Optimization Methods (II)

Hongxin Zhang

2021-04-06

Content

- Linear Programming
- Nonlinear Programming

- Reference：
  - 《线性规划》
    - 张建中，许绍吉，科学出版社
  - 《最优化理论与方法》
    - 袁亚湘，孙文瑜，科学出版社
  - 《数学规划》
    - 黄红选，韩继业，清华大学出版社

1. Linear Programming

- In people’s practice of production， we often meet the problem that how to use existing resources to arrange the production in order to obtain the maximum economic benefit.

$profit driven$

Operations Research →

mathematical programming →

Linear Programming (LP)

- 1947,G. B. Dantzig solved the simplex method for linear programming.
- After the existing computer can handle the LP problems of tens of thousands of constraints and decision variables, the suitable domain of LP became wider.

1. Fundamental Theory

1.1 Problem Definition (linear programming, LP)

- A computer company which has three technical teams A, B, C provides large computer systems
  - High-end system
    - Profit: 4 million yuan per system
    - Require A, B two teams to implement collaboratively
    - Team A needs 2 months， Team B needs 1 month；

  - Intermediate system
    - Profit: 3 million yuan per system
    - Require A, B ,C three teams to implement collaboratively
    - Each team needs 1 month;

  - In 2009, team A can work for 10 months, team B 8 months and team C 7 months.
- Q: what kind of sales strategy can make the biggest profits?

1.1 Problem Definition (linear programming, LP)

- Assume the company can sale x_1 high-end systems and x_2 intermediate systems
- Objective Function

$$\max \quad z = 4x_1 + 3x_2$$

- subject to the constraints

$$\begin{cases} 2x_1 + x_2 \le 10 \\ x_1 + x_2 \le 8 \\ x_2 \le 7 \\ x_1, x_2 \ge 0 \end{cases}$$

1.1 Problem Definition (linear programming, LP)

- Under a set of constraints which must be a linear equality or inequality, maximize or minimize a linear function
- Formal definition:

$$
\begin{aligned}
&\min c_1x_1 + \ldots + c_nx_n \\
&\text{s. t.} \\
&\qquad a_{11}x_1 + \ldots + a_{1n}x_n \le b_1 \\
&\qquad \ldots \\
&\qquad a_{m1}x_1 + \ldots + a_{mn}x_n \le b_m \\
&\qquad x_1, \ldots, x_n \ge 0
\end{aligned}
$$

1.1 Problem Definition (linear programming, LP)

- Under a set of constraints which must be a linear equality or inequality, maximize or minimize a linear function
- Formal definition:

objective function

$$\begin{aligned} &\min \mathbf{c}^T\mathbf{x} \\ &\text{s.t. } A\mathbf{x} \leq \mathbf{b} \\ &\quad\;\; \mathbf{x} \geq \mathbf{0} \end{aligned}$$

c : cost vector
A: constraint matrix
b: right-hand-side vector

1.1 Problem Definition (linear programming, LP)

- **X**： A feasible solution (feasible point)to a linear program is a solution that satisfies all constraints.

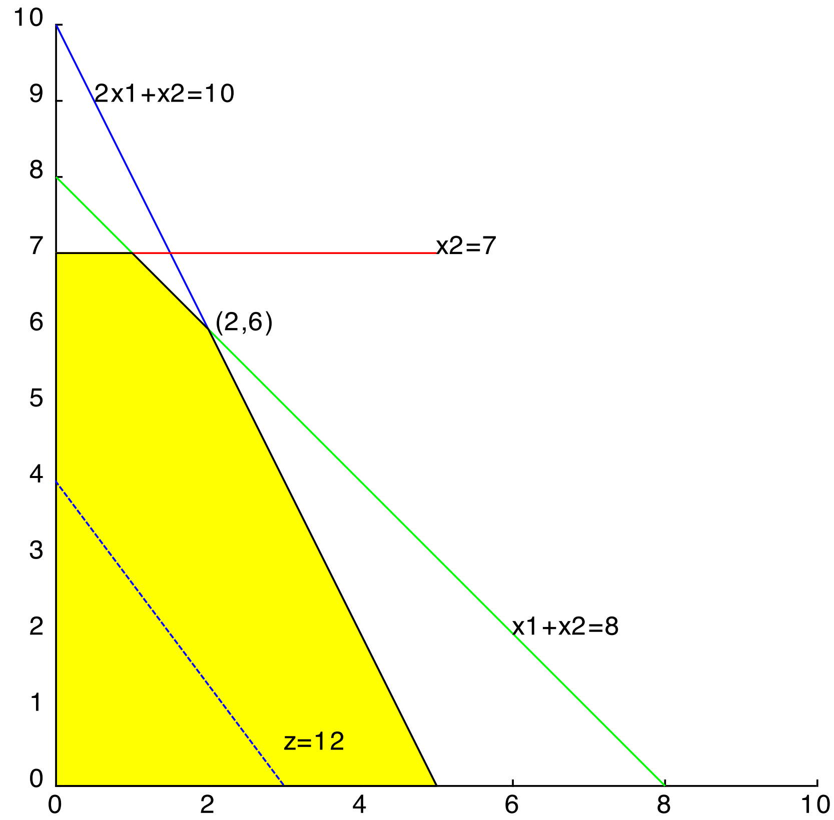
- D： The feasible region in a linear program is the set of all possible feasible solutions.

- LP problem：
  - D = Ø, no solution or infeasible
  - D ≠ Ø， but optimal solution is unbounded in D: unbounded
  - D ≠ Ø， and objective function has limited optimal solutions: unique optimal solution value

Graphical Method for LP

$$\max \quad z = 4x_1 + 3x_2$$

$$\begin{cases} 2x_1 + x_2 \le 10 \\ x_1 + x_2 \le 8 \\ x_2 \le 7 \\ x_1, x_2 \ge 0 \end{cases}$$



1.2 Standard LP Problem

- The form of LP summaried from reality is not exactly the same:
  - The objective function is the maximum or the minimum,
  - The constraints may be equalities or inequalities,
  - The variables may be upper bounded, lower bounded, or unbounded.

$$\boxed{\begin{aligned} &\min z = \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &\text{s.t. } A\mathbf{x} = \mathbf{b} \\ &\qquad\quad \mathbf{x} \geq \mathbf{0} \end{aligned}}$$

Standard form of LP Problem

$$\begin{aligned} \min z &= \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ \text{s.t. } A\mathbf{x} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0} \end{aligned}$$

- Standardization
  - Transformation of objective function max z → min ($-z$)
  - Transformation of constraints (add slack variables)

$$\sum_{j=1}^{n} a_{ij}x_j \leq b_i \Leftrightarrow \begin{cases} \sum_{j=1}^{n} a_{ij}x_j + x_{n+i} = b_i \\ x_{n+i} \geq 0 \end{cases}$$

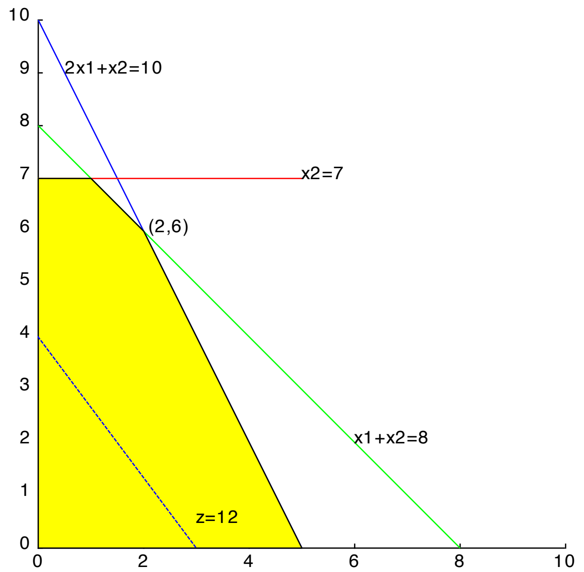
  - Non-negative constraints of variables

$$x_j \geq l_j \Leftrightarrow y_j \geq 0, y_j = x_j - l_j$$

$$x_j \text{ free variables } \Leftrightarrow u_j \geq 0, v_j \geq 0, x_j = u_j - v_j$$

1.3 Feasible Region D

- Now discuss the structure of feasible region $D=\{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x}\geq\mathbf{0}\}$
  - First discuss set $K=\{ \mathbf{x} \mid A\mathbf{x} = \mathbf{b} \}$



1.3 Basic Conception

- Affine Set：

  For a set $S \subseteq E^n$， if for any $\mathbf{x}$, $\mathbf{y} \in S$, $\lambda \in E^1$, there is

$$\lambda \mathbf{x}+(1-\lambda)\mathbf{y} \in S,$$

then set S is an affine set.

- Convex Set：

  If for any $\boldsymbol{x}$, $\boldsymbol{y} \in C$, $\lambda \in (0,1)$， there is

$$\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in C,$$

then set C is a convex set.

=> extreme point， extreme direction

1.4 Basic Feasible Solution

- Theorem
  - Feasible solution **x** is basic feasible solution ⟺ The corresponding column vector of **x** ' positive component is linearly independent
  - Feasible solution **x** is basic feasible solution ⟺ **x** is the extreme point of D

Extreme Direction of Feasible Region

- algebraic property of extreme direction: the direction $\mathbf{d}$ of $D=\{\mathbf{x} \mid A\mathbf{x}=\mathbf{b}, \mathbf{x}\geq\mathbf{0}\}$ has k nonzero components, then $\mathbf{d}$ is the extreme direction of D iff the rank of the corresponding column vectors of $\mathbf{d}$' nonzero components is k-1.
- geometric property of extreme direction: $\mathbf{d}$ is the extreme direction of D iff $\mathbf{d}$ is the direction of a half line interface of D.
- If D has extreme direction, it's obviously D is unbounded set； If D is unbounded set, D has direction, and has extreme direction

1.5 Basic Theorem of LP

Basic Theorem1 (representation theorem)

Theorem: Assume $\mathbf{x}^1, \dots, \mathbf{x}^k$ denote all the extreme points of $D = \{x | Ax = b, x \geq 0\}$, $\mathbf{d}^1, \dots, \mathbf{d}^l$ denote all the extreme directions, then $\mathbf{x} \in D$ iff there are $\lambda_i (i = 1, \dots, k)$ and $\mu_j (j = 1, \dots, l)$ make

$$\mathbf{x} = \sum_{i=1}^{k} \lambda_i \mathbf{x}^i + \sum_{j=1}^{l} \mu_j \mathbf{d}^j,$$

$$\lambda_i \geq 0,\ i = 1, \dots, k;\ \mu_j \geq 0,\ j = 1, \dots, l$$

$$\sum_{i=1}^{k} \lambda_i = 1$$

Basic Theorem 2

Given the LP problem

$$\min z = \mathbf{c}^{\mathrm{T}}\mathbf{x}$$

$$\text{s.t. } A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

(1) If there are limited optimal values (the LP problem has the optimal solution), the optimal value must reach on a extreme point in feasible region D.

(2) Objective function has limited optimal values iff $\mathbf{c}\mathbf{d}^j \geq 0$ for all the extreme directions $\mathbf{d}^j$ of D.

Comments and Understanding

- Theorem illustration:
  - The optimal value of standard LP problem's objective function must reach in a basic feasible solution.
  - Solving standard LP problem, only need to search in the basic feasible solution set.
- It is usually impractical to calculate and compare all the basic feasible solution, and the reason is that when n is larger, the number of basic feasible solution will be bigger.
- The usual method is calculating and searching in a subset of basic feasible solutions according to certain rules, this is simplex method.

2. Simplex Method

2.1 Basic Simplex Method

- Main idea: First find a basic feasible solution and distinguish whether it is the optimal solution. If not, find a better basic feasible solution and test it. Iterate like this until finally find the optimal solution or ensure unbounded.

- Two main problems:
  - Looking for initial solution
  - How to distinguish and iterate (considered first)

Determine initial basic feasible solution

Determining initial basic feasible solution is equivalent to determining initial feasible basis， once initial feasible basis is determined, the corresponding initial basic feasible solution can be determined uniquely. For convenience, assume in the standard linear programming, the first m coefficient column vectors of coefficient matrix **A** can form a feasible basis, which means:

$\mathbf{A} = (\mathbf{BN})$ where

$\mathbf{B} = (P_1, P_2, \ldots, P_m)$ is the feasible basis formed by the coefficient column vectors of basic variables $x_1, x_2, \ldots, x_m$,

$\mathbf{N} = (P_{m+1}, P_{m+2}, \ldots, P_n)$ is the matrix formed by the coefficient column vectors of nonbasic variable $x_{m+1}, x_{m+2}, \ldots, x_n$.

So the constraint equation $AX=b$ can be expressed as :

$$AX=(BN)\begin{pmatrix} X_B \\ X_N \end{pmatrix}=BX_B+NX_N=b$$

Do premultiplication at both ends of the equation by the inverse matrix $\mathbf{B}^{-1}$ of feasible basis **B**, then transposition to achieve:

$$X_B=B^{-1}b-B^{-1}NX_N$$

If we make all the nonbasic variables $X_N=0$, then the basic variables can be

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

Thus we get the initial basic solution $$X=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

$$AX=b \rightarrow BX_B+NX_N=b \rightarrow X_B=B^{-1}b-B^{-1}NX_N \rightarrow X_N=0, X_B=B^{-1}b$$

- Question:

➢ It is not easy to determine whether the m coefficient column vectors can form a base.

The base is formed by m linearly independent coefficient column vectors of coefficient matrix **A**.

But it's not easy to determine whether the m coefficient column vectors is linearly independent.

➢ Even if a basis **B** is found in the coefficient matrix **A**, there is no guarantee that the basis is feasible. Since we can not guarantee that the basic variable $\mathbf{X_B} = \mathbf{B}^{-1}\boldsymbol{b} \geq 0$

➢ In order to obtain the basic feasible solution $\mathbf{X} = \begin{pmatrix} \mathbf{B}^{-1}\mathbf{b} \\ 0 \end{pmatrix}$, it is definite to get the inverse matrix $\mathbf{B}^{-1}$.

But it's also troublesome to obtain $\mathbf{B}^{-1}$.

- **Conclusion:**
- **Try to get a m order unit matrix I as the initial feasible basis B in the standardization process of LP.**

In order to get a m order unit matrix **I** as the initial feasible basis **B**, do the following process in the standardization process of LP:

- Before standardization, if the m constraint equations are all in the form of "$\leq$ ", then just need to add a slack variable $x_{n+i}(i = 1,2, \ldots, m)$ at the left side of each constraint equation during standardization process.

- Before standardization, if there are inequalities "$\geq$ " in constraint equations, then in addition to subtracting the remaining variables at the left side of equation to translate inequality into equality, we must add a new nonnegative variable at the left side, which called artificial variable.

- Before standardization, if there are equalities in constraint equations, then directly add artificial variables at the left side of the equations.

■ Judge whether the current basic feasible solution is optimal

If a basic feasible solution has been obtained $X=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$

Put this basic feasible solution into objective function, then can get the value of objective function

$$Z=CX=(C_B C_N)\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}=C_B B^{-1}b$$

where $C_B=(c_1,c_2,\cdots c_m)$, $C_N=(c_{m+1},c_{m+2},\cdots c_n)$ represent the corresponding value coefficient subvectors of basic variable and nonbasic variable respectively.

In order to judge whether $Z=C_B B^{-1}b$ has reached the maximum, put $X_B=B^{-1}b-B^{-1}NX_N$ into objective function and represent objective function by nonbasic variables, that is:

$$Z=CX=(C_B C_N)\begin{pmatrix} X_B \\ X_N \end{pmatrix}$$

$$=C_B X_B+C_N X_N=C_B(B^{-1}b-B^{-1}NX_N)+C_N X_N$$

$$=C_B B^{-1}b+(C_N-C_B B^{-1}N)X_N$$

$$\triangleq C_B B^{-1}b+\sigma_N X_N=C_B B^{-1}b+(\sigma_{m+1},\sigma_{m+1},\cdots\sigma_n)\begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$

where $\sigma_N=C_N-C_B B^{-1}N=(\sigma_{m+1},\sigma_{m+1},\cdots\sigma_n)$ is called nonbasic variable X_N's test vectors whose each component is called check number. If each check number of σ_N is less than or equal to 0, then the basic feasible solution in current is the optimal solution.

Theorem 1: Optimal solution judgment theorem

For the LP problem $\max Z = CX, \quad D = \{X \in R^n / AX = b, X \geq 0\}$

If the test vector of a certain basic feasible solution $\sigma_N = C_N - C_B B^{-1} N \leq 0$

then the basic feasible solution is the optimal solution

$$Z = C_B B^{-1} b + (\sigma_{m+1,} \sigma_{m+1,} \cdots \sigma_n) \begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$

Theorem 2: Infinite optimal solutions judgment theorem

If $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ is a basic feasible solution, and the corresponding test vector

$\sigma_N = C_N - C_B B^{-1} N \leq 0$, where there is a check number $\sigma_{m+k} = 0$, then the LP problem has infinite optimal solutions.

■ Improvement of basic feasible solution

If the current basic feasible solution **X** is not the optimal solution, namely, there is positive check number in the test vector $\sigma_N = C_N - C_B B^{-1} N$, then we need to seek a new basic feasible solution based on the former basic feasible solution **X**, and improve the objective function. Specific approach is:

- First determine a swap-in variable from the nonbasic variable whose check number is positive, and transform the nonbasic variable into basic variable (its value increases from 0 to a positive value)
- Second determine a swap-out variable from the original basic variable, and transform the basic variable into nonbasic variable (its value decreases from a positive value to 0).

From this, a new basic feasible solution can be obtained, according to

$$Z = C_B B^{-1} b + (\sigma_{m+1}, \sigma_{m+1}, \cdots \sigma_n) \begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$

the transformation must increase the value of objective function.

The determination of swap-in and swap-out variable:

● The determination of swap-in variable— maximum increase principle

Assume the test vector $\sigma_N = C_N - C_B B^{-1} N = (\sigma_{m+1}, \sigma_{m+2}, \cdots \sigma_n)$

If there are two or more check numbers are positive, then in order to make the value of objective function increase faster, "maximal increase principle" is the general choice, namely, choose the corresponding nonbasic variable of the maximal positive check number as the swap-in variable, that is if

$$\max\left\{\sigma_j / \sigma_j > 0, m+1 \le j \le n\right\} = \sigma_{m+k}$$

Then choose the corresponding x_{m+k} as swap-in variable,

Because $\sigma_{m+k} > 0$ and is maximal

So when X_{m+k} increases from 0 to a positive value,

the value of objective function $Z = C_B B^{-1} b + (\sigma_{m+1,} \sigma_{m+1,} \cdots \sigma_n) \begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$ can be increased furthest

- The determination of swap-out variable — minimum ratio principle

If x_{m+k} is determined to be the swap-in variable, equation

$$X_B = B^{-1}b - B^{-1}NX_N \Rightarrow X_B = B^{-1}b - B^{-1}P_{m+k}x_{m+k}$$

where P_{m+k} is the coefficient column vector in **A** corresponding to x_{m+k}.

Now it is required to determine a basic variable as swap-out variable in

$X_B = (x_1, x_2, \cdots x_m)^T$.

When x_{m+k} increases from 0 to a certain value gradually, the non-negativity of X_B may be broken. In order to keep the feasibility of solution, the swap-out variable can be determined according to the minimum ratio principle:

If

$$\min\left\{\frac{(B^{-1}b)_i}{(B^{-1}P_{m+k})_i} \Big/ (B^{-1}P_{m+k})_i > 0, 1 \le i \le m\right\} = \frac{(B^{-1}b)_l}{(B^{-1}P_{m+k})_l}$$

then choose the corresponding basic variable x_l as swap-out variable.

Theorem 3: No optimal solution judgment theorem

If $X=\begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ is a basic feasible solution, there is a check number $\sigma_{m+k} > 0$ but $B^{-1}P_{m+k} \le 0$, then the this LP has no optimal solution.

Proof: Fix $x_{m+k} = \lambda, (\lambda > 0)$ then a new feasible solution can be obtained

Put the equation above into

$$X_B = B^{-1}b - B^{-1}P_{m+k}x_{m+k} = B^{-1}b - B^{-1}P_{m+k}\lambda$$

$$Z = C_B B^{-1}b + (\sigma_{m+1}, \cdots \sigma_{m+k}, \cdots \sigma_n)\begin{pmatrix} x_{m+1} \\ \vdots \\ \lambda \\ \vdots \\ x_n \end{pmatrix} = C_B B^{-1}b + \sigma_{m+k}\lambda$$

Because $\sigma_{m+k} > 0$, when $\lambda \to +\infty$, $Z \to +\infty$.

■ Obtain the basic feasible solution improved by elementary transformation

Assume **B** is the feasible solution of LP $\max Z = CX, AX = b, X \geq 0$, then

$$AX = b \Rightarrow (BN)\begin{pmatrix} X_B \\ X_N \end{pmatrix} = b \Rightarrow (I, B^{-1}N)\begin{pmatrix} X_B \\ X_N \end{pmatrix} = B^{-1}b$$

Fix the nonbasic variable $X_N = 0$, then the basic variable $X_B = B^{-1}b$.

The basic feasible solution $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ can be obtained.

➢ Doing premultiplication at both ends of the constraint equation set by the inverse matrix B^{-1} is equal to doing a series of elementary transformation. The result of transformation is transforming feasible basis **B** in coefficient matrix **A** into identity matrix **I**, transforming the matrix **N** composed of nonbasic variable's coefficient column vectors into $B^{-1}N$, and transforming resource vector b into $B^{-1}b$.

Since the equation set after elementary transformation $(I, B^{-1}N)\begin{pmatrix} X_B \\ X_N \end{pmatrix} = B^{-1}b$

has the same solution with the original constraint equation set $AX=b$ or

$$(B, N)\begin{pmatrix} X_B \\ X_N \end{pmatrix} = b$$

And the basic feasible solution improved X' just replace one of the swap-out variables with a swap-in variable based on X's basic variables, the other basic variables remain unchanged. The coefficient column vectors of these basic variables are the unit vectors of identity I. In order to get the improved basic feasible solution X', we just need to do the elementary transformation to the augmented matrix $(I, B^{-1}N, B^{-1}b)$, and transform the coefficient column vector of swap-in variable into the corresponding unit vector of swap-out variable.

Discrimination and Iteration

- Assume rank(A) = m<n, and assume a nonsingular feasible solution has been found, namely, a basis B has been found, then $A\mathbf{x}=\mathbf{b}$ can be written as

$$\mathbf{x}_B+B^{-1}N\mathbf{x}_N=B^{-1}\mathbf{b} \qquad (1)$$

Note

$B = (a_1, \dots, a_m)$

$\bar{\boldsymbol{a}}_j = B^{-1}\boldsymbol{a}_j = (\bar{a}_{1j}, \dots, \bar{a}_{mj})^T, j = 1, \dots n$

$\bar{\boldsymbol{b}} = B^{-1}\boldsymbol{b} = (\bar{b}_1, \dots, \bar{b}_m)^T$

$\bar{N} = B^{-1}N$

Then the equation (1) can be written as $\mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}}$ (2)

Obviously, the corresponding equation expressions of different basis are different.

Equation (2) is called the canonical equation expression of basis B.

Discrimination and Iteration

$$\mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}} \qquad (2)$$

Obviously, the corresponding equation expressions of different basis are different.

Equation (2) is called the canonical equation expression of basis B.

If $\bar{\mathbf{b}} \geq \mathbf{0}$, equation (2) is corresponding to the basic feasible solution $\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{b}} \\ 0 \end{pmatrix}$.

Do relevant transformation to the objective function,

$$\mathbf{z} = \mathbf{c}\mathbf{x} = \mathbf{c}_B\bar{\mathbf{b}} - (\mathbf{c}_B\bar{N} - \mathbf{c}_N)\mathbf{x}_N = \mathbf{c}_B\bar{\mathbf{b}} - \sum_{j=m+1}^{\boldsymbol{n}} (\mathbf{c}_B\bar{\mathbf{a}}_j - c_j)\mathbf{x}_j \qquad (3)$$

Use z_0 denote the value of objective function at the basic feasible solution $\bar{\mathbf{x}}$.

then $z_0 = \mathbf{c}\bar{\mathbf{x}} = \mathbf{c}_B\bar{\mathbf{b}}$ which is the constant term of the equation above.

Discrimination and Iteration

Since $\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_m$ are linearly independent vectors,

when $j = 1, \dots, m$,

$\mathbf{c}_B \bar{\mathbf{a}}_j - c_j = 0$.

Introduce the symbol

$\xi_j = \mathbf{c}_B \bar{\mathbf{a}}_j - c_j, j = 1, \dots, n$

or use the form of vector

$\xi = \mathbf{c}_B B^{-1} A - \mathbf{c} = (\boldsymbol{\xi}_B, \boldsymbol{\xi}_N) = (0, \mathbf{c}_B B^{-1} N - \mathbf{c}_N)$

then equation (3) can be rewritten as $\mathbf{z} = \mathbf{c}_B \bar{\mathbf{b}} - \xi \mathbf{x}$.

Discrimination and Iteration

$$\mathbf{z} = \mathbf{c}_B\bar{\mathbf{b}} - \boldsymbol{\xi}\mathbf{x}\ , \zeta = \mathbf{c}_B B^{-1}A - \mathbf{c} = (\zeta_B, \zeta_N) = (\mathbf{0}, \mathbf{c}_B B^{-1}N - \mathbf{c}_N)$$

After the transformation of variable, the LP problem can be narrated as

$$\min \mathrm{z} = z_0 - \boldsymbol{\xi}\mathbf{x} \qquad (4)$$

$$s.t. \qquad \mathbf{x}_B + B^{-1}N\mathbf{x}_N = \bar{\mathbf{b}}, \qquad (5)$$

$$\mathbf{x} \geq \mathbf{0}.$$

Theorem 2.1: If $\boldsymbol{\xi} \leq 0$ in (4), then $\bar{\mathbf{x}}$ is the optimal solution.

Theorem 2.2: If a component $\xi_k > 0$ of $\boldsymbol{\xi}$ in (4), and $\overline{\mathbf{a}_k} \leq 0$, then the original problem has no solution.

Theorem 2.3: If $\xi_k > 0$ in (4), and $\overline{\mathbf{a}_k}$ at least exist a positive component, then a basic feasible solution $\hat{\mathbf{x}}$ can be found which makes $\mathbf{c}\hat{\mathbf{x}} < \mathbf{c}\bar{\mathbf{x}}$.

Steps of Simplex Method

1. Find the initial feasible basis;
2. Solve the corresponding canonical equation;
1. Calculate $\xi_k = \max\{\xi_j | j = 1, \dots, \mathrm{n}\}$
2. If $\xi_k \le 0$, stop. The optimal solution is $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix}$ and the optimal value is $z = \mathbf{c}_B \bar{\mathbf{b}}$.
3. If $\bar{\mathbf{a}}_k \le 0$, stop, the original problem is unbounded.
4. Calculate the minimum ratio $\frac{\bar{b}_r}{\bar{a}_{rk}} = \min\{\frac{\bar{b}_i}{\bar{a}_{ik}} | \bar{a}_{ik} > 0\}$

ξ is called test vector.

During iteration process, if ξ has more than one positive components , then for making the objective function descend more faster, generally choose the column vector $\mathbf{a}_k$ corresponded to the maximum component ξ_k into the basis.

Example 1 $\max Z = 5x_1 + 2x_2 + 3x_3 - x_4 + x_5$

$$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 8 \\ 3x_1 + 4x_2 + x_3 + x_5 = 7 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases}$$

$$C = (5, 2, 3, -1, 1)$$

Solution:

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 & 0 \\ 3 & 4 & 1 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

(1) Determine the initial basic feasible solution

$B = (P_4 P_5) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the basic variables x_4, x_5, the nonbasic variables x_1, x_2, x_3.

$$X_B = \begin{pmatrix} x_4 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, \begin{matrix} C_B = (-1, 1) \\ C_N = (5, 2, 3) \end{matrix}, b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

$$X_N = 0 \rightarrow X_B = B^{-1}b = \begin{pmatrix} 8 \\ 7 \end{pmatrix} \Rightarrow X = (0, 0, 0, 8, 7)^T$$

$$Z = C_B B^{-1} b = (-1, 1) \begin{pmatrix} 8 \\ 7 \end{pmatrix} = -1$$

$$X_B=\begin{pmatrix} x_4 \\ x_5 \end{pmatrix}, X_N=\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, \begin{matrix} C_B=(-1,1) \\ C_N=(5,2,3) \end{matrix}, b=\begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

---

(2) Examine whether $X=(0,0,0,8,7)^T$ is the optimal solution.

Test vector $\sigma_N=C_N-C_B B^{-1}N=(5,2,3)-(-1,1)\begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}$

$$=(5,2,3)-(2,2,-1)=(\underset{\sigma_1}{\underset{\uparrow}{3}},\ \underset{\sigma_2}{\underset{\uparrow}{0}},\ \underset{\sigma_3}{\underset{\uparrow}{4}})$$

Because $\sigma_1=3$, $\sigma_3=4$, both of them are greater than zero, so $X=(0,0,0,8,7)^T$ is not the optimal solution.

$$X_B=\begin{pmatrix}x_4\\x_5\end{pmatrix},X_N=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},B=\begin{pmatrix}1&0\\0&1\end{pmatrix},N=\begin{pmatrix}1&2&2\\3&4&1\end{pmatrix},\begin{matrix}C_B=(-1,1)\\C_N=(5,2,3)\end{matrix},b=\begin{pmatrix}8\\7\end{pmatrix}$$

$$\sigma_N=(\sigma_1,\sigma_2,\sigma_3)=(3,0,4)$$

---

(3) The improvement of basic feasible solution $X=(0,0,0,8,7)^T$

① Choose the swap-in variable

Because $\max\{3,4\}=4$, choose x_3 as the swap-in variable.

② Choose the swap-out variable

$$B^{-1}b=\begin{pmatrix}8\\7\end{pmatrix},B^{-1}P_3=\begin{pmatrix}2\\1\end{pmatrix}\nleq 0 \text{ and } \min\left\{\frac{8}{2},\frac{7}{1}\right\}=\frac{8}{2},$$

choose x_4 as the swap-out variable.

$$\begin{pmatrix}x_4\\x_5\end{pmatrix}=B^{-1}b-B^{-1}P_3x_3=\begin{pmatrix}8\\7\end{pmatrix}-\begin{pmatrix}2\\1\end{pmatrix}x_3$$

$$X_B=\begin{pmatrix}x_4\\x_5\end{pmatrix},X_N=\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix},B=\begin{pmatrix}1&0\\0&1\end{pmatrix},N=\begin{pmatrix}1&2&2\\3&4&1\end{pmatrix},\begin{matrix}C_B=(-1,1)\\C_N=(5,2,3)\end{matrix},b=\begin{pmatrix}8\\7\end{pmatrix}$$

---

(4) Solve the improved basic feasible solution X'

Do elementary transformation to the augmented matrix of constraint equation set, and transform the corresponding coefficient column vector $P_3=\begin{pmatrix}2\\1\end{pmatrix}$ of swap-in variable x_3 into the corresponding unit vector $P_4=\begin{pmatrix}1\\0\end{pmatrix}$ of swap-out variable x_4 , Notice keep the coefficient column vector $P_5=\begin{pmatrix}0\\1\end{pmatrix}$ of basic variable x_5 as unit vector.

$$\begin{pmatrix}1&2&\boxed{2}&\boxed{1}&0&8\\3&4&\boxed{1}&\boxed{0}&1&7\end{pmatrix}\xrightarrow{\text{The first line divided by 2}}\begin{pmatrix}\frac{1}{2}&1&1&\frac{1}{2}&0&4\\3&4&1&0&1&7\end{pmatrix}$$

$$\xrightarrow{\text{The second line minus the first line}}\begin{pmatrix}\frac{1}{2}&1&1&\frac{1}{2}&0&4\\\frac{5}{2}&3&0&\frac{-1}{2}&1&3\end{pmatrix}$$

$$C=(5,2,3,-1,1) \qquad C=(5,2,3,-1,1)$$

$$\left(\begin{array}{ccccc|c} 1 & 2 & 2 & 1 & 0 & 8 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & \frac{-1}{2} & 1 & 3 \end{array}\right)$$

The improved basic feasible solution can be obtained.

$B=(P_3 P_5)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the basic variable x_3 x_5, the nonbasic variable x_1, x_2, x_4,

$$X_B=\begin{pmatrix} x_3 \\ x_5 \end{pmatrix}, X_N=\begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & \frac{-1}{2} \end{pmatrix}, \begin{matrix} C_B=(3,1) \\ C_N=(5,2,-1) \end{matrix}, b=\begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

$X_N=0 \to X_B=B^{-1}b=\begin{pmatrix} 4 \\ 3 \end{pmatrix} \Rightarrow$ basic feasible solution $X=(0,0,4,0,3)^T$

The value of objective function $Z=C_B B^{-1} b=(3,1)\begin{pmatrix} 4 \\ 3 \end{pmatrix}=15$

The value of objective function increases compared original Z=-1, then turn to step (2)

$$X_B=\begin{pmatrix} x_3 \\ x_5 \end{pmatrix}, X_N=\begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & \frac{-1}{2} \end{pmatrix}, \begin{matrix} C_B=(3,1) \\ C_N=(5,2,-1) \end{matrix}, b=\begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

---

(2) Examine whether $X=(0,0,4,0,3)^T$ is the optimal solution.

Test vector

$$\sigma_N=C_N-C_B B^{-1}N=(5,2,-1)-(3,1)\begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & \frac{-1}{2} \end{pmatrix}$$

$$=(5,2,-1)-(4,6,1)=(1,\ -4,\ -2)$$

$$\begin{matrix} \uparrow & \uparrow & \uparrow \\ \sigma_1 & \sigma_2 & \sigma_4 \end{matrix}$$

Because $\sigma_1 = 1 > 0$

so $X=(0,0,4,0,3)^T$ is not the optimal solution.

$$X_B=\begin{pmatrix}x_3\\x_5\end{pmatrix},X_N=\begin{pmatrix}x_1\\x_2\\x_4\end{pmatrix},B=\begin{pmatrix}1&0\\0&1\end{pmatrix},N=\begin{pmatrix}\frac{1}{2}&1&\frac{1}{2}\\\frac{5}{2}&3&\frac{-1}{2}\end{pmatrix},\begin{matrix}C_B=(3,1)\\C_N=(5,2,-1)\end{matrix},b=\begin{pmatrix}4\\3\end{pmatrix}$$

---

(3) The improvement of basic feasible solution $X=(0,0,4,0,3)^T$

① Choose the swap-in variable

Because $\sigma_1=1>0$, choose x_1 as the swap-in variable.

② Choose the swap-out variable

$$B^{-1}b=\begin{pmatrix}4\\3\end{pmatrix},B^{-1}P_1=\begin{pmatrix}\frac{1}{2}\\\frac{5}{2}\end{pmatrix}\not\leq 0 \text{ and } \min\left\{\frac{4}{1/2},\frac{3}{5/2}\right\}=\frac{3}{5/2}$$

choose x_5 as the swap-out variable.

$$\begin{pmatrix}x_3\\x_5\end{pmatrix}=B^{-1}b-B^{-1}P_1x_1=\begin{pmatrix}4\\3\end{pmatrix}-\begin{pmatrix}1/2\\5/2\end{pmatrix}x_1$$

$$X_B=\begin{pmatrix} x_3 \\ x_5 \end{pmatrix}, X_N=\begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & \frac{-1}{2} \end{pmatrix}, \begin{matrix} C_B=(3,1) \\ C_N=(5,2,-1) \end{matrix}, b=\begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

---

(4) Solve the improved basic feasible solution X''

Do elementary transformation to the augmented matrix of constraint equation set, and transform the corresponding coefficient column vector $P_1=\begin{pmatrix} \frac{1}{2} \\ \frac{5}{2} \end{pmatrix}$ of swap-in variable x_1 into the corresponding unit vector $P_5=\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ of swap-out variable x_5.

$$\begin{pmatrix} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & \frac{-1}{2} & 1 & 3 \end{pmatrix} \xrightarrow[\text{multiplied by 2/5}]{\text{The second line}} \begin{pmatrix} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ 1 & \frac{6}{5} & 0 & \frac{-1}{5} & \frac{2}{5} & \frac{6}{5} \end{pmatrix}$$

$$\xrightarrow[\text{the second line 1/2 times}]{\text{The first line minus}} \begin{pmatrix} 0 & \frac{2}{5} & 1 & \frac{3}{5} & \frac{-1}{5} & \frac{17}{5} \\ 1 & \frac{6}{5} & 0 & \frac{-1}{5} & \frac{2}{5} & \frac{6}{5} \end{pmatrix}$$

$$C=(5,2,3,-1,1) \quad C=(5,2,3,-1,1)$$

$$\left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & \frac{-1}{2} & 1 & 3 \end{array}\right) \Rightarrow \left(\begin{array}{ccccc|c} 0 & \frac{2}{5} & 1 & \frac{3}{5} & \frac{-1}{5} & \frac{17}{5} \\ 1 & \frac{6}{5} & 0 & \frac{-1}{5} & \frac{2}{5} & \frac{6}{5} \end{array}\right)$$

The improved basic feasible solution can be obtained.

$B=(P_3P_1)=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, the basic variable x_3, x_1, the nonbasic variable x_2, x_4, x_5,

$$X_B=\begin{pmatrix} x_3 \\ x_1 \end{pmatrix}, X_N=\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} \frac{2}{5} & \frac{3}{5} & \frac{-1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \end{pmatrix}, \begin{matrix} C_B=(3,5) \\ C_N=(2,-1,1) \end{matrix}, b=\begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix}$$

$$X_N=0 \rightarrow X_B=B^{-1}b=\begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix} \Rightarrow$$ basic feasible solution $X=(\frac{6}{5},0,\frac{17}{5},0,0)^T$

The value of objective function $Z=C_BB^{-1}b=(3,5)\begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix}=\frac{81}{5}$

increases compared $Z=15$, then turn to step (2).

$$X_B=\begin{pmatrix} x_3 \\ x_1 \end{pmatrix}, X_N=\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}, B=\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N=\begin{pmatrix} \frac{2}{5} & \frac{3}{5} & \frac{-1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \end{pmatrix}, \begin{matrix} C_B=(3,5) \\ C_N=(2,-1,1) \end{matrix}, b=\begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix}$$

---

(2) Examine whether $X''=(\frac{6}{5},0,\frac{17}{5},0,0)^T$ is the optimal solution.

Test vector

$$\sigma_N=C_N-C_BB^{-1}N=(2,-1,1)-(3,5)\begin{pmatrix} \frac{2}{5} & \frac{3}{5} & \frac{-1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \end{pmatrix}$$

$$=(2,-1,1)-(\frac{36}{5},\frac{4}{5},\frac{7}{5})=(\underset{\substack{\uparrow \\ \sigma_2}}{\frac{-26}{5}},\underset{\substack{\uparrow \\ \sigma_4}}{\frac{-9}{5}},\underset{\substack{\uparrow \\ \sigma_5}}{\frac{-2}{5}})$$

Because all of the check numbers are less than zero,

so $X^*=X''=(\frac{6}{5},0,\frac{17}{5},0,0)^T$ is the optimal solution. $Z^*=\frac{81}{5}$

2.2 Simplex Tableau

According to the composition of θ and $\hat{\mathbf{x}}$, the improved basic feasible solution which make the original nonbasic variable x_k become the basic variable choosing positive value, meanwhile, make the original basic variable x_r to be zero (become nonbasic variable).

Equally, basis $B = [\mathbf{a}_1, \dots, \mathbf{a}_m]$ becomes another basis

$\hat{B} = [\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{a}_k, \mathbf{a}_{r+1}, \dots, \mathbf{a}_m]$.

*Basis exchange transformation

*x_k is called entering basic variable, x_r is basifugal variable.

x_r and x_k exchange their status.

The kth column becomes unit vector $\mathbf{e}_r$ through elementary transformation.

Based on the basis B, the coefficient augmented matrix of canonical equation is

$$
\begin{array}{cccccccccc}
x_1 & \dots & x_r & \dots & x_m & x_{m+1} & \dots & x_k & \dots & x_n & \\
1 & \dots & 0 & \dots & 0 & \bar{a}_{1,m+1} & \dots & \bar{a}_{1,k} & \dots & \bar{a}_{1,n} & \overline{b_1} \\
\vdots & \ddots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \dots & 1 & \dots & 0 & \bar{a}_{r,m+1} & \dots & \bar{a}_{r,k} & \dots & \bar{a}_{r,n} & \overline{b_r} \\
\vdots & & \vdots & \ddots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\
0 & \dots & 0 & \dots & 1 & \bar{a}_{m,m+1} & \dots & \bar{a}_{m,k} & \dots & \bar{a}_{m,n} & \overline{b_m}
\end{array}
$$

1)The rth line $\overline{\mathbf{a}_r}^T$ divides $\bar{a}_{rk}$, $\hat{\mathbf{a}}_r{}^T = \overline{\mathbf{a}_r}^T / \bar{a}_{rk}$

2)To the ith line($i \neq r$) $\overline{\mathbf{a}_i}^T$, $\hat{\mathbf{a}}_r{}^T = \overline{\mathbf{a}_i}^T - \hat{\mathbf{a}}_r{}^T \bar{a}_{ik}$

The last line is the basic feasible solution corresponding to the basis $\hat{B}$

$\hat{b}_r = \bar{b}_r / \bar{a}_{rk}$, $\hat{b}_i = \bar{b}_i - (\bar{b}_r / \bar{a}_{rk}) \bar{a}_{ik}$, $i \neq r$

When change the basis, the objective function $z = \mathbf{c}_B \bar{\mathbf{b}} - (\mathbf{c}_B \bar{N} - \mathbf{c}_N)\mathbf{x}_N$ should make some adjustment, regard the equivalent form of the equation above

$$z + (\mathbf{c}_B B^{-1} N - \mathbf{c}_N)\mathbf{x}_N = \mathbf{c}_B \bar{\mathbf{b}}, \text{ or } z + \xi_N \mathbf{x}_N = z_0$$

as a equation, put it into the canonical equation and do elementary transformation together.

	z	x_1	$\cdots$	x_r	$\cdots$	x_m	x_{m+1}	$\cdots$	x_k	$\cdots$	x_n	RHS
z	1	0	$\cdots$	0	$\cdots$	0	ζ_{m+1}	$\cdots$	ζ_k	$\cdots$	ζ_n	z_0
x_1	0	1	$\cdots$	0	$\cdots$	0	$\bar{a}_{1,m+1}$	$\cdots$	$\bar{a}_{1,k}$	$\cdots$	$\bar{a}_{1,n}$	$\bar{b}_1$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_r	0	0	$\cdots$	1	$\cdots$	0	$\bar{a}_{r,m+1}$	$\cdots$	$\bar{a}_{r,k}$	$\cdots$	$\bar{a}_{r,n}$	$\bar{b}_r$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_m	0	0	$\cdots$	0	$\cdots$	1	$\bar{a}_{m,m+1}$	$\cdots$	$\bar{a}_{m,k}$	$\cdots$	$\bar{a}_{m,n}$	$\bar{b}_m$

Simplex Tableau

shaft column

shaft line

shaft unit

Simplex tableau is short for

单纯形表可简记为

	z	$\mathbf{x}_B$	$\mathbf{x}_N$	RHS
z	1	$\mathbf{0}$	$\boldsymbol{\zeta}_N$	z_0
$\mathbf{x}_B$	$\mathbf{0}$	I	$\bar{N}$	$\bar{\mathbf{b}}$

The above-mentioned transformation is called pivot, and is similar to the PCA elimination method of linear equations solving.

□Table simplex method

Through example 1, we find during the solving process of simplex method, there are following indicators:

- Every basic feasible solution's test vector $\sigma_N = C_N - C_B B^{-1} N$

According to the fact that test vector can determine whether the obtained basic feasible solution is the optimal solution. If the solution is not the optimal and can determine the appropriate swap-in variable through the test vector.

- Every basic feasible 's objective function $Z = C_B B^{-1} b$

By means of the objective function, we can observe if every iteration of simplex method can make the value of objective function increase effectively, until obtain the optimal objective function.

➢ During process of simplex method, every basic feasible solution **X** treats the identity matrix **I** of a certain constraint equations which gone through the elementary transformation as the feasible basis.

When $\mathbf{B} = I$, $B^{-1} = \mathbf{I}$, it is easy to know: $\sigma_N = C_N - C_B N \quad Z = C_B b$

Design a simple table as follow for the computation of these important conclusions, that is , simplex tableau:

<table>
<tr><td colspan="3">C</td><td>C_B</td><td>C_N</td><td rowspan="2">θ</td></tr>
<tr><td>C_B</td><td>X_B</td><td>b</td><td>$X_1 \quad X_2 \quad \cdots \quad X_m$</td><td>$X_{m+1} \quad X_{m+2} \quad \cdots \quad X_n$</td></tr>
<tr><td>C_1
C_2
.
.
C_m</td><td>X_1
X_2
.
X_m</td><td>b_1
b_2
.
.
b_m</td><td>I</td><td>N</td><td>θ_1
θ_2
.
.
θ_m</td></tr>
<tr><td colspan="2">Z</td><td>$C_B b$</td><td>0</td><td>$C_N - C_B N$</td><td></td></tr>
</table>

Example 2. Try to use simplex tableau to solve the optimal solution mentioned in Example 1:

$$\max Z = 5x_1 + 2x_2 + 3x_3 - x_4 + x_5 \qquad C=(5,2,3,-1,1)$$

$$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 8 \\ 3x_1 + 4x_2 + x_3 + x_5 = 7 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases} \qquad (Ab)=\begin{pmatrix} 1 & 2 & 2 & 1 & 0 & 8 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{pmatrix}$$

Get the initial simplex tableau:

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
-1	x_4	8	1	2	2	1	0	
1	x_5	7	3	4	1	0	1	
Z		-1	3	0	4	0	0	

$Z=C_B b \qquad \sigma_N = C_N - C_B N$

The initial basic feasible solution $X=(0,0,0,8,7)^T$, **Z= -1 ,**

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
-1	x_4	8	1	2	2	1	0	8/2
1	x_5	7	3	4	1	0	1	7/1
Z		-1	3	0	4	0	0	

x_3 is swap-in variable, x_4 is swap-out variable, 2 as pivot for pivot transformation

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	4	1/2	1	1	1/2	0	
1	x_5	3	5/2	3	0	-1/2	1	
Z		15	1	-4	0	-2	0	

Basic feasible solution $X=(0,0,4,0,3)^T$, **Z= 15 ,**

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	4	1/2	1	1	1/2	0	$4/_{1/2}$
1	x_5	3	5/2	3	0	-1/2	1	$3/_{5/2}$
Z		15	1	-4	0	-2	0	

x_1 is swap-in variable, x_5 is swap-out variable, 5/2 as pivot for pivot transformation

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	17/5	0	2/5	1	3/5	-1/5	
5	x_1	6/5	1	6/5	0	-1/5	2/5	
Z		81/5	0	-26/5	0	-9/5	-2/5	

$\sigma_N = C_N - C_B N \le 0$ optimal solution $X^* = \left(\frac{6}{5}, 0, \frac{17}{5}, 0, 0\right)^T$ optimal value $Z^* = \frac{81}{5}$

2.3 Initial Solution: Two-phase Method

Assume the primal problem is

$$\begin{aligned} &\min \mathbf{cx} \\ &s.t. \quad A\mathbf{x} = \mathbf{b} \ (\mathbf{b} \geq \mathbf{0}) \\ &\qquad\quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Introduce m artificial variables $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, consider the auxiliary problem as follows

$$\begin{aligned} &\min g = \mathbf{1x}_a \\ &s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \\ &\qquad\quad \mathbf{x}, \mathbf{x}_a \geq 0 \end{aligned}$$

Where $\mathbf{1} = (1, \dots, 1)$. Assume the feasible region of primal problem is D, the feasible region of auxiliary problem is D',

Obviously $\mathbf{x} \in D \Leftrightarrow \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D'$, and $\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0$.

2.3 Initial Solution: Two-phase Method

m artificial variables $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, consider the auxiliary problem as follows

$$\begin{aligned} \min\ g &= \mathbf{1}\mathbf{x}_a \\ s.t.\quad A\mathbf{x} + \mathbf{x}_a &= \mathrm{b} \\ \mathbf{x}, \mathbf{x}_a &\geq 0 \end{aligned}$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

To the auxiliary problem, a basic feasible solution is $\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix}$,

So we can get solution by simplex iteration, the result has two possibilities:

1)min g > 0, means $D = \emptyset$.

2)min g $= 0$, It naturally exists $\mathbf{x}_a = 0$, divide it and obtain a feasible solution of primal problem.

2.3 Initial Solution: Two-phase Method

m artificial variables $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, get the auxiliary problem as follows

$$
\begin{aligned}
\min g &= \mathbf{1}\mathbf{x}_a \\
s.t. \quad A\mathbf{x} + \mathbf{x}_a &= \mathbf{b} \\
\mathbf{x}, \mathbf{x}_a &\geq 0
\end{aligned}
$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

*If all of the artificial variables are nonbasic variables by this time, then the solution is basic feasible solution.

*Otherwise, eliminate the artificial variable of basic variable by pivot transformation.

Assume basic variable x_r is artificial variable,

- Select the nonzero element $\bar{a}_{rs}$ of the top n elements in the rth line as the axis unit to transform, then the unartificial variable x_s enter basic variable, eliminate the x_r at the same time.
- If all the top n elements in the rth line are zero, then eliminate this line and the corresponding artificial variable directly.

2.4 Initial Solution: Big M Method

Assume the primal problem is

$$
\begin{aligned}
&\min \mathbf{cx} \\
&s.t. \quad A\mathbf{x} = \mathbf{b} \ (\mathbf{b} \geq \mathbf{0}) \\
&\qquad\quad \mathbf{x} \geq \mathbf{0}
\end{aligned}
$$

Introduce $\mathbf{x}_a \in E^m$ and a big enough positive number M, get the new problem as follows

$$
\begin{aligned}
&\min \mathbf{cx} + M\mathbf{1}\mathbf{x}_a \\
&s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \\
&\qquad\quad \mathbf{x}, \mathbf{x}_a \geq 0
\end{aligned}
$$

Where $\mathbf{1} = (1, \dots, 1)$. Once M is big enough, then

x is the optimal solution of primal problem $\Leftrightarrow \begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix}$ is the optimal solution of new problem. And the new problem has initial solution $\begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix}$, so it always can be solved by simplex method.

2.5 Degeneration and Cycle Prevention

- If there is component is zero among the right-hand vector in simplex tableau, namely, in the case that degenerate solution appears, the cycle may appear. In order to prevent cycle, some pivot rules need to be supplemented.
  - Dictionary sequence method
  - Bland Rule

Dictionary sequence method

Nonzero vector **x**, and the first nonzero component is non-negative, then **x** is called non-negative based dictionary sequence denoted as $\mathtt{x} \succ = 0$. To vector **x, y** , If $\mathtt{x} - \mathtt{y} \succ = 0$, then **x** is called be equal or greater than **y** based dictionary sequence.

If to $\mathbf{x}^i$, exist $\mathbf{x}^t$ making $\mathbf{x}^i$ satisfy $\mathbf{x}^i \succ = \mathbf{x}^t$, then $\mathbf{x}^t$ is called the maximum based dictionary sequence in the this vector set.

Note: $\mathbf{x}^t = lex\ \mathrm{min}\ \mathbf{x}^i$

Dictionary sequence method make the following equation after selecting the entering basic variable x_k

$$\frac{\mathbf{p}_r}{\bar{a}_{rk}} = lex\ \mathrm{min}\left\{ \frac{\mathbf{p}_i}{\bar{a}_{ik}} \middle| \bar{a}_{ik} > 0, i = 1, \ldots, m \right\}$$

where $\mathbf{p}_i$ is the ith line of $(\bar{\mathbf{b}}, B^{-1})$.

Bland Rule

1)Choose the nonbasic variable x_k corresponding to the positive check number ξ_k with the minimum subscript as the entering basic variable.

2)The determination of basifugal variable x_l : If several $\frac{b_r}{\bar{a}_{rk}}$ reach minimum at the same time, choose the basic variable with the minimum subscript as the basifugal variable. That is

$$l = \min\left\{ r \left| \frac{b_r}{\bar{a}_{rk}} = \min\left\{ \frac{b_i}{\bar{a}_{ik}} \middle| \bar{a}_{ik} > 0 \right\} \right. \right\}$$

- Bland Rule's advantage is simple and practicable, but the disadvantage is only considering the minimum subscript rather than the decrease speed of objective function value. So the efficiency is lower than dictionary sequence method or simplex method.
- In reality, degeneration is common, but not always generate cycle. In fact, generating cycle is infrequent.

2.6 Modify Simplex Method

- When the scale of LP problem is quite large, how to reduce the memory space and computation time is a problem must be considered. In reality, modifying simplex method is mostly employed.

Inverse Matrix Method

The general simplex tableau is

$$\begin{array}{c} \\ z \\ \mathbf{x}_B \end{array}\begin{bmatrix} z & \mathbf{x}_B & \mathbf{x}_N & RHS \\ 1 & \mathbf{0} & \mathbf{c}_B B^{-1} N - \mathbf{c}_N & \mathbf{c}_B B^{-1}\mathbf{b} \\ \mathbf{0} & I & B^{-1}N & B^{-1}\mathbf{b} \end{bmatrix}$$

To every simplex tableau, inverse matrix method only store the following data

$$\begin{bmatrix} \mathbf{w} & z_o \\ B^{-1} & \bar{\mathbf{b}} \end{bmatrix} = \begin{bmatrix} \mathbf{c}_B B^{-1} & \mathbf{c}_B B^{-1}\mathbf{b} \\ B^{-1} & B^{-1}\mathbf{b} \end{bmatrix}$$

This table is called the modified simplex tableau.

Do the iterative computation according to this table and raw data. Determine x_k as the entering basic variable by $\xi_k = \max\{\mathbf{wa}_j - c_j | x_j \text{is nonbasic variable}\}$, then solve $\bar{\mathbf{a}}_k = B^{-1}\mathbf{a}_k$, determine the basifugal variable x_r by the minimum ratio to get new basis $\hat{B}$, finally construct the modified simplex tableau corresponding to $\hat{B}$.

If each iteration must construct B^{-1}, the calculated amount is unsatisfactory.

Theorem: Add $\begin{pmatrix} \xi_k \\ \bar{\mathbf{a}}_k \end{pmatrix}$ to the right of modified simplex tableau corresponding to basis B, regard $\bar{a}_{rk}$ as the spindle unit to spin to get basis $\hat{B}$'s corresponding modified simplex tableau.

3. Optimality Condition and Duality Theory

Duality

- For every linear programming P, another linear programming D always exists. There is a close connection between the both of them, people can acquire primal programming P’s optimal solution by means of solving dual program D.

3.1 Karush-Kuhn-Tucker Conditions

- Theorem

Assume A is the matrix of order $m \times n$, $\mathbf{b} \in E^m, \mathbf{c} \in E^n, \mathbf{x} \in E^n$

$$\begin{aligned} &\min \mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &s.t. \quad A\mathbf{x} \geq \mathbf{b} \\ &\qquad\quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$

$\mathbf{x}^*$ is the optimal of above LP problem iff there exists $\mathbf{w} \in E^m, \mathbf{v} \in E^n$ meeting the following *KKT* conditions

$$\begin{aligned} &A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0} \\ &\mathbf{c} - \mathbf{w}^{\mathrm{T}}A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0} \\ &\mathbf{w}^{\mathrm{T}}(A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^{\mathrm{T}}\mathbf{x}^* = 0 \end{aligned}$$

If eliminate $\mathbf{v}$ (use $\mathbf{c} - \mathbf{w}^{\mathrm{T}}A - \mathbf{v} = \mathbf{0}$) ,we can obtain another form of *KKT* conditions

$$\begin{aligned} &A\mathbf{x}^* \geq \mathbf{b}, \quad \mathbf{x}^* \geq \mathbf{0} \\ &\mathbf{w}A \leq \mathbf{c}, \quad \mathbf{w} \geq \mathbf{0} \\ &\mathbf{w}(A\mathbf{x}^* - \mathbf{b}) = 0, \quad (\mathbf{w}A - \mathbf{c})\mathbf{x}^* = 0 \end{aligned}$$

3.1 Karush-Kuhn-Tucker Conditions

- Proof:
- Sufficiency

$$A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0} \Longrightarrow \mathbf{x}^* \text{is feasible solution}$$

$$\mathbf{c} - \mathbf{w}^{\mathrm{T}} A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0} \Longrightarrow (\mathbf{c} - \mathbf{w}^{\mathrm{T}} A - \mathbf{v})^{\mathrm{T}} (\mathbf{x} - \mathbf{x}^*) = 0$$

$$\mathbf{w}^{\mathrm{T}} (A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^{\mathrm{T}} \mathbf{x}^* = 0$$

$$\Downarrow$$

$$(\mathbf{c} - \mathbf{w}^{\mathrm{T}} A - \mathbf{v})^{\mathrm{T}} (\mathbf{x} - \mathbf{x}^*) = \mathbf{c}^{\mathrm{T}} (\mathbf{x} - \mathbf{x}^*) - \mathbf{w}^{\mathrm{T}} (A\mathbf{x} - \mathbf{b}) - \mathbf{v}^{\mathrm{T}} \mathbf{x}$$

$$\Downarrow$$

$$A\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq \mathbf{0} \Longrightarrow \mathbf{c}^{\mathrm{T}} \mathbf{x} \geq \mathbf{c}^{\mathrm{T}} \mathbf{x}^*$$

$$\Downarrow$$

$$\mathbf{x}^* \text{is optimal solution}$$

3.1 Karush-Kuhn-Tucker Conditions

- Proof:
- Necessity

$\mathbf{x}^*$ is optimal solution $\Rightarrow$ $\mathbf{x}^*$ is feasible solution $\Rightarrow$ $\boxed{A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0}}$

$$\mathbf{c} - \mathbf{w}^{\mathrm{T}} A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0}$$

$$\mathbf{w}^{\mathrm{T}}(A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^{\mathrm{T}}\mathbf{x}^* = 0$$

$$I = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}, I_1\mathbf{x}^* > 0, I_2\mathbf{x}^* = 0$$

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} \Rightarrow A_1\mathbf{x}^* = \mathbf{b}_1, A_2\mathbf{x}^* > \mathbf{b}_2$$

$$B = \begin{pmatrix} A_1 \\ I_2 \end{pmatrix}, B\mathbf{d} \geq 0$$

sufficiently small positive number $\lambda > 0$

$$A(\mathbf{x} + \lambda\mathbf{d}) \geq \mathbf{b}, \mathbf{x} + \lambda\mathbf{d} \geq 0$$

$\mathbf{c}^{\mathrm{T}}\mathbf{x}^* \leq \mathbf{c}^{\mathrm{T}}(\mathbf{x}^* + \lambda\mathbf{d})$ $\quad \mathbf{c}^{\mathrm{T}}\mathbf{d} \geq 0$ $\quad$ Inequalities: $\mathbf{c}^{\mathrm{T}}\mathbf{d} < 0, \; B\mathbf{d} \geq 0,$ $\quad$ No solution !

Based *Farkas* lemma, there exists non-negative vector $\mathbf{r} \geq 0$, which makes

$$B^{\mathrm{T}}\mathbf{r} = (A_1^{\mathrm{T}}, I_1^{\mathrm{T}})\mathbf{r} = \mathbf{c}$$

3.1 Karush-Kuhn-Tucker Conditions

- Proof:
- Necessity

$\mathbf{x}^*$is optimal solution $\Rightarrow$ $\mathbf{x}^*$is feasible solution $\Rightarrow$ $\boxed{A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0}}$

Based *Farkas* lemma, there exists non-negative vector $\mathbf{r} \geq 0$, which makes

$$B^{\mathrm{T}}\mathbf{r} = (A_1^{\mathrm{T}}, I_1^{\mathrm{T}})\mathbf{r} = \mathbf{c}$$

$\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2)$ blocked mode is corresponding to B, easy to construct two non-negative vectors

$$\mathbf{w} = \begin{pmatrix} \mathbf{r}_1 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 0 \\ \mathbf{r}_2 \end{pmatrix},$$ lead to :

$$A^{\mathrm{T}}\mathbf{w} + \mathbf{v} = B^{\mathrm{T}}\mathbf{r} = \mathbf{c} \Rightarrow \mathbf{c} - \mathbf{w}^{\mathrm{T}}A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0}$$

Easy to verify $\mathbf{v}^{\mathrm{T}}\mathbf{x}^* = 0$ $\Rightarrow$ $\mathbf{w}^{\mathrm{T}}(A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^{\mathrm{T}}\mathbf{x}^* = 0$

$$A\mathbf{x}^* - \mathbf{b} = \begin{pmatrix} 0 \\ A_2\mathbf{x}^* - \mathbf{b}_2 \end{pmatrix}$$

3.1 Kuhn Tucker Conditions

For the standard form of *LP* problem's *K-T* conditions

As long as rewrite $A\mathbf{x} = \mathbf{b}$ as $\begin{pmatrix} A \\ -A \end{pmatrix}\mathbf{x} \geq \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \end{pmatrix}$, the corresponding *K-T* conditions would be introduced

$$A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$$

$$A^{\mathrm{T}}\mathbf{w} + \mathbf{v} = \mathbf{c},\ \mathbf{v} \geq \mathbf{0}$$

$$\mathbf{v}^{\mathrm{T}}\mathbf{x} = 0$$

or the equal form

$$A\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}$$

$$\mathbf{w}A \leq \mathbf{c}$$

$$(\mathbf{w}A - \mathbf{c})\mathbf{x} = 0.$$

3.1 Kuhn Tucker Conditions

A basic feasible solution $\begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix} (\bar{\mathbf{b}} > \mathbf{0})$ is the optimal iff the test vector $\zeta \leq 0$, and what's the relationship with K-T conditions?

According to $\mathbf{v}\mathbf{x} = \mathbf{v}_B\mathbf{x}_B + \mathbf{v}_N\mathbf{x}_N = 0$, and $\mathbf{x}_B > 0$, we can know $\mathbf{v}_B = 0$.

Condition $\mathbf{c} - \mathbf{w}A - \mathbf{v} = \mathbf{0}$ can be rewritten as

$(\mathbf{c}_B, \mathbf{c}_N) - \mathbf{w}(B, N) - (\mathbf{v}_B, \mathbf{v}_N) = (\mathbf{0}, \mathbf{0})$, so there is

$$\begin{cases} \mathbf{c}_B - \mathbf{w}B = \mathbf{0} \\ \mathbf{c}_N - \mathbf{w}N - \mathbf{v}_N = \mathbf{0} \end{cases} \Rightarrow \begin{cases} \mathbf{w} = \mathbf{c}_B B^{-1} \\ \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1} N \end{cases}$$

In other words, given a basic feasible solution X, if get the K-T conditions of $\mathbf{v}_B = 0, \mathbf{w} = \mathbf{c}_B B^{-1}, \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1} N$, except $\mathbf{v} \geq \mathbf{0}$, the other conditions are met.

Based the selection of *v*, easy to know $v \geq \mathbf{0} \Leftrightarrow \zeta \leq \mathbf{0}$.

3.2 Duality Theory

- Dual problem can be regarded as the "transpose of raw and column" of the primal problem.

$$\boxed{\min \quad \mathbf{c}^T\mathbf{x} \qquad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0}$$

$$\min \quad c^T x \qquad \text{s.t.} \quad \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix}, x \geq 0$$

- The dual problem is :

$$\max \quad \begin{bmatrix} b^T & -b^T \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \qquad \text{s.t.} \quad \begin{bmatrix} A^T & -A^T \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq c$$

- where y_1 and y_2 represent the dual variables set corresponding to the constraints Ax>=b and -Ax<=B respectively.

  Make y=y_1-y_2

$$\boxed{\max \quad \mathbf{b}^T\mathbf{y} \qquad \text{s.t.} \qquad \mathbf{A}^T\mathbf{y} \leq \mathbf{c}}$$

3.2 Duality Theory

Dual programming generative rules:

(1) min $\to$ max;

(2) $\mathbf{c} \to \mathbf{b}$; (3) $A \to A^T$;

(4) Put on the inequality sign according to following rules

Primal problem		Dual problem	
Variable	≥ 0	row constraint	$\leq$
	≤ 0		$\geq$
	Unlimited		$=$
row constraint	$\geq$	Variable	≥ 0
	$\leq$		≤ 0
	$=$		Unlimited

LP problem (P)

$$\begin{aligned} &\min \mathbf{cx} \\ &s.t. \quad A\mathbf{x} \geq \mathbf{b} \\ &\qquad\quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$

Dual problem (D)

$$\begin{aligned} &\max \mathbf{wb} \\ &s.t. \quad \mathbf{w}A \leq \mathbf{c} \\ &\qquad\quad \mathbf{w} \geq \mathbf{0} \end{aligned}$$

3.2 Duality Theory

Dual programming generative rules:

(1) $\min \rightarrow \max$;

(2) $\mathbf{c} \rightarrow \mathbf{b}$; (3) $A \rightarrow A^T$;

(4) Put on the inequality sign according to following rules

Primal problem		Dual problem	
Variable	≥ 0	row constraint	$\leq$
	≤ 0		$\geq$
	Unlimited		$=$
row constraint	$\geq$	Variable	≥ 0
	$\leq$		≤ 0
	$=$		Unlimited

Standard LP problem

$$\min \mathbf{cx}$$
$$s.t. \quad A\mathbf{x} = \mathbf{b}$$
$$\mathbf{x} \geq 0$$

Dual problem

$$\max \mathbf{wb}$$
$$s.t. \quad \mathbf{w}A \leq \mathbf{c}$$

3.2 Duality Theory

Dual programming generative rules:

(1) min $\to$ max;

(2) $\mathbf{c} \to \mathbf{b}$; (3) $A \to A^T$;

(4) Put on the inequality sign according to following rules

Primal problem		Dual problem	
Variable	≥ 0	row constraint	$\leq$
	≤ 0		$\geq$
	Unlimited		$=$
row constraint	$\geq$	Variable	≥ 0
	$\leq$		≤ 0
	$=$		Unlimited

LP problem

$$\min \mathbf{c}_1\mathbf{x}_1 + \mathbf{c}_2\mathbf{x}_2$$

$$\begin{aligned} s.\ t.\quad & A_{11}\mathbf{x}_1 + A_{12}\mathbf{x}_2 \geq \mathbf{b}_1 \\ & A_{21}\mathbf{x}_1 + A_{22}\mathbf{x}_2 = \mathbf{b}_2 \\ & A_{31}\mathbf{x}_1 + A_{32}\mathbf{x}_2 \leq \mathbf{b}_3 \\ & \mathbf{x}_1 \geq 0 \end{aligned}$$

Dual problem

$$\max \mathbf{w}_1\mathbf{b}_1 + \mathbf{w}_2\mathbf{b}_2 + \mathbf{w}_3\mathbf{b}_3$$

$$\begin{aligned} s.\ t.\quad & \mathbf{w}_1A_{11} + \mathbf{w}_2A_{21} + \mathbf{w}_3A_{31} \leq \mathbf{c}_1 \\ & \mathbf{w}_1A_{12} + \mathbf{w}_2A_{22} + \mathbf{w}_3A_{32} = \mathbf{c}_2 \\ & \mathbf{w}_1 \geq 0, \mathbf{w}_3 \leq 0 \end{aligned}$$

Duality Theory

- Assume **x** and **w** respectively are the feasible solution of (P) and (D), then $\mathbf{cx} \geq \mathbf{wb}$.
- Assume **x*** and **w*** respectively are the feasible solution of (P) and (D) , then **x*** and **w*** respectively are the optimal solution of (P) and (D) iff $\mathbf{cx}^*=\mathbf{w}^*\mathbf{b}$.
- Inference: If (P) has the optimal solution **x***, then (D) has the optimal solution **w***, and $\mathbf{cx}^*=\mathbf{w}^*\mathbf{b}$; If (P) is unbounded, (D) has no solution, and vice versa.
- Complementary slackness: Assume **x*** and **w*** respectively are the optimal solution of (P) and (D), then there are $\mathbf{w}^*(A\mathbf{x}^*-\mathbf{b})=\mathbf{0}$, $(\mathbf{c}-\mathbf{w}^*A)\mathbf{x}^*=\mathbf{0}$.

K-T conditions:

$A\mathbf{x}^* \geq \mathbf{b},\ \mathbf{x}^* \geq \mathbf{0}$	(original feasibility)
$\mathbf{w}A \leq \mathbf{c},\ \mathbf{w} \geq \mathbf{0}$	(dual feasibility)
$\mathbf{w}(A\mathbf{x}^* - \mathbf{b}) = 0,\ (\mathbf{w}A - \mathbf{c})\mathbf{x}^* = 0$	(complementary slackness)

Duality Theory

- It can be known from the relationship between the optimal criterions of *K-T* conditions and simplex method: In simplex method, all the *K-T* conditions are met except $\mathbf{v} \geq \mathbf{0}$, namely, simplex method gradually improves basic feasible solution under the condition of keeping original feasibility, complementary slackness and $\mathbf{c}-\mathbf{w}A-\mathbf{v}=\mathbf{0}$, which makes the $\mathbf{v}=\mathbf{c}-\mathbf{w}A\geq\mathbf{0}$ of dual feasibility be met.
- Inspiration: Improve the solution to meet the original feasibility under the condition of meeting dual feasibility and complementary slackness.

3.3 Dual Simplex Method

$$(P)\min \mathbf{cx} \qquad\qquad (D)\max \mathbf{wb}$$
$$s.t. A\mathbf{x} = b \qquad\qquad s.t.\ \ \mathbf{w}A \le \mathbf{c}$$
$$\mathbf{x} \ge 0$$

The basic feasible solution of D : Assume $A=(B, N)$,where B is nonsingular matrix, then $\mathbf{w}B = \mathbf{c}_B$'s solution $\overline{\mathbf{w}} = \mathbf{c}_B B^{-1}$ is the basic solution of (D), if $\overline{\mathbf{w}}N \le \mathbf{c}_N$, then $\overline{\mathbf{w}}$ is the basic feasible solution of (D).

The regular solution of P: If the corresponding test vector of the primal problem (P)'s feasible solution $\mathbf{x} = \begin{pmatrix} B^{-1}\mathbf{b} \\ 0 \end{pmatrix}$ is $\zeta = (\mathbf{0}, \mathbf{c}_B B^{-1}N - \mathbf{c}_N) \le \mathbf{0}$, then $\mathbf{x}$ is the regular solution of problem (P).

Now the basis B is the regular basis. Easy to verify, the basic feasible solution of D and the regular solution of P is one-to-one correspondence .

Same as simplex method, solving dual programming (D) is iterating from a basic feasible solution to another to increase the value of objective function. Equally, solving the primal programming (P) starts from a regular solution and iterates to another regular solution to increase the objective function $z = \mathbf{wb} = \mathbf{c}_B B^{-1}\mathbf{b} = \mathbf{cx}$, until $B^{-1}\mathbf{b} \ge 0$, which means the optimal solution would be found when the regular solution meets the original feasibility. This method is dual simplex method.

3.3 Dual Simplex Method

$$(P)\ \min \mathbf{cx} \qquad\qquad (D)\ \max \mathbf{wb}$$
$$s.t. A\mathbf{x} = b \qquad\qquad s.t.\ \ \mathbf{w}A \le \mathbf{c}$$
$$\mathbf{x} \ge 0$$

1. Find a regular basis B, and build the simplex tableau.
2. If $\bar{\mathbf{b}} = B^{-1}\mathbf{b} \ge 0$, stop, the optimal solution of primal problem has been found; otherwise, $\bar{b}_r = \min\{\bar{b}_i \mid i = 1, \ldots, m\}$.
3. If $\bar{\mathbf{a}}^r \ge 0$, stop, the primal problem has no solution; otherwise, $\frac{\zeta_k}{\bar{a}_{rk}} = \min\{\frac{\zeta_j}{\bar{a}_{rj}} \mid a_{rj} < 0\}$
4. Regard $\bar{a}_{rk}$ as the spindle unit to spin, return to 2.

3.4 Original—Dual Simplex Method

- Solve from both of the dual problem and the first phase of auxiliary problem.
- Begin with the arbitrary feasible solution of dual problem, meanwhile, make $Ax=b$ by getting rid of artificial variable in the case of keeping dual feasibility, complementary slackness and $x \geq 0$ during the iterative process.

$$(P)\ \min \mathbf{cx} \qquad\qquad (D)\ \max \mathbf{wb}$$

$$s.t.\ A\mathbf{x} = b \qquad\qquad s.t.\ \mathbf{w}A \le \mathbf{c}$$

$$\mathbf{x} \ge 0$$

Introduce artificial variable to (P) to get the auxiliary problem:

$$\min g = \mathbf{1x}_a$$

$$s.t\quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \qquad (\mathbf{b} \ge 0)$$

$$\mathbf{x}, \mathbf{x}_a \ge \mathbf{0}$$

If one of (D)'s feasible solutions $\overline{\mathbf{w}}$ is known, for keeping complementary slackness $(\mathbf{x}(\overline{\mathbf{w}}A - \mathbf{c}) = \mathbf{0})$, we make $x_j = 0$ (when $\overline{\mathbf{w}}\mathbf{a}_j \ne c_j$), so the auxiliary problem is

$$(P')\ \min g = \mathbf{1x}_a$$

$$s.t\quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$x_j = 0,\ j \notin Q$$

$$x_j \ge 0,\ j \in Q$$

$$\mathbf{x}_a \ge \mathbf{0}$$

where $Q = \{j \mid \overline{\mathbf{w}}\mathbf{a}_j = c_j, j = 1, \ldots, n\}$. Problem (P') is the limited problem corresponding to $\overline{\mathbf{w}}$. Solve (P') and obtain the optimal solution $\begin{pmatrix}\mathbf{x}^* \\ \mathbf{x}_a^*\end{pmatrix}$ which is the basic feasible solution of (P') 's auxiliary problem.

If $\mathbf{x}_a^* = \mathbf{0}$, then $\mathbf{x}^*$ is the optimal solution of (P). (Because $\mathbf{x}^*$ and $\overline{\mathbf{w}}$ are the feasible solutions of (P) and (D), respectively, and both of them meet complementary slackness).

If $\mathbf{x}_a^* \ne \mathbf{0}$, find another basic feasible solution $\hat{\mathbf{w}}$ of (D) and increase the value of (D)'s objective function. Meanwhile, the optimal value of the limited problem corresponding to $\hat{\mathbf{w}}$ decreases compared with that corresponding to $\overline{\mathbf{w}}$.

Consider the dual problem (D') of (P')

$$(D')\ \max\ \mathbf{vb}$$
$$s.t.\quad \mathbf{va}_j \le 0,\ j \in Q$$
$$\mathbf{v} \le 1$$

Regard $\mathbf{v}^*$ as the optimal solution. If for all the $j = 1, \ldots, n$, there is $\mathbf{v}^*\mathbf{a}_j \le \mathbf{0}$, which means $\mathbf{v}^*$ is the optimal solution of dual problem of auxiliary problem, so $\begin{pmatrix}\mathbf{x}^*\\ \mathbf{x}_a^*\end{pmatrix}$ is the optimal of auxiliary problem. Since $\mathbf{x}_a^* \ne 0$, the primal problem (P) has no solution.

Construct $\hat{\mathbf{w}} = \overline{\mathbf{w}} + \theta\mathbf{v}^*$, where

$$\theta = \min\left\{-\frac{\overline{\mathbf{w}}\mathbf{a}_j - c_j}{\mathbf{v}^*\mathbf{a}_j} \,\middle|\, \mathbf{v}^*\mathbf{a}_j > 0,\ j = 1, \ldots, n\right\}$$

it can be proved that $\hat{\mathbf{w}}$ meet demand. Meanwhile, $\begin{pmatrix}\mathbf{x}^*\\ \mathbf{x}_a^*\end{pmatrix}$ is a basic feasible solution of the limited problem corresponding to $\hat{\mathbf{w}}$, so we can use it as the initial value to solve the limited problem. Repeat this.

3.4 Original—Dual Simplex Method

1. Change the programming being solved into the form of

$$\begin{aligned} &\min \mathbf{cx} \\ &s.t. \quad A\mathbf{x} = b \\ &\qquad\quad \mathbf{x} \geq 0 \end{aligned}$$

Find initial solution $\mathbf{w}$ of it's dual programming, which meets $\mathbf{w}A \leq \mathbf{c}$.(when $\mathbf{c} \geq 0$, take $\mathbf{w}$=0 as the initial solution)

2. Make $Q = \{j \mid \overline{\mathbf{w}}\mathbf{a}_j = c_j, j = 1, \ldots, n\}$, solve the limited problem of

$$\begin{aligned} &\min g = \mathbf{1}\mathbf{x}_a \\ &s.t \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \\ &\qquad x_j = 0, j \notin Q \\ &\qquad x_j \geq 0, j \in Q \\ &\qquad \mathbf{x}_a \geq \mathbf{0} \end{aligned}$$

If it's optimal solution $g = 0$, stop and find the optimal solution of primal problem; otherwise, solve it's dual problem and make the optimal solution be $\mathbf{v}$.

3. If $\mathbf{v}A \leq \mathbf{0}$, then stop and the primal problem has no solution; otherwise

$\theta = \min\left\{-\dfrac{\overline{\mathbf{w}}\mathbf{a}_j - c_j}{\mathbf{v}^*\mathbf{a}_j} \,\middle|\, \mathbf{v}^*\mathbf{a}_j > 0, j = 1, \ldots, n\right\}$ and construct $\mathbf{w} = \mathbf{w} + \theta\mathbf{v}^*$, return to 2.

3.5 Dual Initial Solution

- Both of dual simplex method and original-dual simplex method need a dual feasible solution. And to dual simplex method, the solution also should be the basic feasible solution.

For an arbitrary LP problem, a basis B always can be found by gauss elimination, and transformed as a canonical equation. If $\zeta_N \le 0$, it means a basic feasible solution $w = c_B B^{-1}$ of the dual problem has been found. Otherwise, add a constraint $\mathbf{1x}_N \le M$, where M is a sufficiently big parameter (thus the constraint wouldn't effect the primal problem).

After adding the constraint, the problem is called extended problem. (It can be proved: primal problem has feasible solution $\Longleftrightarrow$ extended problem has feasible solution towards sufficiently big M).

	z	$\mathbf{x}_B$	$\mathbf{x}_N$	x_{n+1}	RHS
z	1	0	ζ_N	0	z_0
$\mathbf{x}_B$	$\mathbf{0}$	I	$\overline{N}$	$\mathbf{0}$	$\overline{\mathbf{b}}$
x_{n+1}	0	$\mathbf{0}$	$\mathbf{1}$	1	M

Assume $\zeta_k = \max\{\zeta_i\}$, treat $\bar{a}_{m+1,k}$ as the spindle to spin and get new simplex tableau. Since $\zeta \le 0$, the new simplex tableau corresponds to a regular solution. So begin to solve by dual simplex method or original-dual simplex method.

The result contains two possibilities:

1. Extended problem has no solution $\Longleftrightarrow$ The primal problem has no solution.

2. Extended problem has the optimal solution $\bar{\mathbf{x}} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M$,the optimal value is $z_0 = z_1 + z_2 M$. Now the primal problem must has feasible $\hat{\mathbf{x}}$ and $z_1 + z_2 M \le \mathbf{c}\hat{\mathbf{x}}$.

Because the M is sufficiently big, so $z_2 \le 0$ (otherwise $\mathbf{c}\hat{\mathbf{x}}$ will be arbitrarily sufficiently big)

(i) $z_2 < 0$ .when $M \to \infty$, $z_0 \to -\infty$ .The problem is unbounded.

(ii) $z_2 = 0$. $z_0 = z_1$ is the optimal value of primal problem, $\bar{\mathbf{x}} \ge 0$ is the optimal solution of primal problem. If $\bar{\mathbf{x}}_2 = 0$, $\bar{\mathbf{X}}$ is a basic feasible solution. Otherwise($\bar{\mathbf{x}}_2 \ne 0$) , make $M_o = \min\{M \mid \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M \ge 0\}$, now $\bar{\mathbf{x}}_0 = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M_0$ also is the basic feasible solution. And $\{\mathbf{x} \mid \mathbf{x} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M,\, M \ge M_0\}$ presents a half line interface of the feasible region, arbitrary point in it is the optimal solution.

SVM and Dual Method

- http://www.cad.zju.edu.cn/home/zhx/csmath/lib/exe/fetch.php?media=svm_cjlin_dm.pdf

Algorithm Complexity

- The worst performance
  - The performance of the algorithm in the worst case.
- Average performance
  - The expected performance of the algorithm in every possible case.
- Problem:
  - The mathematical problem need to solve usually contains a set of parameters and unknown free variables, and describes them:
    - The description of all the parameters
    - The description of the characters which need to be met by the solution.
- Instances
  - The problem instances which given the parameters and description specifically
- Algorithm
  - A set of univocal simple commands, which could be followed to solve a problem.
  - The solvability of a problem indicates that existing a algorithm which can solve any instances of this problem in the limited time and memory space. “Halting problem” is can’t be solved by algorithm.

Algorithm Complexity

- Measure standard : Number of elementary operation required at run time of an algorithm.
- Input scale : Encoding length of a instance
- Algorithm complexity function : $f: N \to R^+$
  - Polynomial-time algorithm, the increment speed is the polynomial function of the scale
    - $f(k) = O(k^3)$
  - Exponential-time algorithm, the increment speed is the exponential function of the scale
    - $f(k) = O(2^k)$

Complexity of Simplex Method

- If it isn't polynomial time, the worst case is exponential time!

$$\max \quad c = \sum_{i=1}^{n} w_i$$

$$s.t. \quad w_i + 2\sum_{j<i} w_i \leq \theta^{\text{i-1}}$$

$$w_i \geq 0, i = 1, \ldots, n$$

- According to MRT rule, the iteration needs $2^n - 1$ times.

Karmarkar Projection Dimension Algorithm

- Two basic facts:
  - If a inner point is located in the center of polytope, along the steepest descent direction of the objective function is a good direction
  - In the condition of keeping the basic characteristics, there exists an appropriate transformation which can place the given inner point in the feasible region in the center transformed.
- Basic idea: inner point → projection scale transform→ center → move inverse transformation → new inner point → projection scale transform → center……

Karmarkar Projection Dimension Algorithm

Consider a Linear Programming problem in matrix form:

maximize $c^T x$

subject to $Ax \leq$ b.

```
Algorithm Affine-Scaling
  Input:  A, b, c, x^0, stopping criterion, γ.
```

$k \leftarrow 0$

do while *stopping criterion* not satisfied

$v^k \leftarrow b - Ax^k$

$D_v \leftarrow \operatorname{diag}(v_1^k, \ldots, v_m^k)$

$h_x \leftarrow (A^T D_v^{-2} A)^{-1} c$

$h_v \leftarrow -A h_x$

if $h_v \geq 0$ then

return unbounded

end if

$\alpha \leftarrow \gamma \cdot \min\{-v_i^k/(h_v)_i \mid (h_v)_i < 0,\ i = 1, \ldots, m\}$

$x^{k+1} \leftarrow x^k + \alpha h_x$

$k \leftarrow k + 1$

end do