Lecture 1 - Simplex Method

Lan Peng, Ph.D.

School of Management, Shanghai University, Shanghai, China

"One step at a time."

1 Preliminaries

1.1 Linear Algebra

Linear combination A vector μ is said to be a linear combination of $v^1, v^2, \dots, v^m$ if

$$\sum_{i=1}^{m} \lambda_i v^i = \mu$$

In addition, μ is a

- conic combination if $\lambda_i \geq 0$
- affine combination if $\sum_{i=1}^{m} \lambda_i = 1$
- convex combination if $\sum_{i=1}^{m} \lambda_i = 1$ and $\lambda_i \geq 0$

Linear independence and affinely independence A collection of vectors $v^1, v^2, \dots, v^m$ of dimension n is called linearly independent if

$$\sum_{j=1}^{k} \lambda_j v^j = 0 \quad \Rightarrow \quad \lambda_j = 0, \forall j = 1, 2, \dots, m$$

A collection of vectors $v^1, v^2, \dots, v^m$ of dimension n is called affinely independent if

$$\sum_{j=1}^{k} \lambda_j v^j = 0 \text{ and } \sum_{j=1}^{k} \lambda_j = 0 \Rightarrow \lambda_j = 0, \forall j = 1, 2, \dots, m$$

All the following statements are equivalent:

- $v^1, v^2, \dots, v^m$ of dimension n are affinely independent
- $v^2 - v^1, v^3 - v^1, \dots, v^m - v^1$ of dimension n are linearly independent
- $\begin{bmatrix} v^1 \\ 1 \end{bmatrix}, \begin{bmatrix} v^2 \\ 1 \end{bmatrix}, \dots, \begin{bmatrix} v^m \\ 1 \end{bmatrix}$ are linearly independent

The difference between linearly independent and affinely independent is indicated in the following figure. For example, this figure is in 2-Dimension space

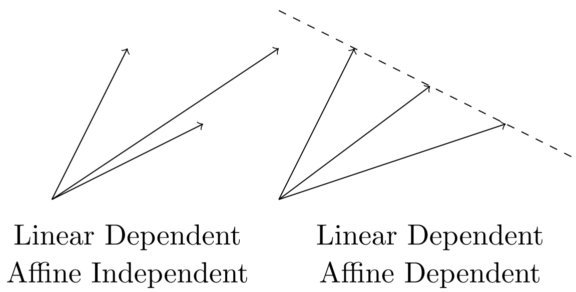

Figure 1: Linearly Independent / Affinely Independent

Spanning set and basis A collection of vectors $v^1, v^2, \ldots, v^m$ of dimension n is said to span $\mathcal{R}^n$ if any vector in $\mathcal{R}^n$ can be represented as a linear combination of $v^1, v^2, \ldots, v^m$. $v^1, v^2, \ldots, v^m$ is said to form a basis of $\mathcal{R}^n$ if the following conditions holds.

- $v^1, v^2, \ldots, v^m$ span $\mathcal{R}^n$.
- If any of these vectors is deleted, the remaining collection of vector does not span $\mathcal{R}^n$.

Rank of a matrix The span of the columns of a matrix A is called the column space or the range, denoted $range(A)$. The span of the rows of a matrix A is called the row space. The dimension of the column space and row space are equal, which is denoted by $rank(A)$. $rank(A) \leq \min\{m, n\}$, if $rank(A) = \min\{m, n\}$, then A is said to have full rank and A is a basis of $\mathcal{R}^n$.

1.2 Polyhedral Sets

Convex sets A set $S \subseteq \mathcal{R}^n$ is convex if $\forall x, y \in S, \lambda \in [0, 1]$, we have $\lambda x + (1 - \lambda)y \in S$

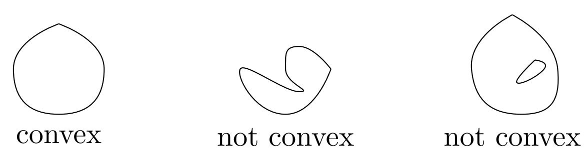

Figure 2: A set is convex iff for any two points in the set, the line segment joining those two points lies entirely in the set

Let $x^1, ..., x^k \in \mathcal{R}^n$ and $\lambda \in \mathcal{R}^k$ be given such that $\lambda^T e = 1$

- The vector $\displaystyle\sum_{i=1}^{k} \lambda_i x^i$ is said to be a convex combination of $x^1, ..., x^k$
- The convex hull of $x^1, ..., x^k$ is the set of all convex combinations of these vectors, denoted $conv(x^1, ..., x^k)$

Polyhedral, hyperplanes and half-spaces

- A polyhedron is a set of the form $\{x \in \mathcal{R}^n | Ax \leq b\} = \{x \in \mathcal{R}^n | a^i x \leq b^i, \forall i \in M\}$, where $A \in \mathcal{R}^{m \times n}$ and $b \in \mathcal{R}^m$

- A polyhedron $P \subset \mathcal{R}^n$ is bounded if there exists a constant K such that $|x_i| < K, \forall x \in P, \forall i \in [1, n]$, in this case the polyhedron is call polytope.
- The lower-bound of K is called diagonal denoted by d
- A hyperplane is $\{x \in \mathcal{R}^n | a^T x = b\}$
- A half-space is $\{x \in \mathcal{R}^n | a^T x \leq b\}$

Open, close sets: boundary and interior

- Denote $N_\epsilon = \{y \in \mathcal{R}^n | \|y - x\| < \epsilon\}$ as the neighborhood of $x \in \mathcal{R}^n$
- Given $S \subseteq \mathcal{R}^n$, x belongs to the interior of S, denoted by $int(S)$ if there is $\epsilon > 0$ such that $N_\epsilon(x) \leq S$
- S is said to be an open set iff $S = int(S)$
- x belongs to the boundary ∂S if $\forall \epsilon > 0$, $N_\epsilon(x)$ contains at least one point in S and a point not in S
- $x \in S$ belongs to the closure of S, denoted $cl(s)$ if $\forall \epsilon > 0$, $N_\epsilon(x) \cap S = \emptyset$
- S is called closed iff $S = cl(S)$
- In IP, LP, MIP, etc. we always work with close set. No "<" or ">"

Dimension of polyhedral

- A polyhedron P is dimension k, denoted $dim(P) = k$, if the maximum number of affinely independent points in P is $k + 1$
- A polyhedron $P \subseteq \mathcal{R}^n$ is full-dimensional if $dim(P) = n$
- Let $M = \{1, 2, ..., m\}$, $M^= = \{i \in M | a_i x = b_i, \forall x \in P\}$, i.e. the equality set, $M^\leq = M \setminus M^=$, i.e. the inequality set. Then, Let $(A^=, b^=)$, $(A^\leq, b^\leq)$ be the corresponding rows of (A, b), If $P \subseteq \mathcal{R}^n$, then $dim(P) = n - rank(A^=, b^=)$. To proof a constraint $(A^=, b^=)$ is an equality constraint, we need to proof all point in the closure of P satisfied the constraint, to proof it is not an equality constraint, we need to find one point that is not in the hyperplane.
- $x \in P$ is called an inner point of P if $a^i x < b_i, \forall i \in M^\leq$
- $x \in P$ is called an interior point of P if $a^i x < b_i, \forall i \in M$
- Every nonempty polyhedron has at least one inner point
- A polyhedron has an interior point iff P is full-dimensional, i.e., there is no equality constraint

1.3 E.P. and B.F.S.

Extreme point A point x in a convex set X is called an extreme point iff x cannot be represented as a strict convex combination of two distinct points of X. In other words, if $x = \lambda x_1 + (1-\lambda)x_2$, then $x_1 = x_2 = x$.

An alternative definition is as follows. Let the hyperplanes associated with the $(m+n)$ defining half-spaces of X be referred to as defining hyperplanes of X. Furthermore, note that a set of defining hyperplanes are linearly independent if the coefficient matrix associated with this set of equations has full row rank. Then a point x is said to be an extreme point if x lies on some n linearly independent defining hyperplanes of X. In addition, if more than n defining hyperplanes pass through an extreme point, then such extreme point is called a degenerated extreme point.

Basic feasible solution Consider the system $\{A_{m\times n}x = b_m, x \geq 0\}$, suppose $rank(A, b) = rank(A) = m$, we can arrange A and have a partition of A. Let $A = [B, N]$ where B is $m \times m$ invertible matrix, and N is a $m \times (n-m)$ matrix. The solution $x = \begin{bmatrix} x_B \\ x_N \end{bmatrix}$ to the equation $Ax = b$, where

$$x_B = B^{-1}b$$

and

$$x_N = 0$$

is called basic solution of system. If $x_B \geq 0$, it is called basic feasible solution. If $x_B > 0$ it is called non-degenerate basic feasible solution. For $x_B \geq 0$, if some $x_j = 0$, those components are called degenerated basic feasible solution. B is called the basic matrix, N is called nonbasic matrix.

Correspondence between B.F.S and E.P.

Theorem 1.1. *x is an E.P. iff x is a B.F.S.*

Proof. $(\Rightarrow)$ If x is an extreme point, by definition, there are (at least) n linearly independent defining hyperplanes at x, since $Ax = b$ provides m linearly independent binding hyperplane, there must by some $p = n - m$ additional binding defining hyperplanes from the non-negativity constraints that together with $Ax = b$ provide n linearly independent defining hyperplanes binding at x. Denoting these p additional hyperplanes by $x_N = 0$, we therefore conclude that the system $Ax = b, x_N = 0$ has x as the unique solution. Now, let N represent the columns of the variables x_N in A, and let B be the remaining columns of A with x_B as the associated variables. Since $Ax = b$ can be written as $Bx_B + Nx_N = b$, this means that B is $m \times m$ and invertible, and moreover, $x_B = B^{-1}b \geq 0$, since $x = (x_B, x_N)$ is a feasible solution. Therefore, x is a basic feasible solution.

$(\Leftarrow)$ If x is a basic feasible solution, by definition, $x = (x_B, x_N)$ where correspondingly $A = (B, N)$ such that $x_B = B^{-1}b \geq 0$ and $x_N = 0$. This means that the n hyperplanes $Ax = b, x_N = 0$ are binding at x and are linearly independent. Thus, x is an extreme point. □

2 Simplex Method

2.1 Search Algorithm

Improving search algorithm A simplex method is a search algorithm, for each iteration it finds a not-worse solution, which can be represented as

$$x^t = x^{t-1} + \lambda_{t-1}d^{t-1}$$

Where x^t is the solution of the tth iteration, λ_t is the step length of tth iteration, and d^t is the direction of the tth iteration.

Optimality test

$$
\begin{aligned}
z &= cx \\
&= \begin{bmatrix} c_B & c_N \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix} \\
&= c_B x_B + c_N x_N \\
\text{and} \quad Ax &= b \\
&\Rightarrow B x_B + N x_N = b, x_B \geq 0, x_N \geq 0 \\
&\Rightarrow x_B = B^{-1}b - B^{-1}N x_N \\
&\Rightarrow z = c_B B^{-1} b - c_B B^{-1} N x_N + c_N x_N
\end{aligned}
$$

for current solution $\hat{x} = \begin{bmatrix} \hat{x_B} \\ 0 \end{bmatrix}$, $\hat{z} = c_B B^{-1} b$, then

$$
z - \hat{z} = \begin{bmatrix} 0 & c_N - c_B B^{-1} N \end{bmatrix} \begin{bmatrix} x_B \\ x_N \end{bmatrix}
$$

The $c_N - c_B B^{-1} N$ is the reduced cost, for a minimized problem, if $c_N - c_B B^{-1} N > 0$ means $z - \hat{z} \geq 0$, it reaches the optimality because we cannot find a solution less than $\hat{z}$.

Find direction Suppose we choose x_k as a candidate to pivot into Basis

$$
x = \begin{bmatrix} B^{-1}b - B^{-1}a_k x_k \\ 0 + e_k x_k \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} + \begin{bmatrix} -B^{-1}a_k \\ e_k \end{bmatrix} x_k
$$

In this form, we can see: x is the result after tth iteration, $\begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix}$ is the result after $(t-1)$th iteration. $\begin{bmatrix} -B^{-1}a_k \\ e_k \end{bmatrix}$ is the iteration direction, x_k is the step length. The only requirement of x_k is $r_k < 0$ where $r_k = c_k - z_k$ is reduce cost, which is the kth entry of $c_N - c_B B^{-1} N$. Generally speaking, we usually take $r_k = \min\{c_j - z_j\}$ (which in fact can not guarantee the efficient of the algorithm.)

Find the step length We need to guarantee the non-negativity, so for each iteration, we need to make sure $x \geq 0$. Which means

$$
B^{-1}b - B^{-1}a_k x_k \geq 0
$$

Denote $B^{-1}b$ as $\bar{b}$, denote $B^{-1}a_k$ as y_k. If $y_k \leq 0$, we can have x_k as large as infinite, which means unboundness. If $y_k > 0$ now we can use the minimum ratio to guarantee non-negativity.

$$
\forall i \in B, \bar{b}_i - y_{ik} x_k \geq 0 \Rightarrow x_k = \min_{i \in B}\{\frac{\bar{b}_i}{y_{ik}}, y_{ik} > 0\}.
$$

2.2 Find Initial Solution

Non-trivial case If some of the constraint is not in $\sum_{i=1}^{n} a_i x_i \leq 0$ form, we cannot add a positive slack variable. In this case, we add an artificial variable other than slack variable.

$$\sum_{i=1}^{n} a_i x_i \geq (or =) 0 \Rightarrow \sum_{j=1}^{n} a_i x_i + x_a = 0$$

Notice that in an optimal solution, $x_a = 0$, otherwise it is not valid.
Artificial variables are only a tool to get the simplex method started. A so-called Two-phase Method is describe in this section.

Two-Phase method

- Phase I: Solve the following program start with a basic feasible solution $x = 0, x_a = b$, i.e., the artificial variable forms the basis.

$$\begin{aligned} \min \quad & 1x_a \\ \text{s.t.} \quad & Ax + x_a = b \\ & x \geq 0 \\ & x_a \geq 0 \end{aligned}$$

  If the optimal $1x_a \neq 0$, infeasible, stop. Otherwise proceed Phase II.

- Phase II: Remove the columns of artificial variables, replace the objective function with the original objective function, proceed to solve simplex method.

Discussion

- Case A: $x_a \neq 0$. Infeasible.

- Case B.1: $x_a = 0$ and all artificial variables are out of the basis. At the end of Phase I, we derive

x_0	x_B	x_N	x_a	RHS
1	0	0	-1	0
0	I	$B^{-1}N$	B^{-1}	$B^{-1}b$

  We can discard x_a columns, (or we can leave it because it keeps track of B^{-1}), and then we do the Phase II

z	x_B	x_N	RHS
1	0	$c_B B^{-1} N - c_N$	$c_B B^{-1} b$
0	I	$B^{-1}N$	$B^{-1}b$

- Case B.2: Some artificial variables are in the basis at zero values. This is because of degeneracy. We pivot on those artificial variables, once they leave the basis, eliminate them.

2.3 Degeneracy and Cycling

Degeneracy If the basic variable x_B is not strictly > 0, i.e. if some basic variable equals to 0, we call it degenerate.

Cycling In the degenerate case, pivoting by the simplex rule does not always give a strict decrease in the objective function value, because it may have $b_r = 0$. It is possible that the tableau may repeat if we use the simplex rule.Geometrically speaking, it means that at the same point - extreme point - it corresponds to more than one feasible solutions, so when we are pivoting, we stays at the same place. However, in computer algorithm, we rarely care about cycling because the data in computer is not precise, it is very hard to get into cycling.

Cycling prevent

- Lexicographic rule
  - For entering variable, same as simplex rule
  - For leaving variable, if there is a tie, choose the variable with the smallest $\dfrac{y_{r1}}{y_{rk}}$.
- Bland's rule
  - For entering variable, choose the variable with smallest index where $z_j - c_j \leq 0$
  - For leaving variable, if there is a tie, choose the variable with smallest index.
- Successive ratio rule
  - Select the pivot column as any column k where $z_k - c_k \leq 0$
  - Given k, select the pivot row r as the minimum successive ratio row associated with column k. In other words, for pivot columns where there is no tie in the usual minimum ratio, the successive ratio rule reduces to the simplex rule.

3 Revised Simplex Method

3.1 Key to Revised Simplex Method

The procedure of Simplex Method is (almost) exactly the same as original simplex method. However, notice that we don't need to use N so for the revised simplex method, we don't calculate any matrix related to N

The original matrix:

z	x_B	x_N	RHS
1	0	$c_B B^{-1} N - c_N$	$c_B B^{-1} b$
0	I	$B^{-1} N$	$B^{-1} b$

The revised matrix:

Basic Inverse	RHS
$w = c_B B^{-1}$	$c_B \bar{b} = c_B B^{-1} b$
B^{-1}	$\bar{b} = B^{-1} b$

For each pivot iteration, calculate $z_j - c_j = wa_j - c_j = c_B B^{-1} a_j - c_j, \forall j \in N$, pivot rules are the same as simplex method, each time find a variable x_k to enter basis

B^{-1}	RHS
w	$c_B \bar{b}$
B^{-1}	$\bar{b}$

x_k
$z_k - c_k$
y_k

Do the minimum ratio rule to find the variable x_r to leave the basis

B^{-1}	RHS	x_k
w	$c_B \bar{b}$	$z_k - c_k$
	$\bar{b}_1$	y_{1k}
	$\bar{b}_2$	y_{2k}
	$\cdots$	$\cdots$
B^{-1}	$\bar{b}_r$	y_{rk}(pivot at here)
	$\cdots$	$\cdots$
	$\bar{b}_m$	y_{mk}

3.2 Comparison between Simplex and Revised Simplex

Advantage of revised simplex

- Save storage memory
- Don't need to calculate N (including $B^{-1}N$ and $c_B B^{-1} N$)
- More accurate because round up errors will not be accumulated

Disadvantage of revised simplex Need to calculate wa_j for all $j \in N$ (in fact don't need to calculated it for the variable just left the basis)

Computation complexity

Method	Type	Operations
Simplex	$\times$	$(m+1)(n-m+1)$
	$+$	$m(n-m+1)$
Revised Simplex	$\times$	$(m+1)^2 + m(n-m)$
	$+$	$m(m+1) + m(n-m)$

When to use?

- When $m >> n$, do revise simplex method on the dual problem
- When $m \simeq n$, revise simplex method is not as good as simplex method
- When $m << n$ perfect for revise simplex method.

3.3 Decomposition of B inverse

Let $B = \{a_{B_1}, a_{B_2}, ..., a_{B_r}, ..., a_{B_m}\}$ and B^{-1} is known. If a_{B_r} is replaced by a_{B_k}, then B becomes $\bar{B}$. Which means a_{B_r} enters the basis and a_{B_k} leaves the basis.
Then $\bar{B}^{-1}$ can be represent by B^{-1}. Noting that $a_k = By_k$ and $a_{B_i} = Be_i$, then

$$\begin{aligned}\bar{B} &= (a_{B_1}, a_{B_2}, ..., a_{B_{r-1}}, a_k, a_{B_{r+1}}, a_m) \\ &= (Be_1, Be_2, ..., Be_{r-1}, By_k, Be_{r+1}, ..., Be_m) \\ &= BT\end{aligned}$$

where T is

$$T = \begin{bmatrix} 1 & 0 & ... & 0 & y_{1k} & 0 & ... & 0 \\ 0 & 1 & ... & 0 & y_{2k} & 0 & ... & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & ... & 1 & y_{r-1,k} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & y_{rk} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & y_{r+1,k} & 1 & ... & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & ... & 0 & y_{mk} & 0 & ... & 1 \end{bmatrix}$$

and

$$E = T^{-1} = \begin{bmatrix} 1 & 0 & ... & 0 & \frac{-y_{1k}}{y_{rk}} & 0 & ... & 0 \\ 0 & 1 & ... & 0 & \frac{-y_{2k}}{y_{rk}} & 0 & ... & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & ... & 1 & \frac{-y_{r-1,k}}{y_{rk}} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & \frac{1}{y_{rk}} & 0 & ... & 0 \\ 0 & 0 & ... & 0 & \frac{-y_{r+1,k}}{y_{rk}} & 1 & ... & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & ... & 0 & \frac{-y_{mk}}{y_{rk}} & 0 & ... & 1 \end{bmatrix}$$

For each iteration, i.e. one variable enters the basis and one leaves the basis, $\bar{B}^{-1} = T^{-1}B^{-1} = EB^{-1}$. Given that the first iteration starts from slack variables, the first basis B_1 is I, then we have

$$B_t^{-1} = E_{t-1}E_{t-2} \cdots E_2E_1I$$

Using E in calculation can simplify the product of matrix where

$$\begin{aligned}cE &= c_1, c_2, ..., c_m \begin{bmatrix} 1 & 0 & ... & g_1 & ... & 0 \\ 0 & 1 & ... & g_2 & ... & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & g_m & ... & 1 \end{bmatrix} \\ &= (c_1, c_2, ..., c_{r-1}, cg, c_{r+1}, ..., c_m)\end{aligned}$$

and

$$Ea = \begin{bmatrix} 1 & 0 & ... & g_1 & ... & 0 \\ 0 & 1 & ... & g_2 & ... & 0 \\ \vdots & \vdots & & \vdots & & \vdots \\ 0 & 0 & ... & g_m & ... & 1 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{bmatrix}$$

$$= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{r-1} \\ 0 \\ a_{r+1} \\ \vdots \\ a_m \end{bmatrix} + a_r \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{r-1} \\ g_r \\ g_{r+1} \\ \vdots \\ g_m \end{bmatrix}$$

$$= \bar{a} + a_r g$$

Then we can calculate w, y_k and $\bar{b}$

$$\begin{aligned} w &= c_B B^{-1} = c_B E_{t-1} E_{t-2} ... E_2 E_1 \\ y_k &= B^{-1} a_k = E_{t-1} E_{t-2} ... E_2 E_1 a_k \\ \bar{b} &= B_{t+1}^{-1} b = E_t E_{t-1} E_{t-2} ... E_2 E_1 b \end{aligned}$$

4 Duality

4.1 Dual Formulation

For any prime problem

$$\begin{aligned} \min \quad & cx \\ \text{s.t.} \quad & Ax \geq b \\ & x \geq 0 \end{aligned}$$

we can have a dual problem

$$\begin{aligned} \max \quad & wb \\ \text{s.t.} \quad & wA \leq c \\ & w \geq 0 \end{aligned}$$

4.2 Mixed Forms of Duality

For the following prime problem

$$\begin{aligned} \text{P(or D)} \quad \min \quad & c_1 x_1 + c_2 x_2 + c_3 x_3 \\ \text{s.t.} \quad & A_{11} x_1 + A_{12} x_2 + A_{13} x_3 \geq b_1 \\ & A_{21} x_1 + A_{22} x_2 + A_{23} x_3 \leq b_2 \\ & A_{31} x_1 + A_{32} x_2 + A_{33} x_3 = b_3 \end{aligned}$$

$$\begin{aligned} &x_1 \geq 0 \\ &x_2 \leq 0 \\ &x_3 \quad \text{unrestricted} \end{aligned}$$

The dual of the problem

$$\begin{aligned} \text{D(or P)} \quad \max \quad & w_1 b_1 + w_2 b_2 + w_3 b||3 \\ \text{s.t.} \quad & w_1 A_{11} + w_2 A_{21} + w_3 A_{31} \leq c_1 \\ & w_1 A_{12} + w_2 A_{22} + w_3 A_{32} \geq c_2 \\ & w_1 A_{13} + w_2 A_{23} + w_3 A_{33} = c_3 \\ & w_1 \geq 0 \\ & w_2 \leq 0 \\ & w_3 \quad \text{unrestricted} \end{aligned}$$

In sum, the relation between primal and dual problems are listed as following

	Minimization		Maximization	
Var	≥ 0	$\longleftrightarrow$	≤ 0	Cons
	≤ 0	$\longleftrightarrow$	≥ 0	
	Unrestricted	$\longleftrightarrow$	$=$	
Cons	≥ 0	$\longleftrightarrow$	≥ 0	Var
	≤ 0	$\longleftrightarrow$	≤ 0	
	$=$	$\longleftrightarrow$	Unrestricted	

4.3 Primal-Dual Relationships

Weak duality property Let x_0 be any feasible solution of a primal minimization problem,

$$Ax_0 \geq b, \quad x_0 \geq 0$$

Let x_0 be any feasible solution of a dual maximization problem,

$$w_0 A \leq c, \quad w_0 \geq 0$$

Therefore, we have

$$cx_0 \geq w_0 A x_0 \geq w_0 b$$

which is called the weak duality property. This property is for any feasible solution in the primal and dual problem.
Therefore, any feasible solution in the maximization problem gives the lower bound of its dual problem, which is a minimization problem, vice versa. We use this to give the bounds in using linear relaxation to solve IP problem.

Fundamental theorem of duality With regard to the primal and dual LP problems, one and only one of the following can be true.

- Both primal and dual has optimal solution x^* and w^*, where $cx^* = w^*b$

- One problem has an unbounded optimal objective value, the other problem must be infeasible
- Both problems are infeasible.

P		D
Optimal Degeneracy (or Multiple optimal)	$\Rightarrow$	Optimal Multiple optimal (or Degeneracy)
Infeasible	$\Rightarrow$	Unbounded or Infeasible
Unbounded	$\Rightarrow$	Infeasible

Strong duality property From KKT condition, we know that in order to make x^* the optimal solution, the following condition should be met.

- Primal Optimal: $Ax^* \geq b$, $x^* \geq 0$
- Dual Optimal: $w^*A \leq c$, $w^* \geq 0$
- Complementary Slackness:

$$\begin{cases} w^*(Ax^* - b) = 0 \\ (c - w^*A)x^* = 0 \end{cases}$$

The first condition means the primal has an optimal solution, the second condition means the dual has an optimal solution. The third condition means $cx^* = w^*b$, which is also called strong duality property.

Complementary slackness theorem Let $\boldsymbol{x^*}$ and $\boldsymbol{w^*}$ be any feasible solutions, they are optimal iff

$$\begin{aligned} (c_j - \boldsymbol{w^* a_j})x_j^* = 0, &\quad j = 1, ..., n \\ w_i^*(\boldsymbol{a^i x^*} - b_i) = 0, &\quad i = 1, ..., m \end{aligned}$$

In particular

$$\begin{aligned} x_j^* > 0 &\Rightarrow \boldsymbol{w^* a_j} = c_j \\ \boldsymbol{w^* a_j} < c_j &\Rightarrow x_j^* = 0 \\ w_i^* > 0 &\Rightarrow \boldsymbol{a^i x^*} = b_i \\ \boldsymbol{a^i x^*} > b_i &\Rightarrow w_i^* = 0 \end{aligned}$$

It means, if in optimal solution a variable is positive (has to be in the basis), the correspond constraint in the other problem is tight. If the constraint in one problem is not tight, the correspond variable in the other problem is zero.

Use dual to solve the primal in the dual problem, we solved some w which is positive, we can know that the correspond constraint in primal is tight, furthermore we can solve the basic variables from those tight constraints, which becomes equality and we can solve it using Gaussian-Elimination.

4.4 Shadow Price

Shadow price under non-degeneracy Let B be an optimal basis for primal problem and the optimal solution x^* is non-degenerated.

$$z = c_B B^{-1} b - \sum_{j \in N} (z_j - c_j) x_j = w^* b - \sum_{j \in N} (z_j - c_j) x_j$$

therefore

$$\frac{\partial z^*}{\partial b_i} = c_B B_i^{-1} = w_i^*$$

w^* is the shadow prices for the right-hand-side vectors. We can also regard it as the incremental cost of producing one more unit of the ith product. Or w^* is the fair price we would pay to have an extra unit of the ith product.

Shadow price under degeneracy For shadow price under degeneracy, the w^* may not be the true shadow price, for it may not be the right basis. In this case, the partial differentiation may not be valid, for component b_i, if $x_i = 0$ and x_i is a basic variable, we can't find the differentiation.

4.5 Dual Simplex Method

Occasionally, an initial feasible solution for the Simplex Method might be difficult to acquired, for example, when **0** is not an initial solution, or, in some other cases, a basic infeasible solution is known in advance. To solve the LP, we can take advantage of the LP Strong Duality Theorem, and solve the dual problem to optimality instead of the prime problem. Such technique is called the Dual Simplex Method. The details of such method is as follows (for minimization problems)

1. Choose an initial basic solution x_B and corresponding basis matrix B so that $c_B B^{-1} A_j - c_j \geq 0$ for all $j \in J$, where J is the set of non-basic variables.

2. Construct a simplex tableau using this initial solution.

3. If $\bar{b} = B^{-1} b \geq 0$, then an optimal solution has been achieved; STOP. Otherwise, the dual problem is feasible, GOTO STEP 4.

4. Choose a leaving variable $x_{B_i} = \bar{b}_i$ so that $\bar{b}_i < 0$

5. Choose the index of the entering variable x_j using the following minimum ratio test:

$$\frac{z_j - c_j}{\bar{a}_{j_i}} = \min\{\frac{z_k - c_k}{\bar{a}_{k_i}} | k \in J, \bar{a}_{j_i} < 0\}$$

6. If no entering variable can be selected ($\bar{a}_{j_i} \geq 0, \forall k \in K$) then the dual problem is unbounded and the prime problem is infeasible. STOP

7. Using a standard simplex pivot, on element $\bar{a}_{j_i}$, thus causing w_{B_i} to become 0 (and thus feasible) and causing x_j to enter the basis. GOTO STEP 3.

In general, the Dual Simplex Method is (almost) the same as a Simplex Method. The Dual Simplex Method starts with an infeasible basic solution, during each iteration, it maintains dual feasibility and work towards primal feasibility. In Dual Simplex Method, the old Right Hand Side becomes new $z_j - c_j$, and the old $z_j - c_j$ become new Right Hand Side.