Math 373: Mathematical Programming and Optimization I

John C. Bowman and Pouria Ramazi
University of Alberta
Edmonton, Canada

September 26, 2018
## Contents

1 Linear Programming ........................................... 5  
   1.A Linear Programming Problems .......................... 9  
      1.A.1 Standard form ..................................... 11  
   1.B Reduction to standard form ............................. 12  

2 The Geometry of Linear Programming ......................... 14  
   2.A Polyhedra .................................................. 14  
   2.B Review of convexity for functions $f : \mathbb{R} \to \mathbb{R}$ .................................................. 15  
   2.C Convexity for functions $f : \mathbb{R}^n \to \mathbb{R}$ .................................................. 16  
   2.D The convex hull ........................................... 16  
   2.E Piecewise linear convex objective functions .............. 18  
   2.F Problems involving absolute values .................... 19  
   2.G Extreme points .......................................... 20  

Bibliography .................................................... 24  

Index ............................................................ 24
Preface

The figures in this text were drawn with the vector graphics language Asymptote (freely available at http://asymptote.sourceforge.net).
Chapter 1
Linear Programming

Optimization is an important area of applied mathematics that is widely used in many areas of society, including industry, business, science, and economics. Some optimization problems involve an infinite number of variables (like those that arise in the calculus of variations that you may encounter in a mechanics course). However, in this course, we focus on optimization problems that involve maximizing or minimizing a function of a finite number of variables subject to finitely many inequality (or equality) constraints. In particular, we will examine the case where both the function and the constraints are linear functions of their arguments. This subject is often confusingly called linear programming, even though it it deals more with linear algebra and geometry than the technical aspects of computer programming. The term “programming” is used here in the sense of detailed military logistics planning, stemming from the work of George Danzig in the US Air Force and at the RAND Corporation just after the end of World War II. The alternative term linear optimization is perhaps a more suitable name.

To illustrate what linear programming is all about, let us begin with a simple optimization problem:

- A farmer has 25 acres of land on which to grow barley or corn. He earns $600 for each acre of barley and $500 for each acre of corn. Harvesting the barley requires 20 hours of labour per acre, and harvesting the corn requires 10 hours of labour per acre. The farmer can afford 200 hours of labor. How can he maximize his profit?

To help us answer this question, let us define the decision variables that express the choices facing the farmer: how many acres of barley to grow \( B \) and how many acres of corn \( C \) to grow.

We are told that the profit is \( f(B, C) = 600B + 500C \).

There are four constraints in the problem. First, the total number of acres \( B + C \leq 25 \). Second, the total amount of labour \( 20B + 10C \leq 200 \). Finally, since you cannot plant negative acres of grain, we know that \( B \geq 0 \) and \( C \geq 0 \).
Remark: It is helpful to illustrate this linear programming problem graphically. In the following graph, the shaded (green+gray) region in the $B$–$C$ plane depicts the constraints $B + C \leq 25$, $B \geq 0$, $C \geq 0$, which is further refined by the constraint $20B + 10C \leq 200$ to the green triangle. The shaded green triangle is the feasible set, which is the collection of all feasible solutions $(B, C)$ that satisfy all four given constraints.

Remark: Let us now consider the behaviour of the objection function $f(B, C) = 600B + 500C$ as $B$ and $C$ vary over the box $[0, 25] \times [0, 25]$ in the $B$–$C$ plane.
One notices that the function takes on its maximum value over the green region at the vertex (0, 20) and its minimum value at the vertex (0, 0). The graph illustrates that the solution to the farmer’s maximization problem occurs at $C = 20$ and $B = 0$. That is, in order to maximize his profit, the farmer should only plant 20 acres of corn; this will use up all 200 hours of his labour but reward him with a tidy profit of $10,000.

**Remark:** Notice that the extrema of the objective function occur at vertices of the feasible set.

**Remark:** Another feature that we notice in this example is the first constraint $B + C \leq 25$ was superfluous and can be removed from the problem. In some applications, the removal of such superfluous constraints can make solving linear programming problems much more efficient!

**Remark:** In general, an optimization problem to an objective function $f$ of $n$ decision variables $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, over the feasible set $F \subset \mathbb{R}^n$ of solutions satisfying all of the constraints, can be written symbolically as

$$\max_{x \in F} f(x)$$

or

$$\min_{x \in F} f(x).$$

• For example, in the problem

$$\max_{x_1^2 + x_2^2 \leq 1} \left( x_1^2 + x_2^2 \right),$$

the feasible set is the unit disk.

**Definition:** The *optimal value* is the desired extreme (maximum or minimum) value (if it exists) of the objective function.

**Definition:** A feasible solution $x^*$ that extremizes the objective function is called an *optimal feasible solution*, or simply an *optimal solution*.

**Definition:** The *optimal set* is the set of optimal solutions; that is, the set of feasible solutions at which the objective function $f$ takes on its optimal value (if it exists).

**Remark:** In the above problem, the optimal value is 1 and the set of optimal solutions is all points on the unit circle.
Remark: In determining optimal solutions, one considers only points $x$ that are feasible.

- For the problem
  \[
  \max_{x_1^2 + x_2^2 \leq 1, \quad x_2 \geq 0} (x_1^2 + x_2^2),
  \]
  the optimal set is the upper half circle.

- For the problem
  \[
  \max_{x \geq -1, \quad x \leq 1} (x^2 + 2),
  \]
  the feasible set is $[-1, 1]$, the optimal value is 3, and $x = -1$ and $x = 1$ are the optimal solutions.

- For the problem
  \[
  \min_{x \geq -1, \quad x \leq 3} (x^2 + y^2),
  \]
  the feasible set $F = \{(x, y) : x \in [-1, 3], y \in (-\infty, 2]\}$, the optimal value is 0 and the optimal solution is $(0, 0)$.

- For the problem
  \[
  \max_{x > 0} \frac{1}{x},
  \]
  the feasible set is $(0, \infty)$, we say that the optimal value is $\infty$, but there are no optimal solutions (the optimal set is empty).

- For the problem
  \[
  \min_{x > 0} \frac{1}{x},
  \]
  the feasible set is again $(0, \infty)$, we say that the optimal value is 0, but there are still no optimal solutions.

Remark: There is no need to study maximization and minimization problems separately, because maximizing $f$ is equivalent to minimizing $-f$.

Remark: In typical real-life problems, determining the optimal set is often even more important that knowing the optimal value. For example, if you are asked to solve an optimization problem to maximize profit, it doesn’t help much to know the optimal value if you don’t know how to achieve it!
Remark: Solving an optimization problem means finding the optimal value and all optimal solutions.

Remark: There are various possible outcomes for an optimization problem:

The problem is

\[
\begin{cases}
\text{infeasible (no feasible solutions)} \\
\text{feasible} \\
\text{no optimal value (unbounded)} \\
\text{an optimal value} \\
\text{finitely many optimal solutions} \\
\text{infinitely many optimal solutions}
\end{cases}
\]

Problem 1.1: Suppose you are asked to produce two kinds of meals: economy and deluxe. The economy meal sells for $3/kg and the deluxe version sells for $4/kg. The ingredients are rice and lamb. The economy version should contain at most 10% lamb, while the deluxe version should contain at least 50% lamb. The cost of rice is $1/kg, whereas the lamb costs $2/kg. You have available 300 kg of rice and 100 kg of lamb. Formulate a linear programming problem that investigates how much rice and lamb should you put into each dish to maximize your profit, assuming that all of the prepared dishes sell.

Let the decision variables \( r_e \) and \( r_d \), represent the weight of rice in the economy and deluxe dishes respectively, and \( l_e \) and \( l_d \) represent the corresponding amounts of lamb. The objective function is \( f(r_e, r_d, l_e, l_d) = 3(r_e + l_e) + 4(r_d + l_d) - (r_e + r_d) - 2(l_e + l_d) \). One wants to maximize this function subject to the constraints

\[
\begin{aligned}
r_e + r_d & \leq 300, \\
l_e + l_d & \leq 100, \\
l_e & \leq 0.1(r_e + l_e), \\
l_d & \geq 0.5(r_d + l_d),
\end{aligned}
\]

and

\[
\begin{aligned}
r_e, l_e, r_d, l_d & \geq 0.
\end{aligned}
\]

1.A Linear Programming Problems

Definition: A function \( f : \mathbb{R}^n \to \mathbb{R} \) is linear if can be expressed as

\[
f(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} c_i x_i,
\]

for some real numbers \( c_i \).
**Remark:** Equivalently, a function $f : \mathbb{R}^n \to \mathbb{R}$ is linear if
\[ f(\alpha \mathbf{x} + \beta \mathbf{y}) = \alpha f(\mathbf{x}) + \beta f(\mathbf{y}) \]
for all vectors $\mathbf{x}$ and $\mathbf{y}$ and real numbers $\alpha$ and $\beta$.

**Definition:** A linear constraint on the variables $x_1, x_2, \ldots, x_n$ has one of the following forms, where $a_1, a_2, \ldots, a_n$ and $b$ are real numbers:

1. linear equalities, such as
   \[ \sum_{i=1}^{n} a_i x_i = b; \]
2. linear inequalities, such as
   \[ \sum_{i=1}^{n} a_i x_i \geq b. \]

**Definition:** Although for $n > 1$, $\mathbb{R}^n$ is not an ordered field, it is nevertheless possible to introduce a partial ordering on $\mathbb{R}^n$ component-wise: given two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ where $x_i \leq y_i$ for all $i$, we say $\mathbf{x} \leq \mathbf{y}$. Similarly, $\mathbf{x} \geq \mathbf{y}$ means that $x_i \geq y_i$ for all $i$. Corresponding definitions hold for strict inequalities, with $\leq$ replaced by $<$ and $\geq$ replaced by $>$. **Definition:** In a linear programming problem, one seeks to minimize a linear cost function $c^\top \mathbf{x} = \sum_{i=1}^{n} c_i x_i$ over all vectors $\mathbf{x} = (x_1, \ldots, x_n)$ in $\mathbb{R}^n$, subject to a set of linear equality and inequality constraints. The vector $\mathbf{c} = (c_1, \ldots, c_n)$ is called a cost vector.

We recall that in a linear programming problem of the form

\begin{align*}
\text{minimize} & \quad c^\top \mathbf{x}, \\
\text{subject to} & \quad a_i^\top \mathbf{x} \geq b_i, \quad i \in M_1, \\
& \quad a_i^\top \mathbf{x} \leq b_i, \quad i \in M_2, \\
& \quad a_i^\top \mathbf{x} = b_i, \quad i \in M_3, \\
& \quad x_j \leq 0, \quad j \in N_1, \\
& \quad x_j \geq 0, \quad j \in N_2,
\end{align*}

where $M_1$, $M_2$, $M_3$, $N_1$, and $N_2$ are finite sets of integers, the variables $\mathbf{x} = (x_1, \ldots, x_n)$ are called decision variables. When $\mathbf{x}$ satisfies all of the given constraints, it is called a feasible solution or feasible vector. If there are no restrictions on the sign of $x_j$, i.e., $j$ does not belong to $N_1 \cup N_2$, we say that $x_j$ is a free or unrestricted variable. The linear function $f(\mathbf{x}) = c^\top \mathbf{x}$ is called the objective function or cost function. A feasible solution $\mathbf{x}^*$ that minimizes the objective function is called an optimal feasible solution or optimal solution. The value $c^\top \mathbf{x}^*$ is the optimal cost.
Problem 1.2: By defining an appropriate matrix $A$ and show that (1.1) can be rewritten in the compact form
\[
\begin{align*}
\text{minimize} & \quad c^\top x, \\
\text{subject to} & \quad Ax \geq b.
\end{align*}
\]

1.A.1 Standard form

A linear programming problem of the form
\[
\begin{align*}
\text{minimize} & \quad c^\top x, \\
\text{subject to} & \quad Ax = b \\
& \quad x \geq 0.
\end{align*}
\]
is said to be in standard form.

Remark: If we let $A_1, \ldots, A_n$ represent the column vectors of the matrix, the constraint $Ax = b$ can be written as
\[
\sum_{i=1}^n A_i x_i = b.
\]
The goal is to synthesize the target vector $b$ from a non-negative amount $x_i$ of each resource $A_i$, such that the cost $\sum_{i=1}^n c_i x_i$ is minimized, where $c_i$ is the unit cost of the $i$th resource.

- The diet problem. Suppose that there are $n$ different foods and $m$ different nutrients. The following table indicates the nutritional content of a unit of each food:

<table>
<thead>
<tr>
<th>nutrient 1</th>
<th>food 1</th>
<th>$\cdots$</th>
<th>food $n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_{11}$</td>
<td>$\cdots$</td>
<td>$a_{1n}$</td>
<td></td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td></td>
</tr>
<tr>
<td>nutrient $m$</td>
<td>$a_{m1}$</td>
<td>$\cdots$</td>
<td>$a_{mn}$</td>
</tr>
</tbody>
</table>

Let $A$ be the $m \times n$ matrix with entries $a_{ij}$, so that its $j$th column $A_j$ represents the nutritional content of the $j$th food. Let $b$ be the vector with the requirements of an ideal diet. The standard form problem aims to mix non-negative quantities $x_i$ of the available foods to achieve an ideal diet at minimal cost. The condition for synthesizing the ideal diet is
\[
\sum_{i=1}^n A_i x_i = b.
\]
CHAPTER 1. LINEAR PROGRAMMING

If instead, $b$ represents the minimal requirements of an adequate diet, then the condition will become

$$\sum_{i=1}^{n} A_i x_i \geq b.$$  

1.B Reduction to standard form

Using the following procedures, it is possible to reduce any linear programming problem to standard form:

1. *Elimination of free variables*: Each free variable $x_j$ can be replaced by the difference $x_j^+ - x_j^-$ of two additional non-negative variables $x_j^+$ and $x_j^-$ (every real number can be as the difference of two non-negative numbers). This removes $x_j$ from the problem but adds two new constraints: $x_j^+ \geq 0$ and $x_j^- \geq 0$.

2. *Elimination of inequality constraints*: Given an inequality constraint of the form

$$\sum_{j=1}^{n} a_{ij} x_j \leq b_i,$$

we introduce a non-negative *slack variable* $s_i$ to obtain the standard form constraints

$$\sum_{j=1}^{n} a_{ij} x_j + s_i = b_i,$$

$$s_i \geq 0.$$  

Likewise, for each constraint of the form

$$\sum_{j=1}^{n} a_{ij} x_j \geq b_i,$$

we introduce a non-negative *slack variable* $s_i$, such that

$$\sum_{j=1}^{n} a_{ij} x_j - s_i = b_i,$$

$$s_i \geq 0.$$
For the linear programming problem
\[
\begin{align*}
\text{minimize} & \quad x_1 + 2x_2 \\
\text{subject to} & \quad x_1 + 3x_2 \leq 2, \\
& \quad 2x_1 + x_2 = 1, \\
& \quad x_1 \geq 0,
\end{align*}
\]
an equivalent problem in standard form is
\[
\begin{align*}
\text{minimize} & \quad x_1 + 2x_2^+ - 2x_2^- \\
\text{subject to} & \quad x_1 + 3x_2^+ - 3x_2^- + x_3 = 2, \\
& \quad 2x_1 + x_2^+ - x_2^- = 1, \\
& \quad x_1 \geq 0, \\
& \quad x_2^+ \geq 0, \\
& \quad x_2^- \geq 0, \\
& \quad x_3 \geq 0,
\end{align*}
\]
For example, for the feasible solution \((x_1, x_2) = (1, -1)\) to the original problem, the standard-form problem has the corresponding feasible solution \((x_1, x_2^+, x_2^-, x_3) = (1, 0, 1, 4)\), which has the same cost. Conversely, the feasible solution \((x_1, x_2^+, x_2^-, x_3) = (2, 1, 4, 9)\) to the standard-form problem corresponds to the solution \((x_1, x_2) = (2, -3)\) to the original problem at the same cost.
Chapter 2

The Geometry of Linear Programming

2.A Polyhedra

Definition: A polyhedron is a set that can be described in the form
\[ \{ x \in \mathbb{R}^n : Ax \geq b \}, \]
where \( A \) is an \( m \times n \) matrix and \( b \) is a vector in \( \mathbb{R}^m \).

Remark: By the same arguments used to reduce a linear programming problem to standard form, a polygon can be equivalently defined as a set
\[ \{ x \in \mathbb{R}^n : Ax = b, x \geq 0 \}. \]

Remark: A polyhedron can either extend to infinity or can be contained in a finite region.

Definition: A set \( S \subseteq \mathbb{R}^n \) is bounded if there exists some constant \( K \) such that \( |x| \leq K \) for every element \( x \in S \). Typically, we will use the Euclidean norm \( |x| = \sqrt{\sum_{i=1}^{n} x_i^2} \).

Definition: Let \( a \in \mathbb{R}^n, a \neq 0 \) and \( b \in \mathbb{R} \). Then

1. The set \( \{ x \in \mathbb{R}^n : a^T x = b \} \) is called a hyperplane.
2. The set \( \{ x \in \mathbb{R}^n : a^T x \geq b \} \) is called a half space.

Notice that the hyperplane is the boundary of the corresponding half space and is perpendicular to \( a \).
2.B. REVIEW OF CONVEXITY FOR FUNCTIONS $f : \mathbb{R} \to \mathbb{R}$

**Definition:** The *graph* of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set of ordered pairs 
$$\{(x, f(x)) : x \in \mathbb{R}^n\}.$$

**Remark:** The graph of $f$ is just the set of values taken on by the function 
$$F : \mathbb{R}^n \to \mathbb{R}^{n+1}, F(x) = (x, f(x)).$$

2.B. Review of convexity for functions $f : \mathbb{R} \to \mathbb{R}$

**Definition:** A function $f : \mathbb{R} \to \mathbb{R}$ is *convex* (sometimes called *concave up*) on an interval $I \subseteq \mathbb{R}$ if its graph lies below or on the secant line segment joining $(p, f(p))$ and $(q, f(q))$.

**Definition:** A function $f$ is *concave* (sometimes called *concave down*) on an interval $I$ if $-f$ is convex on $I$.

**Remark:** Since the equation of the line through $(p, f(p))$ and $(q, f(q))$ is 
$$y = f(p) + \frac{f(q) - f(p)}{q - p} (x - p),$$
the definition of convex says 
$$f(x) \leq f(p) + \frac{f(q) - f(p)}{q - p} (x - p) \quad \forall x \in [p, q], \quad \forall p, q \in I. \quad (2.1)$$

This condition may be rewritten by re-expressing the *linear interpolation* of $f$ between $[p, q]$ on the right-hand side of (2.1):

$$f(x) \leq \left(\frac{q-x}{q-p}\right)f(p) + \left(\frac{x-p}{q-p}\right)f(q) \quad \forall x \in [p, q], \quad \forall p, q \in I. \quad (2.2)$$

It is sometimes convenient to introduce the *parameter* $t = \frac{q-x}{q-p}$, in terms of which we may express $x = q - (q-p)t$ and 
$$\frac{x-p}{q-p} = \frac{(q-p)-(q-p)t}{q-p} = 1-t.$$

This allow us to restate (2.2) in *parametric form*:

$$f(tp + (1-t)q) \leq tf(p) + (1-t)f(q) \quad \forall t \in [0,1], \quad \forall p, q \in I. \quad (2.3)$$
2.C Convexity for functions \( f : \mathbb{R}^n \rightarrow \mathbb{R} \)

**Definition:** A set \( S \) is convex if for every \( p, q \in S \) the line segment joining \( p \) and \( q \) is contained entirely within \( S \).

**Definition:** Given points \( p, q \in \mathbb{R}^n \), the set \( \{ tp + (1-t)q : t \in [0,1] \} \) describes the line segment with endpoints \( p \) and \( q \).

**Definition:** A function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is convex if for every \( p, q \) and \( t \in [0,1] \) we have \( f(tp + (1-t)q) \leq tf(p) + (1-t)f(q) \).

**Remark:** Equivalently, a function is convex if its graph lies below or on the secant line segment joining \( (p,f(p)) \) and \( (q,f(q)) \).

**Definition:** A function \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) is concave if for every \( p, q \) and \( t \in [0,1] \) we have \( f(tp + (1-t)q) \geq tf(p) + (1-t)f(q) \).

**Definition:** An affine function is a function of the form \( f(x) = a^T x + c \) for some constant \( c \).

**Remark:** The only functions that are both convex and concave are affine functions.

**Remark:** A real-valued function \( f \) has a local minimum at \( p \) if \( f(q) \geq f(p) \) for all \( q \) in the vicinity of \( p \).

**Remark:** A real-valued function \( f \) has a global minimum at \( p \) if \( f(q) \geq f(p) \) for all \( q \).

**Theorem 2.1:** For convex functions, local minima are also global minima.

2.D The convex hull

**Definition:** Let \( x_1, \ldots, x_k \in \mathbb{R}^n \), \( t_1, \ldots, t_k \geq 0 \), and \( \sum_{i=1}^k t_i = 1 \). Then

1. The vector \( \sum_{i=1}^k t_i x_i \) is said to be a convex combination of the vectors \( x_1, \ldots, x_k \).

2. The convex hull of the vectors \( x_1, \ldots, x_k \) is the set of all convex combinations of these vectors. Equivalently, it is the smallest convex polygon that contains the points \( x_1, \ldots, x_k \).
Remark: Every half space \( H = \{ x \in \mathbb{R}^n : a^T x \geq b \} \) is convex: let \( x, y \in H \). For every \( t \in [0, 1] \) we see that \( a^T(tx + (1 - t)y) \geq tb + (1 - t)b = b \). That is, \( a^T(tx + (1 - t)y) \in H \). Thus \( H \) is convex.

**Theorem 2.2 (Properties of Convex Sets):**

a) The intersection of convex sets is convex.

b) Every polyhedron \( \{ x \in \mathbb{R}^n : Ax \geq b \in \mathbb{R}^m \} \) is a convex set.

c) A convex combination of a finite number of elements of a convex set also belongs to that set.

d) The convex hull of a finite number of points is a convex set.

Proof:

a) Consider a collection of convex sets \( S_i \), where \( i \) belongs to some index set \( I \). If \( x, y \in \cap_{i \in I} S_i \), then for each \( i \in I \) and \( t \in [0, 1] \), we know from the convexity of \( S_i \) that \( tx + (1 - t)y \in S_i \). Therefore \( tx + (1 - t)y \in \cap_{i \in I} S_i \). Thus \( \cap_{i \in I} S_i \) is convex.

b) Since a polyhedron is an intersection of convex half spaces, which we have seen are convex, this follows from a).

c) By the definition of convexity, a convex combination of two elements of a convex set \( S \) lies in that set. Suppose that every convex combination of \( k \) elements of \( S \) belongs to \( S \), where \( k \in \{2, 3, \ldots \} \). Given a nontrivial convex combination \( \sum_{i=1}^{k+1} t_i x_i \) of \( k + 1 \) elements \( x_i \) of \( S \), where \( i = 1, \ldots, k + 1 \) and \( t_i \in (0, 1) \) such that \( \sum_{i=1}^{k+1} t_i = 1 \), we express

\[
\sum_{i=1}^{k+1} t_i x_i = t_{k+1} x_{k+1} + \sum_{i=1}^{k} \frac{t_i}{1 - t_{k+1}} x_i.
\]

The summation in the second term belongs to \( S \) since it is a linear combination of \( k \) elements of \( S \): \( \sum_{i=1}^{k+1} t_i = \sum_{i=1}^{k+1} t_i - t_{k+1} = 1 - t_{k+1} \). Thus \( \sum_{i=1}^{k+1} t_i x_i \) is a linear combination of two elements of \( S \) and is therefore also in \( S \). By induction, a convex combination of any finite number of elements of a convex set also belongs to that set.

d) Let \( y = \sum_{i=1}^{k} \alpha_i x_i \) and \( z = \sum_{i=1}^{k} \beta_i x_i \) be two elements of the convex hull \( S \) of points \( x_1, \ldots, x_k \) and \( \tau \in [0, 1] \). Then

\[
ty + (1 - t)z = t \sum_{i=1}^{k} \alpha_i x_i + (1 - t) \sum_{i=1}^{k} \beta_i x_i = \sum_{i=1}^{k} [t \alpha_i + (1 - t) \beta_i] x_i.
\]
is a convex combination of \( x_1, \ldots, x_k \), noting that

\[
\sum_{i=1}^{k} [t\alpha_i + (1 - t)\beta_i] = t \sum_{i=1}^{k} \alpha_i + (1 - t) \sum_{i=1}^{k} \beta_i = t + (1 - t) = 1.
\]

It follows from c) that \( S \) is convex.

### 2.E Piecewise linear convex objective functions

**Definition:** A function \( f : \mathbb{R}^n \to \mathbb{R} \) is **piecewise linear** if it is linear on each of a finite number of intervals of \( \mathbb{R}^n \).

- The absolute value function

\[
|x| = \begin{cases} 
  x & \text{if } x \geq 0, \\
  -x & \text{if } x < 0.
\end{cases}
\]

is piecewise linear.

**Remark:** Although we stated earlier that we will focus on objective functions that are linear, one can easily generalize the methods we will develop to functions that are **piecewise linear**. We will then be able to apply linear programming to problems of the form

\[
\begin{align*}
\text{minimize} & \quad \max_{i=1,\ldots,m} (c_i^T x + d_i) \\
\text{subject to} & \quad Ax \geq b.
\end{align*}
\]

Since \( \max_{i=1,\ldots,m} (c_i^T x + d_i) \) is equal to the smallest number \( M \) such that \( M \geq c_i^T x + d_i \) for all \( i \), the above optimization problem is equivalent to the linear programming problem

\[
\begin{align*}
\text{minimize} & \quad M \\
\text{subject to} & \quad M \geq c_i^T x + d_i, \quad i = 1, \ldots, m, \\
& \quad Ax \geq b.
\end{align*}
\]

**Theorem 2.3** (Piecewise convex functions): Let \( f_1, \ldots, f_m : \mathbb{R}^n \to \mathbb{R} \) be convex functions. Then the function \( f(x) = \max_{i=1,\ldots,m} f_i(x) \) is also convex.

**Proof:** Since

\[
f(tp + (1 - t)q) = \max_{i=1,\ldots,m} f_i(tp + (1 - t)q)
\leq \max_{i=1,\ldots,m} [tf_i(p) + (1 - t)f_i(q)]
\leq t \max_{i=1,\ldots,m} f_i(p) + (1 - t) \max_{i=1,\ldots,m} f_i(q)
= tf(p) + (1 - t)f(q),
\]

for all \( p \) and \( q \), we see that \( f \) is indeed convex.
• Since the absolute value function \( f(x) = |x| = \max\{x, -x\} \), we see from Theorem 2.3 that the absolute value function is convex (as well as piecewise linear).

**Remark:** Piecewise linear convex functions can be used to approximate more general functions.

**Remark:** In addition to now being able to handle piecewise linear convex objective functions, we can also handle piecewise linear affine constraints such as
\[
\max_{i=1,\ldots,m} (e_i^\top x + d_i) \leq b
\]
by rewriting them as \( e_i^\top x + d_i \leq b \).

## 2.F Problems involving absolute values

Consider problems of the form
\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{n} c_i |x_i| \\
\text{subject to} & \quad Ax \geq b,
\end{align*}
\]
where \( x = (x_1, \ldots, x_n) \) and the cost coefficients \( c_i \) are non-negative. Although the objection function here is easily shown to be piecewise linear and convex (being the sum of piecewise linear convex functions) and therefore can be handled in the manner just described, a more efficient method is available. We can use the fact that \( |x_i| \) is the smallest number \( M_i \) that bounds both \( x_i \) and \( -x_i \) from above to rewrite the linear programming problem as:
\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{n} c_i M_i \\
\text{subject to} & \quad Ax \geq b, \\
& \quad x_i \leq M_i, \quad i = 1, \ldots, n, \\
& \quad -x_i \leq M_i, \quad i = 1, \ldots, n.
\end{align*}
\]

An alternative method is to replace each occurrence of \( |x_i| \) with \( x_i^+ + x_i^- \) and then each of the remaining occurrences of \( x_i \) with \( x_i^+ - x_i^- \), where \( x_i^+ \) and \( x_i^- \) are new non-negative decision variables:
\[
\begin{align*}
\text{minimize} & \quad \sum_{i=1}^{n} c_i (x_i^+ + x_i^-) \\
\text{subject to} & \quad Ax^+ - Ax^- \geq b, \\
& \quad x^+, x^- \geq 0,
\end{align*}
\]
where \( x^+ = (x_1^+, \ldots, x_n^+) \) and \( x^- = (x_1^-, \ldots, x_n^-) \). The equivalence of these two formulations follows from the observation that at an optimal solution, for each \( i \) one of \( x_i^+ \) or \( x_i^- \) must be zero (otherwise we could decrease both variables, preserving feasibility and reducing the cost), from which it follows that \( |x_i| = x_i^+ + x_i^- \).

2.G Extreme points

**Definition:** Let \( P \) be a polyhedron. A point \( x \in P \) is an extreme point of \( P \) if we cannot find two vectors \( y, z \in P \), both different from \( x \), and a scalar \( t \in [0, 1] \) such that \( x = ty + (1 - t)z \). That is, a point \( x \in P \) is an extreme point if it cannot be expressed as a convex combination of two other points of \( P \).

**Definition:** Let \( P \) be a polyhedron. A vector \( x \in P \) is a vertex of \( P \) if there exists some \( c \) such that \( c^\top x < c^\top y \) for all \( y \in P, y \neq x \). That is, \( x \) is a vertex of \( P \) if \( P \) lies entirely on one side of the hyperplane \( \{ y : c^\top y = c^\top x \} \), intersecting the hyperplane only at \( x \).

Consider a polyhedron \( P \in \mathbb{R}^n \) defined in terms of
\[
\begin{align*}
\mathbf{a}_i^\top \mathbf{x} & \geq b_i & i \in M_1, \\
\mathbf{a}_i^\top \mathbf{x} & \leq b_i & i \in M_2, \\
\mathbf{a}_i^\top \mathbf{x} & = b_i & i \in M_3.
\end{align*}
\]

**Definition:** If a vector \( \mathbf{x}^* \) satisfies \( \mathbf{a}_i^\top \mathbf{x}^* = b_i \) for some \( i \) in \( M_1, M_2 \) or \( M_3 \), we say that the corresponding constraint is active or binding at \( \mathbf{x}^* \).

- Given the polyhedron \( P = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1, x_1, x_2, x_3 \geq 0\} \) shown in Figure 2.1, which constraints are active at each of the following points?
  A: \( x_1 + x_2 + x_3 = 1, x_2 = 0, x_3 = 0 \)
  B: \( x_1 + x_2 + x_3 = 1, x_1 = 0, x_3 = 0 \)
  C: \( x_1 + x_2 + x_3 = 1, x_1 = 0, x_2 = 0 \)
  D: \( x_1 = 0, x_2 = 0, x_3 = 0 \)
  E: \( x_1 + x_2 + x_3 = 1, x_3 = 0 \)

**Theorem 2.4 (Active constraints):** Let \( \mathbf{x}^* \in \mathbb{R}^n \) and \( I = \{i : \mathbf{a}_i^\top \mathbf{x}^* = b_i\} \) be the set of indices of active constraints at \( \mathbf{x}^* \). Then the following are equivalent:

1. There exists \( n \) vectors in the set \( \{\mathbf{a}_i : i \in I\} \), which are linearly independent.
2. The span of the vectors \( \{\mathbf{a}_i : i \in I\} \) is \( \mathbb{R}^n \).
3. The system of equations \( \mathbf{a}_i^\top \mathbf{x} = b_i, i \in I \), has a unique solution.
Figure 2.1: Polyhedron $P = \{(x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1, x_1, x_2, x_3 \geq 0\}$.

(a) $\iff$ (b): Suppose that the vectors $\{a_i : i \in I\}$ span $\mathbb{R}^n$. Then $n$ of these vectors form a basis of $\mathbb{R}^n$ and therefore must be linearly independent. Conversely, suppose that $n$ of the vectors $a_i, i \in I$, are linearly independent. The subspace spanned by these vectors is $n$ dimensional and must therefore span all of $\mathbb{R}^n$.

(b) $\iff$ (c): Suppose that $x_1$ and $x_2$ are solutions to the system of equations $a_i^T x = b_i, i \in I$. Then $d = x_1 - x_2$ satisfies $a_i^T d = 0$ for all $i \in I$. That is, $d \in \mathbb{R}^n$ is orthogonal to every $a_i, i \in I$ and therefore cannot be expressed as a linear combination of these vectors. That is, $\{a_i : i \in I\}$ do not span all of $\mathbb{R}^n$. Conversely, if the vectors $a_i, i \in I$, do not span all of $\mathbb{R}^n$, choose a nonzero vector $d$ orthogonal to the subspace they span. For every solution $x$ to the system of equations $a_i^T x = b_i, i \in I$, the vector $x + d$ will then be a distinct solution to the same system of equations.

**Definition**: Consider a polyhedron $P$ defined by linear equality and inequality constraints, and let $x^* \in \mathbb{R}^n$. Then

1. The vector $x^*$ is a *basic solution* if
   
   (a) all equality constraints are active;
   
   (b) of the constraints that are active at $x^*$, $n$ of them are linearly independent.

2. If $x^*$ is a basic solution that satisfies all of the constraints, we say that it is a *basic feasible solution*. 

CHAPTER 2. THE GEOMETRY OF LINEAR PROGRAMMING

Figure 2.2: Polyhedron $P = \{ (x_1, x_2, x_3) \mid x_1 + x_2 + x_3 = 1, x_1, x_2, x_3 \geq 0 \}$.

- Given the polyhedron $P$ shown in Figure 2.2, determine which of the following points are feasible, basic or basic feasible solutions:

  A:

  B:

  C:

  D:

  E:

  What would happen if the equality constraint $x_1 + x_2 + x_3 = 1$ were to be replaced by the constraints $x_1 + x_2 + x_3 \leq 1$ and $x_1 + x_2 + x_3 \geq 1$?

  According to the above example, whether a point is a basic solution or not may depend on the way that a polyhedron is represented.

- Given the polyhedron in Figure 2.3, determine which of the following points are feasible, basic or basic feasible solutions:

  A:

  B:

  C:

  D:

  E:
2.G. EXTREME POINTS

Figure 2.3:

Theorem 2.5 (Vertices): Let $P$ be a nonempty polyhedron and let $x^* \in P$. Then the following are equivalent:

1. $x^*$ is a vertex;
2. $x^*$ is an extreme point;
3. $x^*$ is a basic feasible solution.