Revision notes - MA3252

Ma Hongqiang

February 20, 2018

Contents

1 Introduction to Linear Programming

1.1 Learnt optimisation methods

1.
$$\min_{x\in\mathbb{R}} f(x), \text{where } f:\mathbb{R}\to\mathbb{R}$$
Solution: Find x such that $f'(x)=0$.

2.
$$\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) = f(x_1,\dots,x_n), \text{where } f:\mathbb{R}^n\to\mathbb{R}$$
Solution: Find $\mathbf{x}$ such that $\nabla f(\mathbf{x}) = \mathbf{0}$

3.
$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \\ \text{s.t.}\quad & g_k(\mathbf{x}) = g_k(x_1,\dots,x_n) = 0 \text{ for } k=1,\dots,m \end{aligned}$$
Solution: Lagrange multiplier (See `MA1104.pdf`)

Definition 1.1 (Optimisation Problem)**.**
A general optimisation problem is of the form

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^n} f(\mathbf{x}) \\ \text{s.t.}\quad & x\in P \end{aligned}$$

where $P \subset \mathbf{R}^n$ is a feasible region.

1.2 Prerequisite for Linear Programming

- For linear programming, the inner product is the standard inner product(See `MA2101.pdf`) over $\mathbf{R}^n$.
- In $\mathbb{R}^2$(resp. $\mathbb{R}^3$, $\mathbb{R}^n$), $\{\mathbf{x} \mid \mathbf{a}^{\mathrm{T}}\mathbf{x} = b\}$ is line(resp. plane, hyperplane) with normal vector $\mathbf{a}$.
- For $\{\mathbf{x} \mid \mathbf{a}^{\mathrm{T}}\mathbf{x} = b\}$, vector $\mathbf{a}$ corresponds to direction of increasing $\mathbf{c}^{\mathrm{T}}\mathbf{x}$.
- The inequality $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \le b_i$ represents a halfspace. We refer to $\mathbf{a}_i$ as the **normal vector** for half spaces specified in this form.

To draw a halfspace of the form $\{\mathbf{x} \mid \mathbf{a}^{\mathrm{T}}\mathbf{x} \le b\}$,

1. Draw the line $\mathbf{a}^{\mathrm{T}}\mathbf{x} = b$.
2. Draw the normal vector $\mathbf{a}$.
3. Colour the region in the opposite direction of $\mathbf{a}$.

1.3 Linear Programming in Low Dimensions

Definition 1.2 (Linear Programming)**.**
A **linear programming** problem is of the form

$$\begin{aligned}
\min_{\mathbf{x}\in\mathbb{R}^n} (\text{or } \max) \mathbf{c}^{\mathrm{T}} \mathbf{x} & \\
\text{s.t.} \qquad\qquad & \mathbf{a}_i^{\mathrm{T}} \mathbf{x} \geq b_i \text{ for some } i \\
& \mathbf{a}_i^{\mathrm{T}} \mathbf{x} \leq b_i \text{ for some } i \\
& \mathbf{a}_i^{\mathrm{T}} \mathbf{x} = b_i \text{ for some } i \\
& x_j \geq 0 \text{ for some } j \\
& x_j \leq 0 \text{ for some } j \\
& x_j \in \mathbb{R} \text{ for some } j
\end{aligned}$$

where $\mathbf{a}_i = (a_{i,1}, a_{i,2}, \ldots, a_{i_n})^{\mathrm{T}} \in \mathbb{R}^n, b_i \in \mathbb{R}$. We require the constraints to be **finite**.
Notation wise,

- $\mathbf{c} = (c_1, \ldots, c_n)^{\mathrm{T}} \in \mathbb{R}^n$ is called **cost** vector.
- The function $\mathbf{c}^{\mathrm{T}} \mathbf{x}$ is called the **objective function**.
- The n variables x_i are called **decision variables**/
- A vector $\mathbf{x}$ satisfying all constraints is a **feasible solution**.
- The **feasible set** is the set of all feasible solutions.
- A feasible solution $\mathbf{x}^*$ that minimises objective function is an **optimal solution**.
- The value $\mathbf{c}^{\mathrm{T}} \mathbf{x}^*$ is **optimal objective value**.

To solve a linear minimisation problem in $\mathbb{R}^2$,

1. Sketch the feasible region
2. Find an optimal solution graphically by shifting $\mathbf{c}^{\mathrm{T}} \mathbf{x}$ in the direction of $-\mathbf{c}$.

Remark:

- For a minimisation problem, the cost is **unbounded** if for every $K \in \mathbb{R}$, there is a feasible $\mathbf{x}$ such that $\mathbf{c}^{\mathrm{T}} \mathbf{x} < K$. Equivalently, we say that the optimal cost is $-\infty$. The case for maximisation problem is similar.
- For a LP problem, the problem is **infeasible** if the feasible set is empty. Equivalently, the objective value is ∞ for minimisation problem and $-\infty$ for maximisation problem.
- There are four possibilities for an LP: (1) Unique solution (2) Multiple optimal solutions (3) Optimal cost unbounded and (4) Infeasible
- $\max \mathbf{c}^{\mathrm{T}} \mathbf{x} = -\min -\mathbf{c}^{\mathrm{T}} \mathbf{x}$

From	To
$\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \leq b_i$	$(-\mathbf{a}_i)^{\mathrm{T}}\mathbf{x} \geq -b_i$
$x_j \geq 0$	$\mathbf{e}_j^{\mathrm{T}}\mathbf{x} \geq 0$
$\mathbf{a}_i^{\mathrm{T}}\mathbf{x} = b$	$\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \geq b_i$ and $(-\mathbf{a}_i)^{\mathrm{T}}\mathbf{x} \geq -b_i$

1.4 Compact Form

Every constraint can be translated into the form of $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \geq b_i$. As such, we can write the constraint in the form of $\mathbf{A}\mathbf{x} = \mathbf{b}$ where $\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^{\mathrm{T}} \\ \vdots \\ \mathbf{a}_m^{\mathrm{T}} \end{pmatrix}$ and $\mathbf{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$.

Definition 1.3 (Compact Form)**.**
A LP is in compact form when it is of the form

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} \geq \mathbf{b} \end{aligned}$$

1.5 Standard Form

Definition 1.4 (Standard Form)**.**
A LP of the form

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0} \end{aligned}$$

is said to be in the standard form.

Three defining characteristics of standard form is

- Minimisation objective
- Equality constraints
- Nonnegative variables

Theorem 1.1 (Conversion to Standard Form)**.**
To convert an LP to standard form

1. Eliminate inequality constraint by **introducing slack variable**.
   For example, $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \leq b_i$ becomes $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} = b_i$ with $s_i \geq 0$.
2. Eliminate nonpositive variables by **replacing** x_i **by** $-x_i^-$ for all $x_i \leq 0$.
3. Eliminate free variables by **replacing** x_i **by** $x_i^+ - x_i^-$, where $x_i^+, x_i^- \geq 0$.

Remark: In general, feasible set $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ where $\mathbf{R} \in \mathbb{R}^{m \times n}, m < n$ of a standard form LP is the intersection between

- a $(n-m)$-dimensional subset $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{b}\}$, and
- constraints: $x_i \geq 0, i = 1, \dots, m$.

Next, we show some convex objective functions can also be modelled using LP.

Definition 1.5 (Convex sets)**.**
A set $S \subset \mathbb{R}^n$ is convex if for every $\mathbf{x}, \curvearrowright \in S$ and every $\lambda \in [0,1]$, we have $\lambda\,\mathbf{x} + (1-\lambda)\mathbf{y} \in S$.

Definition 1.6 (Convex Combination)**.**
$\mathbf{x} \in \mathbb{R}^n$ is a convex combination of $\mathbf{x}^1, \dots, x^k \in \mathbb{R}^n$ if and only if

$$x = \sum_{i=1}^{k} \lambda_i\,\mathbf{x}^i$$

where $\lambda_i \in [0,1]$ are such that $\sum_{i=1}^{k} \lambda_i = 1$.

Definition 1.7 (Convex Hull)**.**
The convec hull of $\mathbf{x}^1, \dots, \mathbf{x}^k$ is the set

$$\operatorname{conv}(\mathbf{x}^1, \dots, \mathbf{x}^k) := \{x \in \mathbb{R}^n : \mathbf{x} \text{ is a convex combination of} x^1, ldots, x^k\}$$

Definition 1.8 (Convex Function)**.**
A convex function satisfies

$$f(\lambda\,\mathbf{x} + (1-\lambda\mathbf{y})) \leq \lambda f(\mathbf{x}) + (1-\lambda) f(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n, \lambda \in [0,1]$.

Remark: f is concave if and only if $-f$ is convex.

Theorem 1.2 (Convexity of Affine function)**.**
Affine function $f(\mathbf{x}) = d + \mathbf{c}^{\mathrm{T}}\,\mathbf{x}$ is both convex and concave.

Theorem 1.3.
Let $f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}^n$ be convex functions. Then the function

$$f(\mathbf{x}) := \max_{i=1,\dots,m} f_i(\mathbf{x})$$

is also convex.

Therefore, piecewise affine function $\max_{i=1,\dots,m}(\mathbf{c}_i^{\mathrm{T}}\,\mathbf{x} + d_i)$ is convex.
In the case of optimsation of the form $\min \max(f(\mathbf{x}), g(\mathbf{x}))$, we transform the optimisation into

$$\begin{aligned} \min\ & t \\ \text{s.t.}\ & t \geq f(\mathbf{x}) \\ & t \geq g(\mathbf{x}) \\ & \text{Rest constraints} \end{aligned}$$

2 Geometry of Linear Programming

Definition 2.1 (Polyhedron)**.**
A polyhedron is a set of the form $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ where $\mathbf{A} \in \mathbf{R}^{m \times n}$ and $\mathbf{b} \in \mathbf{R}^m$.

Geometrically, a polyhedron is a finite intersection of half spaces specified by half spaces. From definition of standard form LP, the feasible region for a standard form LP is a polyhedron.

Definition 2.2 (Extreme Point)**.**
Consider a convex set $P \in \mathbb{R}^n$. A point $\mathbf{x}^* \in \P$ is a extreme point of P if whenever points $\mathbf{y}$ and $\mathbf{z} \in P$ and scalar $\lambda \in (0, 1)$ are such that $\mathbf{x}^* = \lambda\mathbf{y} + (1 - \lambda)\mathbf{z}$, we have $\mathbf{y} = \mathbf{z} = \mathbf{x}$.

Definition 2.3 (Vertex)**.**
A point $\mathbf{x}^* \in P$ is a vertex of P if there is a $\mathbf{c} \in \mathbb{R}$ such that

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}^* > \mathbf{c}^{\mathrm{T}}\mathbf{y} \text{ for all } \mathbf{y} \in P \setminus \{\mathbf{x}\}$$

Definition 2.4 (Basic Feasible Solution(BFS))**.**
Consider $\mathbf{x}^* \in \mathbb{R}$ and constraint $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} < / = / > b_i$. If $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} = b_i$, we say that this constraint is **tight at $\mathbf{x}^*$**.
Constraints $\mathbf{a}_i^{\mathrm{T}}\mathbf{x} \geq b_i, i \in I$ are said to be linearly independent if the corresponding vectors $\mathbf{a}_i, i \in I$ are linearly independent.
$\mathbf{x}^*$ is a basic feasible solution of a polyhedron P if $\mathbf{x}^* \in P$ and n linearly independent constraints are active at $\mathbf{x}^*$.

Remark:

- A vector $x^* \in \mathbb{R}^n$ is said to be of rank k if the span of collections of $\{\mathbf{a}_i : \mathbf{a}_i^{\mathrm{T}}\mathbf{x}^* = b_i\}$ has dimension k.
  Thus, $\mathbf{x}$ is a BFS iff it has rank n and $\mathbf{x} \in P$.
- A vector $x^* \in \mathbb{R}^n$ is a basic solution if it has rank n.
- To check whether some $\mathbf{x} \in \mathbb{R}^n$ is BFS, we check that
  1. $\mathbf{x}$ is feasible
  2. $\mathbf{x}$ is tight for n constraints
  3. which are linearly independent
- A basic solution $\mathbf{x} \in \mathbb{R}^n$ is degenerate if more than n constraints are active at $\mathbf{x}$.

The next theorem shows the equivalence of these three concepts.

Theorem 2.1.
Let P be a non-empty polyhedron and $\mathbf{x} \in P$. Then the three defintions are equivalent.

2.1 Basic Feasible Solution for Standard Polyhedra

Definition 2.5 (Standard Form Polyhedra)**.**
The polyhedra in the standard form is of the form

$$P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\} \text{ where } \mathbf{A} \in \mathbb{R}^{m \times n}$$

We further assume

1. m rows of matrix A are linearly independent, and
2. $m \leq n$

Next is an equivalent definition of basic solution of standard form polyhedron.

Theorem 2.2 (Equivalent Basic Solution Definition)**.**
We say that a vector $\mathbf{x}^* \in \mathbb{R}^n$ is a **basic solution** of the standard form polyhedron P as speicified above if

1. $\mathbf{A}\mathbf{x}^* = \mathbf{b}$ and
2. There exist indices $B(1), B(2), \ldots, B(m)$ such that:
   (a) The columns $\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}$ are linearly independent, and
   (b) $\mathbf{x}_i^* = 0$ for $i \notin \{B(1), \ldots, B(m)\} := \mathrm{Im}(B)$.

Definition 2.6 (More Terminologies)**.**
We call variable $x_{B(1)}, \ldots, x_{B(m)}$ **basic variables** and $x_i, i \notin \mathrm{Im}(B)$ nonbasic variables.
B in subscript denotes the image of B and N denotes $1, \ldots, n \setminus B$.
The vector $\mathbf{x}_B \in \mathbb{R}^m$ is defined with the basic variables

$$\mathbf{x}_B = \begin{pmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{pmatrix}$$

The matrix $\mathbf{B} \in \mathbb{R}^{m \times m}$ is obtained by concatenating the m basic columns and is called a **basis matrix**

$$\mathbf{B} = \begin{pmatrix} \mathbf{A}_{B(1)} & \cdots & \mathbf{A}_{B(m)} \end{pmatrix}$$

Note that $\mathbf{B}$ is invertible and $\mathbf{B}\mathbf{x}_B = \mathbf{b}$, so $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$.
We define the basis to be the set of independent columns: $\{\mathbf{A}_{B(1)}, \ldots, \mathbf{A}_{B(m)}\}$, where $\mathbf{A}_i$ denotes the ith column.
Matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ can be partitioned as $(\mathbf{B} \mid \mathbf{N})$.

Theorem 2.3 (Procedure for Finding Basic Solution)**.**
Any basic solution can be found using the following steps:

1. Choose an admissible basis index set $B = \{b_1, \ldots, b_m\}$ where $|B| = m$ and $\mathbf{A}_{b_i}, b_i \in B$ are linearly independent.

2. Write down $\mathbf{B} = \begin{pmatrix}\mathbf{A}_{b_1} & \dots & \mathbf{A}_{b_m}\end{pmatrix}$ and solve for $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b}$.

3. Extract information of $(\mathbf{x}_B)_{b_i}$ and write down $\mathbf{x}^*$ where $\mathbf{x}^*_{b_i} = (\mathbf{x}_B)_{b_i}$ for $b_i \in B$ and $(\mathbf{x}_B)_j = 0$ otherwise.

4. $\mathbf{x}^*$ is a basic solution.

5. If $\mathbf{x}^* \geq 0$, then $\mathbf{x}^*$ is a basic feasible solution.

Definition 2.7 (Adjacent Basic Solutions)**.**
Two distinct basic solutions are **adjacent** *if and only if* one of the following *equivalent* properties hold:

1. The corresponding bases share all but one basic column; or

2. There are a common $n-1$ (but not n) linearly independent constraints that are active at both of them.

Geometrically, adjacent BFS are extreme points connected by an edge on the boundary.

Definition 2.8.
A polyhedron $P \subset \mathbb{R}^n$ **contains a line** if there exists $\mathbf{x}^* \in P$ and a nonzero $\mathbf{d} \in \mathbb{R}^n$, such that $\mathbf{x}^* + \lambda\mathbf{d} \in P$ for all $\lambda \in \mathbb{R}$.

Geometrically, a polyhedron containing an infinite line does not contain an extreme point.

Theorem 2.4.
Suppose $P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} \geq \mathbf{b}\} \neq \varnothing$. The following are equivalent:

1. P does not contain a line.

2. P has a BFS.

3. P has n linearly independent contraints.

Implication: Every nonempty bounded polyhedron and every nonempty standard form polyhedron has at least one BFS.

Theorem 2.5 (Optimality of Basic Feasible Solution)**.**
Consider the LP

$$\begin{aligned} &\min_{\mathbf{x} \in \mathbb{R}^n} \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &\text{s.t. } \; x \in P, \; \textbf{a polyhedron} \end{aligned}$$

Suppose P has at least one BFS and LP has an optimal solution. Then there is an optimal solution which is a BFS.

Implication: To find optimal solutions, it suffices to check BFS.

3 Simplex Method

Consider the standard form LP

$$\begin{aligned} \min_{\mathbf{x}\in\mathbb{R}^n} & \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t.} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq 0 \end{aligned}$$

We further assume $\mathbf{A} \in \mathbb{R}^{m\times n}$, $\mathrm{rank}(\mathbf{A}) = m$ with $m \leq n$. Let

$$P = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$$

Then P does not contain a line due to $\mathbf{x} \geq 0$ constaint. Then by Theorem 2.4, either

- There is an optimal solution which is a BFS(Theorem 2.5), or
- The optimal value is unbounded, either $-\infty$ unbounded or ∞ infeasible

We adopt terminologies and notations from the previous chapter.

Definition 3.1 (Feasible Direction)**.**
For a polyhedron P, and a point $\mathbf{x} \in P$, a vector $\mathbf{d}$ is a **feasible direction** if $\mathbf{x} + \theta\mathbf{d} \in P$ for some $\theta > 0$.

The standard procedure of solving standard LP involves

1. calculate feasible directions $\mathbf{d}$ connecting 2 adjacent BFS,
2. calculate $\bar{\theta}$ so that from BFS $\mathbf{x}_a$ we get new BFS $\mathbf{x}_a + \bar{\theta}\mathbf{d}$,
3. determine if feasible direction lowers objective

Theorem 3.1 (Computation of $\mathbf{d}^j$)**.**
Let B be a basis for $\mathbf{x}$. For $j \in N$ an index outside basis, let $\mathbf{d}^j := (\mathbf{d}^j_B, \mathbf{d}^j_N)^{\mathrm{T}}$ be such that

1. $\mathbf{A}\mathbf{d}^j = \mathbf{0}$, which guarantees $\mathbf{A}(\mathbf{x} + \theta\mathbf{d}^j) = \mathbf{b}$.
2. $\mathbf{d}^j_j = 1$ and $\mathbf{d}^j_i = 0$ for $i \in N \setminus \{j\}$.

Then

$$\mathbf{d}^j_B = -\mathbf{B}^{-1}\mathbf{A}_j$$

The above theorem solve (1). The theorem below tackles (2).

Theorem 3.2 (Computation of $\bar{\theta}$)**.**
The largest possible $\bar{\theta}_j$ such that $\mathbf{x} + \bar{\theta}_j\mathbf{d}^j \geq \mathbf{0}$ is

$$\bar{\theta}_j = \min\left\{ \frac{(\mathbf{B}^{-1}\mathbf{b})_i}{(\mathbf{B}^{-1}\mathbf{A}_j)_i} \mid i \in B, (\mathbf{B}^{-1}\mathbf{A}_j)_i > 0 \right\}$$

Should the last condition $\mathbf{B}^{-1}\mathbf{A}_j > 0$ not hold, then $\mathbf{d}^j = (-\mathbf{B}^{-1}\mathbf{A}_j, \mathbf{d}_N^j) \geq 0$, and thus $\mathbf{x} + \theta\mathbf{d}^j \geq 0$ for all $\theta \geq 0$ and $\bar{\theta} = \infty$.
Remark: Note that if l is such that $\bar{\theta}_j = \frac{(\mathbf{B}^{-1}\mathbf{b})_l}{(\mathbf{B}^{-1}\mathbf{A}_j)_l}$, then $(\mathbf{x} + \bar{\theta}_j\mathbf{d}^j)_l = 0$ and $\mathbf{x} + \bar{\theta}_j\mathbf{d}^j$ is a BFS. Furthermore, j enters basis at that point and l leaves basis.

Definition 3.2 (Reduced Cost $\bar{c}_j$)**.**
Let $\mathbf{x}$ be a basic solution. Let $\mathbf{c} = (\mathbf{c}_B, \mathbf{c}_N)$. For each $j \in \{1, \ldots, n\}$, the **reduced cost** $\bar{c}_j$ of variable x_j is defined by

$$\bar{c}_j = \mathbf{c}^{\mathrm{T}}\mathbf{d}^j = c_j - \mathbf{c}_B^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{A}_j$$

Remark:

1. If $j \in B$, then $\bar{c}_j = 0$.
2. If $\bar{c}_j < 0$ and $\bar{\theta}_j > 0$ then $\mathbf{x} + \bar{\theta}_j\mathbf{d}^j$ is a guaranteed improved BFS.
3. If $\bar{c}_j < 0$ and $\bar{\theta}_j = \infty$, then problem is unbounded.
4. Reduced cost indicates if the adjacent BFS improves objective.

Definition 3.3 (Degeneracy of BFS)**.**
A BFS is **nondegenerate** if $\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} > 0$.

A BFS is **degenerate** if some elements of $\mathbf{x}_B$ is zero.[1]

Theorem 3.3 (Optimality Conditions)**.**
Consider a BFS $\mathbf{x}$ associated with basis matrix $\mathbf{B}$, and let $\bar{\mathbf{c}}$ be corresponding vector of reduced costs. We have the following results:

1. If $\bar{\mathbf{c}} \geq 0$, then $\mathbf{x}$ is optimal.
2. If $\mathbf{x}$ is optimal and non-degenerate, then $\bar{\mathbf{c}} \geq 0$.

Remark:

- Since $\bar{\mathbf{c}_B} = 0$ for all basic variables, to verify whether a BFS is optimal, we only need to check whether $\bar{\mathbf{c}_N} \geq 0$ for all non-basic variables.
- For optimality for a maximisation problem, check $\bar{\mathbf{c}} \leq 0$ instead.
- For degenerate case, optimal BFS need not have $\bar{\mathbf{c}} \geq 0$.

3.1 Simplex Method

The simplex method starts at a BFS, which is guaranteed existence for a feasible standard form LP, and continues with the following iterations:

1. Start with basis B and its basic columns $\mathbf{A}_B$ and BFS $\mathbf{x}$.

[1] All $\mathbf{x}_B \geq 0$ due to standard form LP.

2. Compute reduced costs $\bar{c}_j := c_j - \mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1} \mathbf{A}_j$ for all $j \in N$.

   - If $\bar{c}_j \geq 0$ for all $j \in N$, the **current BFS optimal**. END.
   - Otherwise, choose some j_* for which $\bar{c}_* < 0$
     The corresponding x_{j_*} is the **entering variable**.

3. Compute $\mathbf{u} = \mathbf{B}^{-1} \mathbf{A}_{j_*}$.

   - If $\mathbf{u} \leq 0$, then problem is **unbounded**, and algorithm ENDs.
   - Otherwise, let $\theta^* = \min\{\frac{\mathbf{x}_{B(i)}}{\mathbf{u}_i} \mid u_i > 0\}$

4. Let l be such that $\theta^* = \frac{\mathbf{x}_{B(l)}}{\mathbf{u}_l}$.
   The corresponding $\mathbf{x}_{B(l)}$ is the **leaving variable**.

5. Form a new basis by replacing $\mathbf{A}_{B(l)}$ with $\mathbf{A}_j$.

6. The other basic variables are $\mathbf{x}_{B(i)} - \theta^* \mathbf{u}_i$ for all $i \neq l$.

7. The entering variable $\mathbf{x}_j$ assumes $\theta^* = \frac{\mathbf{x}_{B(l)}}{u_l}$. Go to Step (1).

The Simplex Method can be carried out via a tableau. A generic simplex tableau has its initial form of Here, the first column entries can be obtained by the following method.

Basic	$\mathbf{x}$	Solution
$\bar{\mathbf{c}}$	$\mathbf{c}^{\mathrm{T}} - \mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1} \mathbf{A}$	$-\mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1} \mathbf{b}$
$\mathbf{x}_B$	$\mathbf{B}^{-1} \mathbf{A}$	$\mathbf{B}^{-1} \mathbf{b}$

Note that

$$\begin{aligned} \bar{\mathbf{c}} &= \mathbf{c} - \mathbf{p}\mathbf{A} \text{ where } \mathbf{p} = \mathbf{c}_B^{\mathrm{T}} \mathbf{B}^{-1} \\ &= \mathbf{c} - \sum_i p_i \mathbf{a}_i \end{aligned}$$

and we have $\mathbf{c}_B = 0$ for all basic variable. Therefore, we can solve $\mathbf{p}$ and then obtain $\bar{\mathbf{c}}$.
Remark: $\mathbf{B}^{-1}$ is obtained by the columns of the $\mathrm{Perm}(I)$, where $\mathbf{B}$ is the final basis matrix.

3.1.1 Finding the first BFS to start Simplex Method

To find the first BFS that enables Simplex Method, we need to introduce and solve **auxiliary** LP.

Definition 3.4 (Auxiliary LP)**.**
A standard form LP can be transformed into a auxiliary LP by

- Inverse the sign for appropriate rows of $\mathbf{A}$ and $\mathbf{b}$ so that $\mathbf{b} \geq 0$.

- Add a set of m artificial variables y's to constraints without positive slack[2], and use Simplex method on auxiliary LP:

$$
\begin{aligned}
\min \sum_{i=1}^{m} y_i & \qquad (1)\\
\text{s.t. } \mathbf{A}\mathbf{x} + \mathbf{y} = \mathbf{b} & \qquad (2)\\
\mathbf{x} \geq 0 & \qquad (3)\\
\mathbf{y} \geq 0 & \qquad (4)
\end{aligned}
$$

- Initialise auxiliary LP with $\mathbf{x} = 0$ and $\mathbf{y} = \mathbf{b}$.
- Starting basis matrix is the identity matrix.

After we solve the optimal solution for auxiliary LP as phase I, transfer the column and rows corresponding to basic variables and solution to the phase II, where the original LP is solved. Equivalently speaking, these end basic variables from phase I will be the starting basic variables in phase II.

Theorem 3.4.
We have the following:

- If LP is infeasible, it is detected at Phase I.
- If LP has optimum, it is detected at Phase II.
- If LP is unbounded, it is detected at Phase II.

Specifically, if the minimum value of auxiliary LP $\sum y \geq 0$, then at least one of the y cannot be non-basic variables, which indicates infeasibility of original LP.

3.2 Big-M Method

Big-M Method is similar to the simplex method, except that, after the auxiliary LP is constructed, we solve

$$\min \mathbf{c}^{\mathrm{T}} \mathbf{x} + M \sum_{i=1}^{m} y_i, \text{ where } M \gg 0$$

If the original LP is feasible and has finite optimal value, then

- all aritificial variables are eventually driven to zero, and
- the original objective function is minimised

[2]so either no slack or have negative slack like $-s_1$.

3.3 Special Cases

3.3.1 Degeneracy

According to the definition of degeneracy, a BFS with one or more zero basic variables is degenerate.
In Simplex Algorithm, a tie in the minimum ratio test leads to degeneracy. This is revealed by a 0 entry in the solution column for one or more of the basic variables.
Degeneracy may lead to cycling of BFS vistation. Therefore, anti-cycling rules like Bland's rule can be applied.

3.3.2 Alternative Optima

An LP will have more than one optimal solution, when objective function is parallel to a binding constraint.
In Simplex Algorithm, alternative optima is revealed by a 0 entry in the reduced cost of a non-basic variable, and some permissible choices of leaving variable[3].

Theorem 3.5 (Number of Alternative Optima)**.**
If an LP has $k \geq 2$ optimal BFS $\mathbf{x}^1, \ldots, \mathbf{x}^k$, then the LP has infinitely many optimal solutions. Moreover, if the set of optimal solutions is bounded, it will be

$$\operatorname{conv}(\mathbf{x}^1, \ldots, \mathbf{x}^k)$$

3.3.3 Unbounded Solution

Suppose the feasible set is unbounded, then there are three possibilities

- Finite objective value, bounded optimal set
- Finite objective value, unbounded optimal set
  $\min x_1$ s.t. $x_1 > 0, x_2 > 0$
- Optimum value $-\infty$ $\min -x_1 - x_2$ s.t. $x_1 > 0, x_2 > 0$

Unboundedness can be detected if constraint coefficients of a nonbasic variable x_j are all non-positive, which means x_j can be increased to infinity without violating $\mathbf{x} \geq 0$. We can say that $\mathbf{x}$ can move infinitely along the direction of d^j.
Moreover, if in addition, $\bar{c}_j < 0$, then objective value is $-\infty$.

3.3.4 Infeasibility

Infeasibility occurs when not all constraints can be satisfied.
Detection infeasibility in two-phase method is done in phase I. Detection in big-M method is to check, after an optimal tableau is obtained, $\mathbf{y} > 0$.

[3] The ratio test should have at least one denominator > 0

4 Duality Theory

Definition 4.1 (Dual Problem)**.** *Given a LP, denoted as* ***primal*** *LP, we associate it with another LP, denoted as* ***dual*** *LP, by following a set of mechanical rules.*
Specifically, for given matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ *with rows* $\mathbf{a}_i^{\mathrm{T}}$ *and columns* $\mathbf{A}_j$*, the primal problem on the left has a dual defined on the right.*

$$
\begin{array}{llll}
\text{(P)} & \min & \mathbf{c}^{\mathrm{T}}\mathbf{x} & \\
& & \mathbf{a}^{\mathrm{T}}\mathbf{x} \geq b_i & i \in M_+ \\
& & \mathbf{a}^{\mathrm{T}}\mathbf{x} \leq b_i & i \in M_- \\
& & \mathbf{a}^{\mathrm{T}}\mathbf{x} = b_i & i \in M_0 \\
& & x_j \geq 0 & j \in N_+ \\
& & x_j \leq 0 & j \in N_- \\
& & x_j \text{ free} & j \in N_{\mathbb{R}}
\end{array}
\qquad
\begin{array}{llll}
\text{(D)} & \max & \mathbf{p}^{\mathrm{T}}\mathbf{b} & \\
& \text{s.t.} & \mathbf{p}_i \geq 0 & i \in M_+ \\
& & \mathbf{p}_i \leq 0 & i \in M_+ \\
& & \mathbf{p}_i \text{ free} & i \in M_0 \\
& & \mathbf{p}^{\mathrm{T}}\mathbf{A}_j \leq c_j & j \in N_+ \\
& & \mathbf{p}^{\mathrm{T}}\mathbf{A}_j \geq c_j & j \in N_+ \\
& & \mathbf{p}^{\mathrm{T}}\mathbf{A}_j = c_j & j \in N_{\mathbb{R}}
\end{array}
$$

The above is the conversion rule from the primal LP to the dual LP. Specifically,

- Dual is “max” if primal is “min” and vice versa;
- Each constraint corresponds to a dual viariables
- The objective function is formed the 1-norm from RHS of contraints and dual variables
- To each primal variable, form dual constraints as follows:
  - Take the column in constraints and multiply with each dual variable to form LHS of constraints
  - Take the component in objective function to form RHS
  - Sign determined from the above conversion rule.
- The state of each dual variable is determined from table

Theorem 4.1 (The dual of the dual is the primal)**.**
If we transform the dual into an equivalent minimisation LP, and form its dual, then final LP is equivalent to original LP.

$$P \to D \to D' \to DD' \equiv P$$

Remark: An LP can be manipulated into equivalent forms by

- Introducing slack variables,
- Replacing free variables by difference of nonnegative variables

These techniques are introduced before during Standard Form LP.

Theorem 4.2.
Duals of equivalent problems are equivalent.
In other words, suppose $P_1 \equiv P_2$ and D_1 is dual of P_1 and D_2 dual of P_2. Then , $D_1 \equiv D_2$.

Theorem 4.3 (Weak Duality).
Suppose the primal LP (P) is a minimisation problem. Then,

- If $\mathbf{x}$ is feasible in (P) and $\mathbf{p}$ is feasible in (D), then
$$\mathbf{p}^{\mathrm{T}}\mathbf{b} \leq \mathbf{c}^{\mathrm{T}}\mathbf{x}$$
In other words, for general primal-dual LP pairs, objective of the minimisation problem is at least the objective of the maximisation problem.

- Thus,
$$\sup_{\mathbf{p}\text{ dual feasible}} \mathbf{p}^{\mathrm{T}}\mathbf{b} \leq \inf_{\mathbf{x}\text{ dual feasible}} \mathbf{c}^{\mathrm{T}}\mathbf{x}$$
i.e., the objective of (P) is at least that of (D).

Theorem 4.4.
Let $\mathbf{x}$ and $\mathbf{p}$ be primal and dual feasible, and $\mathbf{p}^{\mathrm{T}}\mathbf{b} = \mathbf{c}^{\mathrm{T}}\mathbf{x}$. Then $\mathbf{x}$ and $\mathbf{p}$ are primal and dual optimal respectively.

Theorem 4.5.
Unboundedness in one problem implies infeasibility in the other.

Remark: We have the possibilities of primal-dual pairs:

	Finite Opt.	Unbounded	Infeasible
Finite Opt.	$*$		
Unbounded			$*$
Infeasible		$*$	$*$

Theorem 4.6 (Strong Duality).
If an LP has an optimum, so does it dual, and both optimal costs are equal.

Theorem 4.7 (Computation of $\mathbf{p}^{\mathrm{T}}$).
From the previous theorem, we can calculate optimal dual $\mathbf{p}^{\mathrm{T}}$ in standard form LP by

- $\mathbf{p}^{T} = \mathbf{c}_B^{\mathrm{T}}\mathbf{B}^{-1}$, where B is an optimal basis for the primal LP. This optimal basis can be derived at the termination of Simplex Method.

- If there is a basis B_0 such that $\mathbf{A}_{B_0} = \mathbf{I}$, then
$$\bar{\mathbf{c}}_{B_0}^{\mathrm{T}} = \mathbf{c}_{B_0}^{\mathrm{T}} - \mathbf{c}_B^{\mathrm{T}}\mathbf{B}^{-1}\mathbf{A}_{B_0} = \mathbf{c}_{B_0}^{\mathrm{T}} - \mathbf{p}^{\mathrm{T}}$$
Thus, an optimal dual solution is $\mathbf{p}^{\mathrm{T}} = \mathbf{c}_{B_0}^{\mathrm{T}} - \bar{\mathbf{c}}_{B_0}^{\mathrm{T}}$.

Theorem 4.8 (Complementary Slackness).
Complementary Slackness is useful in checking whether a primal dual pair is optimal.
Let $\mathbf{x}$ and $\mathbf{p}$ be primal and dual feasible respectively. Then $\mathbf{x}$ and $\mathbf{p}$ are optimal if and only if

$$p_i(\mathbf{a}_i^{\mathrm{T}}\mathbf{x} - b_i) = 0 \text{ for all } i, \ \textbf{and}$$
$$(c_j - \mathbf{p}^{\mathrm{T}}\mathbf{A}_j)x_j = 0 \text{ for all } j$$

Theorem 4.9.
A consequence of the above theorem: suppose $\mathbf{x}$ is feasible. Then $\mathbf{x}$ is primal optimal if and only if there is a dual feasible $\mathbf{p}$ such that $(\mathbf{x}, \mathbf{p})$ satisfies complementary slackness.