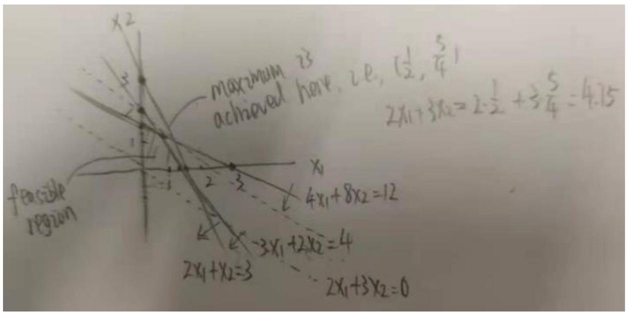
Part 3 Linear Program & Dual

(**Example 1**) (**Linear Program**, LP) Consider the following optimization problem with constraints:

$$\begin{aligned} &\max\ 2x_1+3x_2 \\ &\text{s.t. } 4x_1+8x_2\le 12 \\ &\quad\ \ 2x_1+x_2\le 3 \\ &\quad\ \ 3x_1+2x_2\le 4 \\ &\quad\ \ x_1,x_2\ge 0 \end{aligned},$$

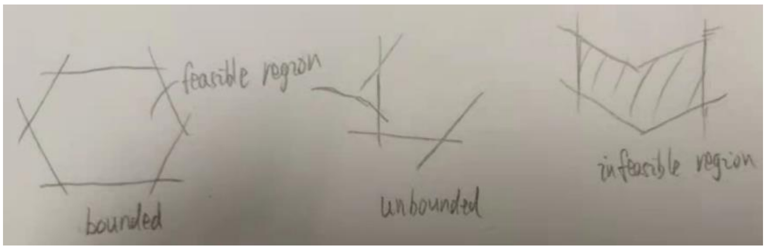
where $2x_1+3x_2$ is the objective function. The aforementioned LP can be visualized as follow:



Notes: LP only contains linear constraints, e.g., $4x_1+8x_2\le 12$. We don't consider the constraints like $x_1^2+x_2^2\ge 2$ or $x_1x_2\ge 2$.

Notes: For LP with 2 variables, e.g., (x_1, x_2), each equation like $2x_1+3x_2=C$ (with C as an arbitrary constant) defines a **line**. A linear constraint like $2x_1+3x_2\ge C$ defines a **half-plane**. For LP with 3 variables, e.g., (x_1, x_2, x_3), a linear constrain like $3x_1+2x_2+7x_3\le 10$ defines a **half-space**.

Notes: A **feasible region** of LP (defined by the linear constraints) should be **convex polygon/polyhedron** (凸多边形/多面体), which can be **bounded** or **unbounded**.



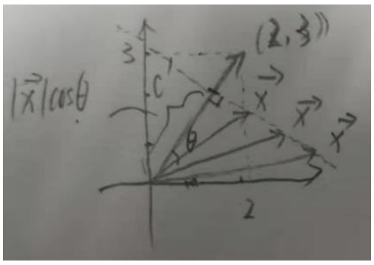
(**Example 2**) For **Example 1**, let the objective function $2x_1+3x_2$ be equal to a constant C. It can be rewritten into the dot product of vectors, i.e.,

$$\begin{bmatrix}2\\3\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\end{bmatrix}=2x_1+3x_2=C\cdot$$

For the vector $[2, 3]^T$, the dot product between $[2, 3]^T$ and $[x_1, x_2]^T$ is also defined as

$$\begin{bmatrix}2\\3\end{bmatrix}\cdot\begin{bmatrix}x_1\\x_2\end{bmatrix}=\sqrt{13}\cdot|\mathbf{x}|\cos\theta=C,$$

such that $|\mathbf{x}|\cos\theta=C'$. Especially, $|\mathbf{x}|\cos\theta=C'$ can be regarded as the **projection** from $[x_1, x_2]^T$:



It indicates an interested property that all the vectors represented by the points on the **orthogonal (正交)** line/plane/hyperplane have the same dot product value with the coefficient vector (e.g., $[2, 3]^T$). In fact, the **orthogonal line/plane/hyperplane** is the line/plane/hyperplane represented by **x**. (!!)

(**Definition 1**) Let $[x_1, x_2]^T=[c_1, c_2]^T$ iff $x_1=c_1$ and $x_2=c_2$. Let $[x_1, x_2]^T\leq[c_1, c_2]^T$ iff $x_1\leq c_1$ and $x_2\leq c_2$. Let $[x_1, x_2]^T\geq[c_1, c_2]^T$ iff $x_1\geq c_1$ and $x_2\geq c_2$.

(**Example 3**) Let $\mathbf{c}=[2, 3]^T$, $\mathbf{x}=[x_1, x_2]^T$, $\mathbf{b}=[12, 3, 4]^T$, and

$$\mathbf{A}=\begin{bmatrix} 4 & 8 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}.$$

Example 1 can be rewritten into the following matrix form:

$$\begin{aligned} &\max\ \mathbf{c}^T\mathbf{x} \\ &\text{s.t. } \mathbf{Ax}\geq\mathbf{b},\ \mathbf{x}\geq\mathbf{0} \end{aligned}.$$

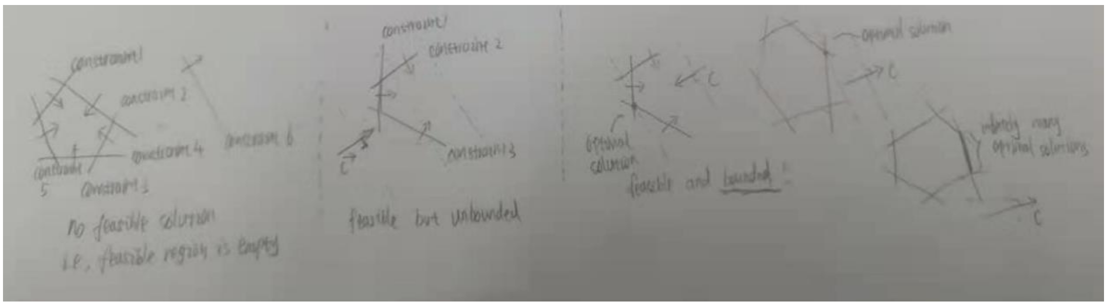
(**Definition 2**) For n variables $(x_1, x_2, \cdots, x_n)$ in LP, any vector $\mathbf{x}\in\Re^n$ **satisfies all the constrains** of LP is called a **feasible solution**. Any vector $\mathbf{x}^*\in\Re^n$ that gives the maximum possible value of $\mathbf{c}^T\mathbf{x}$ among all the feasible **x** called an **optimal solution**. Further, an LP is **infeasible** if its feasible region is empty. Otherwise, LP is **feasible**.

(**Theorem 1**) LP has the following 3 types:

(1) LP is **infeasible**;

(2) LP is **feasible** and **unbounded** (i.e., the objective function can be made as large as you like);

(3) LP is **feasible** and **bounded**, where LP may have **a single optimal solution** or **infinitely many optimal solutions**.



The number of vertices (of the feasible region, i.e., the convex polygon/polyhedron) can be **exponential** in the **input size**, but LP can be solved in **polynomial time**.

(**Theorem 2**) In general, an LP can have 2 other kinds of constraints, e.g.,

$$2x_1+3x_2\ge 10,\ (1)$$

$$2x_1+3x_2=10.\ (2)$$

One can rewrite the inequality (1) into the following form:

$$-2x_1-3x_2\le -10.$$

One can rewrite the equation (2) into the following form:

$$2x_1+3x_2\ge 10,$$

$$2x_1+3x_2\le 10,$$

which can be further rewritten into the following form:

$$-2x_1-3x_2\le -10,$$

$$2x_1+3x_2\le 10.$$

Hence, the **general form** of LP can be represented as

$$\max\ \mathbf{c}^{\mathrm{T}}\mathbf{x} \text{ s.t. } \mathbf{Ax}\le \mathbf{b}, \mathbf{x}\ge 0.$$

(**Example 4**) Recall **Example 1**.

$$\begin{aligned}
\max\ & 2x_1+3x_2\\
\text{s.t. } & 4x_1+8x_2\le 12 \text{ - (1)}\\
& 2x_1+x_2\le 3 \text{ - (2)}\\
& 3x_1+2x_2\le 4 \text{ - (3)}\\
& x_1,x_2\ge 0
\end{aligned}$$

Since $x_1,x_2\ge 0$, by comparing the objective function with inequality (1), we have

$$2x_1+3x_2\le 4x_1+8x_2\le 12,$$

where 12 can be a possible upper bound.

By comparing the objective function with the 1/2 of inequality (1), we have

$$2x_1+3x_2\le 2x_1+4x_2\le 6,$$

where 6 is a better upper bound.

By comparing the objective function with $\frac{1}{3}((1)+(2))$, we have

$$2x_1+3x_2=\frac{1}{3}((4x_1+8x_2)+(2x_1+x_2))\le\frac{1}{3}\times 15=5,$$

where 5 is a better upper bound.

The aforementioned example indicates that the **upper bound** of the objective function can be estimated based on the **linear combination** of the **constraints of LP**. For example, we have

$$2x_1+3x_2\le d_1x_1+d_2x_2\le h,$$

$$d_1\ge 2, d_2\ge 3,$$

where h can be an upper bound of the objective function.

(**Example 5**) Recall **Example 1**.

$$\begin{aligned}
\max\ & 2x_1+3x_2\\
\text{s.t. } & 4x_1+8x_2\le 12 \text{ - (1)}\\
& 2x_1+x_2\le 3 \text{ - (2)}\\
& 3x_1+2x_2\le 4 \text{ - (3)}\\
& x_1,x_2\ge 0
\end{aligned}$$

According to **Example 4**, one can use the **linear combination** of constraints (1), (2), and (3) to estimate the upper bound of the objective function. Let y_1, y_2, and y_3 be the **coefficient** w.r.t. (1), (2), and (3) in the **linear combination** (i.e., $y_1\times(1)+y_2\times(2)+y_3\times(3)$), we have

$$y_1(4x_1+8x_2)+y_2(2x_1+x_2)+y_3(3x_1+2x_2)\le 12y_1+3y_2+4y_3,$$

which can be rearranged as

$$(4y_1+2y_2+3y_3)x_1+(8y_1+y_2+2y_3)x_2\le 12y_1+3y_2+4y_3,$$

where we also need to ensure

$$2\le 2y_1+2y_2+3y_3,$$
$$3\le 8y_1+y_2+2y_3.$$

In this case, we also have

$$2x_1+3x_2\le 12y_1+3y_2+4y_3. \quad (\#)$$

We can extract <u>another LP</u> w.r.t. variables (y_1, y_2, y_3):

$$\begin{aligned} &\min\ 12y_1+3y_2+4y_3 \\ &\text{s.t. } 2y_1+2y_2+3y_3\ge 2, \\ &\quad 8y_1+y_2+2y_2\ge 3, \\ &\quad y_1\ge 0, y_2\ge 0, y_3\ge 0. \end{aligned}$$

Let **Example** 1 be the **Primal LP**. Then, the rearranged LP w.r.t. (y_1, y_2, y_3) is defined as the corresponding **Dual LP**.

Notes: The Dual LP ‘guards’ the original LP (i.e., Primal LP), i.e., <u>every **feasible solution** (y_1, y_2, y_3) of **Dual LP** provides an **upper bound** on the **maximum of the objective function** of **Primal LP**.</u>

Notes: There’s a big problem: How well does it guard?

(**Example 6**) For the following **Primal LP** (in the matrix form):

$$\max\ \mathbf{c}^{\mathrm{T}}\mathbf{x} \text{ s.t. } \mathbf{A}\mathbf{x}\le\mathbf{b}, \mathbf{x}\ge 0,$$

we have the following **Dual LP**:

$$\min\ \mathbf{y}^{\mathrm{T}}\mathbf{b} \text{ s.t. } \mathbf{y}^{\mathrm{T}}\mathbf{A}\ge\mathbf{c}^{\mathrm{T}}, \mathbf{y}\ge 0.$$

Alternatively, we also have

$$\min\ \mathbf{y}^{\mathrm{T}}\mathbf{b} \text{ s.t. } \mathbf{A}^{\mathrm{T}}\mathbf{y}\ge\mathbf{c}, \mathbf{y}\ge 0.$$

For the constant matrix **A**, <u>**Ax** represents the linear combination of **A**’s column vectors</u>, while <u>$\mathbf{y}^{\mathrm{T}}\mathbf{A}$ represents the linear combination of **A**’s row vectors</u>.

(**Theorem 3**) (**Weak Duality Theorem**) For each **feasible solution x** of **Primal LP** and each **feasible solution y** of **Dual LP**, we have

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}\le\mathbf{y}^{\mathrm{T}}\mathbf{b}.$$

Proof of Theorem 3. For the **Primal LP**, since $\mathbf{A}\mathbf{x}\le\mathbf{b}$ and $\mathbf{y}\ge\mathbf{0}$, we have

$$\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}\le\mathbf{y}^{\mathrm{T}}\mathbf{b}.$$

For the **Dual LP**, since $\mathbf{y}^{\mathrm{T}}\mathbf{A}\ge\mathbf{c}^{\mathrm{T}}$ and $\mathbf{x}\ge 0$, we have

$$\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}\ge\mathbf{c}^{\mathrm{T}}\mathbf{x}.$$

Hence, we have

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}\le\mathbf{y}^{\mathrm{T}}\mathbf{A}\mathbf{x}\le\mathbf{y}^{\mathrm{T}}\mathbf{b}.$$

Namely, we have

$$\mathbf{c}^{\mathrm{T}}\mathbf{x}\le\mathbf{y}^{\mathrm{T}}\mathbf{b},$$

which finishes the proof.

(**Corollary 1**) If Primal LP is <u>unbounded</u>, then Dual LP is <u>infeasible</u>.

Proof of Corollary 1 (by contradiction). For the sake of contradiction, assume that Dual LP is feasible and $\mathbf{y}$ is a feasible solution. According to the **Weak Duality Theorem**, $\mathbf{y}^{\mathrm{T}}\mathbf{b}$ is an upper bound of Primal LP, which contradicts with the fact that Primal LP is unbounded.

Hence, Dual LP if infeasible, when Primal LP is unbounded.

(**Corollary 2**) If Dual LP is <u>unbounded</u>, then Primal LP is <u>infeasible</u>.

(**Theorem 4**) (**Strong Duality Theorem**) If <u>either **Primal LP** or **Dual LP** is **feasible** and **bounded**</u>, then so in each other and the **optimal values** of both LPs are the **same**.

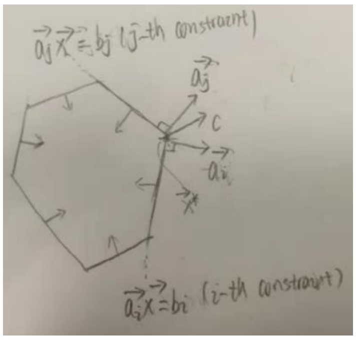
(**Intuition** for the Proof of **Theorem 4**) Let's restrict to a 2D space first. Assume Primal LP is feasible and bounded. In general, for a Primal LP

$$\max\ \mathbf{c}^{\mathrm{T}}\mathbf{x} \text{ s.t. } \mathbf{A}\mathbf{x} \le \mathbf{b}, \mathbf{x} \ge 0,$$

let $\mathbf{A} = [\mathbf{a}_1^{\mathrm{T}}, \mathbf{a}_2^{\mathrm{T}}, \cdots, \mathbf{a}_n^{\mathrm{T}}]^{\mathrm{T}}$ and $\mathbf{b} = [b_1, b_2, \cdots, b_n]^{\mathrm{T}}$, where $\mathbf{a}_i$ is the i-th <u>row vector</u> of $\mathbf{A}$ and b_i is the i-th element of $\mathbf{b}$. For the i-th row of $\mathbf{A}$ (w.r.t. the i-th constraint of Primal LP), we have

$$\mathbf{a}_i\mathbf{x} \le b_i.$$

Without the loss of generality, consider the following convex polygon described by the Primal LP:



Suppose the optimal solution of Primal LP (denoted as $\mathbf{x}^*$) is achieved at the intersection point w.r.t the i-th and j-th constraint. We have

$$\mathbf{a}_i\mathbf{x}^* = b_i,\ \ \mathbf{a}_j\mathbf{x}^* = b_j.$$

Further, <u>$\mathbf{c}$ can be derived from the following **linear combination** between $\mathbf{a}_i$ and $\mathbf{a}_j$</u>:

$$\mathbf{c}^{\mathrm{T}} = y_i^*\mathbf{a}_i + y_j^*\mathbf{a}_j,$$

where $y_i^*, y_j^* \ge 0$ are 2 non-negative numbers. Consider the solution to Dual LP $\mathbf{y}^*$, where all the entries are zeros, except y_i^* and y_j^* as described above. Namely, we have

$$\begin{aligned} \mathbf{y}^{*\mathrm{T}}\mathbf{A} &= y_1^*\mathbf{a}_1 + y_2^*\mathbf{a}_2 + \cdots + y_i^*\mathbf{a}_i + \cdots + y_j^*\mathbf{a}_j + \cdots + y_n^*\mathbf{a}_n \\ &= y_i^*\mathbf{a}_i + y_j^*\mathbf{a}_j \\ &= \mathbf{c}^{\mathrm{T}} \end{aligned},$$

i.e., $\mathbf{y}^{*\mathrm{T}}\mathbf{A} \ge \mathbf{c}^{\mathrm{T}}$, which indicates that <u>$\mathbf{y}^*$ is a **feasible solution** to **Dual LP**</u> and the **feasible value** is $\mathbf{y}^{*\mathrm{T}}\mathbf{b} = y_i^*b_i + y_j^*b_j$.

For the **optimal solution** to Primal LP, we have

$$
\begin{aligned}
\mathbf{c}^{\mathrm{T}}\mathbf{x}^* &= (y_i^*\mathbf{a}_i + y_j^*\mathbf{a}_j)\mathbf{x}^* \\
&= y_i^*\mathbf{a}_i\mathbf{x}^* + y_j^*\mathbf{a}_j\mathbf{x}^*, \text{ (\#\#)} \\
&= y_i^* b_i + y_j^* b_j \\
&= \mathbf{y}^{*\mathrm{T}}\mathbf{b}
\end{aligned}
$$

which is the value of the objective function of Dual LP at $\mathbf{y}^*$.

According to the **Weak Duality Theorem** (i.e., **Theorem 3**), the optimal value of Primal LP is less than or equal to (i.e., ≤) any feasible values of the Dual LP. Hence, $\mathbf{y}^*$ is the optimal solution to Dual LP, i.e., the Primal and Dual LP have the **same optimal values** $\mathbf{c}^{\mathrm{T}}\mathbf{x}^* = \mathbf{y}^{*\mathrm{T}}\mathbf{b}$.