Introduction to Optimization: Linear Programming

Mingheng Su

November 2024

Contents

List of Figures

1 Basic Definitions of Linear Programming

A general optimization problem can be defined as the following model:

$$\max / \min\{f(x) : g_i(x) \leq b_i, h_j(x) = b_j,\ \forall i = 1, ..., k, j = k+1, ..., m\}$$

where $b_i, b_j \in \mathbb{R}$ and $f(x), g_i(x), h_i(x)$ are functions on variable x, such that $f(x)$ is the objective function and $g_i(x), h_i(x)$ are inequality/equality constraints.

A *Linear Program* LP (in SEF) can be described as the following model:

$$\textbf{(LP)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

where $c^T x$ *s.t.* $c, x \in \mathbb{R}^n$ is the objective function, $Ax = b$ *s.t.* $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$ is the constraints, along with the non-negativity requirements. We aim to maximize the objective value $c^T x$ by choosing the maximizer x from the feasible region such that x satisfies $Ax = b, x \geq 0$, where the feasible region of **LP** is defined as $\mathcal{F} = \{x : Ax = b, x \geq 0\}$.

Theorem 1.0.1. *The Fundamental Theorem of Linear Programming (FTLP)*

1. *A LP is either infeasible, unbounded, or optimal.*
2. *If a LP is optimal, then there must exist an optimal solution that is basic, which is a vertex of the feasible region. If there exist at least two optimal solutions, then there will be infinite optimal solutions which include all points on the line segment between them.*

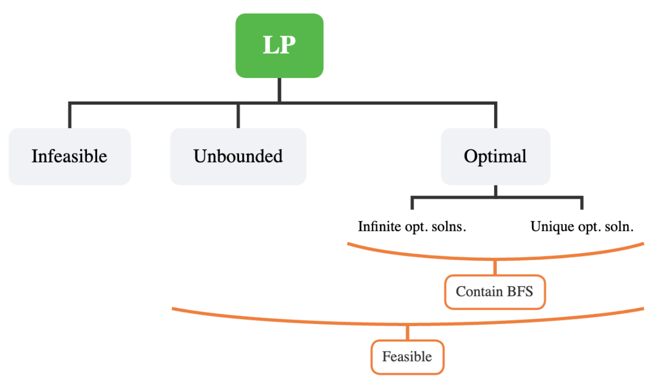

Figure 1.1: Types of Linear Programs

1.1 Standard Equality Form (SEF)

Suppose we have a general linear program

$$\textbf{(GLP)}\ \min\{c^T x : A_1 x?b_1, x?0\}$$

Recall that the standard equality form is

$$\textbf{(LP)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

We aim to transform the GLP to LP. Now assume the condition is messy such that the constraints consist of both equalities and inequalities, the objective is minimization, and some variables are non-negative, some are non-positive or free. We specify the approach to tidy them up step by step.

Objective Function When dealing with minimization problems, directly reverse the direction of optimization by multiplying a -1 to the objective function,

$$\min c^T x \implies \max -c^T x$$

But remember that this objective value is the negative of the original objective value, so once obtain an objective z, the real value is $-z$.

Constraints If a constraint that is already in equality then no change is required. For those inequality constraints, we introduce the *slack variable* for balance. For example,

$$x_1 + x_2 + x_3 \leq 5$$

Add a new non-negative slack variable s_1 such that

$$x_1 + x_2 + x_3 + s_1 = 5$$

Similar for $\geq$ constraints,

$$x_1 + x_2 + 2x_3 \geq 3 \implies x_1 + x_2 + 2x_3 - s_2 = 3$$

Note that if there are multiple inequality constraints, we need a slack variable for each of them. Thus if $Ax?b$ consists of inequalities only, we have

$$(A| \pm I)z = b, A \in \mathbb{R}^{m\times n}, (A| \pm I) \in \mathbb{R}^{m\times(n+m)}, z = \begin{bmatrix} x \\ s \end{bmatrix} \in \mathbb{R}^{n+m}, s, b \in \mathbb{R}^m$$

where the $\pm I$ is a $m \times m$ diagonal matrix whose diagonal entry $i = 1, ..., m$ is

$$\pm I_{ii} = \begin{cases} 1 & \text{if } i\text{th constraint is } \leq. \\ -1 & \text{if } i\text{th constraint is } \geq. \end{cases}$$

This ensures the equality constraints.

Variables If a variable x_i is already non-negative, no change is required. If a variable $x_i \leq 0$, reverse the corresponding objective entry c_i to $-c_i$ as well as the ith column of A, then change the variable to $x_i \geq 0$. It is more complicated if x_i is free such that, it can be negative, positive, or zero. To handle this, split and turn the variable into the difference of two non-negative variables such that $x_i = x^+ - x^-, x^+, x^- \geq 0$, and then augment the objective function and constraint as well. For example,

$$
\begin{gathered}
c = \begin{bmatrix} 1 & 2 & 3 \end{bmatrix} \\
A = \begin{bmatrix} 1 & 1 & 4 \\ 2 & 1 & 3 \end{bmatrix} \\
x_1 \leq 0, x_2 \geq 0, x_3 \text{ free}
\end{gathered}
$$

then

$$
\begin{gathered}
c = \begin{bmatrix} -1 & 2 & 3 & -3 \end{bmatrix} \\
A = \begin{bmatrix} -1 & 1 & 4 & -4 \\ -2 & 1 & 3 & -3 \end{bmatrix} \\
x_1, x_2, x^+, x^- \geq 0
\end{gathered}
$$

Now we have completed the transformation. The SEF is significant in various aspects, which we will see later.

Exercise 1.1.1. *Transform the following system into SEF.*

$$
\begin{gathered}
\min \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \end{bmatrix} x \\
\begin{bmatrix} 3 & 2 & 8 & 4 & 2 \\ 1 & 0 & 2 & 7 & 3 \\ 5 & 3 & 1 & 9 & 4 \end{bmatrix} x \begin{bmatrix} \geq \\ \leq \\ = \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 2 \end{bmatrix} \\
x_1 \geq 0, x_2, x_3 \leq 0, x_4, x_5 \textit{ free.}
\end{gathered}
$$

(Don't worry if the system is inconsistent, just obtain the SEF).

1.2 Feasibility

A LP is said to be feasible if there exists a feasible solution, that said the feasible region $\mathcal{F} \neq \emptyset$. Thus a certificate of feasibility is any feasible solution $\bar{x} \in \mathcal{F}$ such that $A\bar{x} = b, \bar{x} \geq 0$. The feasible region $\mathcal{F}$ is a convex polytope which is the intersection of finitely many half-spaces, each defined by an inequality constraint if assuming that the LP is not in SEF. We now look at the boundedness of $\mathcal{F}$,

The feasible region $\mathcal{F}$ is bounded if, and only if, $\|x\|_2 \leq M \ \forall x \in \mathcal{F}$, for some $M \geq 0 \in \mathbb{R}$.

If $\mathcal{F}$ is bounded then the objective value will be limited in some fixed range, which will never exceed $M\|c\|$ or below $-M\|c\|$ by the Cauchy-Schwartz Inequality $|c^T x| \leq \|c\| \|x\|$, but the bound might not be tight.

We call a LP infeasible if no feasible solution exists such that $\mathcal{F} = \emptyset$.

Proposition 1.2.1 (Certificate of Infeasibility)**.** *A certificate of infeasibility is*

$$y^T A \geq 0^T, y^T b < 0$$

for some $y \in \mathbb{R}^m$.

Note that y is free. Assume the LP is feasible where there exists feasible solution x satisfying $Ax = b, x \geq 0$. Then $y^T Ax = y^T b$. If $y^T A \geq 0^T$ then $y^T Ax \geq 0$ thus it is impossible to have $y^T b < 0$. In fact, the result comes from Farkas' Lemma,

Theorem 1.2.2 (Farkas' Lemma)**.** *For any LP, exactly one of the two following conditions is true:*

1. *(feasible)* $\exists x \in \mathbb{R}^n$ *such that* $Ax = b, x \geq 0$.
2. *(infeasible)* $\exists y \in \mathbb{R}^m$ *such that* $y^T A \geq 0^T, y^T b < 0$.

Example 1.2.1. $\mathcal{F} = \{x \in \mathbb{R}^2 : x_1 + x_2 \leq 5, x_2 \geq 7, x \geq 0\}$ *is infeasible.*
Set up the matrix in SEF

$$A = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & -1 \end{bmatrix}, b = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

A possible certificate is $y = (1, -1)^T$, *for which* $y^T A = (1, 0, 1, 1) \geq 0^T, y^T b = -2 < 0$.

Remark. *Ideally, the matrix system $Ax = b$ is consistent. When talking about infeasibility, we are referring to the violation of non-negative constraint $x \geq 0$ such that solution x contains negative entries. However, if $Ax = b$ is inconsistent and we still want to solve it, replace it with the normal equation and find an optimal least squares solution A.5.*

1.3 Basis

As stated in FTLP 1.0.1, we are interested in finding basic optimal solutions. In fact, when performing the simplex algorithm, we are actually looking for an *optimal basis*, whose corresponding basic feasible solution carries an optimal value. We first look at the definition of a basis.

Suppose the LP is in SEF with full row rank (delete redundant rows). Normally the constraint matrix in SEF is underdetermined, i.e. has more columns than rows. Although the matrix itself cannot be non-singular, we can pick some linearly independent columns to form a basis for $\mathbb{R}^m$.

Definition 1.3.1 (Bases). *A basis of A is a $m \times m$ non-singular matrix formed with m linearly independent columns of A, where m is the number of rows of A as well as its rank.*

The basis of A is indeed a basis for $\mathbb{R}^n$. Recall the basis of $\mathbb{R}^m$ can map all vectors in $\mathbb{R}^m$, as stated in Appendix A.1.3. If A has m rows, then there should be exactly m columns in the basis. Thus the number of bases is bounded which is at most $\binom{n}{m}$ as well as the maximum number of simplex iterations.

After determining the basis, we can use it to find the corresponding *basic solution*.

Definition 1.3.2 (Basic Solutions). *The basic solution with respect to basis B is defined as:*

$$\begin{aligned} x_{\mathcal{B}} &= A_{\mathcal{B}}^{-1} b \\ x_N &= 0 \end{aligned}$$

where $A_{\mathcal{B}}$ is the submatrix formed with columns in the basis B, x_B is the subvector of the basic entries (entries corresponding to the basic columns) of x, and x_N is the subvector of the rest entries of x.

Remark. *The basic entries might be 0, but non-basic entries are enforced to be 0. Thus we can conclude that the basic solution with respect to basis of size m has at most m non-zero entries.*

As $A_{\mathcal{B}}$ is non-singular, the corresponding basic solution is unique. According to the definition, the basic solution is actually using the basic columns to map b while remaining other entries 0. This will be emphasized in the following example.

Example 1.3.1.

$$A = \begin{bmatrix} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix}, b = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{bmatrix}$$

We see that it can have infinitely many solutions, but we are only interested in basic solutions. We start with picking a valid basis. Recall that a basis only consists of linearly independent columns, so for example, $\mathcal{B} = \{1, 2, 3\}$ is a valid basis, with corresponding basic solution

$$\begin{bmatrix} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix} x = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$

$$x = \begin{bmatrix} \underline{2} & \underline{4} & \underline{8} & 0 & 0 & 0 \end{bmatrix}$$

The underlined entries are the basic entries of x, which is $x_{\mathcal{B}}$, and the following entries are x_N. Now it is clear that we are finding a solution using only the basic columns. There are *many choices of bases, as shown below,*

$$\mathcal{B} = \{4,5,6\}$$
$$\begin{bmatrix} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix} x = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$
$$x = \begin{bmatrix} 0 & 0 & 0 & \underline{1} & \underline{2} & \underline{4} \end{bmatrix}$$

$$\mathcal{B} = \{1,5,6\}$$
$$\begin{bmatrix} 1 & 0 & 0 & 2 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2 \end{bmatrix} x = \begin{bmatrix} 2 \\ 4 \\ 8 \end{bmatrix}$$
$$x = \begin{bmatrix} \underline{2} & 0 & 0 & 0 & \underline{2} & \underline{4} \end{bmatrix}$$
$$\ldots$$

However, $\mathcal{B} = \{1,4,5\}$ is not a valid basis since columns $1,4$ are linearly dependent.

Exercise 1.3.1. *Find some other valid/invalid bases in 1.3.1.*

For a LP, a basic solution is feasible if and only if it is non-negative, however, in a general program it is feasible as long as it satisfies all constraints. We only focus on LP,

Definition 1.3.3. *Suppose the system $Ax = b$ is consistent and $\bar{x}$ is a basic solution with respect to basis $\mathcal{B}$. Then it is a basic feasible solution (BFS) if and only if $\bar{x} \geq 0$.*

One can easily understand the basic infeasible, non-basic feasible, non-basic infeasible solutions as well. As $x_N = 0$ is enforced, we can fit the above definition to

Proposition 1.3.4. *A basic solution $\bar{x}$ is feasible if and only if $\bar{x}_{\mathcal{B}} \geq 0$.*

The feasibility and other properties of a basis can be very apparent if we convert the LP to canonical form, which we will discuss below.

1.4 Canonical Form

Suppose $\mathcal{B}$ is a valid basis of LP. Recall the LP

$$\textbf{(LP)} \ \max\{c^T x : Ax = b, x \geq 0\}$$

We define its canonical form with respect to $\mathcal{B}$:

Definition 1.4.1 (Canonical Form). *The canonical form of LP with respect to basis $\mathcal{B}$ is in the shape of*

$$\begin{aligned}\bar{c}^T\bar{x} + \bar{z}\\ \bar{A}\bar{x} = \bar{b}\end{aligned}$$

where $\bar{c}_{\mathcal{B}} = 0, \bar{A}_{\mathcal{B}} = I_m, \bar{x}_{\mathcal{B}} = \bar{b}, \bar{x}_N = 0$, with objective value $\bar{z}$.

Example 1.4.1. *The following LP is in canonical form w.r.t. basis $\mathcal{B} = \{1, 2, 6\}$:*

$$\begin{aligned}\max\ &\begin{bmatrix}0 & 0 & -1 & 3 & 2 & 0\end{bmatrix} x + 5\\ &\begin{bmatrix}1 & 0 & 7 & 4 & -2 & 0\\ 0 & 1 & 3 & 1 & 5 & 0\\ 0 & 0 & 2 & 8 & 3 & 1\end{bmatrix} x = \begin{bmatrix}2\\1\\4\end{bmatrix}\\ \bar{x} = &\begin{bmatrix}2 & 1 & 0 & 0 & 0 & 4\end{bmatrix}\end{aligned}$$

with value 5. Observe that the basic entries of objective $c_{\mathcal{B}} = 0$, along with $A_{\mathcal{B}} = I, x_{\mathcal{B}} = b, x_N = 0$.

Remark. *Refer to the definition of canonical form in 1.4.1, note that*

1. *$\bar{b}$ is the basic entries of basic solution $\bar{x}$ with respect to $\mathcal{B}$.*
2. *$\bar{c}^T\bar{x} = 0$, thus the objective value of $\mathcal{B}$ is $\bar{z}$, which is 5.*

It is clear that if $A_{\mathcal{B}} = I$ then obviously $x_{\mathcal{B}} = b$ such that the augmented b is already a basic solution. For this convenience, we need to convert $A_{\mathcal{B}}$ to I. A straightforward approach will be multiply a $A_{\mathcal{B}}^{-1}$ (which always exists as $A_{\mathcal{B}}$ is non-singular) to the constraint

$$A_{\mathcal{B}}^{-1}Ax = A_{\mathcal{B}}^{-1}b$$

The new matrix $\bar{A}_{\mathcal{B}}$ will then be I, thus the basic solution is $x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}b$. Recall from the previous proposition 1.3.4, we can verify the feasibility of the basis:

Lemma 1.4.2. *A basis $\mathcal{B}$ is (primal) feasible if and only if $A_{\mathcal{B}}^{-1}b$ is non-negative.*

That said, the basis is infeasible if $A_{\mathcal{B}}^{-1}b$ contains negative entries. It is also essential to enforce $\bar{c}_{\mathcal{B}} = 0$ since

Lemma 1.4.3. $\bar{c}^T\bar{x} = \bar{c}_{\mathcal{B}}^T\bar{x}_{\mathcal{B}} + \bar{c}_N^T\bar{x}_N = 0$ *where $\bar{c}_{\mathcal{B}} = 0, \bar{x}_N = 0$.*

That said the inner product of canonical $\bar{c}, \bar{x}$ is always 0, since whatever what $\bar{x}_{\mathcal{B}}, \bar{c}_N$ is, they will be mapped to 0 by $\bar{c}_{\mathcal{B}}, \bar{x}_N$, as shown in example 1.4.1. The reason why we are making this assertion is because that based on this, we can come up with a necessary condition for optimality,

Theorem 1.4.4 (Canonical Optimality Condition)**.** *Assume the canonical form with respect to basis $\mathcal{B}$ has objective function $\bar{c}^T x + \bar{z}$. If $\bar{c} \leq 0$, then $\mathcal{B}$ is optimal as well as the basic solution $\bar{x}$, and $\bar{z}$ which is the constant following $\bar{c}^T \bar{x}$ is the optimal value.*

Proof. Assume $\bar{c} \leq 0$ holds, and as $x \geq 0$, we have $\bar{c}^T x \leq 0$ thus 0 is the maximum value it can reach, that said $\bar{c}^T \bar{x} + \bar{z} = 0 + \bar{z}$ (1.4.3) $= \bar{z}$ is the optimal value. $\square$

This is also referred to be the *dual feasibility* such that, if a basis is both primal and dual feasible, then it is an optimal basis. It is also the sign of the termination of simplex algorithm, and can lead to further optimality conditions such as complementary slackness theorem.

After understanding the importance of canonical form, we now continue to focus on how to transform c in this way. Keep in mind that, when converting c, we cannot change its attribute but transform it to our desired shape, that said the objective functions before and after transformation must remain equivalent such that $c^T x = \bar{c}^T x + \bar{z}$. Recall that the basic entries $\bar{c}_{\mathcal{B}} = 0$, so to achieve this, we still use the constraints. Define a vector $y \in \mathbb{R}^m$, which is referred to be the dual variable, then

$$y^T A x = y^T b \implies y^T A x - y^T b = 0$$

Remark. *A side note is that $y^T A$ is the linear combination of the rows of A, which results in a row vector.*

We first notice that, as $y^T A x - y^T b$ has a constant value 0, there will be no penalty if we add or subtract this from the objective function. Thus

$$c^T x = c^T x - (y^T A x - y^T b) = (c^T - y^T A)x + y^T b$$

is exactly the same function. Make use of the basic columns of $y^T A$, as we want to make $\bar{c}_{\mathcal{B}} = 0$, we have

$$y^T A_{\mathcal{B}} = c_{\mathcal{B}}^T$$

such that if this equality holds, then $\bar{c} = (c^T - y^T A)$ is exactly the shape we want (i.e. $\bar{c}_{\mathcal{B}} = 0$), as we do not care about the non-basic entries $\bar{c}_N$. Since $A_{\mathcal{B}}$ is non-singular, there exists a unique y such that $y = A_{\mathcal{B}}^{-T} c_{\mathcal{B}}$. Substitute this y back to the equation,

$$\begin{aligned} \bar{c}_N^T &= c_N^T - y^T A_N \\ &= c_N^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A_N \\ \bar{c}_{\mathcal{B}}^T &= 0^T \end{aligned}$$

Look back to the optimality conditions 1.4.4, since $\bar{c}_{\mathcal{B}} = 0$ is enforced, it can be fit to

Proposition 1.4.5. *In the canonical form with respect to primal feasible basis $\mathcal{B}$, if $\bar{c}_N \leq 0$ then $\mathcal{B}$ is an optimal basis.*

The converse also holds. The proof will be presented in later chapters on optimality. Note that the canonical form in example 1.4.1 is not maximum since the transformed $\bar{c}$ contains positive entries. Now we update and conclude the transformation of canonical form,

Proposition 1.4.6 (Canonical Form Transformation)**.** *For a LP*

$$\textbf{\textit{(P)}}\ \max\{c^T x : Ax = b, x \geq 0\},$$

The canonical form of LP with respect to basis $\mathcal{B}$ is in the shape of

$$\begin{aligned} \bar{c}^T \bar{x} + \bar{z} &= (c^T - \bar{y}^T A)\bar{x} + \bar{y}^T b \\ \bar{A}\bar{x} &= \bar{b} \end{aligned}$$

where $\bar{x}_{\mathcal{B}} = \bar{b} = A_{\mathcal{B}}^{-1} b, \bar{x}_N = 0, \bar{A} = A_{\mathcal{B}}^{-1} A, \bar{y} = A_{\mathcal{B}}^{-T} c_{\mathcal{B}}, \bar{c}_B = 0.$

Example 1.4.2. *s*

Remark. *Note that:*

1. *As its canonical $\bar{b}$ is non-negative, $\mathcal{B}$ is a primal feasible basis.*
2. *Observe that $\bar{c}^T \bar{x} = \bar{b}^T \bar{y}$, with unique $\bar{x}, \bar{y}$.*
3. *As $\bar{c} \leq 0$, this is an optimal state in canonical form, with value $\bar{z} =$.*

Exercise 1.4.1. *Convert the LP obtained in 1.1.1 into canonical form with any basis.*

1.5 Duality

For a primal LP, P,

$$\textbf{(P)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

its dual LP is defined as

$$\textbf{(D)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

Observe that the y system in 1.4.6 when transforming into canonical form is indeed the dual. This is the general conversion table,

Table 1: Primal-Dual Conversion Rules

MAX LP	MIN LP
$\textbf{(P)}\ \max\{c^T x : Ax = b, x \geq 0\}$	$\textbf{(D)}\ \min\{b^T y : A^T y \geq c, y \text{ free}\}$
$\leq$ constraint	≥ 0 variable
$=$ constraint	free variable
$\geq$ constraint	≤ 0 variable
≥ 0 variable	$\geq$ constraint
free variable	$=$ constraint
≤ 0 variable	$\leq$ constraint

Example 1.5.1 (Dual). *For a primal LP*

$$\textbf{\textit{(P)}}\ \max \begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix}x$$
$$\begin{bmatrix}1 & 3 & 5 & 7\\ 2 & 4 & 6 & 8\end{bmatrix}x = \begin{bmatrix}0\\1\end{bmatrix}$$
$$x \geq 0$$

its dual LP is

$$\textbf{\textit{(D)}}\ \min \begin{bmatrix}0 & 1\end{bmatrix}y$$
$$\begin{bmatrix}1 & 2\\ 3 & 4\\ 5 & 6\\ 7 & 8\end{bmatrix}y \geq \begin{bmatrix}1\\2\\3\\4\end{bmatrix}$$
$$y \text{ free}$$

Remark. *Note that:*

1. *To get the dual, the LP is first rotated in the right direction.*
2. *With the primal non-negativity constraint $x \geq 0$, the dual constraints has a $\geq$ sign.*
3. *With the primal equality constraint $Ax = b$, the dual variable y is free.*

We can now conclude the properties of a basis.

Theorem 1.5.1. *A basis $\mathcal{B}$ of A*

1. *is non-singular;*
2. *carries unique primal&dual solutions $\bar{x}$ and $\bar{y}$;*
3. *has unique primal&dual objective values $c^T\bar{x}$ and $b^T\bar{y}$*

The reference table of primal-dual pair

The above table could be concluded into the following sentences:

1. If the primal is optimal, then the dual must be optimal (by 2.1.3);
2. If the primal is unbounded, then the dual must be infeasible;
3. If the primal is infeasible, then the dual is either infeasible or unbounded.

The further properties of primal-dual pair on optimality will be discussed later.

2 Optimality

2.1 Duality and Optimality

Prior to optimality analysis, we first introduce some basic duality theory. Recall that for a primal LP,

$$\textbf{(P)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

its dual LP is defined as

$$\textbf{(D)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

A significant assertion is that $c^T x \leq b^T y$, that the primal objective value is always less than or equal to the dual's. To show this,

$$y^T b = y^T Ax \geq c^T x$$

The reason why we define a dual is to construct a game, such that $\max c^T x \leq \min b^T y$ which can help us find an optimal solution. That said, when $z = c^T x = b^T y$ such that they are equal, then z is the maximum objective value of P and the minimum objective value of D. Thus we can define that,

Proposition 2.1.1. *A certificate of optimality is*

$$c^T x = b^T y$$

such that x, y are a feasible solution to P, D respectively.

This states that x (resp. y) is also a maximizer(resp. minimizer) to P (resp. D), which is the optimal solution. This result comes from weak duality,

Theorem 2.1.2 (Weak Duality Theorem)**.** *Let x, y be a feasible solution to P, D respectively,*

1. $c^T x \leq b^T y$
2. *If $c^T x = b^T y$ then x, y are an optimal solution to P, D respectively.*

Note that the necessary condition is that x, y must be feasible, otherwise the equality does not guarantee an optimal value. What follows is strong duality,

Theorem 2.1.3 (Strong Duality Theorem)**.** *Consider a LP,*

1. *If both primal and dual are feasible, then they are optimal.*
2. *If one of the primal-dual pair is optimal, then the other one is also optimal, and their optimal values are equal.*

Let us go back to the optimality condition in 1.4.4, In the canonical form with respect to primal feasible basis $\mathcal{B}$, if $\bar{c} \leq 0$ then $\mathcal{B}$ is an optimal basis.

We now prove it using weak duality theorem. Assume $\bar{c}_N \leq 0$. First we claim that,

Lemma 2.1.4. *$\bar{c}^T = (c^T - y^T A) \leq 0$ is the dual constraint of the primal LP such that $c^T \leq y^T A$. If there exists any y that satisfies this constraint, then y is a feasible solution to the dual LP.*

As y is free, it is feasible as long as the dual constraint (the inequality) holds. Thus the $\bar{y}$ obtained in the canonical form is feasible. We also show that

Theorem 2.1.5. *Let $\bar{x}, \bar{y}$ be the basic primal&dual solutions obtained from the canonical form with respect to some basis $\mathcal{B}$, then*

$$c^T \bar{x} = b^T \bar{y}$$

Proof. $c^T \bar{x} = (c^T - \bar{y}^T A)\bar{x} + \bar{y}^T b = 0 + \bar{y}^T b = \bar{y}^T b$ where $\bar{c}^T \bar{x} = 0$ by 1.4.3. $\square$

(Note: Feasibility is not necessary in this theorem)

Hence by weak duality theorem, they are optimal. Notice that, this equality states that during the transformation of canonical form with respect to any basis, $c^T \bar{x}$ always equals $b^T \bar{y}$. Thus once find that $\bar{x}, \bar{y}$ are both feasible, we can conclude the optimality. Rearrange these we obtain a stronger theorem,

Lemma 2.1.6. *If a basis $\mathcal{B}$ is both primal and dual feasible, then it is an optimal basis.*

In fact this is similar to the proof of weak duality theorem. The converse is also true. The strong duality theorem states that, when optimal, both the primal-dual pair must be optimal. This enforces the feasibility as well. Therefore, we conclude that

Theorem 2.1.7 (Optimality Condition)**.** *A basis $\mathcal{B}$ is optimal if and only if it is both primal and dual feasible.*

We claim that a LP is optimal if and only if there exists at least one optimal basis.

2.2 Amount of Optimal Solutions

As stated in the FTLP 1.0.1, if a LP has at least two optimal solutions, then all points on the line segment (which also lies in the feasible region by convexity) between them are optimal. Here comes the following theorem,

Theorem 2.2.1. *For a LP*

$$\textbf{\textit{(LP)}} \ \max\{c^T x : Ax = b, x \geq 0\}$$

if there exist at least two optimal solutions x_1, x_2, then all points on the line segment between *them are optimal.*

Proof. Say the optimal value is $c^T x_1 = c^T x_2 = z$. For all $\lambda \in [0, 1]$, let $\lambda x_1 + (1 - \lambda)x_2$ denote the line segment between x_1, x_2.

Feasibility The line segment is feasible such that $A(\lambda x_1 + (1-\lambda)x_2) = \lambda A x_1 + (1-\lambda)Ax_2 = \lambda b + (1 - \lambda)b = b$, by assuming that optimal solutions are feasible.

Optimality The line segment is optimal such that $c^T(\lambda x_1 + (1 - \lambda)x_2) = \lambda c^T x_1 + (1 - \lambda)c^T x_2 = \lambda z + (1 - \lambda)z = z$. □

Thus, an optimal LP only has two conditions: either unique optimal solution or infinite optimal solutions.

2.3 Complementary Slackness Conditions

The C.S. conditions are strongly related to optimality.

Definition 2.3.1 (Complementary Slackness Conditions)**.** *For a LP, its complementary slackness conditions are*

1. $x_j = 0$ *or the corresponding* j*th dual constraint is tight.*
2. $y_i = 0$ *or the corresponding* i*th primal constraint is tight.*

A fastest way to understand these messy conditions is, this is exactly the shape of canonical form. Recall from 2.1.5 and 2.1.6, if the LP is in this shape and has corresponding feasible solution $\bar{x}, \bar{y}$, then they are optimal. Thus this leads to

Theorem 2.3.2 (Complementary Slackness Theorem)**.** *For a LP, if there exist feasible solutions* $\bar{x}, \bar{y}$ *such that the complementary slackness conditions 2.3.1 hold, then they are optimal solutions.*

Again, note the importance of feasibility.

3 Unboundedness

3.1 Unboundedness of LP

A LP is unbounded if and only if it is feasible but does not have an optimal value, such that the objective value can go to infinity. Note the unboundedness of feasible region does not guarantee the unboundedness of the LP. The unboundedness can be verified using the certificate.

Proposition 3.1.1 (Certificate of Unboundedness)**.**

$$\begin{aligned} c^T d > 0 \ (\boldsymbol{a.k.a}\ d \textit{ is an objective direction}) \\ Ad = 0 \ (\boldsymbol{a.k.a}\ d \in \operatorname{Null} A) \\ d \geq \neq 0 \ (\boldsymbol{a.k.a}\ d \textit{ is a feasible direction}) \end{aligned}$$

If a LP is unbounded, then we can find two non-negative feasible solutions $x_1 \leq \neq x_2$ such that $Ax_1 = b, Ax_2 = b$ with $c^T x_1 < c^T x_2$. That said, x_2 has at least one entry that is greater than that of x_1 and the objective value $c^T x_2$ is also greater than $c^T x_1$. If we keep increasing x in this direction, the objective value is likely increasing as well which is desired in the max problem. Rearrange above, we call $\vec{d} := x_2 - x_1 (d \geq \neq 0)$ the *direction of increment* such that

$$Ad = A(x_2 - x_1) = Ax_2 - Ax_1 = b - b = 0$$

by linearity and given that x_1, x_2 are feasible, and

$$c^T d = c^T (x_2 - x_1) = c^T x_2 - c^T x_1 > 0.$$

This demonstrates the derivation of the unboundedness certificate. The following example shows how d looks.

Example 3.1.1.

$$A = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

Note the existence of positive and negative columns such that there might exist some non-negative null vector. That said, as x_3 grows infinitely large, x_1, x_2 can always be increased to cancel out x_3. For example, $d = t(1,1,1)^T$ is a feasible direction for any $t \geq 0$ including $\lim_{t \to \infty} t(1,1,1)^T$. Hence the LP is unbounded once $c^T d > 0$ is satisfied.

Example 3.1.2.

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$$

The nullspace is $\{(1,1,-1)^T\}$ which does not contain a feasible d, thus the feasible region is bounded (Note that $d := (1,1,-1)^T$ is also an unbounded direction if the $x \geq 0$ constraint is not present). An optimal solution exists if the LP is feasible. The proof will be presented below.

Condition 1 The first necessary condition of unboundedness is that the feasible region is unbounded. To illustrate this, we focus on d itself. Normalize $\hat{d}$ to demonstrate the direction only, and define a special line $\vec{x} = \bar{x} + td, t \geq 0 \in \mathbb{R}, \bar{x} \in \mathcal{F}$, where t is the length of the line and $\bar{x}$ is any point in the feasible region. That said, $\vec{x}$ is a line in direction d starting at any feasible point $\bar{x}$ which is obviously unbounded and feasible (i.e. $\vec{x} \in \mathcal{F}$ where any point on $\vec{x}$ is feasible),

$$A\vec{x} = A(\bar{x} + td) = A\bar{x} + tAd = b + t \cdot 0 = b, \vec{x} \geq 0.$$

This proves that $\mathcal{F}$ must be unbounded. In fact, if there exists any feasible direction d, then $\mathcal{F}$ is unbounded. The contrapositive states that any bounded $\mathcal{F}$ must contain an optimal solution, assuming that $\mathcal{F} \neq \emptyset$. Furthermore, a LP could have various feasible directions, where in this case are the non-negative null vectors, that said there might exist multiple lines $\vec{x}$.

Condition 2 Assuming that the first necessary condition holds, the second necessary condition of unboundedness is that the objective value $c^T x$ keeps increasing in the desired direction (e.g. growing in a max problem or declining in a min problem) as x continuously moves forward along the line $\vec{x}$, which we call an *objective direction*. For convenience we focus on max problems. Say $x := \bar{x} + td$ is a point on the line $\vec{x}$ where x moves along $\vec{x}$ as t increases, and we can see the change in objective value as t grows $\bar{t}$ is $c^T(\bar{t}d) = \bar{t}(c^T d)$. Hence, once $c^T d > 0$, any increment in t will result in an increase in the objective value. Equivalently, as t grows infinitely, the objective value will be

$$\lim_{t \to \infty} c^T(\bar{x} + td) = c^T \bar{x} + \lim_{t \to \infty} t(c^T d) = \lim_{t \to \infty} t(c^T d)$$

ignoring the constant $c^T \bar{x}$ and, if $c^T d > 0$, the objective value goes to ∞ which is unbounded.

Exercise 3.1.1. *Analyze the optimality and boundedness of the cases $c^T d \leq 0$. Where will the optimal solution be at? What about min problems?*

Exercise 3.1.2. *Assume that the LP model in 3.1.1 is feasible. Determine some objective vectors c such that the LP is optimal/unbounded. Do the same for*

$$\begin{bmatrix} 1 & 0 & 0 & -1 & -1 & 1 \\ 0 & 1 & 0 & -3 & 1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \end{bmatrix}$$

Start with finding all unbounded directions, and determine the optimal/unbounded objective functions for each of them.

3.2 Unboundedness of Free LP

The next more general model provides a better insight into the understanding of unboundedness. Obtain the new model by removing the non-negativity constraint,

$$\textbf{(LPfree)} \ \max\{c^T x : Ax = b\}.$$

Without the $x \geq 0$ bounds, x can go any direction. Assume that x_1, x_2 are two feasible solutions such that $Ax_1 = b, Ax_2 = b$. Then both $x_2 \geq x_1, x_2 \leq x_1$ are fine, just make sure $x_1 \neq x_2$, and thus the new feasible direction will be $d := x_2 - x_1 \neq 0$, where the angle θ of d is not limited to be in the range $\theta \in [0, \frac{\pi}{2}]$ only. We claim that,

Lemma 3.2.1. *Any non-zero vector in the nullspace of A is a feasible direction.*

That said if the nullity of A is n, then there will be n linearly independent feasible directions d. The following also come from the fundamental linear algebra,

Lemma 3.2.2. *If d is a feasible direction, so does $-d$. As well, $c^T d > 0$ iff $c^T(-d) < 0$.*

We still use the definition of the special line from above, $\vec{x} := \bar{x} + td$. Recall that $\bar{x} \in \mathcal{F}$, $t \geq 0$, and d is normalized. The only difference is that now $\vec{x}$ could be negative but is still feasible in this model such that $\vec{x} \in \mathcal{F}$.

Assume that the above first necessary condition holds. We analyze the change in the objective value $c^T x$ as x moves forward along the line $\vec{x}$. Recall, the ending objective value depends on

$$\lim_{t \to \infty} c^T(\bar{x} + td) = \lim_{t \to \infty} t(c^T d)$$

ignoring the constant $c^T \bar{x}$. If $c^T d > 0$ then the LP is obviously unbounded. However, the above lemma states that, when $c^T d < 0$, pick $d^- = -d$ is still an objective direction such that $c^T d^- > 0$, where we know d^- is also an unbounded direction by 3.2.2. Rearrange these we have that any $c^T d \neq 0$ will guarantee the unboundedness. To conclude,

Proposition 3.2.3 (Certificate of Unboundedness of LPfree)**.**

$$\begin{aligned} c^T d &\neq 0 \\ Ad &= 0 \\ d &\neq 0 \end{aligned}$$

Exercise 3.2.1. *Repeat 3.1.2 using the model **LPfree**, such that the $x \geq 0$ constraint is removed.*

Exercise 3.2.2. *Suppose a LPfree only has one linearly independent feasible direction d. If $c^T d = 0$, where will the optimal solution be at?*

4 Solving Optimization Problems

Now we focus on how to solve the optimization problems. As stated in FTLP 1.0.1, there must exist an optimal solution that is basic. Thus we aim to check each basis as well as their values and feasibility, until a basis satisfying the optimality conditions is found. Instead of comparing the values of every single basis, we introduce a powerful algorithm which strongly relies on optimality conditions.

4.1 Auxiliary Problems

Before finding the optimal solution, we first try to find out an starting basic feasible solution.

Definition 4.1.1 (Auxiliary Problems)**.** *For a LP*

$$\textbf{\textit{(LP)}}\ \max\{c^T x : Ax = b, x \geq 0\}$$

where $A \in \mathbb{R}^{m \times n}, x \in \mathbb{R}^n, b \in \mathbb{R}^m$*, its corresponding auxiliary problem is defined as*

$$\min\ \{t_1 s_1 + ... + t_n s_n : (A|I_m) \begin{bmatrix} x_1 \\ ... \\ x_n \\ s_1 \\ ... \\ s_m \end{bmatrix} = b, s \geq 0\}$$

for some $t_1, ..., t_m \in \mathbb{R}_+$*.*

Normally the $t \in \mathbb{R}^m_+$ is chosen to be $\mathbb{1} \in \mathbb{R}^m$ which is a vector of 1's, but in fact any positive numbers could work.

Example 4.1.1. *For an original LP*

$$\begin{aligned} \max\ & \begin{bmatrix} 2 & 5 & 1 & 3 \end{bmatrix} x \\ & \begin{bmatrix} 1 & 6 & 4 & 5 \\ 1 & 3 & 1 & 4 \end{bmatrix} x = \begin{bmatrix} 25 \\ 13 \end{bmatrix} \\ & x \geq 0 \end{aligned}$$

its auxiliary problem is

$$\begin{gathered} \min\ s_1 + s_2 \\ \begin{bmatrix} 1 & 6 & 4 & 5 & 1 & 0 \\ 1 & 3 & 1 & 4 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 25 \\ 13 \end{bmatrix} \\ (x, s) \geq 0 \end{gathered}$$

Note that the two problems have the same size of basis, which is 2 in this case since the number of rows is not changed. If we solve the auxiliary optimization problem and get an optimal basic feasible solution with value 0, then eliminating the extra s entries will give a basic feasible solution to the original LP.

Proposition 4.1.2 (Feasibility via Auxiliary)**.** *A LP has a feasible solution if and only if its auxiliary problem has feasible solution with value* 0*, which is optimal.*

Proof. For a LP, its ALP is defined as $\min\ \{t^T s : \bar{A}(x, s) = b\}$.
$(\implies)$ Assume the LP has a feasible solution $\bar{x}$. Then append m zero entries after $\bar{x}$ will be a feasible solution to the ALP, and is also optimal with value 0.
$(\impliedby)$ Assume the ALP has an optimal and feasible solution $(\bar{x}, \bar{s})$ with value 0 where $A(\bar{x}, \bar{s}) = b, (\bar{x}, \bar{s}) \geq 0$. Note that the objective vector of ALP t only contains positive entries, that said the auxiliary variables s must be 0. Given $\bar{x}$ is also feasible as $A\bar{x} = b, \bar{x} \geq 0$ since $\bar{A}(\bar{x}, \bar{s}) = A\bar{x}, (\bar{x}, \bar{s}) \geq 0$, it is a feasible solution to the original LP. □

Remark. *The contrapositive states that a LP is infeasible if and only if its auxiliary problem has an optimal value that is greater than* 0*. This can be used to detect the feasibility of LP.*

This proposition holds for general feasible solutions, but we can focus on basic feasible solutions only.

Corollary 4.1.3. $(\bar{x}, \bar{s})$ *is a basic feasible solution to ALP with value* 0 *if and only if* $\bar{x}$ *is a basic feasible solution to LP.*

Proof. By 4.1.2, $\bar{s} = 0$. As they have the same size of basis, $(\bar{x}, \bar{s})$ has at most m positive entries, and as $\bar{s} = 0$, $\bar{x}$ has at most m positive entries. Thus $\bar{x}$ is a basic feasible solution to LP. The converse can be proved by appending m zero entries to the basic feasible solution to LP. □

Example 4.1.2. *Recall example 4.1.1,* $(1, 4, 0, 0, 0, 0)^T$ *is an optimal and basic feasible solution to ALP with vale* 0*, and thus* $(1, 4, 0, 0)^T$ *is a basic feasible solution to LP.*

The auxiliary problems work well with simplex method. As stated before, we need to start with a basic feasible solution prior to performing the simplex algorithm. Note that the appended identity matrix is a great choice of basis, thus a starting basic feasible solution is choosing $s = b$ while remaining other entries zero, and then perform network simplex algorithm to find the optimal solution with value 0, or deduce the infeasibility of LP. For example, in 4.1.1, we can start with $(0, 0, 0, 0, 25, 13)^T$ as a basic feasible solution to ALP, and then apply simplex method on ALP to obtain an optimal basic feasible solution. If the optimal value is 0, then the first four entries of the optimal solution is a basic feasible solution to LP, otherwise LP is infeasible. Now we can officially start the simplex method on LP. We conclude the procedure of solving a LP:

Method to solving a LP (two-phase simplex method):

1. Use the auxiliary problem ALP to find a basic feasible solution of LP with its basis, or deduce that it is infeasible.

2. Use simplex method to find an optimal solution of LP, or prove it is unbounded.

4.2 Tableau

Consider the LP

$$\textbf{(LP)}\ \max\{c^T x : Ax = b, x \geq 0\}$$

Recall that $A \in \mathbb{R}^{m\times n}, x, c \in \mathbb{R}^n, b \in \mathbb{R}^m$. Now consider the general case that there is a constant, say $k \in R$, following the objective function where $c^T x + k$ is the objective value. Take the variable z to represent the objective value, $z = c^T x + k$, then $z - c^T x = k$. Add z as an extra variable to the matrix, and stack this extra constraint on the matrix A

$$\begin{bmatrix} 1 & -c^T \\ \underline{0} & A \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} k \\ b \end{bmatrix}$$

The $\underline{0}$ is a vector of 0's and $\underline{0} \in \mathbb{R}^m$. Apparently the new matrix is in $\mathbb{R}^{(m+1)\times(n+1)}$. Note that $k \in \mathbb{R}$ is the constant following the original objective function $c^T x$. Rearrange these we have an augmented matrix (a.k.a tableau)

$$\begin{bmatrix} 1 & -c^T & k \\ \underline{0} & A & b \end{bmatrix}$$

Example 4.2.1. *Consider the LP*

$$\begin{aligned} \max\ & \begin{bmatrix} 1 & 2 & 3 & 1 \end{bmatrix} x + 5 \\ s.t.\ & \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} \\ & x \geq 0 \end{aligned}$$

The tableau is

$$\begin{bmatrix} \mathbf{1} & -1 & -2 & -3 & -1 & \underline{5} \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

We can also convert it into canonical form. For example, take basis $\{1, 2, 3\}$.

$$\begin{bmatrix} \mathbf{1} & 0 & 0 & 0 & 5 & \underline{9} \\ 0 & 1 & 0 & 0 & 1 & 2 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

which is optimal by the optimality condition 1.4.4 as the first row except for the first and last entries, which is $-\bar{c}^T$, is non-negative (note that their signs are reversed), and thus 9 is the optimal value.

When changing basis, normally we just pivot the new basic columns (using matlab or other pivoting calculators) and the whole matrix will be automatically transformed into the canonical form for the new basis, which is a great superiority of tableau.

4.3 Simplex Algorithm

The simplex algorithm switches among the valid bases until finding the optimal basis. First introduce the basic ideas of the algorithm.

Start with a feasible basis.

1. Transform the LP into the canonical form with respect to this basis.
2. Check the optimality of this basis. If the optimality conditions are not satisfied then switch to another basis and repeat step 1.
3. Stop if the basis is optimal.

We illustrate the method by starting with an example. To maintain simplicity, a totally unimodular matrix is used.

$$\begin{aligned}
\max\ &\begin{bmatrix}2 & 1 & 3 & 0\end{bmatrix}x \\
s.t.\ &\begin{bmatrix}1 & 1 & 1 & 0\\ 1 & 0 & -1 & 1\end{bmatrix}x = \begin{bmatrix}12\\ 6\end{bmatrix} \\
&x \geq 0
\end{aligned}$$

Recall we first find a starting basic feasible solution obtained from the auxiliary problem, if necessary. Start with Basis $\{2,4\}$. Convert the LP into canonical form with respect on this basis. The basic submatrix is already an identity matrix so no action is needed to transform the matrix.

$$\begin{aligned}
y &= A_{\mathcal{B}}^{-1}c_{\mathcal{B}} = (1,0)^T \\
c^T - y^T A &= (2,1,3,0) - (1,1,1,0) = (1,0,2,0) \\
y^T b &= (1,0)(12,6)^T = 12
\end{aligned}$$

The new matrix is

$$\begin{aligned}
\max\ &\begin{bmatrix}1 & 0 & 2 & 0\end{bmatrix}x + 12 \\
s.t.\ &\begin{bmatrix}1 & 1 & 1 & 0\\ 1 & 0 & -1 & 1\end{bmatrix}x = \begin{bmatrix}12\\ 6\end{bmatrix} \\
x &= \begin{bmatrix}0 & 12 & 0 & 6\end{bmatrix}
\end{aligned}$$

Note that entries $\bar{c}_1$ and $\bar{c}_3$ are positive and the optimality condition 1.4.4 is violated. That said increasing these two variables x_1, x_3 might potentially increase the objective value. This means one of them is going to enter the basis, while some already existing basic variable is going to be kicked out of the basis, as the basis cannot contain more than 2 variables.

Remark. *Once obtained the canonical form with respect to some basis, the non-basic variables with positive objective entries (a.k.a infeasible dual variables) are eligible to enter the basis. One of them will be chosen to enter the basis, and exactly one of the current basic variables will leave.*

Our interest is, which variable should enter the basis, and which one should leave? Although $\bar{c}_3$ has a larger value, increasing x_3 might not guarantee a larger value than increasing x_1, which depends on the constraints that limit how much they can increase. Thus we follow a rule for choosing the entering variable.

Remark. *When changing the basis, if there are multiple variables that are eligible to enter the basis, always pick the one with the smallest index, which is the leftmost one.*

So here, x_1 is eligible to enter the basis. Set $x_1 = t$ and increase it as much as possible. Note the constraints,

$$\begin{aligned} t + x_2 = 12 &\implies x_2 = 12 - t \geq 0 \\ t + x_4 = 6 &\implies x_4 = 6 - t \geq 0 \end{aligned}$$

Note that as currently x_3 is not a candidate basic variable, we can ignore it which keeps 0. When increasing t, the other two variables x_2, x_4 are decreased by the constraints, and as they cannot be reduced to be less than 0, for the sake of primal feasibility, there is a limit on the increase in t. One can imagine that x_1 is fighting against x_2 and x_4, and ultimately one of x_2, x_4 will lose and leave, giving up its place to x_1. From the two inequalities, the maximum value that t can reach is 6, and now $x_4 = 6 - t = 0, x_2 = 12 - t = 6$, such that x_4 leaves the basis. Therefore the new basis is $\{1, 2\}$. Convert the LP into canonical form with respect to this basis,

$$\begin{aligned} \max \ & \begin{bmatrix} 0 & 0 & 3 & -1 \end{bmatrix} x + 18 \\ s.t. \ & \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 2 & -1 \end{bmatrix} x = \begin{bmatrix} 6 \\ 6 \end{bmatrix} \\ & x = \begin{bmatrix} 6 & 6 & 0 & 0 \end{bmatrix} \end{aligned}$$

The objective value has increased to 18, but we are not done yet, since the optimality conditions are not met. Now the only choice of entering variable is x_3. Again, set $x_3 = t$, and look at the constraints

$$\begin{aligned} x_1 - t = 6 &\implies x_1 = 6 + t \geq 0 \\ x_2 + 2t = 6 &\implies x_2 = 6 - 2t \geq 0 \end{aligned}$$

The first inequality does not provide an upper bound for t, but the second one does, which is 3. Thus $x_1 = 6 + t = 9, x_2 = 6 - 2t = 0$, and x_2 leaves the basis. Transform the LP to canonical form w.r.t $\{1, 3\}$.

$$\begin{aligned} \max \ & \begin{bmatrix} 0 & -3/2 & 0 & 1/2 \end{bmatrix} x + 27 \\ s.t. \ & \begin{bmatrix} 1 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1 & -1/2 \end{bmatrix} x = \begin{bmatrix} 9 \\ 3 \end{bmatrix} \\ & x = \begin{bmatrix} 9 & 0 & 3 & 0 \end{bmatrix} \end{aligned}$$

Again, x_4 enters. Set $x_4 = t$.

$$\begin{aligned} x_1 + (1/2)t = 9 &\implies x_1 = 9 - (1/2)t \geq 0 \\ x_3 - (1/2)t = 3 &\implies x_3 = 3 + (1/2)t \geq 0 \end{aligned}$$

So $t = 18$, and now $x_1 = 6 - (1/2)t = 0, x_2 = 6 + (1/2)t = 12$. x_1 leaves. New basis $\{3, 4\}$.

$$\begin{aligned} \max \ & \begin{bmatrix} -1 & -2 & 0 & 0 \end{bmatrix} x + 36 \\ s.t. \ & \begin{bmatrix} 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 12 \\ 18 \end{bmatrix} \\ x = & \begin{bmatrix} 0 & 0 & 12 & 18 \end{bmatrix} \end{aligned}$$

Check the optimality conditions 1.4.4, we beautifully ended up at this basis, with an optimal value 36. For summary, we tide up the critical steps we have used.

Starting Feasible Basis Observe or obtain a feasible basis from the corresponding auxiliary problem, and convert to its canonical form. Deduce the optimality, otherwise change to another feasible basis.

Entering Variable In the canonical form, the leftmost non-basic variable with a positive objective entry enters the basis.

Leaving Variable On each row i, the entering variable x_j is battling with the pivot x_i (exactly one basic variable on/dominating each row) on that row such that $x_i + a_{ij}x_j = b_i \implies x_k = b - a_{ij}x_j$. Ideally, their coefficients have the same signs, and increasing x_j will decrease the pivot variable x_i on that row. To keep the primal feasibility, x_j stops at the value when one of the current basic variables is reduced to 0, which we call 'the first basic variable that dies as x_j increases', such that $x_i = b_i - a_{ij}x_j \geq 0 \implies x_j \leq \frac{b_i}{a_{ij}}, \forall i = 1, ..., m$. Then x_j replaces the pivot dominating that row to be the new basic variable, while the original basic variable leaves. This can be summarized to be: Basic variable x_i leaves where $i \in \operatorname{argmin}\{\frac{b_i}{a_{ij}} : a_{ij} > 0, b_i > 0, \forall i = 1, ..., m\}$, i.e. pick the variable with min ration. If there are many choices, just pick any one that is eligible to leave. We conclude the procedure of the algorithm.

Algorithm 4.3.1 (Simplex Algorithm)**.** *For a LP, starting at any feasible basis,*

1. *Convert the LP w.r.t this basis. Deduce the optimality, otherwise go to step 2.*
2. *Change the basis where:*
   (a) *The leftmost non-basic variable with a positive objective entry, x_j, enters the basis.*
   (b) *One of the basic variable x_i satisfying $i \in \operatorname{argmin}\{\frac{b_i}{a_{ij}} : a_{ij} > 0, b_i > 0, \forall i = 1, ..., m\}$ leaves. Deduce the unboundedness if the set is empty, otherwise repeat step 1.*

4.4 Simplex with Tableau

As mentioned before, life gets easier with simplex tableau. We can now avoid the tedious procedures 1.4.6 for converting the LP into canonical form, just pivot the basic columns.

4.5 Correctness of Simplex Algorithm

4.5.1 Basis Change

Recall the method for changing basis during simplex, we now prove the validity of the new picked basis. Refer to the definition 1.3.1, a valid basis has exactly m linearly independent columns of A, where m is the number of rows in A. As each time we add in a new column and kick out a column from the original basis, the number of columns in the new basis is still m. Now we prove the linear independence of the new basis' columns. Let us say the non-basic variable x_j is entering the basis, along with a_j, the jth column of A. Extract the basic columns of current canonical form (which is the identity matrix with size m) and a_j, $(I_m|a_j)\begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = b$, and in particular,

$$I_m x_{\mathcal{B}} + x_j \mathbf{a_j} = b$$

Say a_{ij} is the ith entry of $\mathbf{a_j}$, and x_i is the ith basic variable, $i = 1, ..., k$, $k \leq m$, and assume we have dropped those $(m-k)$ rows with $a_{ij} \leq 0$, then it becomes

$$\begin{gathered} x_1 + a_{1j} x_j = b_1 \\ \dots \\ x_k + a_{kj} x_j = b_k \end{gathered}$$

When maximizing x_j, the x_i with smallest $\frac{b_i}{a_{ij}}$ is reduced to 0, giving its pivot position to x_j, which we call x_i is 'killed'. Now we see that the number of pivots is unchanged, which is still m, that said the rank is unchanged, thus the new basis is still non-singular. Note that when there are multiple x_i's are killed, just pick any of them to leave the basis, or it is equivalent to say that the new basic solution $x_{\mathcal{B}}$ contains 0. These will be illustrated by the following examples.

Example 4.5.1. *Consider the following extracted system* $(I_m|a_j)\begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = b$,

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$
$$x = \begin{bmatrix} 6 & 6 & 6 & 0 \end{bmatrix}$$

where the first three columns are the basic columns, and the last column is $\mathbf{a_j}$, *the column for the candidate basic variable* x_j. *As we can see that* $\mathbf{a_j}$ *is linearly dependent and battling with all three basic columns, the maximum* x_j *can reach is* 2, *for where* x_3 *is killed, which becomes*

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix} \begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$
$$x = \begin{bmatrix} 4 & 2 & 0 & 2 \end{bmatrix}$$

In this case x_j *defeats and replaces exactly one linearly dependent column, which is column* 3, *and take the pivot on row 3. The rank of the new candidate basis* $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$ *is 3, which is non-singular and a valid basis.*

Example 4.5.2. *Consider the following extracted system* $(I_m|a_j)\begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = b$*,*

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$
$$x = \begin{bmatrix} 6 & 6 & 6 & 0 \end{bmatrix}$$

In this case there are multiple basic variables to be 'killed'. We first maximize x_j which is 2, and it becomes

$$\begin{bmatrix} 1 & 0 & 0 & 3 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$
$$x = \begin{bmatrix} 0 & 0 & 4 & 2 \end{bmatrix}$$

Then any one of the original basic variables, x_1 and x_2, has to leave the basis. In fact with the canonical form in this shape, there are multiple choices of bases, i.e. $\mathcal{B} = \{1, 3, j\}$ or $\mathcal{B} = \{2, 3, j\}$ (Recall that j is the last column).

Example 4.5.3. *Consider the following extracted system* $(I_m|a_j)\begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = b$*,*

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_{\mathcal{B}} \\ x_j \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \\ 6 \end{bmatrix}$$
$$x = \begin{bmatrix} 6 & 6 & 6 & 0 \end{bmatrix}$$

In this case we do not drop the rows with $a_{ij} \leq 0$, and see what happens. First notice the second row, as $a_{2j} = 0$, there is no way for x_j to defeat and take over the pivot x_2 on this row, thus it is useless in determining the maximum of x_j. The third row is more interesting such that it is an unbounded row, where there is also no limit on the increase in x_j. Thus the supreme of x_j only depends on the first row, which is 6. If the first row is not present such that there are no rows with positive a_{ij}, then there exists an 'unbounded direction' 3.1 which suggests a potential unbounded program, whose unboundedness finally depends on the objective function 3.1. This will be discussed later.

4.5.2 Feasibility of New Basis

When picking a new basis, the simplex method avoids the primal infeasible bases, to fulfill the optimality conditions 2.1.7. As we've seen in previous sections, when increasing the candidate basic variable, the supreme from the pivot rows is actually protecting the current basic variables from being negative. Hence once we follow the procedure of the simplex method, we will always find a primal feasible basis.

4.6 Termination of Simplex Algorithm

As stated in the Algorithm 4.3.1, the simplex method will ultimately stop, with various consequences.

4.6.1 Simplex on Optimal LP

As simplex only works with basic feasible solutions, it stops at the final solution where the optimality conditions are met. As a LP always has a basic optimal solution 1.0.1, the simplex will always find it.

4.6.2 Simplex on Unbounded LP

As a great insight on the simplex algorithm 4.3.1, the LP is unbounded if the set mentioned in the table is empty. We start with an example from the section of unboundedness 3.1.1, where

4.7 Dual Simplex

5 Polyhedra

Focus on geometry properties of the feasible region $\mathcal{F}$ of LP.

5.1 Convexity

5.1.1 Convex Sets

Definition 5.1.1 (Line Segment). *The line segment between two points x, y is defined as $[x, y] = \lambda x + (1 - \lambda)y, \forall \lambda \in [0, 1]$.*

This is indeed the closed interval between x and y, but extended to the multi-dimensional space.

Definition 5.1.2 (Convex Sets). *A set $\mathcal{S}$ is convex if and only if $\forall x, y \in \mathcal{S}, \forall \lambda \in [0, 1], \lambda x + (1 - \lambda)y \in \mathcal{S}$.*

That said the line segment between them is contained in $\mathcal{S}$ as well. In addition, any interior inside a convex set is also convex.

Example 5.1.1. *A classic example of a convex set is the open ball $\{\|x\| < r : r \in \mathbb{R}\}$. For any two points in the ball, drawing a line connecting them also lies in the ball. Another example is polyhedron, shown below on the left, which is the intersection of finitely many halfspaces, as we will discuss later.*

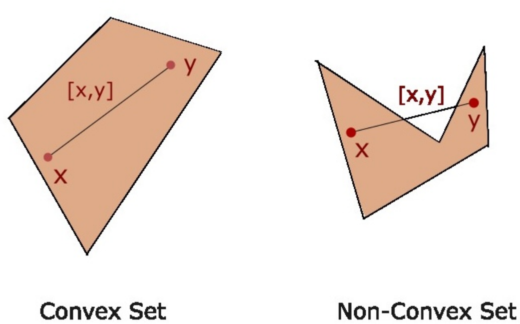

Figure 5.1: Convex and Non-Convex Sets

Proposition 5.1.3. *For two convex sets A, B, their intersection $A \cap B$ is also convex.*

Proof. If $A \cap B = \emptyset$ then it is trivially convex. Otherwise, $A \cap B$ is in both A and B, thus all points in $A \cap B$ are in both A and B, so do the line segments. $\square$

Remark. *The union of two convex sets $A \cup B$ is not guaranteed to be convex. Consider the figure above on the right 5.1, both left and right halves are convex, but their union is not.*

5.1.2 Convex Functions

Definition 5.1.4 (Epigraphs). *The epigraph of a function* $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ *is defined as*

$$\operatorname{epi} f = \{(x, r) \in \mathbb{R}^n \times \mathbb{R} : f(x) \leq r\}$$

Definition 5.1.5 (Convex Functions). $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ *is said to be a convex function if* $\operatorname{epi} f$ *is a convex set.*

That said the region of a convex function under any $r \in \mathbb{R}$ is a convex set. Furthermore, a function is said to be concave if it is a vertically reversed convex function.

Proposition 5.1.6. *The function* $f : \mathbb{R}^n \to \overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty\}$ *is convex if and only if* $f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$.

5.2 Polyhedra

Definition 5.2.1 (Polyhedra). *A polyhedron is a set* $\mathcal{P}$ *in the form*

$$\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$$

for some $A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^m$.

Lemma 5.2.2. *The feasible region of a LP is always a polyhedron.*

Recall the constraints of LP are linear inequalities and equations. Reverse the direction of the $\geq$ constraints to make them consistent,

$$x_1 + x_2 + x_3 \geq 9 \implies -x_1 - x_2 - x_3 \leq -9$$

Remember to add the non-negativity constraint $x \geq 0$ as well. For the equalities, split them into two inequalities such that

$$3x_1 + x_2 + 2x_3 = 4 \implies \begin{bmatrix} 3x_1 + x_2 + 2x_3 \leq 4 \\ 3x_1 + x_2 + 2x_3 \geq 4 \end{bmatrix} \implies \begin{bmatrix} 3x_1 + x_2 + 2x_3 \leq 4 \\ -3x_1 - x_2 - 2x_3 \leq -4 \end{bmatrix}$$

Or just remove the slack variables if any. Thus the feasible region of any LP is a polyhedron. Refer to the definition of a convex set,

Theorem 5.2.3. *A polyhedron* $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$ *is a convex set.*

Proof. Let $x, y \in \mathcal{P}$ be arbitrary two points in the polyhedron. Then $Ax \leq b, Ay \leq b$. Let $\lambda \in [0, 1]$, then $A[x, y]_\lambda = A(\lambda x + (1 - \lambda)y) \leq \lambda b + (1 - \lambda)b = b$ thus $[x, y]_\lambda \in \mathcal{P}$, which is convex. ☐

This states that the feasible region $\mathcal{F}$ of LP is convex.

Definition 5.2.4 (Hyperplanes)**.** *A hyperplane is a set in the form* $\mathcal{H}_y = \{x \in \mathbb{R}^n : \alpha^T x = \beta\}$ *for some* $\alpha \in \mathbb{R}^n \setminus \{0\}, \beta \in \mathbb{R}$.

An example will be a line $ax + by = \beta$ in $\mathbb{R}^2$, or a plane $ax + by + cz = \beta$ in $\mathbb{R}^3$. The vector α is referred to be the normal (or direction) of the hyperplane and is orthogonal to the hyperplane itself, while β determines the position of the plane. For example, controlling α will bevel the hyperplane to a different angle, and increase β will move the hyperplane upwards or downwards. Two hyperplanes are parallel if they have parallel normals, and orthogonal if their normals are orthogonal. In general, the angle between two hyperplanes is determined by the angle between their normals. If $\beta = 0$, then it is a hyperplane with normal α that passes through the origin, and is also a subspace of $\mathbb{R}^n$.

Definition 5.2.5 (Halfspaces)**.** *A halfspace is a set in the form* $\mathcal{H}_a = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta\}$ *for some* $a \in \mathbb{R}^n \setminus \{0\}, \beta \in \mathbb{R}$.

The literal meaning is that a halfspace is the vector space being cut in half by the hyperplane $\{x : a^T x = \beta\}$, and we can see the hyperplane is a boundary as well as a subset of the halfspace.

Definition 5.2.6 (Translates)**.** $\mathcal{S}'$ *is a translate of a set* $\mathcal{S}$ *if there exists* $p \in \mathbb{R}^n$ *such that* $\mathcal{S}' = \{x + p : x \in \mathcal{S}\}$.

Sometimes it can be denoted by $\mathcal{S}' = \mathcal{S} + p$.

Corollary 5.2.7. *Define the null hyperplane* $\mathcal{H}_0 = \{x : \alpha^T x = 0\}$. *Then any hyperplane* $\mathcal{H}_y = \{x \in \mathbb{R}^n : \alpha^T x = \beta\}$ *is a translate of* $\mathcal{H}_0$.

Proof. $\mathcal{H}_y = \mathcal{H}_0 + \beta$. ☐

Corollary 5.2.8. *Define the null halfspace* $\mathcal{H}_0 = \{x : \alpha^T x \leq 0\}$. *Then for any halfspace* $\mathcal{H}_a = \{x \in \mathbb{R}^n : \alpha^T x \leq \beta\}$, *if* $\alpha^T (x - p) \leq \beta - \beta = 0$ *for any* $p \in \mathcal{H}_a$, *then* $\mathcal{H}_a$ *is a translate of* $\mathcal{H}_0$.

Exercise 5.2.1. *Find out which direction and how far the hyperplane/halfspace will move or extend to as* β *changes.*

That said a translate is obtained by moving the original plane or space. In addition, if we know the normal of the hyperplane and want to move it to pass some point x, we can pick $\beta = a^T x$.

Proposition 5.2.9. *A hyperplane with normal* α *that passes some point* $\bar{x}$ *is in the form of* $\{x : \alpha^T x = \alpha^T \bar{x}\}$.

Proof. As the normal α is determined and orthogonal to the line between x and $\bar{x}$ which lies in the hyperplane, it satisfies $\alpha^T(x-\bar{x})=0 \implies \alpha^T x = \alpha^T \bar{x}$. □

Proposition 5.2.10. *Hyperplanes and halfspaces are convex sets.*

Proof. Are special cases of 5.2.3, where the matrix A only has one row α. □

The feasible region of LP, which is a polyhedron, is the intersection of finitely many halfspaces, each defined by an inequality constraint. We can see that the feasible region is thus convex.

5.3 Polyhedron LP

As the feasible region of a LP is a polyhedron, if we add an objective function to it then it will be an optimization problem, $\max\{c^T x : Ax \leq b\}$. We are interested in the optimality of the polyhedron LP.

Definition 5.3.1 (Cone of Vectors)**.** *For a set of vectors $\{x_1, ..., x_n\}$, the cone is defined as $\mathcal{C} = \{\lambda_1 x_1 + ... + \lambda_n x_n : \lambda_1, ..., \lambda_2 \geq 0\}$.*

That said the cone is a linear combination of the vectors with non-negative coordinates. Thus the cone is the region bounded by the vectors, and is convex.

Proposition 5.3.2 (Convexity of Cone)**.** *A cone is a convex set.*

Proof. Consider the non-negativity of the addition of non-negative numbers λ. □

Definition 5.3.3 (Tight Constraints)**.** *For a polyhedron LP $\max\{c^T x : Ax \leq b\}$ with feasible solution $\bar{x}$, the tight constraints of A for $\bar{x}$ are the equality constraints such that $\text{Row}_i(A)x = b_i$.*

We study the optimality via duality.

Proposition 5.3.4 (Dual of Polyhedron LP)**.** *For a LP in polyhedron form $\max\{c^T x : Ax \leq b\}$, the dual LP is $\min\{b^T y : A^T y = c, y \geq 0\}$.*

Recall the optimality conditions 2.1.3, a LP is optimal if and only if there exist feasible x, y such that $c^T x = b^T y$. Suppose there exist feasible $\bar{x}, \bar{y}$ such that $A\bar{x} \leq b, A^T\bar{y} = c, \bar{y} \geq 0$. Observe that if y is feasible, then c is in the cone of rows of A, which are the primal constraints. Note that $c^T x = y^T Ax \leq y^T b$. They are equal if $y^T Ax = y^T b$, where: for tight constraints (whose indices are contained in T) of A, $\bar{y}_T \geq 0$; for non-tight constraints (whose indices are contained in N) of A, $\bar{y}_N = 0$. As well, recall the C.S. conditions 2.3.1, if the ith constraint of A is not tight then $y_i = 0$. Thus it is sufficient to show that if $\bar{x}, \bar{y}$ are optimal, then c is in the cone of tight constraints for $\bar{x}$.

Theorem 5.3.5 (Polyhedra Optimality Conditions)**.** *A polyhedron LP* $\max\{c^T x : Ax \le b\}$ *is optimal if and only if the objective vector c is in the cone of tight constraints.*

Proof. As stated above, it satisfies C.S. conditions if and only if c is in the cone of tight constraints for $\bar{x}$ for feasible $\bar{x}, \bar{y}$, such that $A^T \bar{y} = c$. □

Remark. *When referring to cone of tight constraints, it means the cone of the tight constraints' (hyperplanes) normals which are orthogonal to the hyperplanes respectively. An example will be presented later.*

5.4 Geometry

For any polyhedron, the extreme points are its vertices. We will show a more formal definition for them.

Definition 5.4.1 (Proper Containment)**.** *A point x is properly contained in a line segment L if x lies on L and is not an endpoint of L.*

That said if $L = [a, b]$ where $L \subset \mathcal{S}, a, b \in \mathcal{S}$ and x is properly contained in L, then x is indeed in the interior of L, (a, b).

Definition 5.4.2 (Extreme Points)**.** *Let $\mathcal{S} \in \mathbb{R}^n$ be a convex set. A point x is an extreme point if there does not exist a line segment $L \subset \mathcal{S}$ such that x is properly contained in L.*

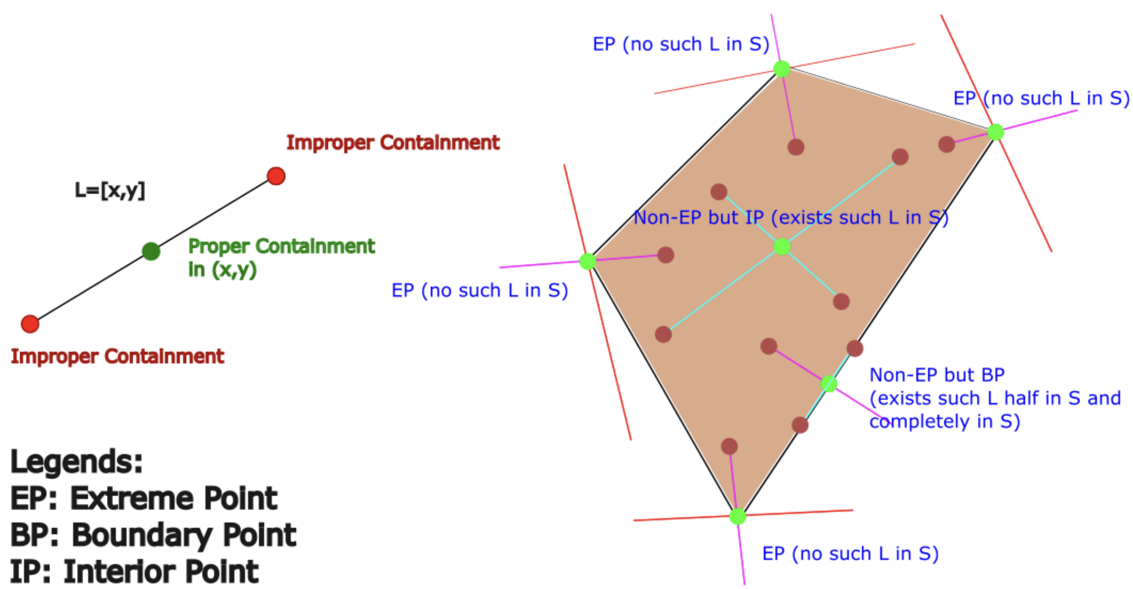

Figure 5.2: Extreme Points

The following requires the reading of Appendix A. Say $\bar{A}\bar{x} = \bar{b}$ is defined as the tight constraints for some solution $\bar{x}$. In a polyhedron $\mathcal{P} \subset \mathbb{R}^n$, for some solution $\bar{x}$, if it has n linearly independent and consistent tight constraints, then $\bar{x}$ is fixed at some point, which is an extreme point. The tight constraint is the boundary of this halfspace, which is the outmost place x can reach in the corresponding normal direction. If $\bar{x}$ has multiple tight constraints, then its degree of freedom will be reduced such that its feasible region is reduced to some lower dimension (i.e. the more tight constraints, the less freedom of $\bar{x}$). Once there are enough (which are n) tight constraints such that the feasible dimension is reduced to 0-D (single point), then it is fixed. This will be demonstrated in the geometry section of Appendix A.2.

Remark. *The tight constraints we talk about here are assumed to be linearly independent and consistent.*

Theorem 5.4.3. *Let $\mathcal{P} = \{x \in \mathbb{R}^n : Ax \leq b\}$ be a polyhedron. For all $\bar{x} \in \mathcal{P}$, $\operatorname{rank}(\bar{A}) = n$ if and only if $\bar{x}$ is an extreme point.*

Proof. By A.2.1, if $\bar{x}$ has n linearly independent and consistent constraints then it has reached n boundaries of the halfspaces such that it is fixed on the intersection of n hyperplanes (which is a point), thus it is an extreme point. □

Note that if there are $r < n$ tight constraints, the degree of freedom of $\bar{x}$ is $n - r$, which is the nullity of $\bar{A}$ whose rank is r.

Theorem 5.4.4. *For all $\bar{x} \in \mathcal{P}$, $\bar{x}$ is an EP if and only if $\bar{x}$ is a basic feasible solution of the SEF of $\mathcal{P}$.*

Proof. Say A has m rows. By 5.4.3, $\operatorname{rank}(\bar{A}) = n$, that there are n tight constraints and $m-n$ non-tight constraints. $n > m$ makes no sense, so assume $m \geq n$. Extend $\mathcal{P}$ to SEF where the non-negativity constraints are merged into the equality constraints, thus there are now m rows and $m+n$ columns. Call the new solution $x' \in \mathbb{R}^{m+n}$, and the size of the basis is m. The n slack variables for the n tight constraints are 0 which are the non-basic entries of x', and the $m-n$ slack variables are greater than 0 which are part of the new basic entries of x'. With the original solution $\bar{x} \in \mathbb{R}^n$ that provides n basic entries (these entries might be 0 as $\bar{x} \geq 0$), there are $(m-n)+n = m$ basic entries, which align with the size of the basis m. □

Remark. *One thing that is worth to think about is, we have assumed the columns of the original polyhedron A to be linearly independent, otherwise these columns are not eligible to be in the basis of SEF, which has been forced to do so. Why it works? As we have seen previously, if $\bar{x}$ is a extreme point, it is the unique solution to the tight matrix $\bar{A}x = \bar{b}$. Then as long as $\bar{x}$ satisfies other non-tight constraints which might be redundant to the tight constraints, the columns of A are linearly independent, by the theorem of unique solution A.1.3.*

This means that any basic feasible solution is an extreme point, and vice-versa. A polyhedron has k feasible bases if and only if it has k corresponding extreme points. Note that a basic infeasible solution is beyond the smallest feasible region (i.e. does not satisfy some other constraints), thus cannot be a extreme point. We demonstrate the above proof using an example.

Example 5.4.1. *Given the polyhedron*

$$\max\ c^T x$$
$$\begin{bmatrix} 2 & 2 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} x \leq \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$

Pick $\bar{x} = (1,0)^T$ *which is an extreme point with the first two tight constraints. Note that another choice of extreme point is* $(0,1)^T$*, tightening the first and third constraints. The SEF for* $\bar{x} = (1,0)^T$ *is*

$$\max\ (c,0)^T x$$
$$\begin{bmatrix} 2 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$
$$x' = \begin{bmatrix} \underline{1} & \underline{0} & 0 & 0 & \underline{1} \end{bmatrix}$$

Note that although the second entry of $\bar{x}$ *is* 0*, we still take it as a basic entry. Now* x' *is a basic feasible solution with respect to basis* $\{1,2,5\}$*. Note that as there are no* $x' \geq 0$ *constraints (might be merged into the constraint matrix),* x' *is feasible as long as it is a feasible solution to the system of linear equations. For the other extreme point* $(0,1)^T$*,*

$$\max\ (c,0)^T x$$
$$\begin{bmatrix} 2 & 2 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix} x = \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}$$
$$x' = \begin{bmatrix} \underline{0} & \underline{1} & 0 & \underline{1} & 0 \end{bmatrix}$$

Thus this is a basic feasible solution with respect to basis $\{1,2,4\}$*.*

Theorem 5.4.5. *One optimal solution of* $\mathcal{P}$ *is among the extreme points, depending on* c*.*

Proof. By FTLP 1.0.1 and 5.4.4, at least one of the extreme points of $\mathcal{P}$ is the optimal solution, and by 5.3.5, an EP $\bar{x}$ is optimal if and only if c is in the cone of tight constraints for $\bar{x}$. $\square$

That said one of the extreme points must be optimal, which will vary among EPs as long as c changes. Note that there could be multiple optimal EPs.

Example 5.4.2. *Refer to the last example 5.4.1, if* $c = (3,2)^T$ *then* $\bar{x} = (1,0)^T$ *is the optimal EP, while if* $c = (2,4)$ *then* $\bar{x} = (0,1)^T$ *is the optimal EP, where* c *is in the cone of their corresponding tight constraints. If* c *is in the intersection of their cone of tight constraints then both of them are optimal (e.g.* $c = 0$*).*

5.5 Union of Polyhedra

Suppose we are given two polyhedra $\mathcal{P}_1, \mathcal{P}_2 \subset \mathbb{R}^n$ where

$$\mathcal{P}_1 = \{x : A_1 x \leq b_1\}$$
$$\mathcal{P}_2 = \{x : A_2 x \leq b_2\}$$

and want to optimize the problem

$$\textbf{(UP)} \ \max \{c^T x : x \in \mathcal{P}_1 \cup \mathcal{P}_2\}$$

The region is bounded which means an optimal value exists, however, unfortunately, the union of polyhedra is **not always convex**. The issue is that we cannot directly set up an LP for this problem. One resolution is to express it as a Mixed Linear Program (MLP).

Definition 5.5.1 (MLP for Union of Polyhedra)**.** *For a union of polyhedra problem where $x_1 \in \mathcal{P}_1, x_2 \in \mathcal{P}_2$ and for some lower and upper bounds $l \leq x_1, x_2 \leq u$, the union of polyhedra $\{x : x \in \mathcal{P}_1 \cup \mathcal{P}_2\}$ is defined as*

$$\begin{aligned}
&y_1 + y_2 = 1 \\
&x = x_1 + x_2 \\
&A_1 x_1 \leq y_1 b_1 \\
&A_2 x_2 \leq y_2 b_2 \\
&y_1 l \leq x_1 \leq y_1 u \\
&y_2 l \leq x_2 \leq y_2 u \\
&y_1, y_2 \in \{0, 1\} (binary)
\end{aligned}$$

The bounds of a variable can be used as a 'constraint switch', such that it can be turned off or on by some conditional variable. For example if $l \leq x \leq u$ is some bound for x and we have a constraint $x \leq b$ where $b < u$, define the conditional variable

$$y = \begin{cases} 1 & \text{if the constraint is active} \\ 0 & \text{if the constraint is not active} \end{cases}$$

and set up the new constraint

$$x \leq yb + (1 - y)u$$

That said, if $y = 1$ then the constraint $x \leq b$ is active, and otherwise there's no effect on x. It is essential to include the upper bound since if we have

$$x \leq yb$$

only, then once set $y = 0$ to turn off the constraint, it will be $x \leq 0$ which is not desired. The similar technique is used in the MLP for union of polyhedron such that, we set up the conditional variables to control which polyhedron is used. The constraint $y_1 + y_2 = 1$ limits that exactly one of y_1, y_2 is active, as they are binary variables. For example, if $y_1 = 1$ then $y_2 = 0$ and the constraints from polyhedron $\mathcal{P}_1$ are active, while those from polyhedron $\mathcal{P}_2$ are not, that said x is in $\mathcal{P}_1$. To show this note that $x_1 \in \mathcal{P}_1, x_2 \in \mathcal{P}_2$ represent the points in each polyhedron respectively. As $y_2 = 0$, x_2 is limited by $0l \leq x_2 \leq 0u$ which is 0 and not violating the $A_2x_2 \leq y_2b_2$ since they are now both 0, but x_1 is active, thus $x = x_1 + 0 = x_1$. Now it is equivalent to say $A_1x \leq b_1, l \leq x \leq u$, which is exactly the constraints of $\mathcal{P}_1$ where $x \in \mathcal{P}_1$. Similarly, if take $y_2 = 1$, then the active constraints are $A_2x \leq b_2, l \leq x \leq u$ where $x \in \mathcal{P}_2$. This perfectly fulfills the union of polyhedra using linear functions.

Definition 5.5.2 (Relaxation)**.** *Suppose (P),(Q) are optimization problems where*

1. *P, Q have the same optimization direction (min or max)*
2. *P, Q have the same objective function*
3. *The feasible region of P is a subset of that of Q*

Then Q is a relaxation of P.

For example, deleting a constraint in P would result in a relaxation of P, as the feasible region extends.

Proposition 5.5.3. *If Q is a relaxation of P, then* $\max\{c^T x : x \in P\} \leq \max\{c^T x : x \in Q\}$.

Proof. As P is included in Q, P's maximum value z is a part of values of Q. If Q\P has a larger maximum value k then $z < k$, otherwise z is the maximum value among Q. □

Theorem 5.5.4. *If all extreme points of Q are integral and Q has an optimal solution, then Q has an optimal integral solution.*

Proof. Follows 1.0.1 and 5.4.4. □

Definition 5.5.5 (Integral Polyhedra)**.** *A polyhedron Q is integral if all extreme points have integer coordinates.*

6 Sensitivity Analysis

We aim to analyze the stability of an optimal basis after modifications to the constraints and object function.

A Linear Algebra

A.1 Basics of System of Linear Equations

A system of linear equations (SLES) is defined as

$$Ax = b, x \in \mathbb{R}^n, A \in \mathbb{R}^{m \times n}, b \in \mathbb{R}^n$$

which consists of m linear equations (i.e. hyperplanes). The solution set to this system is the intersection of the hyperplanes. However, it is infeasible if there exist two parallel and non-coincident hyperplanes since there will be no intersection, such that $\alpha^T x = \beta_1, \alpha^T x = \beta_2, \beta_1 \neq \beta_2$. For example,

$$\begin{aligned} x + y + z &= 1 \\ x + y + z &= 2 \end{aligned}$$

which can be written as

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} x = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$

It is clear that the left matrix A has rank 1, and the augmented matrix $[A|b]$ has rank 2. Thus

Theorem A.1.1. *A system $Ax = b$ is inconsistent if and only if* $\operatorname{rank} A < \operatorname{rank}[A|b]$.

Proof. As stated above,
($\Longrightarrow$) Then the left matrix A has at least one linearly dependent row, which can be row reduced to be a zero row. However, as their right values b do not align with the elimination rule of the left rows, the b value corresponding to the zero row is not eliminated to 0, which becomes a pivot on that row. Thus $\operatorname{rank} A < \operatorname{rank}[A|b]$.
($\Longleftarrow$) That said the left matrix A could be row reduced to have a zero row $0^T x$, but $[A|b]$ has a non-zero number on the corresponding row. As the zero row $0^T x$ should have a value of 0, it is not consistent. □

Now suppose the system is consistent. The following are some facts and fundamental theorems in linear algebra, but we will not focus on themselves too much, just know how to use them in optimization.

Lemma A.1.2. *The nullity of A, denoted by $n - \operatorname{rank}(A)$, is the dimension of the solution set to $Ax = b$ (i.e. number of free parameters in the solution set).*

That said, if the nullity of A is 0, then A is square and $Ax = b$ has a unique solution. This is a more complete criteria for the basis we used in optimization,

Theorem A.1.3 (Non-Singular Matrix)**.** *Consider the system $Ax = b, A \in \mathbb{R}^{n \times n}, x, b \in \mathbb{R}^n$. The following statements are equivalent.*

1. *A is non-singular (invertible) and* $\det(A) \neq 0$.

2. *A's columns and rows are linearly independent, respectively.*

3. RREF(A) $= I_n$, *that said* $Ax = b$ *has a unique solution.*

4. Null(A) $= \{0\}$ *and nullity of* A *is* 0.

5. Range(A) $= \mathbb{R}^n$ *and rank of* A *is* n.

6. A *is both injective and surjective (isomorphism).*

However, we cannot guarantee the diagonality, eigenvalues and singular value decomposition of a non-singular matrix. It is possible for a non-singular matrix to be non-diagonalizable.

A.2 Geometry of System of Linear Equations

Recall that the solution set is the intersection of the hyperplanes, so instead of studying the equations, we look at the geometry facts of it. Suppose we are in $\mathbb{R}^n$. The intersection of two hyperplanes is a (n-1)-dimensional space. That said, each time we add an equality constraint (linearly independent and consistent), which adds a pivot, the dimension of the intersection is reduced by 1. For example,

$$2x + 3y + z = 1$$

has a 2-D solution set $\{t(1, 0, -2) + k(0, 1, -3) + (0, 0, 1) : t, k \in \mathbb{R}\}$, which is a plane, but

$$\begin{aligned} 2x + 3y + z &= 1 \\ 3x + y + z &= 4 \end{aligned}$$

has a 1-D solution set $\{(11/7, -5/7, 0) + t(-2/7, -1/7, 1) : t \in \mathbb{R}\}$, since the intersection of two planes is a line, and

$$\begin{aligned} 2x + 3y + z &= 1 \\ 3x + y + z &= 4 \\ 4x + 3y + z &= 3 \end{aligned}$$

has a unique solution $(1, -1, 2)$, which is the intersection of a line and a plane. From this example, and by the criteria A.1.3 as well as the geometry properties shown above, we can see that to obtain a unique solution in $\mathbb{R}^n$, we need exactly n linearly independent and consistent equations.

Theorem A.2.1. *In* $\mathbb{R}^n$, *the system of linear equations with exactly* n *linearly independent and consistent equations has a unique solution.*

The converse is also true. Note that when there are more than n linear equations and if the system is still consistent, then the extra linear equations must be redundant which are coincident to some existing equations. That is why we are always eliminating redundant constraints in optimization as they are not providing extra information.

A.3 Basis

Definition A.3.1 (Basis for $\mathbb{R}^n$). *A set of* n *vectors* $V = \{v_1, ..., v_n\}$ *is a basis for* $\mathbb{R}^n$ *if it is linearly independent and* $\text{Span}(V) = \mathbb{R}^n$.

A basis for $\mathbb{R}^n$ can map to all vectors in the vector space. We assume the basis is ordered, so it can be represented in the matrix form $[V] = \begin{bmatrix} v_1 & ... & v_n \end{bmatrix}$ which is a non-singular linear operator A.1.3, such that $\forall y \in \mathbb{R}^n$, the system $[V]x = y$ has a unique solution $x = [V]^{-1}y$. Note that x is also referred to be the coordinates of y with respect to the basis V, and is formally denoted as $x = [y]_{\mathcal{B}}$.

Theorem A.3.2. *In* $\mathbb{R}^n$*, if a set consists of more than* n *vectors, then it must be linearly dependent.*

Proof. Say $V = \{v_1, ..., v_k\}$ for some $k > n \in \mathbb{Z}^+$ and $v_i \in \mathbb{R}^n$ for $i = 1, ..., k$. Consider $[V]t = 0$ for $t \in \mathbb{R}^k$. As $[V]$ is a $n \times k$ underdetermined matrix, it contains at least $k - n > 0$ free variables which guarantees a non-trivial kernel, that said it has a non-zero solution. Thus it is linearly dependent. $\square$

Theorem A.3.3 (Rank-Nullity Theorem). *For matrix* $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) + \text{nullity}(A) = n$.

A.4 Diagonality

sss

A.5 Normal Equations

A common problem in the real world is, the system $Ax = b$ is not always consistent. Instead, we can find an approximate solution to it.

Lemma A.5.1. *The matrix* X^TX *is invertible (i.e. non-singular) if and only if the columns of* X *are linearly independent.*

Proof. We aim to show $\text{Null}(X^TX) = \text{Null}(X)$.

1. $\text{Null}(X) \subseteq \text{Null}(X^TX)$: Suppose $x \in \text{Null}(X)$, then $Xx = 0$. Thus $X^TXx = X^T0 = 0$ which means $x \in \text{Null}(X^TX)$.

2. $\text{Null}(X^TX) \subseteq \text{Null}(X)$: Suppose $x \in \text{Null}(X^TX)$, then $X^TXx = 0$. Thus $x^TX^TXx = x^T0 = 0 \implies \|Xx\|^2 = 0 \implies Xx = 0$ which means $x \in \text{Null}(X)$.

If X's columns are linearly independent then $\text{Null}(X) = \{0\}$ and thus $\text{Null}(X^TX) = \text{Null}(X) = \{0\}$, that said X^TX is invertible. $\square$

This lemma will be used to determine the uniqueness of solution to the normal equation, or equivalently, **Least Squares Estimate**.

Proposition A.5.2 (Normal Equations). *$A^TAx = A^Tb$ is the normal equation of $Ax = b$. Moreover, if A^TA is invertible such that A's columns are linearly independent A.5.1, then the normal equations have a unique solution $\hat{x} = (A^TA)^{-1}A^Tb$.*

This is useful in many fields such as non-linear optimization and statistics, especially in linear regression. The column space of matrix A, denoted by $\mathrm{Col}(A)$, is equivalent to its range such that $\mathrm{Col}(A) = \mathrm{Range}(A)$, which is the set of all vectors Ax can produce, for all x in the domain $\mathbb{R}^n$. The system $Ax = b$ is consistent if and only if b is in $\mathrm{Col}(A)$. For an inconsistent system $Ax = p$, its closest approximate solution comes from its projection onto $\mathrm{Col}(A)$ where $Ax = \mathrm{proj}_{\mathrm{Col}\,A}(b)$, which is solvable as the $\mathrm{proj}_{\mathrm{Col}\,A}(b)$ lies in $\mathrm{Col}(A)$, whose solutions $\hat{x}$ are referred to be the **Least Squares Solutions**.
Consider the matrix system $Ax = b$, which might be inconsistent. Express it in the following form,

$$\begin{gathered}\mathrm{Row}_1(A)x = b_1 \\ \dots \\ \mathrm{Row}_m(A)x = b_m\end{gathered}$$

For each $\mathrm{Row}_i(A) = b_i$ and any solution $\bar{x}$, the error is given by $\epsilon_i = b - A\bar{x}$. Taking the square ϵ^2 would give information only on the distance between $\mathrm{Row}_i(A)\bar{x}$ and b_i, regardless the sign of the error. We want to find an approximate solution $\hat{x}$ that minimizes the sum of squares of error for all rows,

$$\hat{x} \in \text{argmin } \sum_{i=1}^{n} \epsilon_i^2 = \sum_{i=1}^{n} (b_i - \mathrm{Row}_i(A)\bar{x})^2$$

or equivalently,

$$\hat{x} \in \text{argmin } \|Ax - b\|^2$$

Express the equation,

$$\begin{aligned}\|Ax - b\|^2 &= (Ax - b)^T(Ax - b) \\ &= (x^TA^T - b^T)(Ax - b) \\ &= x^TA^TAx - b^TAx - x^TA^Tb + b^Tb \\ &= x^TA^TAx - 2b^TAx + b^Tb := S(x)\end{aligned}$$

Note that x^TA^TAx is a quadratic form on x with the symmetric associate matrix A^TA, and b^TAx is a linear function on x. To minimize $S(x)$, we take the derivative and set to 0,

$$\begin{aligned}&\frac{\mathrm{d}S(x)}{\mathrm{d}x} = 2x^TA^TA - 2b^TA = 0 \\ &\implies A^TAx = A^Tb\end{aligned}$$

Here comes the normal equation. As the norm function is convex, any of its stationary point guarantees a minimizer. Recall that if A has linearly independent columns then the normal equation has a unique solution $\hat{x} = (A^TA)^{-1}A^Tb$ (i.e. Unique Least Squares Solution, ULSS). Note that $A\hat{x} = A(A^TA)^{-1}A^Tb = \mathrm{proj}_{\mathrm{Col}(A)}(b)$ (i.e. fitted value), which projects b onto the closest vector in $\mathrm{Col}(A)$. Here comes the definition of projection matrices,

Definition A.5.3 (Projection Matrices)**.** *For any system $Ax = b$, the projection matrix is $H = A(A^T A)^{-1} A^T$.*

That said, Hb projects b onto $\mathrm{Col}(A)$.

Definition A.5.4 (Idempotent Matrices)**.** *A matrix X is idempotent if $XX = X$.*

Remark. *By induction, $H^n = H$.*

Lemma A.5.5. *An idempotent matrix's eigenvalues can only be 0 or 1.*

Proof.

$$\begin{aligned} HH &= H \\ HHx &= Hx, \forall x \in \mathrm{domain}(H) \\ \lambda^2 x &= \lambda x, \text{for corresponding eigenvalue } \lambda \\ (\lambda^2 - \lambda) &= 0 \\ \lambda(\lambda - 1) &= 0 \end{aligned}$$

Thus λ can only be 0 or 1. □

The projection matrix H is idempotent.

Corollary A.5.6. *The projection matrix H is idempotent.*

Proof.

$$\begin{aligned} HH &= A(A^T A)^{-1} A^T A (A^T A)^{-1} A^T \\ &= A(A^T A)^{-1} A^T \\ &= H \end{aligned}$$

□

A.6 Rayleigh Quotient

Alphabetical Index