{F}$, it must be the case that $\|\alpha \mathbf{x}\| = \mid \alpha \mid \|\mathbf{x}\|$.

(c) For $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ in S, it must be the case that $\left\| \sum_{i=1}^{k} \mathbf{x}_i \right\| \leq \sum_{i=1}^{k} \|\mathbf{x}_i\|$.

Definition B.1.13.
A *normed vector space* is a pair $(S, \|\cdot\|)$ where S is a vector space and $\|\cdot\|$ is a norm of S.

Example B.1.3.
$(\mathbb{R}^n, \|\cdot\|)$ is a normed vector space where the norm $\|\cdot\|$ can be defined in several ways.

(a) Let p be a finite positive integer. Then the l_p norm or p-norm of any vector $\mathbf{x} = \begin{bmatrix} x_1, x_2, x_3, \cdots, x_n \end{bmatrix}^T \in \mathbb{R}^n$ is denoted by $\|\mathbf{x}\|_p$ and is defined as

$$\|\mathbf{x}\|_p = \sqrt[p]{\sum_{i-1}^{n} \mid x_i \mid^p}$$

(b) the ∞-norm is denoted by $\|\mathbf{x}\|_\infty$ and is defined as

$$\|\mathbf{x}\|_\infty = \max\{x_i \mid 1 \leq i \leq n\}.$$

Among various norms of $\mathbb{R}^n$, the usual choice of norm is the l_2 norm. The l_2 norm of any vector $\mathbf{x} = \left[x_1, x_2, x_3, \cdots, x_n\right]^T \in \mathbb{R}^n$ is simply denoted by $\|\mathbf{x}\|$ and is given by

$$\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$$

This norm computes the magnitude of the vector $\mathbf{x}$ and is known as the *standard norm* on $\mathbb{R}^n$. It is easy to see that $\|x\| = \sqrt{\langle x, x\rangle}$.

Facts B.1.1.

1. *The l_1 norm of any vector $\mathbf{x} \in \mathbb{R}^n$ is the sum of the absolute values of the individual components of $\mathbf{x}$. In particular, if $\mathbf{x}$ is a non-negative vector, then l_1 norm of $\mathbf{x}$ is the sum of the individual components of $\mathbf{x}$.*

2. *The l_2 norm of any non-negative vector $\mathbf{x} \in \mathbb{R}^n$ is at most the l_1 norm of $\mathbf{x}$.*

3. *All norms in $\mathbb{R}^n$ are equivalent. That is for any $p, q \in \mathbb{N}$ and $\mathbf{x} \in \mathbb{R}^n$, there exist positve real numbers α and β such that $\alpha \|x\|_p \leq \|x\|_q \leq \beta \|x\|_p$.*

Definition B.1.14.
A *metric space* is a pair (S, d) where S is a vector space and $d : S \times S \to \mathbb{R}$ is a function that satisfies the following properties.

(a) $d(\mathbf{x}, \mathbf{y}) \geq 0$ for each $\mathbf{x}$ and $\mathbf{y}$ in S.

(b) $d(\mathbf{x}, \mathbf{y}) = 0$ if and only if $\mathbf{x} = \mathbf{y}$.

(c) $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ for each $\mathbf{x}, \mathbf{y}$ and $\mathbf{z}$ in S.

(d) $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$ for each $\mathbf{x}$ and $\mathbf{y}$ in S

Example B.1.4.
$(\mathbb{R}, d)$ is a metric space where $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is a function defined by $d(x, y) =\mid x - y \mid$. We denote this matric space by the pair $(\mathbb{R}, | \dot{} |)$
$(\mathbb{R}^n, d)$ is a metric space where $d : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a function defined by $d(\mathbf{x}, \mathbf{y}) = \|x - y\|_p$ for some $p \in \mathbb{N}$. If $p = 1$, then $d(\mathbf{x}, \mathbf{y})$ computes the manhattan distance between $\mathbf{x}$ and $\mathbf{y}$. If $p = 2$, then $d(\mathbf{x}, \mathbf{y})$ computes the eucledian distance between the vectors $\mathbf{x}$ and $\mathbf{y}$.

B.2 Linear Transformations

Definition B.2.1.
Given vector spaces S_1 and S_2 over the field $\mathbb{F}$. Then a function $T : S_1 \to S_2$ is said to be a *linear transformation* if for any vectors $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \cdots, \mathbf{x}_k$ in S, we must have $T\left(\sum_{i=1}^{k} \alpha_i \mathbf{x}_i\right) = \sum_{i=1}^{k} \alpha_i T(\mathbf{x}_i), \alpha_i \in \mathbb{F}, 1 \leq i \leq k$.

Example B.2.1.
The function $T : \mathbb{R}^n \to \mathbb{R}^m$ defined by $T(\mathbf{x}) = \mathbf{A}\mathbf{x}$ for some $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a linear transformation.

Facts B.2.1.
For any $m, n \in \mathbb{N}$, any linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ is continuous.

Definition B.2.2.

Given vector spaces S_1 and S_2 over the field $\mathbb{F}$ and linear transformation $T : S_1 \to S_2$. Then

(a) the *kernel* or *null space* of T is denoted by $Ker(T)$ and is defined as the set

$$Ker(T) = \{\mathbf{x} \mid \mathbf{x} \in S_1 \text{ and } T(\mathbf{x}) = \mathbf{0_{S_2}}$$

(b) the *range* of T and is denoted by $Range(T)$ and is defined as the set

$$Range(T) = \{\mathbf{y} \mid \mathbf{y} \in S_2 \text{ such that } \mathbf{y} = T(\mathbf{x}) \text{ for some } \mathbf{x} \in S_1\}$$

Facts B.2.2.

Given vector spaces S_1 and S_2 over the field $\mathbb{F}$ and a linear transformation $T : S_1 \to S_2$. Then

(a) $\mathbf{0_{S_1}} \in Ker(T)$.

(b) $Ker(T)$ and $Range(T)$ are subspaces of S_1 and S_2 respectively.

Definition B.2.3.

Given vector spaces S_1 and S_2 over the field $\mathbb{F}$ and linear transformation $T : S_1 \to S_2$. Then the *rank* of T is denoted by $rank(T)$ and is defined as the dimension of the range of T. The *nullity* of T is denoted by $nullity(T)$ and is defined as the dimension of the kernel of T. That is $rank(T) = dim(Range(T))$ and $nullity(T) = dim(ker(T))$.

Fact B.2.1. ***(Rank-Nullity Theorem / Dimension Theorem)***

Given finite dimensional vector spaces S_1 and S_2 over the field $\mathbb{F}$ and linear transformation $T : S_1 \to S_2$. Then we must have $rank(T) + nullity(T) = dim(S_1)$.

Appendix C

Fundamentals of Real Analysis

In this Appendix, we outline the notions of closed sets, open sets, bounded sets and compact sets. It also contains the theorems and results from real analysis which we used in the previous chapters. These results are stated without any proofs and illustrative examples [4]. A detailed discussion of these topics can be seen in any standard textbook on Topology such as the book *Introduction to Topology and Modern Analysis* by *G.F. Simmons*.

C.1 Closed Sets, Open Sets and Bounded Sets

Definition C.1.1.
Let $\mathbf{x}$ be a point in a metric space (S, d) and $\delta \in \mathbb{R}^+$. Then the δ - neighbourhood of $\mathbf{x}$ is denoted by $N_\delta(\mathbf{x})$ and is defined as $N_\delta(\mathbf{x}) = \{\mathbf{y} \mid \mathbf{y} \in S \text{ and } d(\mathbf{x}, \mathbf{y}) < \delta\}$.

Definition C.1.2.
Let (S, d) be a metric space and $T \subseteq S$. Then a sequence in T is a mapping $\phi : \mathbb{N} \to T$ and is denoted by $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$. If $\phi(k) = \mathbf{x}_k$, we say that $\mathbf{x}_k$ is the k^{th} term of the sequence.

Definition C.1.3.
Let (S, d) be a metric space and $T \subseteq S$. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence in $\{n_k\}_{k\in\mathbb{N}}$ be a sequence of increasing natural numbers (i.e $n_i > n_j \Leftrightarrow i > j$). Then $\{\mathbf{x}_{n_k}\}_{k\in\mathbb{N}}$ is called a *subsequence* of the sequence $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$.

Definition C.1.4.
A sequence $\{x_n\}_{n\in\mathbb{N}}$ in the metric space $(\mathbb{R}, |\dot{}|)$ is said to be *non-decreasing* if $x_{n+1} \geq x_n \forall n \in \mathbb{N}$. Similarly the sequence is said to be *non-increasing* if $x_{n+1} \leq x_n \forall n \in \mathbb{N}$. In particular, the sequence is said to be *strictly increasing* if $x_{n+1} > x_n \forall n \in \mathbb{N}$ whereas the sequence is said to be *strictly decreasing* if $x_{n+1} < x_n \forall n \in \mathbb{N}$. A sequence is said to be *monotonic* if it is either non-decreasing or non-increasing.

Definition C.1.5.
Let (S, d) be a metric space and $T \subseteq S$. Let $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be a sequence in T. Then any point $\mathbf{x}_0 \in T$ is said to be the *limit* of the sequence if for all $\delta > 0$, there exists $n_0 > 0$ such that $n \geq n_0 \Rightarrow \mathbf{x}_n \in N_\delta(\mathbf{x}_0)$. If $\mathbf{x}_0$ is the limit of the sequence $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$, we write $\lim_{n\to\infty} \mathbf{x}_n = \mathbf{x}_0$ or simply $\{\mathbf{x}_n\} \to \mathbf{x}_0$.

Fact C.1.1.
Let (S, d) be a metric space and $T \subseteq S$. Let $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be a sequence in T such that $\{\|\mathbf{x}_n\|\} \to \infty$. Then for any vector $\mathbf{x} \in S$, the sequence $\{\mathbf{x}_n - \mathbf{x}\}_{n\in\mathbb{N}}$ belongs to S and $\{\|\mathbf{x}_n - \mathbf{x}\|\} \to \infty$.

Fact C.1.2.
Let $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be a sequence in $\mathbb{R}^n$ defined by $\mathbf{x}_n = \mathbf{x} + n\mathbf{y}$ for some $\mathbf{x}, \mathbf{y}$ in $\mathbb{R}^n$. Then $\{\|\mathbf{x}_n\|\} \to \infty$.

Definition C.1.6.
Let (S, d) be a metric space and $T \subseteq S$. Let $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be a sequence in T. Then the sequence is said to be *convergent* if and only if there exists $\mathbf{x}$ in T such that $\{\mathbf{x}_n\} \to \mathbf{x}$. If there is no such $\mathbf{x}$ in T, then the sequence is called *divergent*.

Fact C.1.3.
Let (S, d) be a metric space and $T \subseteq S$. Let $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ be a sequence in T such that $\{\mathbf{x}_n\} \to \mathbf{x}$ for some $\mathbf{x} \in S$. Then for every subsequence $\{\mathbf{y}_n\}_{n\in\mathbb{N}}$ of $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$, we must have $\{\mathbf{y}_n\} \to \mathbf{x}$.

Definition C.1.7.
Let (S, d) be a metric space and $T \subseteq S$. Then

(a) any point $\mathbf{x} \in S$ is said to be a *limit point* of T if for all $\delta > 0$, there exists $\mathbf{y}$ in T such that $\mathbf{x} \neq \mathbf{y}$ and $\mathbf{y} \in N_\delta(\mathbf{x})$.

(b) any point $\mathbf{x} \in S$ is said to be an *isolated point* of T if there exists $\delta > 0$ such that $(N_\delta(\mathbf{x}) \setminus \{\mathbf{x}\}) \cap T = \phi$.

(c) any $\mathbf{x} \in S$ is said to be a *closure point* of T if for all $\delta > 0$, there exists $\mathbf{y} \in T$ such that $\mathbf{y} \in N_\delta(\mathbf{x})$.

Definition C.1.8.
Let (S, d) be a metric space and $T \subseteq S$. Then

(a) any point $\mathbf{x} \in T$ is said to be an *interior point* of T if there exists some $\delta > 0$ such that $N_\delta(\mathbf{x}) \subseteq T$.

(b) any point $\mathbf{x} \in T$ is said to be a *boundary point* of T if for all $\delta > 0$, it must be the case that $N_\delta(\mathbf{x}) \cap T^c \neq \phi$.

Definition C.1.9.
Let (S, d) be a metric space and $T \subseteq S$. Then the *closure of* T in S is denoted by $\overline{T}$ and is defined as the set of all closure points of T.

Definition C.1.10.
Let (S, d) be a metric space and $T \subseteq S$. Then T is said to be *closed* if T contains all of its limit points. T is said to be *open* if for all $\mathbf{x}$ in T, there exists $\delta > 0$ such that $N_\delta(\mathbf{x}) \subseteq T$.

Fact C.1.4.
Let (S, d) be a metric space and $T \subseteq S$. Then the following statements are equivalent.

(a) T is closed.

(b) $T = \overline{T}$.

(c) Corresponding to each sequence $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in T such that $\{\mathbf{x}_n\} \to \mathbf{x}_0$, it must be the case that $\mathbf{x}_0 \in T$.

(d) T^c is open.

Fact C.1.5.
Let (S, d) be a metric space. Then

(a) the union of any number of open sets over S is open.

(b) the intersection of a finite number of open sets over S is open.

Fact C.1.6.
Let (S, d) be a metric space. Then

(a) *the union of a finite number of closed sets over S is closed.*

(b) *the intersection of any number of closed sets over S is closed.*

Definition C.1.11.
Let (S, d) be a metric space. Let $\mathbf{x}$ be any point in S and r be a positive real number.Then

(a) the *open ball with center* $\mathbf{x}$ *and radius* r is denoted by $B_r(\mathbf{x})$ and is defined as

$$B_r(\mathbf{x}) = \{\mathbf{y} \mid \mathbf{y} \in S \text{ and } d(\mathbf{x}, \mathbf{y}) < r\}.$$

(b) the *closed ball with center* $\mathbf{x}$ *and radius* r is denoted by $\overline{B}_r(\mathbf{x})$ and is defined as

$$\overline{B}_r(\mathbf{x}) = \{\mathbf{y} \mid \mathbf{y} \in S \text{ and } d(\mathbf{x}, \mathbf{y}) \leq r\}.$$

Definition C.1.12.
Let (S, d) be a metric space and $T \subseteq S$. Then T is said to be *bounded* if there exists some $\mathbf{x} \in S$ and $r > 0$ such that $T \subseteq B_r(\mathbf{x})$. Otherwise T is said to be *unbounded*.

Fact C.1.7.
Let $T \subseteq \mathbb{R}$. Then T is bounded if and only if T has finite supremum and finite infimum.

Fact C.1.8.
Let $T \subseteq \mathbb{R}^n$. Then T is unbounded if and only if there exists a sequence $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in T such that the sequence $\{\|\mathbf{x}_n\|\}_{n\in\mathbb{N}}$ is a strictly increasing sequence in $\mathbb{R}$ and $\{\|\mathbf{x}_n\|\} \to \infty$.

Definition C.1.13.
Let (S, d) be a metric space. Then a sequence $\{\mathbf{x}_n\}_{n\in\mathbb{N}}$ in S is said to be *bounded* if the set $\{\mathbf{x}_n \mid n \in \mathbb{N}\}$ is a bounded set in S.

Fact C.1.9. ***(Bolzano-Weistrass Theorem)***
Every bounded sequence over $\mathbb{R}^n$ has a convergent subsequence.

C.2 Compact Sets

Definition C.2.1.
Let (S, d) be a metric space and $T \subseteq S$. Let $I \subseteq \mathbb{N}$ be an index set. Then the set $C = \{C_i \mid C_i \in S \text{ and } i \in I\}$ of S is said to be a *cover* T if $T \subseteq \bigcup_{C_i \in C} C_i$. In particular, the cover is said to be an *open cover* if each set in C is an open set.

Definition C.2.2.
Let (S, d) be a metric space and $T \subseteq S$. Let $I \subseteq \mathbb{N}$ be an index set and the set $C = \left\{C_i{}^1 \mid i \in I\right\}$ be a cover of T. Then $C' \subset C$ is said to be a *subcover* of T if $T \subseteq \bigcup_{C_i \in C'} C_i$. In particular, the subover is said to be an *open subcover* if each set in C' is an open set.

Definition C.2.3.
Let (S, d) be a metric space and $T \subseteq S$. Then T is said to be *compact* if every open cover of T has a finite open subcover.

Fact C.2.1. ***(Heine Borel Theorem)***
Let $T \subseteq \mathbb{R}^n$. Then T is compact if and only if T is closed and bounded. [12]

Fact C.2.2.
Let T be a compact subset of $\mathbb{R}$. Then T has a finite maximum and finite minimum.

Fact C.2.3.
A continuous function maps compact sets to compact sets.

Fact C.2.4. ***(Extreme Value Theorem)***

Let (S, d) *be a metric space and* T *be a compact subset of* S. *Let* $f : T \to \mathbb{R}$ *be a continuous function. Then there exists vectors* $\mathbf{a}$ *and* $\mathbf{b}$ *in* T *such that* $f(\mathbf{a}) = \max(\{f(\mathbf{x}) \mid \mathbf{x} \in T\})$ *and* $f(\mathbf{b}) = \min(\{f(\mathbf{x}) \mid \mathbf{x} \in T\})$. [12]

References

[1] George B. Dantzig. Linear Programming. http://bioinfo.ict.ac.cn/~dbu/AlgorithmCourses/Lectures/Dantzig2002.pdf, 1991.

[2] Vijay V. Vazirani. *Approximation Algorithms*. Springer Publishers, 2001.

[3] Christos H. Papadimitriou and Kenneth Steiglitz. *Combinatorial Optimization:Algorithms and Complexity*. Prentice Hall of India, second edition, 2001.

[4] George F. Simmons. *Introduction to Topology and Modern Analysis*. Krieger Publishing Company, 2014.

[5] PierGianLuca PortaMana. Notes on affine and convex spaces. http://arxiv.org/pdf/1104.0032.pdf, 31 March 2011.

[6] Geir Dahl. An Introduction to Convexity. http://heim.ifi.uio.no/~geird/conv.pdf, November 2010.

[7] Christopher Griffin. Linear Programming: Penn State Math 484 - Lecture Notes - Version 1.8.3. http://www.personal.psu.edu/cxg286/Math484_V1.pdf, 2009-2014.

[8] Proof of Strong Duality, Complementary Slackness and Marginal Values. https://www.math.ubc.ca/~anstee/math340/340strongduality.pdf, February 20th, 2009.

[9] Anders FORSGREN. AN ELEMENTARY PROOF OF OPTIMALITY CONDITIONS FOR LINEARPROGRAMMING. https://people.kth.se/~andersf/doc/optimality.pdf, June 2008.

[10] Peter D. Lax. *Linear Algebra - Pure and Applied Mathematics*. Wiley-Interscience Publication, 2004.

[11] Kenneth Hoffman and Ray Kunze. *Linear Algebra*. Prentice Hall of India, second edition, 2013.

[12] Ohn Holler. Professor Smith Math 295 Lecture Notes. http://www.math.lsa.umich.edu/~kesmith/nov1notes.pdf, 2010.

[13] Metric Spaces and Compactness. http://wolfweb.unr.edu/homepage/jabuka/Classes/2004_fall/Handouts/05%20-%20Metric%20spaces%20and%20compactness.pdf.

[14] Discrete Geometry. http://www.math.bgu.ac.il/~shakhar/teaching/combinatorial_geometry_files/lec-notes.pdf, November 18, 2008.

[15] Stephen Boyd and Lieven Vandenberghe. Convex Optimization. http://web.stanford.edu/~boyd/cvxbook/bv_cvxbook, 2004.

[16] David Karger. 6.854 Advanced Algorithms . http://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-854j-advanced-algorithms-fall-2005/lecture-notes/dualitynotes.pdf, October 13, 2004.

[17] Hung Q. Ngo. CSE 594: Combinatorial and Graph Algorithms. `http://www.cse.buffalo.edu/~hungngo/classes/2006/594/notes/LP-intro.pdf`, September 29, 2006.

[18] Michel X. Goemans. 18.433: Combinatorial Optimization. `http://www-math.mit.edu/~goemans/18433S09/polyhedral.pdf`, February 20th, 2009.

Index