OR

S.Olivia

September 2023

目录

Chapter 1

Introduction

1.1 数学基础

1.1.1 梯度与黑塞矩阵

简便记法

1.1.1.1 梯度

$$\nabla f(x_1, x_2, x_3) = (\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_3})^T \tag{1.1}$$

1.1.1.2 黑塞矩阵

$$\nabla^2 f(x_1, x_2, x_3) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_1} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_3 \partial x_2} \\ \frac{\partial^2 f}{\partial x_1 \partial x_3} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \frac{\partial^2 f}{\partial x_3^2} \end{pmatrix} \tag{1.2}$$

简而言之，式 (??) 即为式 (??) 逐行关于参数（即 $x_1, x_2, x_3, \cdots$）求偏导。

1.1.2 正定矩阵

对一个 n 阶方阵 $\boldsymbol{M}$，如果其对任意非零向量 $\boldsymbol{Z}$ 都有 $\boldsymbol{Z}^{\boldsymbol{T}}\boldsymbol{M}\boldsymbol{Z} > 0$，其中 $\boldsymbol{Z}^{\boldsymbol{T}}$ 为 $\boldsymbol{Z}$ 的转置，则 $\boldsymbol{M}$ 为正定矩阵。
定理：对于 n 阶实对称矩阵 $\boldsymbol{M}$，下列条件是等价的：

1. $\boldsymbol{M}$ 是正定矩阵。
2. $\boldsymbol{M}$ 的特征值均为正。
3. $\boldsymbol{M}$ 的一切顺序主子式均为正。

1.2 基本理论

1.2.1 数学规划模型的一般形式

$$(fS) = \begin{cases} min & f(x) \\ s.t. & x \in S \end{cases} \tag{1.3}$$

1.2.2 凸集、凸函数和凸规划

1.2.2.1 凸集

定义：设 $S \subset \boldsymbol{R}^n$，如果 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S, \lambda \in [0,1]$，均有

$$\boldsymbol{\lambda x}^{(1)} + (1 - \boldsymbol{\lambda})\boldsymbol{x}^{(2)} \in S \tag{1.4}$$

则称 S 为凸集

集合中任意两点连成的线段必属于该集合；规定空集 $\varnothing$ 为凸集，单点集 $\{x\}$ 为凸集。

性质：

1. 凸集的交集是凸集
2. 凸集的内点集是凸集
3. 凸集的闭包是凸集
4. 分离和支撑：凸集边界上任意点存在支撑超平面；两个相互不交的凸集之间存在分离超平面

有一特殊的凸集：凸锥

定义：设非空集合 $C \subset \boldsymbol{R}^n$，如果 $\forall \boldsymbol{x} \in C$ 对 $\forall \lambda > 0$ 有 $\boldsymbol{\lambda x} \in \boldsymbol{C}$，则称 C 为以 0 为顶点的锥（不一定含 0 点）。若 C 又是凸集，则称 C 为凸锥。

1.2.2.2 凸函数

定义 2-7：设 $S \subset \boldsymbol{R}^n$，非空，凸集，函数 $f : S \to \boldsymbol{R}$，如果对 $\forall x^{(1)}, x^{(2)} \in S, \forall \lambda \in (0,1)$ 恒有

$$f(\boldsymbol{\lambda x^{(1)} + (1 - \lambda)x^{(2)}) \leq \lambda f(x^{(1)}) + (1 - \lambda)f(x^{(2)})}) \tag{1.5}$$

则称 f 为 S 上的凸函数。如果式 (??) 恒以严格不等式成立，则称 f 为 S 上的严格凸函数。

几何意义：任意两个点连线在函数曲线的上方。

1.2.2.2.1 水平集

定义 2-8：设 $S \subset \boldsymbol{R}^n$，非空，$f : S \to \boldsymbol{R}, a \in \boldsymbol{R}$, 则称

$$S_\alpha = \{\boldsymbol{x} | f(\boldsymbol{x}) \leq \alpha, \boldsymbol{x} \in S\} \tag{1.6}$$

为 f 的水平集。

水平集的概念相当于在地形图中，海拔高度不高于某一数值的区域。

注意：容易证明，当 f 为凸函数时，$\forall \alpha \in \boldsymbol{R}, S_\alpha$ 是凸集。但是它的逆不成立。

1.2.2.2.2 凸函数的性质

1. 定理 $f(x)$ 为凸集 S 上的凸函数 $\Leftrightarrow S$ 上任意有限点的凸组合的函数值不大于各点函数值的凸组合
2. 设 f_1, f_2 为凸函数，$\lambda_1, \lambda_2 > 0$，则：
   - $\lambda_1 f_1 + \lambda_2 f_2$ 是凸函数
   - $f(x) = max\{f_1(x), f_2(x)\}$ 是凸函数
   - $g(x) = min\{f_1(x), f_2(x)\}$ 不一定是凸函数
3. 若 f 在 S 上凸，那么 f 在 S 的内点集（$intS$）上连续（注：f 在 S 上不一定连续）
4. 若 f 在非空凸集 S 上凸，则对任意方向的方向导数存在

   凸函数常用判定条件：
5. 设 S 非空，凸集，开集，f 在 S 上可微，则：
   - f 在 S 上凸 $\rightleftharpoons \forall \bar{x} \in S$，有 $f(x) \geq f(\bar{x}) + \nabla f^T(\bar{x})(x - \bar{x}), \forall x \in S$。
   - f 在 S 上严格凸 $\rightleftharpoons \forall \bar{x} \in S$，有 $f(x) > f(\bar{x}) + \nabla f^T(\bar{x})(x - \bar{x}), \forall x \in S, x \neq \bar{x}$。
6. 设 S 非空，凸集，开集，f 在 S 上可微，则：
   - f 在 S 上凸 $\rightleftharpoons \forall x^{(1)}, x^{(2)} \in S, (\nabla f(x^{(1)}) - \nabla f(x^{(2)}))^T(x^{(1)} - x^{(2)}) \geq 0$
   - f 在 S 上严格凸 $\rightleftharpoons \forall x^{(1)}, x^{(2)} \in S, x^{(1)} \neq x^{(2)}, (\nabla f(x^{(1)}) - \nabla f(x^{(2)}))^T(x^{(1)} - x^{(2)}) > 0$
7. 设 S 是开集，f 在 S 上二次可微，则：
   - f 在 S 上凸 $\rightleftharpoons \forall x \in S, \nabla^2 f(x)$ 半正定。
   - 如果 $\forall x \in S, \nabla^2 f(x)$ 正定，则 f 在 S 上严格凸。

但是，上课讲的时候，一般是使用**黑塞矩阵是否正定**判断

1. 当 H 为半正定时，f 为凸函数；若 H 是正半定的，当且仅当 H 的每一个主子式都大于等于 0
2. 当 H 为正定时，f 为严格凸函数；若 H 是正定矩阵，当且仅当 H 的 n 个顺序主子式 (严格) 为正。
3. 当 H 为半负定时，f 为凹函数；若 H 是负半定的，当且仅当 H 的每一个奇阶的主子式小于等于 0，每一个偶阶的主子式大于等于 0
4. 当 H 为负定时，f 为严格凹函数；若 H 是负定矩阵，当且仅当 H 的 N 个顺序主子式以如下方式交替出现：$|A_1| < 0, |A_2| > 0, |A_3| < 0, \cdots$
5. 当 H 为不定时，f 既非凸也非凹函数。

(! 注意：只能一个方向判断，不能由是否凸函数来判断矩阵是否正定）

具体可看：数学选读 01：矩阵的主子式与顺序主子式

1.2.2.3 凸规划

定义：

- 当 (fS) 中，S 为凸集，f 是 S 上的凸函数（求 min 时），称 (fS) 为凸规划。
- 对于 (fgh)，当 f, g_i 为凸函数，h_i 为线性函数时，(fgh) 为凸规划。

定理 2-4：设 $S \subset \boldsymbol{R}^n$，非空，凸，$f : S \to \boldsymbol{R}$ 是凸函数。x^* 为问题 (fS) 的 l.opt.，则 x^* 为 g.opt.；又如果 f 是严格凸函数，那么 x^* 是问题 (fS) 的唯一 g.opt.。

1.2.3 多面体、极点、极方向

（看书上 p29 开始的图更好理解）

多面体：有限个半闭空间的交为多面体

极点：$x \in S$，不存在 S 中另外两个点 $x^{(1)}$ 和 $x^{(2)}$，及 $\lambda \in (0, 1)$，使 $x = \lambda x^{(1)} + (1 - \lambda)x^{(2)}$ 根据定义，闭球体的表面上每一点都是极点；一般的闭凸锥有唯一极点，即顶点；平面没有极点；

极方向：

- 方向：$d \in \boldsymbol{R}^n, d \neq 0$ 及对于任意 $x \in S, \lambda > 0$，总有 $x + \lambda d \in S$（可行方向）。其中，当 $d^{(1)} = \lambda d^{(2)} (\lambda > 0)$ 时，称 $d^{(1)}$ 和 $d^{(2)}$ 同方向。
- 极方向：方向 d 不能表示为两个不同方向的组合 $(d = d^{(1)} + d^{(2)})$。

1.2.3.1 极点特征定理

设 A 行满秩，x 是 S 极点的充要条件是：

存在分解 $A = (B, N)$, 其中 B 为 m 阶非奇异矩阵, 使 $x^T = (x_B^T, x_N^T)$, 这里 $x_B = B^{-1}b \geq 0, x_N = 0$

S 中必定存在有限多个极点 $(\leq C_n^m)$

Chapter 2

Linear Programming Problem

2.1 Graphical Method

图解法

具体可看 ppt 或：【运筹学】线性规划图解法 (唯一最优解 | 无穷最优解 | 无界解 | 无可行解)

基本解：各个等式约束直线的交点，外加与坐标轴的交点

基本可行解：基本解里面在可行域范围的那些基本解，可行域的顶点

最优解：基本可行解里面使目标函数最大（最小）的基本可行解

2.2 Simplex Method

一些必要概念：(建议先了解解题步骤，再回过头来看概念)

线性规划的标准形式

$$(LP)\begin{cases} max & \boldsymbol{c^T x} \\ s.t. & \boldsymbol{Ax = b} \\ & \boldsymbol{x \geq 0} \end{cases} \tag{2.1}$$

其中，$\boldsymbol{A}$ 是 $m \times n$ 的矩阵 $(m < n)$

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & \cdots & a_{1m} & a_{1(m+1)} & \cdots & a_{1n} \\ \vdots & & \vdots & \vdots & & \vdots \\ a_{m1} & \cdots & a_{mm} & a_{m(m+1)} & \cdots & a_{mn} \end{pmatrix}$$

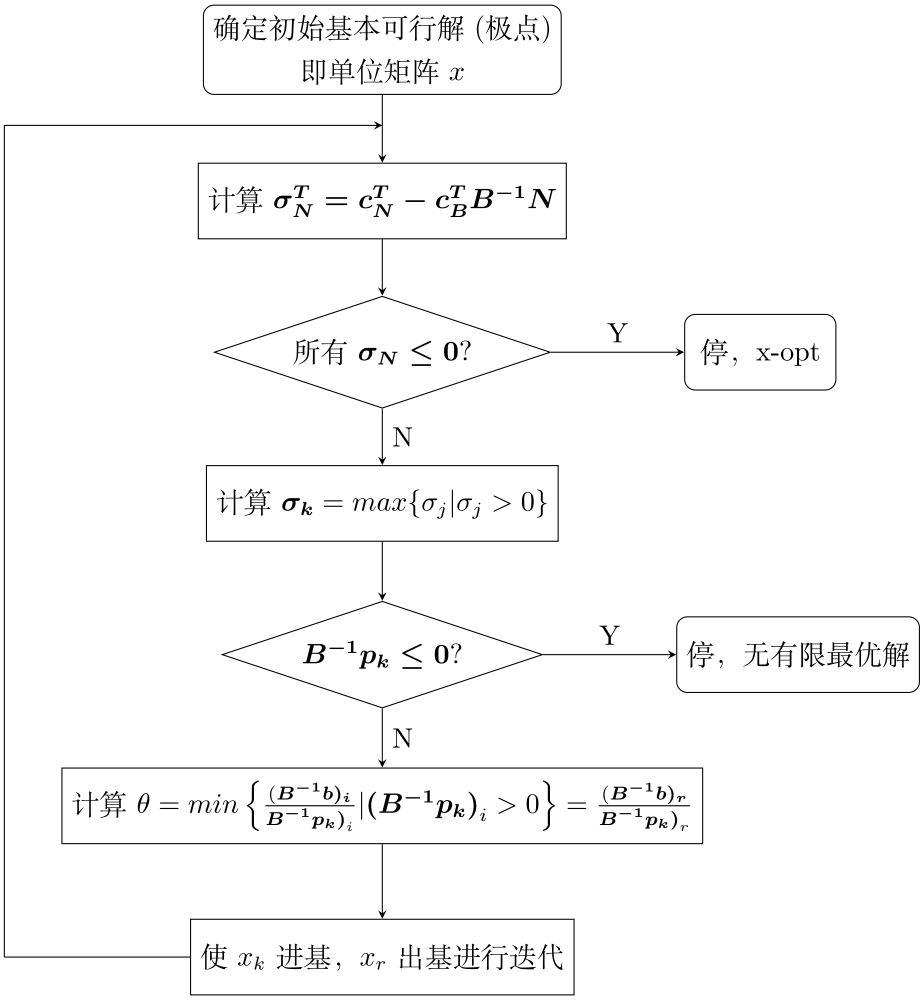
可行解：在公式 (??) 中，$\boldsymbol{x} = (x_1, \cdots, x_n)^T$ 即为**可行解**，可以理解为多面体的极点（顶点?）

基本可行解：可以理解为满足非负约束条件的可行解?（非负约束条件大概就是公式 (??) 中 $\boldsymbol{x} \geq \boldsymbol{0}$）

基：$\boldsymbol{B}$ 是线性规划问题的一个基

$$\boldsymbol{B} = \begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mm} \end{pmatrix}$$

$\boldsymbol{B}$ 中的每个列向量 $\boldsymbol{p}_j = (a_{1j}, a_{2j}, \cdots, a_{mj})^T$ 为**基向量**，基向量中的 a_{mj} 们即为**基变量**。一般的，我们讲基变量为 m 个线性无关的变量。(一般是标准化后的松弛变量所对应的列们的系数，也是**单位矩阵**）



需要注意的是，σ 在表中的计算没有那么复杂，可以先略过。

2.2.1 单纯形法的表格计算

考虑规范形式的线性规划问题：$b_i > 0, i = 1, \cdots, m$

$$
\begin{aligned}
max \quad & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\
s.t. \quad & \begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & \leq b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & \leq b_2 \\
\quad \vdots & \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & \leq b_m \\
x_1, x_2, \cdots, x_n \geq 0 &
\end{cases}
\end{aligned}
$$

加入松弛变量，化为标准形：

$$\begin{cases} max & z = c_1x_1 + c_2x2 + \cdots + c_nx_n \\ s.t. & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + x_{n+1} = b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \quad + x_{n+2} = b_2 \\ & \quad \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \qquad + x_{n+m} = b_m \\ & x_1, x_2, \cdots, x_n, x_{n+1}, \cdots, x_{n+m} \geq 0 \end{cases} \tag{2.2}$$

STEPS：

1. 根据公式 (??) 构造初始单纯形表：

$\boldsymbol{c_B}$	$\boldsymbol{x_B}$	$\boldsymbol{b}$	c_1	c_2	$\cdots$	c_n	c_{n+1}	c_{n+2}	$\cdots$	c_{n+m}	$\boldsymbol{\theta}$
			x_1	x_2	$\cdots$	x_n	x_{n+1}	x_{n+2}	$\cdots$	x_{n+m}	
c_{n+1}	x_{n+1}	b_1	a_{11}	a_{12}	$\cdots$	a_{1n}	1	0	$\cdots$	0	θ_1
c_{n+2}	x_{n+2}	b_2	a_{21}	a_{22}	$\cdots$	a_{2n}	0	1	$\cdots$	0	θ_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
c_{n+m}	x_{n+m}	b_m	a_{m1}	a_{m2}	$\cdots$	a_{mn}	0	0	$\cdots$	1	θ_m
$-z$		$-z'$	σ_1	σ_2	$\cdots$	σ_n	0	0	$\cdots$	0	

表 2.1: 单纯形表

表 (??) 中；

- 列 $\boldsymbol{c_B}$ 填入目标函数中基变量 x 的系数
- 列 $\boldsymbol{x_B}$ 填入基变量
- 列 $\boldsymbol{b}$ 填入约束方程右端的常数
- 中间 4-11 列填入约束方程中 x 的系数，其中从 x_{n+1} 开始的列组成单位矩阵

2. 求出 $-z' = -\sum_{i=1}^{m} c_{n+i}b_i$(初始状态下一般为 0)

3. 求出**检验数** $\boldsymbol{\sigma}_j = c_j - \sum_{i=1}^{m} c_{n+i}a_{ij}$，可以通俗理解为 $\sigma_1 =$ 第一行中的 c_1 减去（c_B 的每一项与 c_1 列对应行的 a_1 相乘的求和）。
需要注意的是，这里的检验数不需要考虑基（即单位矩阵）对应的列，否则会影响下一步。

4. 判断是否所有检验数 $\sigma_j \leq 0$，如果全部小于等于 0，则当前的基本可行解是最优解；如果有一检验数大于 0，那么进行下一步的计算。

5. 求入基变量 x_k：

- 先求出大于 0 的检验数中最大的**检验数** $\boldsymbol{\sigma_k}$
- 得到的下标 k，$\boldsymbol{\sigma_k}$ 所在的列就是主元列

- 那么所对应的 x_k 就是入基变量

6. 求出基变量 x_r：

   - 找出主元列（x_k 列）对应元素 $a_{ik} > 0$ 的
   - 使每个大于 0 的 a_{ik} 被对应行中的 b_i 除，得到 $\theta_i = \frac{b_i}{a_{ik}}$
     （若 $a_{ik} \leq 0$，则 $\boldsymbol{\theta_i} = \infty$）
   - 找出 $min\{\boldsymbol{\theta_i}\} = \boldsymbol{\theta_r}$
   - 此时 $\boldsymbol{\theta_r}$ 所在的行 r 即为主元行，x_r 为出基变量

7. 求出入基变量 x_k 和出基变量 x_r 后，就可以构建下一张表，这张表中：

   - $\boldsymbol{x_B}$ 上的出基变量被入基变量所取代，同时需要更改列 $\boldsymbol{c_B}$ 中与之相对应的行的元素
   - 在表右侧中间部分，使用高斯消元法将 x_k 对应的 a 与其他基变量对应的 a 组成一个单位矩阵，同时列 $\boldsymbol{b}$ 也跟着改变
   - 表 2 基本构建完成，重复第二步，直至出现流程表中 stop 的情况

2.2.2 一般线性规划问题的处理

在某些情况下，如基本初始可行解不明显，即很难在标准形的问题下找到单位矩阵时，可以考虑使用**大 M 法**与**二阶段法**

2.2.2.1 大 M 法

STEPS:

1. 将线性规划问题转化为标准型

2. 观察变量，若初始基本可行解明显，直接进行单纯形法；否则引入人工变量 $x_{n+i} \geq 0 (i = 1, ..., m)$ 及**充分大正数** M，改写原目标函数，进行单纯形法

3. 若得到的最优解满足:

$$x_{n+i} = 0 \ (i = 1, ..., m)$$

   则是原问题的最优解；否则，原问题无可行解

例：使用大 M 法求解下面的问题

$$max \quad z = 5x_1 + 2x_2 + 3x_3 - x_4$$

$$s.t. \quad \begin{cases} x_1 + 2x_2 + 3x_3 & = 15 \\ 2x_1 + x_2 + 5x_3 & = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 & = 26 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}$$

标准化并引入人工变量:

$$
\begin{aligned}
max \quad & z = 5x_1 + 2x_2 + 3x_3 - x_4 - Mx_5 - Mx_6 \\
s.t. \quad & \begin{cases} x_1 + 2x_2 + 3x_3 + x_5 = 15 \\ 2x_1 + x_2 + 5x_3 + x_6 = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 = 26 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{cases}
\end{aligned}
$$

最后使用单纯形法计算

c_B	x_B	b	5	2	3	-1	$-M$	$-M$	θ
			x_1	x_2	x_3	x_4	x_5	x_6	
$-M$	x_5	15	1	2	3	0	1	0	5
$-M$	x_6	20	2	1	[5]	0	0	1	4
-1	x_4	26	1	2	4	1	0	0	6.5
$-z$		$35M+26$	$3M+6$	$3M+4$	$8M+7$	0	0	0	
$-M$	x_5	3	-1/5	[7/5]	0	0	1	-3/5	15/7
3	x_3	4	2/5	1/5	1	0	0	1/5	20
-1	x_4	10	-3/5	6/5	0	1	0	-4/5	25/3
$-z$		$3M-2$	$-M/5+16/5$	$7/5M+13/5$	0	0	0	-8/5M-7/5	
2	x_2	15/7	-1/7	1	0	0	5/7	-3/7	
3	x_3	25/7	[3/7]	0	1	0	-1/7	2/7	25/3
-1	x_4	52/7	-3/7	0	0	1	-6/7	-2/7	
$-z$		-53/7	25/7	0	0	0	-M-13/7	-M-2/7	
2	x_2	10/3	0	1	1/3	0	2/3	-1/3	
5	x_1	25/3	1	0	7/3	0	-1/3	2/3	
-1	x_4	11	0	0	1	1	-1	-0	
$-z$		-112/3	0	0	-25/3	0	-M-2/3	-M+8/3	

表 2.2: 大 M 法例题

得到基本可行解：$(25/3, 10/3, 0, 11)^T$ 为最优解
得到最优值 $z = 112/3$

2.2.2.2 二阶段法

STEPS:

1. 将线性规划问题转化为标准型
2. 观察变量，若初始基本可行解明显，直接进行单纯形法；否则引入人工变量 $x_{n+i} \geq 0 (i = 1, ..., m)$ 构造辅助问题（LP - 1）
3. 第一阶段，求解辅助问题（LP - 1），若得到的最优解满足 $x_{n+i} = 0 (i = 1, ..., m)$，则是原问题的基本可行解；**否则，原问题无可行解**。

4. 第二阶段，得到原问题的基本可行解后，直接删除人工变量，求解原问题

例：使用二阶段法求解下面的问题

$$
\begin{aligned}
max \quad & z = 5x_1 + 2x_2 + 3x_3 - x_4 \\
s.t. \quad & \begin{cases} x_1 + 2x_2 + 3x_3 & = 15 \\ 2x_1 + x_2 + 5x_3 & = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 & = 26 \\ x_1, x_2, x_3, x_4 \geq 0 \end{cases}
\end{aligned}
$$

标准化、引入人工变量并构造第一阶段问题（LP - 1）：

$$
\begin{aligned}
max \quad & z' = -x_5 - x_6 \\
s.t. \quad & \begin{cases} x_1 + 2x_2 + 3x_3 \qquad + x_5 & = 15 \\ 2x_1 + x_2 + 5x_3 \qquad\qquad + x_6 & = 20 \\ x_1 + 2x_2 + 4x_3 + x_4 & = 26 \\ x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{cases}
\end{aligned}
$$

建立第一阶段的单纯形表：

$\boldsymbol{c_B}$	$\boldsymbol{x_B}$	$\boldsymbol{b}$	0	0	0	0	-1	-1	$\boldsymbol{\theta}$
			x_1	x_2	x_3	x_4	x_5	x_6	
-1	x_5	15	1	2	3	0	1	0	5
-1	x_6	20	2	1	[5]	0	0	1	4
0	x_4	26	1	2	4	1	0	0	6.5
$-z$		35	3	3	8	0	0	0	
-1	x_5	3	-1/5	[7/5]	0	0	1	-3/5	15/7
0	x_3	4	2/5	1/5	1	0	0	1/5	20
0	x_4	10	-3/5	6/5	0	1	0	-4/5	25/3
$-z$		3	-1/5	7/5	0	0	0	-8/5	
2	x_2	15/7	-1/7	1	0	0	5/7	-3/7	
3	x_3	25/7	[3/7]	0	1	0	-1/7	2/7	25/3
-1	x_4	52/7	-3/7	0	0	1	-6/7	-2/7	
$-z$		0	0	0	0	0	-1	-1	

表 2.3: 二阶段法例题 1

得到原问题的基本可行解 $(0, 15/7, 25/7, 52/7)^T$

第二阶段删除人工变量，并把基本可行解填入表中

得到基本可行解：$(25/3, 10/3, 0, 11)^T$ 为最优解

得到最优值 $z = 112/3$

(表见下页)

$\boldsymbol{c_B}$	$\boldsymbol{x_B}$	$\boldsymbol{b}$	0	0	0	0	$\boldsymbol{\theta}$
			x_1	x_2	x_3	x_4	
2	x_2	15/7	-1/7	1	0	0	
3	x_3	25/7	[3/7]	0	1	0	25/3
-1	x_4	52/7	-3/7	0	0	1	
$-z$		-53/7	25/7	0	0	0	
2	x_2	10/3	0	1	1/3	0	
5	x_1	25/3	1	0	7/3	0	
-1	x_4	11	0	0	1	1	
$-z$		-112/3	0	0	-25/3	0	

表 2.4: 二阶段法例题 2

2.2.3 线性规划的对偶问题

给定一个优化问题，我们去理解它的时候，或者设计算法的时候，可以研究它的对偶。
有时原问题不好解，但它的对偶相对容易。这个时候，可以从对偶问题出发，进而寻求原问题的解。

2.2.3.1 对偶问题的形式

1. 对称形式的对偶问题

$$(P)\begin{cases} MAX & z=\boldsymbol{c^T x} \\ s.t. & Ax\le b \\ & x\ge 0 \end{cases} \qquad (D)\begin{cases} MIN & f=\boldsymbol{b^T y} \\ s.t. & A^T y\ge c \\ & y\ge 0 \end{cases}$$

一对对称形式的对偶规划之间具有下面的对应关系

- “$Max\ \le$” 和 “$Min\ \ge$” 相对应
- 从约束系数矩阵看：一个模型中为 ，则另一个模型中为 A^T；一个模型是 m 个约束、n 个变量，则它的对偶模型为 n 个约束、m 个变量
- 从数据 b、c 的位置看：在两个规划模型中，b 和 c 的位置对换
- 两个规划模型中的变量皆非负

2. 非对称形式的对偶问题

（下式并不囊括所有情况）

$$(P)\begin{cases} MAX & z=\boldsymbol{c^T x} \\ s.t. & Ax=b \\ & x\ge 0 \end{cases} \qquad (D)\begin{cases} MIN & f=\boldsymbol{b^T y} \\ s.t. & A^T y\ge c \end{cases}$$

一般称不具有对称形式的一对线性规划为非对称形式的对偶规划。对于非对称形式的规划，可以按照下面的对应关系直接给出其对偶规划。

- 将模型统一为“$Max\ \leq$”或“$Min\ \geq$”的形式
- 若原规划的某个约束条件为等式约束，则在对偶规划中与此约束对应的那个变量取值没有非负限制
- 若原规划的某个变量的值没有非负限制，则在对偶问题中与此变量对应的那个约束为等式

原问题与对偶问题的对应关系

原问题（对偶问题）		对偶问题（原问题）	
min		max	
变量	n 个变量	约束	n 个约束
	变量 ≥ 0		约束 $\leq$
	变量 ≤ 0		约束 $\geq$
	无正负限制		约束 $=$
约束	m 个约束	变量	m 个变量
	约束 $\leq$		变量 ≤ 0
	约束 $\geq$		变量 ≥ 0
	约束 $=$		无正负限制
约束条件右端项		目标函数中的变量系数	
目标函数中的变量系数		约束条件右端项	

表 2.5: 原问题与对偶问题的对应关系

2.2.3.2 对偶定理

设有一对互为对偶的线性规划

$$(P)\begin{cases} MAX & z=\boldsymbol{c^T x} \\ s.t. & Ax \leq b \\ & x \geq 0 \end{cases} \qquad (D)\begin{cases} MIN & f=\boldsymbol{b^T y} \\ s.t. & A^T y \geq c \\ & y \geq 0 \end{cases}$$

定理 3-3：若 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别为原规划 (P) 和 (D) 对偶规划，则

$$\boldsymbol{c}^T\boldsymbol{x} \leq \boldsymbol{b}^T\boldsymbol{y}$$

推论 3-1：设 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 分别为原规划 (P) 和 (D) 的可行解，当 $\boldsymbol{c}^T\boldsymbol{x}=\boldsymbol{b}^T\boldsymbol{y}$ 时，$\boldsymbol{x},\boldsymbol{y}$ 分别是两个问题的最优解

推论 3-2：若规划 (P) 有可行解，则规划 (P) 有最优解的充分必要条件是规划 (D) 有可行解

推论 3-3：若规划 (D) 有可行解，则规划 (D) 有最优解的充分必要条件是规划 (P) 有可行解

定理 3-4：若原规划 (P) 有最优解，则对偶规划 (D) 也有最优解，反之亦然，且两者的目标函数值相等。

2.2.3.3 对偶单纯形法

对偶单纯形法是求解原规划的一种方法。

原理:

作为原规划的一个解，会有两个性质等待满足：可行性和最优性。而原问题的可行性和最优性恰好对应对偶问题的最优性和可行性。

单纯形法的思路是，先满足可行性，再逐渐逼近最优性；而对偶单纯形法的思路是，先找到最优性，再逐渐逼近可行性。

也就是说，先找到对偶问题的可行解，再找到原问题的可行解（即对偶问题的最优解)。

最优性：看检验数 σ_j 可行性：看右端项 b_i

从原规划的一个基本解出发，此基本解不一定可行，但它对应着一个对偶可行解；就是说可以从一个对偶可行解出发，然后检验原规划的基本解是否可行，即是否有负的分量。如果有负的分量，则进行迭代，求另一个基本解，此基本解对应着另一个对偶可行解（检验数非正)；而得到的基本解的分量皆非负，则该基本解为最优解。

也就是说，对偶单纯形法在迭代过程中始终保持对偶解的可行性（检验数非正)，使得原规划的基本解由不可行变为可行，当同时得到对偶规划与原规划的可行解时，得到原规划的最优解。

STEPS:

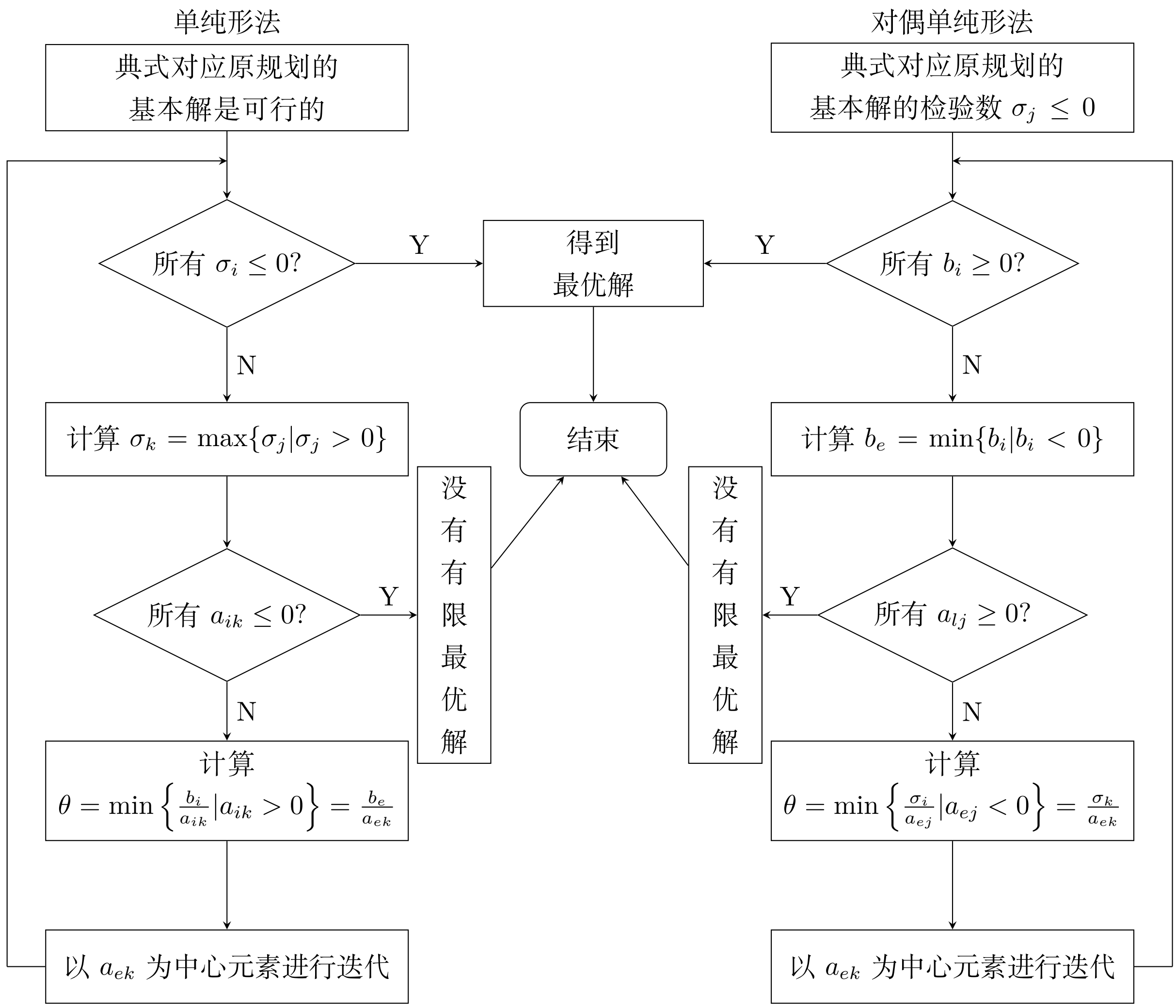
1. 根据线性规划典式形式，建立初始单纯形表（就是还按照单纯形法填初始单纯形表)。此表对应原规划的一个基本解。表要求：检验数数行各元素一定非正，原规划的基本解可以有小于零的分量。

2. 若基本解的所有分量皆非负，则得到原规划的最优解，停止计算；若基本解中有小于零的分量 b_i，并且 b_i 所在行各系数 $a_{ij} \geq 0$，则原规划无可行解，停止计算；若 $b_i < 0$，并且存在 $a_{ir} < 0$，则确定 x_r 为出基变量，并计算

$$\theta = min\left\{\frac{\sigma_j}{a_{rj}}|a_{rj} < 0\right\} = \frac{\sigma_k}{a_{rk}}$$

   确定 x_k 为进基向量。若有多个 $b_i < 0$，则选择最小的进行分析计算。

3. 以 b_{rk} 为中心元素，按照与单纯形法类似的方法，在表中进行迭代计算，返回第 2 步。



2.2.4 灵敏度分析

前提：在求灵敏度分析和影子价格中，常常会遇到要使用对偶问题的最优解（或最优基 $\boldsymbol{B}$）求解问题，这里解释一下怎么使用最优单纯形表求得。

1. 首先利用单纯形法得到最优单纯形表
2. 得到 b 列的 n 个数值，最优单纯形表从后往前数 n 列（就是松弛变量对应的列），有矩阵：

$$\left(\begin{array}{cccc}
a_{1(m-n)} & a_{1(m-n+1)} & \cdots & a_{1m} \\
a_{2(m-n)} & a_{2(m-n+1)} & \cdots & a_{2m} \\
\vdots & & & \vdots \\
a_{n(m-n)} & a_{n(m-n+1)} & \cdots & a_{nm} \\
\hline
\sigma_{m-n} & \sigma_{m-n} & \cdots & \sigma_{m-n}
\end{array}\right)$$

其中：

- 最后一行的检验数 $\sigma^T = -\boldsymbol{c_B^T}\boldsymbol{B}^{-1}$，各检验数取相反数（即 $\boldsymbol{c_B^T}\boldsymbol{B}^{-1}$）即为对偶问题的最优解。
- 除最后一行外（即线上方的）矩阵即为 $\boldsymbol{B}^{-1}$

2.2.4.1 影子价格

影子价格是一个向量，它的分量表示最优目标值随相应资源数量变化的变化率。

若 x^*, y^* 分别为 (LP) 和 (DP) 的最优解，那么有

$$\boldsymbol{c^T x^*} = \boldsymbol{b^T y^*}$$

根据 $f = \boldsymbol{b^T y^*} = b_1 y_1^*, b_2 y_2^*, \cdots, b_m y_m^*$ 可知

$$\frac{\partial f}{\partial b_i} = y_i^*$$

y_i^* 表示 b_i 变化一个单位对目标 f 产生的影响，称 y_i^* 为 b_i 的影子价格

需要指出，影子价格不是固定不变的，当约束条件、产品利润等发生变化时，有可能使影子价格发生变化。另外，影子价格的经济含义，是指资源在一定范围内增加时的情况，当某种资源的增加超过了这个“一定的范围”时，总利润的增加量则不是按照影子价格给出的数值线性地增加。

如何求解影子价格：

1. 求出对偶问题的最优解（可以利用单纯形表，在求出最优解的情况下，将松弛变量对应的各检验数取负 $-\sigma_i$ 就能得到对偶问题的最优解）
2. 对偶问题最优解中的数字依次对应的就是原问题的各资源影子价格

经济意义：每增加一单位的某资源，最终收益增加多少单位

判断资源是否有剩余：$\begin{cases} y_i^* = 0 & \text{有剩余} \\ y_i^* > 0 & \text{无剩余} \end{cases}$

2.2.4.2 目标函数系数 c 变化*

若只有一个系数 c_j 变化，其他系数不变。c_j 的变化只影响检验数 σ_j，而不影响解的非负性。

$$\sigma_j = c_j - \boldsymbol{c_B}^T \boldsymbol{B}^{-1} \boldsymbol{p}_j, j = 1, 2, \cdots, n$$

1. c_k 是非基变量的系数

   非基系数变化只影响与 c_k 有关的一个检验数 σ_k 的变化，对其他无影响，故只需要考虑 σ_k。

   设 $c_k \to \bar{c}_k = c_k + \Delta c_k$，有 σ_k 的变化:

   $$\bar{\sigma}_k = c_k + \Delta c_k - \boldsymbol{c_B}^T \boldsymbol{B}^{-1} \boldsymbol{p}_j = \sigma_k + \Delta c_k$$

   为了保持最优解不变，σ_k 必须满足 $\bar{\sigma}_k = \sigma_k + \Delta c_k \leq 0$。也就是说：

   $$\Delta c_k \leq -\sigma_k, \bar{c}_k = c_k + \Delta c_k \leq c_k - \sigma_k \tag{2.3}$$

   $c_k - \sigma_k$ 是 c_k 变化的上限，若 c_k 不超出上限，最优解不变；否则，将最优单纯形表中的检验数 σ_k 用 $\bar{\sigma}_k$ 取代，取 x_k 为进基变量，继续单纯形的表格计算。

2. c_l 是基变量的系数

   设 $c_l \to c_l + \Delta c_l$，引入 $\Delta c = (0, \cdots, 0, \Delta c_l, 0, \cdots, 0)$，有

   $$\begin{aligned} \sigma_j \to \bar{\sigma}_j &= c_j - [\boldsymbol{c_B}^T + (\Delta \boldsymbol{c})^T] \boldsymbol{B}^{-1} \boldsymbol{p}_j, j \neq l \\ &= \sigma_j - \Delta c_l a'_{lj} \end{aligned}$$

（注意：上式中 a 的 l 与左边基变量的下标对应）为保证最优解不变，Δc_l 要满足

$$\max\left\{\frac{\sigma_j}{a'_{ij}}|a'_{ij}>0\right\}\leq\Delta c_l\leq\min\left\{\frac{\sigma_j}{a'_{ij}}|a'_{ij}<0\right\}$$

若 Δc_l 超出此范围，应求出新的检验数 $\bar{\sigma}_j$，选择其中大于零的检验数对应的变量 x_j 为进基变量，继续迭代。

2.2.4.3 右端常数 b 变化 $\star$

b_r 的变化影响解的可行性，但不影响检验数的符号变化。由 $\boldsymbol{x_B}=\boldsymbol{B}^{-1}\boldsymbol{b}$ 可知 b_r 的变化必会引起最优解数值变化。

最优解的变化分为以下两类：

1. 保持 $\boldsymbol{B}^{-1}\boldsymbol{b}\geq\boldsymbol{0}$，即最优基 $\boldsymbol{B}$ 不变（影子价格不变，也就是对偶问题的最优解不变）
只需要将变化后的 b_r 带入 $\boldsymbol{B}^{-1}\boldsymbol{b}$ 的表达式重新计算即可

2. $\boldsymbol{B}^{-1}\boldsymbol{b}$ 出现负分量，这使最优基 $\boldsymbol{B}$ 变化
需要通过迭代求解新的最优基和最优解

综合一下，可以利用下面的步骤计算：
设 $b_r\to\bar{b}_r=b_r+\Delta b_r$，$\Delta b_r$，此时有

$$\boldsymbol{x_B}\to\bar{\boldsymbol{x}}_{\boldsymbol{B}}=\boldsymbol{B}^{-1}\begin{pmatrix}b_1\\\vdots\\b_r+\Delta b_r\\\vdots\\b_m\end{pmatrix}=\boldsymbol{B}^{-1}\boldsymbol{b}+\boldsymbol{B}^{-1}\begin{pmatrix}0\\\vdots\\\Delta b_r\\\vdots\\0\end{pmatrix}=\boldsymbol{x_B}+\Delta b_r\begin{pmatrix}\beta_{1r}\\\vdots\\\beta_{mr}\end{pmatrix}=\begin{pmatrix}b'_1\\\vdots\\b'_m\end{pmatrix}+\Delta b_r\begin{pmatrix}\beta_{1r}\\\vdots\\\beta_{mr}\end{pmatrix}\geq\begin{pmatrix}0\\\vdots\\0\end{pmatrix}$$

（上式只需要随便算其中的一个，然后看是否大于等于 0 就行，不需要特别算下面的 Δb_r）

其中，为 $\boldsymbol{x_B}^{-1}\boldsymbol{b}$ 原最优解，b'_i 为 $\boldsymbol{x_B}=\boldsymbol{B}^{-1}\boldsymbol{b}$ 的第 i 个分量，β_{ir} 为 $\boldsymbol{B}^{-1}$ 的第 i 行第 r 列元素

要满足

$$\max\left\{\frac{-b'_i}{\beta_{ir}}|\beta_{ir}>0\right\}\leq\Delta b_r\leq\min\left\{\frac{-b'_i}{\beta_{ir}}|\beta_{ir}<0\right\}$$

当 Δb_r 超过此范围时，将使最优解中某个分量小于零，使最优基发生变化。此时可用对偶单纯形法继续迭代新的最优解。

就是说，如果没超范围，直接把 $\bar{\boldsymbol{x_B}}$ 当最优解得出就行；如果超范围的话，就把 $\bar{\boldsymbol{x_B}}$ 放到最优单纯形表的 b 列，然后求对偶单纯形。

2.2.4.4 约束条件系数 a 变化

假设只有一个 a_{ij} 变化，其他数据不变，且只讨论 a_{ij} 为非基变量 x_j 的系数的情况。那么此时 a_{ij} 的变化只影响一个检验数 σ_j。

设 $a_{ij} \to a_{ij} + \Delta a_{ij}$，由检验数的另一种表示形式

$$\sigma_j \to \bar{\sigma}_j = c_j - \boldsymbol{y^T} \begin{pmatrix} a_{1j} \\ \vdots \\ a_{ij} + \Delta a_{ij} \\ \vdots \\ a_{mj} \end{pmatrix} = c_j - \boldsymbol{y^T} \boldsymbol{p}_j - \boldsymbol{y^T} \begin{pmatrix} 0 \\ \vdots \\ \Delta a_{ij} \\ \vdots \\ 0 \end{pmatrix} = \sigma_j - y_i^* \Delta a_{ij}$$

其中，$\boldsymbol{y}$ 为对偶最优解，y_i^* 为 $\boldsymbol{y}$ 的第 i 个变量

为使最优解不变，要使 $\sigma_j \le 0$，即

$$\begin{aligned} \sigma_j &\le y_i^* \Delta a_{ij} \\ \Delta a_{ij} &\ge \frac{\sigma_j}{y_i^*}, y_i^* > 0 \\ \Delta a_{ij} &\le \frac{\sigma_j}{y_i^*}, y_i^* < 0 \end{aligned}$$

2.2.4.5 新增变量 x 分析

增加变量 x_{n+1}，则有相应的约束条件 $\boldsymbol{p}_{n+1}$，目标函数系数 c_{n+1}，那么，计算出

$$\sigma_{n+1} = c_{n+1} - \boldsymbol{c_B}^T \boldsymbol{B}^{-1} \boldsymbol{p}_{n+1}$$

填入最优单纯形表, 若 $\sigma_{n+1} \le 0$，则最优解不变；否则，进一步用单纯形法求解

2.2.4.6 新增约束条件

增加一个约束之后，应把最优解代入新的约束，若满足，则最优解不变；否则，填入最优单纯形表作为新的一行，引入一个新的非负变量（原约束若是小于等于形式，可引入非负松弛变量；否则，引入非负人工变量），并通过矩阵行变换把对应基变量的元素变为 0，进一步用单纯形法或对偶单纯形法求解。

Chapter 3

最优化搜索算法的结构和一维搜索

从此章开始就是求非线性规划问题，与线性规划问题可以在有限步数内得到解不同，非线性规划问题不一定可以在有限的迭代步数内收敛。

3.1 常用的搜索算法结构

3.1.1 收敛性概念

在规划问题的的求解过程中，由于迭代算法是以产生一系列迭代点为目的的，因此算法的收敛性表现在产生的点列 $\{\boldsymbol{x}^{(k)}\}$ 上。

1. 理想的收敛性概念：设 $\boldsymbol{x}^*$ 为问题 (fS) 的全局最优解。当 $\boldsymbol{x}^* \in \{\boldsymbol{x}^{(k)}\}$ 或 $\boldsymbol{x}^{(k)} \neq \boldsymbol{x}^*$ 对于所有 k 成立，并且满足 $\lim\limits_{k\to\infty} \boldsymbol{x}^{(k)} = \boldsymbol{x}^*$ 时，称算法收敛到最优解 $\boldsymbol{x}^*$。
2. 然而，由于实际问题较为复杂，很难达到理想状态，因此需要引入一些实用的收敛性概念：

   首先定义解集 Ω ，是具有某种可接受性质的点集，通常取下面的几种集合：

   (a) $\Omega = \{\boldsymbol{x}^* | \boldsymbol{x}^* \ is\ g.opt\}$

   (b) $\Omega = \{\boldsymbol{x}^* | \boldsymbol{x}^* \ is\ l.opt\}$

   (c) $\Omega = \{\boldsymbol{x}^* | \boldsymbol{x}^*$满足某种最优条件，或者可以说$\nabla f(\boldsymbol{x}^*) = 0\}$

   (d) $\Omega = \{\boldsymbol{x}^* | \boldsymbol{x}^* \in S, f(\boldsymbol{x}^*) \leq B\}$，其中 B 为可接受的目标函数值的一个上界

   设算法产生的点列为 $\boldsymbol{x}^{(k)}$ ，满足下列任一情况时，称算法收敛:

   (a) $\{\boldsymbol{x}^{(k)}\} \cap \Omega \neq \varnothing$

   (b) $\{\boldsymbol{x}^{(k)}\}$ 的任意收敛子列的极限点属于 Ω

求解问题时，需要考虑初始点的影响，因此引入下面的概念：

1. 全局收敛性：若算法对任意的初始点 $\boldsymbol{x}^{(1)}$ 都收敛，则称算法全局收敛。
2. 局部收敛性：若算法只有限制初始点 $\boldsymbol{x}^{(1)}$ 到解集 Ω 附近（当 Ω 为非连通时，指在 Ω 的某个连通子集附近）时，才有收敛性，则称算法局部收敛。

3.1.2 收敛准则（停机条件）

1. $||\boldsymbol{x}^{(k+m)}-\boldsymbol{x}^{(k)}||<\varepsilon$（最常用的是 $m=1$ ）
2. $\frac{||\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^{(k)}||}{||\boldsymbol{x}^{(k)}||}<\varepsilon$
3. $|f(\boldsymbol{x}^{(k+1)})-f(\boldsymbol{x}^{(k)})|<\varepsilon$

3.1.3 收敛速度

设 $\{\boldsymbol{x}^{(k)}\}$ 为算法产生的点列，收敛于解 $\boldsymbol{x}^*$ ，且 $\boldsymbol{x}^{(k)}\neq\boldsymbol{x}^*$，有 $\forall k$

1. 线性收敛：$\exists\alpha\in(0,1)\frac{||\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^*||}{||\boldsymbol{x}^{(k)}-\boldsymbol{x}^*||}\leq\alpha,k$ 充分大时 $\boldsymbol{x}^{(k)}$ 收敛于 $\boldsymbol{x}^*$
表示每迭代一次，比上一次更接近 $\boldsymbol{x}^*$；收敛的阶数 $p_0=1$
2. 超线性收敛：$\lim\limits_{k\to\infty}\frac{||\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^*||}{||\boldsymbol{x}^{(k)}-\boldsymbol{x}^*||}=0$
迭代速度比线性收敛更快，收敛的阶数 $p_0\geq1$
3. 二阶收敛：$\exists\alpha>0$，使得 k 充分大时，有 $\frac{||\boldsymbol{x}^{(k+1)}-\boldsymbol{x}^*||}{||\boldsymbol{x}^{(k)}-\boldsymbol{x}^*||^2}\leq\alpha$
收敛的阶数 $p_0=2$，说明二阶收敛是超线性收敛的一个特例

3.1.4 下降算法模型