

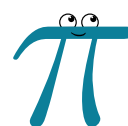
运筹学 - 刁在筠 - 笔记

作者：若水

邮箱：ethanmxzhou@163.com

主页：helloethanzhou.github.io

时间：July 18, 2024



致谢

感谢 勇敢的 自己

目录

第一章 线性规划	1
1.1 线性规划模型	1
1.2 可行区域与基本可行解	2
1.2.1 图解法	2
1.2.2 可行区域的几何结构	2
1.2.3 基本可行解与线性规划基本定理	4
1.3 单纯形方法	6
1.3.1 单纯形方法	6
1.4 初始解	10
1.5 对偶性及对偶单纯形法	11
1.5.1 对偶线性规划	11
1.5.2 对偶理论	12
1.6 灵敏度分析	13
1.6.1 改变价值向量	13
1.6.2 改变右端向量	13
第二章 整数线性规划	14
2.1 整数线性规划问题	14
2.2 Gomory 割平面法	14
2.2.1 Gomory 割平面法的基本思想	14
2.2.2 Gomory 割平面法计算步骤	14
2.3 分枝定界法	15
第三章 非线性规划	16
3.1 基本概念	16
3.1.1 非线性规划问题	16
3.1.2 非线性规划方法概述	16
3.2 凸函数和凸规划	17
3.2.1 凸函数及其性质	17
3.2.2 凸规划及其性质	19
3.3 一维搜索方法	20
3.3.1 0.618 法	20
3.3.2 Newton 法	20
3.4 无约束最优化方法	21
3.4.1 无约束问题的最优性条件	21
3.4.2 最速下降法	22
3.4.3 共轭方向法	22
3.5 约束最优化方法	23
3.5.1 约束最优化问题的最优性条件	23
3.5.2 简约梯度法	25
3.5.3 惩罚函数法	26
3.5.3.1 罚函数法	26

3.5.3.2 障碍函数法	27
第四章 动态规划	29
4.1 多阶段决策问题	29
4.2 最优化原理	29
4.3 确定性的定期多阶段决策问题	30
4.3.1 递推法	30
4.3.2 最短路径问题	30
4.3.3 旅行售货员问题	31
4.3.4 多阶段资源分配问题	31
4.3.5 可靠性问题	33
4.4 确定性的不定期多阶段决策问题	34
4.4.1 最优线路问题	34
第五章 图与网络分析	38
5.1 图的基本概念与矩阵表示	38
5.1.1 图的基本概念	38
5.1.2 图的关联矩阵与邻接矩阵	40
5.1.3 子图	41
5.2 图的连通性	41
5.2.1 图的连通	41
5.2.2 图的割集	42
5.3 树与支撑树	42
5.3.1 树及其基本性质	42
5.3.2 支撑树及其基本性质	44
5.4 最小树问题	44
5.4.1 最小树及其性质	44
5.4.2 Kruskal 算法	45
5.4.3 Dijkstra 算法	47
5.5 最短有向路问题	49
5.6 最大流问题	52
第六章 排队论	55
6.1 基本模型	55
6.2 $M/M/1/\infty$ 模型	56
第七章 决策分析	57
7.1 模型建立	57
7.2 风险型决策分析	57
7.3 不确定型决策分析	57
附录 A 期末复习	61
A.1 填空与判断	61

第一章 线性规划

1.1 线性规划模型

定义 1.1.1 (线性规划问题的一般形式)

线性规划 (Linear Programming, LP) 问题的一般形式为

$$\begin{aligned} \min \quad & z = c_1x_1 + \cdots + c_nx_n \\ \text{s.t.} \quad & \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p \\ a_{p+1,1}x_1 + \cdots + a_{p+1,n}x_n \geq b_{p+1} \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \geq b_m \\ x_1, \cdots, x_n \geq 0 \end{cases} \end{aligned} \quad (*)$$

其中 x_1, \cdots, x_n 为待定的决策变量；系数 a_{ij} 构成的矩阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$$

称为约束矩阵，其第 j 个列向量为 A_j ； $z = c_1x_1 + \cdots + c_nx_n$ 称为目标函数；向量 $c = (c_1, \cdots, c_n)^T$ 称为价值向量；向量 $b = (b_1, \cdots, b_m)^T$ 称为右端向量；条件 $x_1, \cdots, x_n \geq 0$ 称为非负约束，其余决策变量 x_{q+1}, \cdots, x_n 称为自由变量。满足线性规划问题 (*) 的向量 $x = (x_1, \cdots, x_n)^T$ 称为可行解，所有可行解构成可行区域 D 。对于给定 LP 问题，成立如下情况其一。

1. $D = \emptyset$ ：称该问题无解 或不可行。
2. $D \neq \emptyset$ 且目标函数在 D 上无界：称该问题无界。
3. $D \neq \emptyset$ 且目标函数在 D 上有界：称该问题存在最优解。

定义 1.1.2 (线性规划问题的规范形式)

LP 问题的规范形式为

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax \geq b \\ x \geq 0 \end{cases} \end{aligned} \quad (**)$$

定义 1.1.3 (线性规划问题的标准形式)

LP 问题的标准形式为

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned} \quad (***)$$

定理 1.1.1 (线性规划问题的形式的等价性)

$$\text{一般形式} \iff \text{规范形式} \iff \text{标准形式}$$



证明 只需证明

$$\text{一般形式} \implies \text{规范形式}$$

$$\text{一般形式} \implies \text{标准形式}$$

一般形式 \implies 规范形式：对于等式约束

$$a_1x_1 + \cdots + a_nx_n = b_n$$

可等价表示为

$$\begin{cases} a_1x_1 + \cdots + a_nx_n \geq b_n \\ (-a_1)x_1 + \cdots + (-a_n)x_n \geq (-b_n) \end{cases}$$

对于自由变量 x ，可引入变量 x^+ 与 x^- ，使得成立

$$x = x^+ - x^-, \quad x^+, x^- \geq 0$$

一般形式 \implies 标准形式：对于不等式约束

$$a_1x_1 + \cdots + a_nx_n \geq b_n$$

可引入**松弛变量** y ，使得成立

$$a_1x_1 + \cdots + a_nx_n - y = b_n, \quad y \geq 0$$

对于自由变量 x ，可引入变量 x^+ 与 x^- ，使得成立

$$x = x^+ - x^-, \quad x^+, x^- \geq 0$$

1.2 可行区域与基本可行解

1.2.1 图解法

定理 1.2.1

1. LP 问题的可行域 D 形成有界或无界的凸集。
2. 如果 LP 问题存在最优解，那么最优解存在于 D 的某个顶点。



1.2.2 可行区域的几何结构

定义 1.2.1 (凸集)

称集合 $S \subset \mathbb{R}^n$ 为凸集，如果成立

$$\forall x, y \in S, \forall \lambda \in [0, 1] \implies \lambda x + (1 - \lambda)y \in S$$

**定理 1.2.2 (可行域为凸集)**

LP 问题的标准形式 (***) 的可行域

$$D = \{x \in \mathbb{R}^n : Ax = b, x \geq 0\}$$

为凸集。



证明 任取 $\mathbf{x}, \mathbf{y} \in D$, 以及 $\lambda \in [0, 1]$, 由于 $\mathbf{x}, \mathbf{y} \geq \mathbf{0}$, 那么

$$\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \geq \mathbf{0}$$

又由于 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 且 $\mathbf{A}\mathbf{y} = \mathbf{b}$, 那么

$$\mathbf{A}(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) = \mathbf{b}$$

进而

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

从而 D 为凸集。

引理 1.2.1 (凸集的交为凸集)

任意一族凸集的交为凸集。



证明 这几乎是显然的!

定义 1.2.2 (超平面)

对于非零向量 $\mathbf{a} \in \mathbb{R}^n$ 与 $b \in \mathbb{R}$, 称集合

$$H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} = b\}$$

为 \mathbb{R}^n 中的超平面。 H 将 \mathbb{R}^n 分为两个半空间

$$H^+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} \geq b\}, \quad H^- = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} \leq b\}$$



引理 1.2.2

超平面与其分割的半空间为凸集。



证明 这几乎是显然的!

定义 1.2.3 (多面凸集)

定义 \mathbb{R}^n 中的多面凸集为集合

$$S = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}_i^T \mathbf{x} = b_i, \mathbf{a}_j^T \mathbf{x} \geq b_j, 1 \leq i \leq p, p+1 \leq j \leq p+q\}$$



定义 1.2.4 (多面体)

称非空有界多面凸集为多面体,



定理 1.2.3 (多面凸集为凸集)

多面凸集为凸集。



证明 由引理 1.2.1 与 1.2.2, 命题得证!

定义 1.2.5 (凸集的顶点)

对于凸集 $S \subset \mathbb{R}^n$, 称 $\mathbf{x} \in S$ 为其顶点, 如果对于任意 $\mathbf{y} \neq \mathbf{z} \in S$ 以及 $\lambda \in [0, 1]$, 成立

$$\mathbf{x} \neq \lambda \mathbf{y} + (1 - \lambda) \mathbf{z}$$



1.2.3 基本可行解与线性规划基本定理

定义 1.2.6 (基本可行解)

对于 LP 问题的标准形式

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned}$$

其中 $A \in \mathbb{R}^{m \times n}$ 。假设可行域非空，因此不妨假设 $\text{rank}(A) = m \leq n$ 。令 $A = (B, N)$, $x = (x_B, x_N)^T$ 与其对应，其中 B 为 m 阶非退化方阵，因此原方程化为典型方程

$$Ax = b \iff Bx_B + Nx_N = b \iff x_B = B^{-1}b - B^{-1}Nx_N$$

称 B 为基， B 中的列向量为基向量；称 x_B 为基变量， x_N 为非基变量；称 $x = (B^{-1}b, 0)^T$ 为对应 B 的基本解；若 $B^{-1}b \geq 0$ ，则称 $x = (B^{-1}b, 0)^T$ 为对应 B 的基本可行解；若 $B^{-1}b > 0$ ，则称 $x = (B^{-1}b, 0)^T$ 为非退化基本可行解。



引理 1.2.3

对于 LP 问题的标准形式

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned}$$

其中 $A \in \mathbb{R}^{m \times n}$ ，且 $\text{rank}(A) = m \leq n$ ，如果 \bar{x} 为可行解，且其正分量位置为 n_1, \dots, n_k ，那么

$$\bar{x} \text{ 为基本可行解} \iff A_{n_1}, \dots, A_{n_k} \text{ 线性无关}$$



证明 不妨

$$\bar{x} = (\bar{x}_1, \dots, \bar{x}_k, 0, \dots, 0)^T, \quad \bar{x}_i > 0$$

必要性显然！对于充分性，由于 A_{n_1}, \dots, A_{n_k} 线性无关，那么 $k \leq m$ 。由于 \bar{x} 为可行解，那么

$$\sum_{i=1}^k \bar{x}_i A_i = b$$

其中 A_i 为 A 的第 i 个列向量。如果 $k = m$ ，那么 $B = (A_1, \dots, A_k)$ 为基， \bar{x} 为 B 对应的基本可行解。如果 $k < m$ ，那么存在（不妨设为） A_{k+1}, \dots, A_m ，使得 $B = (A_1, \dots, A_m)$ 为基， \bar{x} 为 B 对应的基本可行解。

定理 1.2.4 (基本可行解 \iff 顶点)

对于 LP 问题的标准形式

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned}$$

其中 $A \in \mathbb{R}^{m \times n}$ ，且 $\text{rank}(A) = m \leq n$ ，如果 \bar{x} 为可行解， D 为可行域，那么

$$\bar{x} \text{ 为基本可行解} \iff \bar{x} \text{ 为 } D \text{ 的顶点}$$



证明 如果 \bar{x} 不为基本可行解，不妨 \bar{x} 的前 k 个分量为正，并记作 $\bar{x} = (\bar{x}_1, \dots, \bar{x}_k, 0, \dots, 0)^T$ ，那么由引理 1.2.3，

A_1, \dots, A_k 线性相关, 其中 A_i 为 A 的第 i 个列向量, 从而存在非零向量 $y = (y_1, \dots, y_k, 0, \dots, 0)^T$, 使得成立

$$\sum_{i=1}^k y_i A_i = 0$$

对于任意 $\delta > 0$, 成立

$$\sum_{i=1}^k (\bar{x}_i \pm \delta y_i) A_i = b$$

因此取 $u = \bar{x} + \delta y$ 与 $v = \bar{x} - \delta y$, 就成立 $Au = Av = b$ 。令 δ 充分小, 使得 $u, v \geq 0$, 因此 $u, v \in D$ 。由于 $y \neq 0$, 那么 $u \neq v$ 。由于

$$\bar{x} = \frac{1}{2}u + \frac{1}{2}v$$

那么 \bar{x} 不为顶点。

如果 \bar{x} 不为 D 的顶点, 那么存在 $u \neq v \in D$ 与 $0 < \lambda < 1$ 成立

$$\bar{x} = \lambda u + (1 - \lambda)v$$

其中 $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)^T$ 且 $u = (u_1, \dots, u_n)^T, v = (v_1, \dots, v_n)^T$ 。不妨 \bar{x} 的前 k 个分量为正, 又由于当 $i \geq k+1$ 时, 成立 $\bar{x}_i = 0$ 且 $u_i, v_i \geq 0$, 那么 $u_i = v_i = 0$ 。由于 $Ax = Ay = b$, 那么

$$\sum_{i=1}^k (u_i - v_i) A_i = 0$$

其中 A_i 为 A 的第 i 个列向量。由于 $x \neq y$, 那么存在 $1 \leq i \leq k$, 使得成立 $u_i \neq v_i$, 从而 A_1, \dots, A_k 线性相关。由引理 1.2.3, \bar{x} 不为基本可行解。

定理 1.2.5 (基本可行解的存在性)

对于 LP 问题的标准形式

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned}$$

其中 $A \in \mathbb{R}^{m \times n}$, 且 $\text{rank}(A) = m \leq n$, 如果存在可行解, 那么存在基本可行解。



证明 任取可行解 $x = (x_1, \dots, x_n)$, 不妨其前 k 个分量为正。如果 A 的前 k 个列向量 A_1, \dots, A_k 线性无关, 那么由引理 1.2.3, x 为基本可行解。如果 A_1, \dots, A_k 线性相关, 那么存在非零向量 $\delta = (\delta_1, \dots, \delta_n, 0, \dots, 0)$, 使得成立

$$\sum_{i=1}^k \delta_i A_i = 0$$

因此

$$A\delta = \sum_{i=1}^k \delta_i A_i = 0$$

取

$$\varepsilon = \min \left\{ \frac{x_i}{|\delta_i|} : \delta_i \neq 0, 1 \leq i \leq k \right\}$$

那么 $x \pm \varepsilon \delta \geq 0$ 。由于

$$A(x \pm \varepsilon \delta) = Ax \pm \varepsilon A\delta = b$$

从而 $x \pm \varepsilon \delta$ 均为可行解。

注意到存在 $1 \leq i \leq k$, 使得成立或 $x_i + \varepsilon \delta_i = 0$, 或 $x_i - \varepsilon \delta_i = 0$, 从而得到可行解 $x + \varepsilon \delta$ 与 $x - \varepsilon \delta$ 之一, 使得其非零分量数 $\leq k-1$ 。如果其仍不为基本可行解, 重复上述过程。容易知道该过程有限, 从而可得到基本

可行解。

定理 1.2.6 (基本可行解的存在性)

对于 LP 问题的标准形式

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

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, 且 $\text{rank}(\mathbf{A}) = m \leq n$, 如果存在最优解, 那么最优解在基本可行解处取到。



证明 假设 \mathbf{x} 为最优解, 如果 \mathbf{x} 为基本可行解, 那么问题得证; 否则由定理 1.2.6, 存在 $\varepsilon > 0$ 与 δ , 使得 $\mathbf{x} \pm \varepsilon \delta$ 均为可行解, 且 $\mathbf{x} + \varepsilon \delta$ 与 $\mathbf{x} - \varepsilon \delta$ 之一的非零分量数比 \mathbf{x} 严格小。其目标函数为

$$\mathbf{c}^T(\mathbf{x} \pm \varepsilon \delta) = \mathbf{c}^T \mathbf{x} \pm \varepsilon \mathbf{c}^T \delta \geq \mathbf{c}^T \mathbf{x}$$

从而 $\varepsilon \mathbf{c}^T = 0$, 因而

$$\mathbf{c}^T(\mathbf{x} \pm \varepsilon \delta) = \mathbf{c}^T \mathbf{x}$$

由定理 1.2.6, 重复上述过程, 可得基本可行解 $\bar{\mathbf{x}}$, 使得成立 $\mathbf{c}^T \bar{\mathbf{x}} = \mathbf{c}^T \mathbf{x}$, 进而最优解在基本可行解处取到。

1.3 单纯形方法

1.3.1 单纯形方法

对于 LP 问题的标准形式

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

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, 且 $\text{rank}(\mathbf{A}) = m \leq n$, 如果 $\bar{\mathbf{x}}$ 为基本可行解, 那么记其对应的基为 \mathbf{B} , 此时 $\mathbf{A}\mathbf{x} = \mathbf{b}$ 化为

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

不妨 $\mathbf{B} = (\mathbf{A}_1, \dots, \mathbf{A}_m)$, 其中 \mathbf{A}_i 为 \mathbf{A} 的第 i 个列向量。记

$$\bar{\mathbf{A}}_i = \mathbf{B}^{-1} \mathbf{A}_i = (\bar{a}_{1i}, \dots, \bar{a}_{mi})^T, \quad 1 \leq i \leq n$$

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

那么

$$\mathbf{x}_B + \sum_{i=m+1}^n x_i \bar{\mathbf{A}}_i = \bar{\mathbf{b}} \quad \text{或} \quad \mathbf{x}_B + \mathbf{B}^{-1} \mathbf{N} \mathbf{x}_N = \bar{\mathbf{b}}$$

引入记号

$$\zeta_i = \mathbf{c}_B^T \bar{\mathbf{A}}_i - c_i, \quad 1 \leq i \leq n$$

或以向量表示

$$\boldsymbol{\zeta}^T = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{A} - \mathbf{c}^T = (\zeta_1, \dots, \zeta_n) = (\boldsymbol{\zeta}_B^T, \boldsymbol{\zeta}_N^T) = (\mathbf{0}, \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{N} - \mathbf{c}_N^T)$$

目标函数化为

$$\begin{aligned}
 \mathbf{c}^T \mathbf{x} &= \mathbf{c}_B^T \mathbf{x}_B + \mathbf{c}_N^T \mathbf{x}_N \\
 &= \mathbf{c}_B^T (\bar{\mathbf{b}} - \mathbf{B}^{-1} \mathbf{N} \mathbf{x}_N) + \mathbf{c}_N^T \mathbf{x}_N \\
 &= \mathbf{c}_B^T \bar{\mathbf{b}} - (\mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{N} - \mathbf{c}_N^T) \mathbf{x}_N \\
 &= \mathbf{c}_B^T \bar{\mathbf{b}} - \sum_{i=m+1}^n (\mathbf{c}_B^T \bar{\mathbf{A}}_i - c_i) x_i \\
 &= \mathbf{c}_B^T \bar{\mathbf{b}} - \zeta^T \mathbf{x} \\
 &= \mathbf{c}^T \bar{\mathbf{x}} - \zeta^T \mathbf{x}
 \end{aligned}$$

从而原 LP 问题化为

$$\begin{aligned}
 \min \quad & \mathbf{c}^T \bar{\mathbf{x}} - \zeta^T \mathbf{x} \\
 \text{s.t.} \quad & \begin{cases} \mathbf{x}_B + \mathbf{B}^{-1} \mathbf{N} \mathbf{x}_N = \bar{\mathbf{b}} \\ \mathbf{x} \geq \mathbf{0} \end{cases}
 \end{aligned}$$

定理 1.3.1 (最优性准则)

如果 $\zeta \leq \mathbf{0}$, 那么 $\bar{\mathbf{x}}$ 为原 LP 问题的最优解。



证明 这当然是显然的!

定理 1.3.2

如果存在 $m+1 \leq k \leq n$, 使得成立 $\zeta_k > 0$, 且 $\bar{\mathbf{A}}_k = \mathbf{B}^{-1} \mathbf{A}_k \geq \mathbf{0}$, 那么原 LP 问题无界。



证明 令

$$\mathbf{d} = \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \mathbf{e}_k$$

其中 \mathbf{e}_k 为第 k 个分量为 1, 其余分量为 0 的 n 维向量。由于 $\bar{\mathbf{A}}_k \leq \mathbf{0}$, 那么 $\mathbf{d} \geq \mathbf{0}$, 而

$$\begin{aligned}
 \mathbf{A} \mathbf{d} &= \begin{pmatrix} \mathbf{B} & \mathbf{N} \end{pmatrix} \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \mathbf{A} \mathbf{e}_k \\
 &= \begin{pmatrix} \mathbf{B} & \mathbf{N} \end{pmatrix} \begin{pmatrix} -\mathbf{B}^{-1} \mathbf{A}_k \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
 &= -\mathbf{A}_k + \mathbf{A}_k \\
 &= \mathbf{0}
 \end{aligned}$$

对于充分大的 $\theta > 0$, 考察 $\bar{\mathbf{x}} + \theta \mathbf{d}$, 此时有

$$\mathbf{A}(\bar{\mathbf{x}} + \theta \mathbf{d}) = \mathbf{A} \bar{\mathbf{x}} + \theta \mathbf{A} \mathbf{d} = \mathbf{b}, \quad \bar{\mathbf{x}} + \theta \mathbf{d} \geq \mathbf{0}$$

因此 $\bar{\mathbf{x}} + \theta \mathbf{d}$ 为原 LP 问题的可行解，其目标函数值为

$$\begin{aligned}
 \mathbf{c}^T(\bar{\mathbf{x}} + \theta \mathbf{d}) &= \mathbf{c}^T \bar{\mathbf{x}} + \theta \mathbf{c}^T \mathbf{d} \\
 &= \mathbf{c}^T \bar{\mathbf{x}} + \theta \begin{pmatrix} \mathbf{c}_B^T & \mathbf{c}_N^T \end{pmatrix} \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \theta \begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
 &= \mathbf{c}^T \bar{\mathbf{x}} - \theta(\mathbf{c}_B^T \bar{\mathbf{A}}_k - c_k) \\
 &= \mathbf{c}^T \bar{\mathbf{x}} - \theta \zeta_k
 \end{aligned}$$

由于 $\zeta_k > 0$ ，而 $\theta > 0$ 可任意大，因此原 LP 问题无下界。

定理 1.3.3

对于非退化的基本可行解 $\bar{\mathbf{x}}$ ，如果存在 $m+1 \leq k \leq n$ ，使得成立 $\zeta_k > 0$ ，且 $\bar{\mathbf{A}}_k$ 中存在正分量，那么存在基本可行解 $\hat{\mathbf{x}}$ ，使得成立 $\mathbf{c}^T \hat{\mathbf{x}} < \mathbf{c}^T \bar{\mathbf{x}}$ 。



证明 令

$$\mathbf{d} = \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \mathbf{e}_k$$

其中 \mathbf{e}_k 为第 k 个分量为 1，其余分量为 0 的 n 维向量。由于 $\bar{\mathbf{A}}_k \leq \mathbf{0}$ ，那么 $\mathbf{d} \geq \mathbf{0}$ ，而

$$\begin{aligned}
 \mathbf{A}\mathbf{d} &= \begin{pmatrix} \mathbf{B} & \mathbf{N} \end{pmatrix} \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \mathbf{A}\mathbf{e}_k \\
 &= \begin{pmatrix} \mathbf{B} & \mathbf{N} \end{pmatrix} \begin{pmatrix} -\mathbf{B}^{-1}\mathbf{A} \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\
 &= -\mathbf{A}_k + \mathbf{A}_k \\
 &= \mathbf{0}
 \end{aligned}$$

令

$$\hat{\mathbf{x}} = \bar{\mathbf{x}} + \theta \mathbf{d} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix} + \theta \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \theta \mathbf{e}_k = \begin{pmatrix} \bar{\mathbf{b}} - \theta \bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \theta \mathbf{e}_k$$

容易知道

$$\mathbf{A}\hat{\mathbf{x}} = \mathbf{A}\bar{\mathbf{x}} + \theta \mathbf{A}\mathbf{d} = \mathbf{b}$$

取

$$\theta = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} : \bar{a}_{ik} > 0, 1 \leq i \leq m \right\} = \frac{\bar{b}_r}{\bar{a}_{rk}}$$

从而 $\hat{\mathbf{x}} \geq \mathbf{0}$ ，进而 $\hat{\mathbf{x}}$ 为可行解。

考察 $\hat{\mathbf{x}}$, 其各分量为

$$\begin{aligned}\hat{x}_i &= \bar{b}_i - \frac{\bar{b}_r}{\bar{a}_{rk}} \bar{a}_{ik}, & 1 \leq i \leq m, i \neq r \\ \hat{x}_i &= 0, & m+1 \leq i \leq n, i \neq k \\ \hat{x}_r &= 0 \\ \hat{x}_k &= \frac{\bar{b}_r}{\bar{a}_{rk}}\end{aligned}$$

若要证明 $\hat{\mathbf{x}}$ 为基本可行解, 由引理 1.2.3, 只需证明 $\mathbf{A}_1, \dots, \mathbf{A}_{r-1}, \mathbf{A}_k, \mathbf{A}_{r+1}, \dots, \mathbf{A}_m$ 线性无关。若不然, $\mathbf{A}_1, \dots, \mathbf{A}_{r-1}, \mathbf{A}_k, \mathbf{A}_{r+1}, \dots, \mathbf{A}_m$ 线性相关, 由于 $\mathbf{A}_1, \dots, \mathbf{A}_m$ 线性无关, 因此向量 \mathbf{A}_k 可由其余 $m-1$ 个向量线性表出, 即存在 $m-1$ 个数 y_i , 其中 $1 \leq i \leq m$ 且 $i \neq r$, 使得成立

$$\mathbf{A}_k = \sum_{\substack{1 \leq i \leq m \\ i \neq r}} y_i \mathbf{A}_i$$

又因为 $\bar{\mathbf{A}}_k = \mathbf{B}^{-1} \mathbf{A}_k$, 因此

$$\mathbf{A}_k = \mathbf{B} \bar{\mathbf{A}}_k = \begin{pmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_m \end{pmatrix} \begin{pmatrix} \bar{a}_{1k} \\ \vdots \\ \bar{a}_{mk} \end{pmatrix} = \sum_{i=1}^m \bar{a}_{ik} \mathbf{A}_i$$

两式作差

$$\bar{a}_{rk} \mathbf{A}_r + \sum_{\substack{1 \leq i \leq m \\ i \neq r}} (\bar{a}_{ik} - y_i) \mathbf{A}_i = \mathbf{0}$$

由于 $\bar{a}_{rk} \neq 0$, 因此 $\mathbf{A}_1, \dots, \mathbf{A}_m$ 线性相关, 产生矛盾! 进而 $\hat{\mathbf{x}}$ 为基本可行解。

由非退化假设 $\bar{\mathbf{b}} > \mathbf{0}$, 因此 $\theta = \bar{b}_r / \bar{a}_{rk} > 0$, 从而

$$\begin{aligned}\mathbf{c}^T \hat{\mathbf{x}} &= \mathbf{c}^T (\bar{\mathbf{x}} + \theta \mathbf{d}) \\ &= \mathbf{c}^T \bar{\mathbf{x}} + \theta \mathbf{c}^T \mathbf{d} \\ &= \mathbf{c}^T \bar{\mathbf{x}} + \theta \left(\mathbf{c}_B^T \quad \mathbf{c}_N^T \right) \begin{pmatrix} -\bar{\mathbf{A}}_k \\ \mathbf{0} \end{pmatrix} + \theta \begin{pmatrix} c_1 & \cdots & c_n \end{pmatrix} \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \\ &= \mathbf{c}^T \bar{\mathbf{x}} - \theta (\mathbf{c}_B^T \bar{\mathbf{A}}_k - c_k) \\ &= \mathbf{c}^T \bar{\mathbf{x}} - \theta \zeta_k \\ &< \mathbf{c}^T \bar{\mathbf{x}}\end{aligned}$$

定理 1.3.4 (单纯形方法)

1. 选择基 \mathbf{B} , 使得对应 $\mathbf{b} \geq \mathbf{0}$ 。
2. 求出 $\zeta_k = \max_{1 \leq i \leq n} \zeta_i$ 。
3. 若 $\zeta_k \leq 0$, 停止, 已找到最优解; 否则转至第 4 步。
4. 若 $\bar{\mathbf{A}}_{\cdot, k} \leq \mathbf{0}$, 原问题无界; 否则转至第 5 步。
5. 求出 $\frac{\bar{b}_r}{\bar{a}_{rk}} = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} : \bar{a}_{ik} > 0, 1 \leq i \leq m \right\}$ 。

6. 以 \bar{a}_{rk} 为转轴元作旋转变换，转至第 2 步。



定理 1.3.5 (对偶单纯形方法)

1. 选择基 B ，使得对应 $b \leq 0$ 。
2. 求 $\bar{b}_r = \min_{1 \leq i \leq m} \bar{b}_i$ 。
3. 若 $\bar{b}_r \geq 0$ ，停止，已找到最优解；否则转至第 4 步。
4. 若 $\bar{A}_{r,\cdot} \geq 0$ ，原问题无解；否则转至第 5 步。
5. 求出 $\frac{\zeta_k}{\bar{a}_{rk}} = \min \left\{ \frac{\zeta_j}{\bar{a}_{rj}} : \bar{a}_{rj} < 0, 1 \leq j \leq n \right\}$ 。
6. 以 \bar{a}_{rk} 为转轴元作旋转变换，转至第 2 步。



序号	单纯形方法	对偶单纯形方法
1	选择基 B ，使得对应 $b \geq 0$	选择基 B ，使得对应 $b \leq 0$
2	求出 $\zeta_k = \max_{1 \leq i \leq n} \zeta_i$	求 $\bar{b}_r = \min_{1 \leq i \leq m} \bar{b}_i$
3	若 $\zeta_k \leq 0$ ，停止，已找到最优解；否则转至第 4 步	若 $\bar{b}_r \geq 0$ ，停止，已找到最优解；否则转至第 4 步
4	若 $\bar{A}_{\cdot,k} \leq 0$ ，原问题无界；否则转至第 5 步	若 $\bar{A}_{r,\cdot} \geq 0$ ，原问题无解；否则转至第 5 步
5	求出 $\frac{\bar{b}_r}{\bar{a}_{rk}} = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} : \bar{a}_{ik} > 0, 1 \leq i \leq m \right\}$	求出 $\frac{\zeta_k}{\bar{a}_{rk}} = \min \left\{ \frac{\zeta_j}{\bar{a}_{rj}} : \bar{a}_{rj} < 0, 1 \leq j \leq n \right\}$
6	以 \bar{a}_{rk} 为转轴元作旋转变换，转至第 2 步	以 \bar{a}_{rk} 为转轴元作旋转变换，转至第 2 步

表 1.1: 单纯形表

	x_1	\cdots	x_n	RHS
	ζ	\cdots	ζ_n	z
x_{r_1}	\bar{A}_1	\cdots	\bar{A}_n	\bar{b}_1
\vdots				\vdots
x_{r_n}				\bar{b}_m

1.4 初始解

对于 LP 问题的标准形式

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & \begin{cases} Ax = b \\ x \geq 0 \end{cases} \end{aligned}$$

其中 $A \in \mathbb{R}^{m \times n}$ ，不妨 $b \geq 0$ 。增加辅助变量 $x_\alpha = (x_{n+1}, \cdots, x_{n+m})^T$ ，考察辅助 LP 问题

$$\begin{aligned} \min \quad & \sum_{k=1}^m x_{n+k} \\ \text{s.t.} \quad & \begin{cases} Ax + x_\alpha = b \\ x, x_\alpha \geq 0 \end{cases} \end{aligned}$$

记原 LP 问题的可行域为 D ，辅助 LP 问题的可行域为 D' ，那么

$$\mathbf{x} \in D \iff \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \in D' \iff \mathbf{x}_\alpha = \mathbf{0} \iff \min \sum_{k=1}^m x_{n+k} = 0$$

辅助 LP 问题为 $m+n$ 个变量的标准形式的线性规划，且人工变量对应的 m 列构成 m 阶单位矩阵，又因为 $\mathbf{b} \geq \mathbf{0}$ ，那么辅助 LP 问题的存在基本可行解 $\mathbf{x} = \mathbf{0}, \mathbf{x}_\alpha = \mathbf{b}$ ，又因为目标函数存在下界，那么辅助 LP 问题存在最优解。

1. 最优值 $\sum_{k=1}^m x_{n+k} = 0$ 且人工变量 \mathbf{x}_α 均不为基变量： $\bar{\mathbf{x}}$ 为原 LP 问题的基本可行解。
2. 最优值 $\sum_{k=1}^m x_{n+k} > 0$ ：原 LP 问题无可行解 $D = \emptyset$ 。
3. 最优值 $\sum_{k=1}^m x_{n+k} = 0$ 且存在人工变量 $x_{n+k}, 1 \leq k \leq m$ 为基变量：考察基变量 x_{n+k} 所在第 r 行。
 - 若存在 $1 \leq i \leq n$ ，使得成立 $a_{ik} \neq 0$ ，则将 a_{ik} 作为转轴元作旋转变换。
 - 若对于任意 $1 \leq i \leq n$ ，使得成立 $a_{ik} = 0$ ，则将 r 行删除。

1.5 对偶性及对偶单纯形法

1.5.1 对偶线性规划

定义 1.5.1 (一般形式线性规划问题的对偶线性规划问题)

$$\begin{array}{ll}
 \min & c_1 x_1 + \cdots + c_n x_n \\
 \text{s.t.} & \begin{cases} \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{p1}x_1 + \cdots + a_{pn}x_n = b_p \\ a_{p+1,1}x_1 + \cdots + a_{p+1,n}x_n \geq b_{p+1} \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \geq b_m \end{cases} \\ \begin{cases} x_1 \geq 0 \\ \vdots \\ x_q \geq 0 \\ x_{q+1} \text{ 自由} \\ \vdots \\ x_n \text{ 自由} \end{cases} \end{cases}
 \end{array}
 \iff
 \begin{array}{ll}
 \max & b_1 w_1 + \cdots + b_m w_m \\
 \text{s.t.} & \begin{cases} \begin{cases} w_1 \text{ 自由} \\ \vdots \\ w_p \text{ 自由} \\ w_{p+1} \geq 0 \\ \vdots \\ w_m \geq 0 \end{cases} \\ \begin{cases} a_{11}w_1 + \cdots + a_{m1}w_m \leq c_1 \\ \vdots \\ a_{1q}w_1 + \cdots + a_{mq}w_m \leq c_q \\ a_{1,q+1}w_1 + \cdots + a_{m,q+1}w_m = c_{q+1} \\ \vdots \\ a_{1n}w_1 + \cdots + a_{mn}w_m = c_n \end{cases} \end{cases}
 \end{array}$$



定义 1.5.2 (一般形式线性规划问题的对偶线性规划问题的矩阵表示)

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \begin{cases} \mathbf{Ax} \sim \begin{pmatrix} = b_p \\ \geq b_{m-p} \end{pmatrix} \\ \begin{pmatrix} \mathbf{x}_q \\ \mathbf{x}_{n-q} \end{pmatrix} \sim \begin{pmatrix} \geq \mathbf{0} \\ \text{自由} \end{pmatrix} \end{cases} \end{array} \iff \begin{array}{ll} \max & \mathbf{b}^T \mathbf{w} \\ \text{s.t.} & \begin{cases} \begin{pmatrix} \mathbf{w}_p \\ \mathbf{w}_{m-p} \end{pmatrix} \sim \begin{pmatrix} \text{自由} \\ \geq \mathbf{0} \end{pmatrix} \\ \mathbf{A}^T \mathbf{w} \sim \begin{pmatrix} \leq \mathbf{c}_q \\ = \mathbf{c}_{n-p} \end{pmatrix} \end{cases} \end{array}$$



定义 1.5.3 (规范形式线性规划问题的对偶线性规划问题)

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \begin{cases} \mathbf{Ax} \geq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{array} \iff \begin{array}{ll} \max & \mathbf{b}^T \mathbf{w} \\ \text{s.t.} & \begin{cases} \mathbf{A}^T \mathbf{w} \leq \mathbf{c} \\ \mathbf{w} \geq \mathbf{0} \end{cases} \end{array}$$



定义 1.5.4 (标准形式线性规划问题的对偶线性规划问题)

$$\begin{array}{ll} \min & \mathbf{c}^T \mathbf{x} \\ \text{s.t.} & \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{array} \iff \begin{array}{ll} \max & \mathbf{b}^T \mathbf{w} \\ \text{s.t.} & \mathbf{A}^T \mathbf{w} \leq \mathbf{c} \end{array}$$



1.5.2 对偶理论

定理 1.5.1

如果 LP 问题存在最优解，那么其对偶问题存在最优解，且最有值相等。



推论 1.5.1

如果 \mathbf{x} 与 \mathbf{w} 分别为原始问题与其对偶问题的可行解，那么

$$\mathbf{x} \text{ 与 } \mathbf{w} \text{ 分别为原始问题与其对偶问题的最优解} \iff \mathbf{c}^T \mathbf{x} = \mathbf{w}^T \mathbf{b}$$



定理 1.5.2

对偶问题的对偶问题为原始问题。



定理 1.5.3

对于原始问题与其对偶问题，成立且仅成立如下情况之一。

1. 原始问题与对偶问题均存在最优解。
2. 原始问题与对偶问题均不存在可行解。
3. 原始问题无界且对偶问题不存在可行解。
4. 原始问题不存在可行解且对偶问题无界。



定理 1.5.4 (互补松紧性)

如果 x 与 w 分别为原始问题与其对偶问题的可行解, 那么

x 与 w 分别为原始问题与其对偶问题的最优解

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

其中 \cdot 表示矩阵的点乘。

**定理 1.5.5 (对偶单纯形方法)**

1. 列出初始单纯形表。
2. 求 $\bar{b}_r = \min_{1 \leq i \leq m} \bar{b}_i$ 。
3. 若 $\bar{b}_r \geq 0$, 停止, 已找到最优解。
4. 若对于任意 $1 \leq j \leq n$, $\bar{a}_{rj} \geq 0$, 原问题无解。
5. 求出 $\frac{\zeta_k}{\bar{a}_{rk}} = \min \left\{ \frac{\zeta_j}{\bar{a}_{rj}} : \bar{a}_{rj} < 0, 1 \leq j \leq n \right\}$ 。
6. 以 \bar{a}_{rk} 为转轴元作旋转变换, 转至第 2 步。



1.6 灵敏度分析

1.6.1 改变价值向量

当价值向量 c 仅存在一个分量 c_k 变为 c'_k 时, 可依下述方法处理。

1. x_k 为非基变量: 只有 ζ_k 变化, 新的检验数为

$$\zeta'_k = \zeta_k + (c_k - c'_k)$$

- $\zeta'_k \leq 0$: x 仍为新问题的最优解。
- $\zeta'_k > 0$: 进行单纯形迭代。

2. x_k 为基变量: 将基变量 x_k 所在的行 $\times (c'_k - c_k)$ 加至第 0 行, 再令 $\zeta'_k = 0$, 最后进行单纯形迭代。

1.6.2 改变右端向量

当右端 b 改为 b' 时, 找到原规划问题的最优解对应的基矩阵 B , 那么最后的单纯性表变为 $\bar{b}' = B^{-1}b'$, $z'_0 = c_B^T \bar{b}'$ 。

第二章 整数线性规划

2.1 整数线性规划问题

定义 2.1.1 (整数线性规划 Integer Linear Programming, ILP)

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \text{ 中某分量为整数} \end{cases} \end{aligned}$$

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ 。



2.2 Gomory 割平面法

2.2.1 Gomory 割平面法的基本思想

考虑纯整数线性规划问题

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \in \mathbb{N}^n \end{cases} \end{aligned} \quad (\mathbf{P})$$

其中 $\mathbf{A} \in \mathbb{Z}^{m \times n}$, $\mathbf{c} \in \mathbb{Z}^n$, $\mathbf{b} \in \mathbb{Z}^m$ 。考虑其松弛问题

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{aligned} \quad (\mathbf{P}_0)$$

记 (\mathbf{P}) 的可行域为 D , (\mathbf{P}_0) 的可行域为 D_0 , 那么成立

1. $D \subset D_0$
2. 若 (\mathbf{P}_0) 无可行解, 则 (\mathbf{P}) 无可行解。
3. (\mathbf{P}_0) 的最优值 $\leq (\mathbf{P})$ 的最优值
4. 若 (\mathbf{P}_0) 的最优解 $\mathbf{x}^0 \in \mathbb{N}^n$, 那么 \mathbf{x}^0 是 (\mathbf{P}) 的最优解。

2.2.2 Gomory 割平面法计算步骤

定理 2.2.1 (Gomory 割平面法)

1. 使用单纯形法求解 ILP 问题 (\mathbf{P}) 的松弛问题 (\mathbf{P}_0) :
 - 若 (\mathbf{P}_0) 不存在最优解, 则计算停止。
 - 若 (\mathbf{P}_0) 存在最优解 \mathbf{x}^0 , 且 \mathbf{x}^0 为整数向量, 则 \mathbf{x}^0 为 (\mathbf{P}) 的最优解。
 - 若 (\mathbf{P}_0) 存在最优解 \mathbf{x}^0 , 且 \mathbf{x}^0 不为整数向量, 则转至步骤 2。

2. 求割平面方程：任取 \mathbf{x}^0 的非整数分量 \bar{b}_k ，得到割平面方程

$$-\sum_{j \in S} f_{kj} x_j + s = -f_k$$

3. 将割平面方程加到步骤 1 所得最优单纯形表中，用对偶单纯形法求解此松弛问题：

- 若存在最优解且最优解为整数解，则其为 (P) 的最优解。
- 若存在最优解且最优解不为整数解，则重新记其最优解为 \mathbf{x}^0 ，返回步骤 2。
- 若不存在最优解，则 (P) 不可行。



2.3 分枝定界法

考虑整数线性规划问题

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \in \mathbb{N}^n \end{cases} \end{aligned} \quad (\text{P})$$

其中 $\mathbf{A} \in \mathbb{Z}^{m \times n}$ ， $\mathbf{c} \in \mathbb{Z}^n$ ， $\mathbf{b} \in \mathbb{Z}^m$ 。考虑其松弛问题

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{aligned} \quad (\text{P}_0)$$

记 (P_0) 的最优解为 \mathbf{x}^0 ，若 \mathbf{x}^0 的分量 x_k^0 不为整数，那么 (P) 的最优解的第 k 个分量 x_k 或成立 $x_k \leq [x_k^0]$ ，或成立 $x_k \geq [x_k^0] + 1$ 。因此 (P) 可分解为两个子问题

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \in \mathbb{N}^n \\ x_k \leq [x_k^0] \end{cases} \end{aligned} \quad (\text{P}_2)$$

和

$$\begin{aligned} \min \quad & z = \mathbf{c}^T \mathbf{x} \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} \leq \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \\ \mathbf{x} \in \mathbb{N}^n \\ x_k \geq [x_k^0] + 1 \end{cases} \end{aligned} \quad (\text{P}_2)$$

重复上述过程。

第三章 非线性规划

3.1 基本概念

3.1.1 非线性规划问题

定义 3.1.1 (数学规划 Mathematical Programming, MP)

定义数学规划为

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{aligned} \quad (*)$$

定义其可行域为

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \right\}$$

定义 3.1.2 (整体最优解)

对于数学规划 (*), 称 $\mathbf{x}^* \in X$ 为其整体最优解, 如果对于任意 $\mathbf{x} \in X$, 成立

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$

定义 3.1.3 (局部最优解)

对于数学规划 (*), 称 $\mathbf{x}^* \in X$ 为其局部最优解, 如果存在 $\delta > 0$, 使得对于任意 $\mathbf{x} \in B_\delta(\mathbf{x}^*) \cap X$, 成立

$$f(\mathbf{x}^*) \leq f(\mathbf{x})$$

3.1.2 非线性规划方法概述

定义 3.1.4 (下降方向)

称 $\mathbf{p} \in \mathbb{R}^n$ 为函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$ 在点 $\bar{\mathbf{x}} \in \mathbb{R}^n$ 处的下降方向, 如果存在 $\delta > 0$, 使得对于任意 $0 < t < \delta$, 使得成立

$$f(\bar{\mathbf{x}} + t\mathbf{p}) < f(\bar{\mathbf{x}})$$

特别的, 若 f 在点 $\bar{\mathbf{x}}$ 处可偏导, 那么 $-\nabla f(\bar{\mathbf{x}})$ 为 f 在点 $\bar{\mathbf{x}}$ 处下降最快的方向。

定义 3.1.5 (可行方向)

对于数学规划 (*), 称 $\mathbf{p} \in \mathbb{R}^n$ 为点 $\bar{\mathbf{x}}$ 处关于可行域 X 的可行方向, 如果存在 $t > 0$, 使得成立

$$\bar{\mathbf{x}} + t\mathbf{p} \in X$$

定义 3.1.6 (可行下降方向)

称 \mathbf{p} 为可行下降方向, 如果既为下降方向, 又为可行方向。

定义 3.1.7 (基本迭代格式)

1. 选取初始点 \mathbf{x}^0 。
2. 如果已得到可行点 \mathbf{x}^k 。
3. 构造可行下降方向 \mathbf{p}^k 。
4. 确定步长 t_k 。
5. 令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ 。

**定义 3.1.8 (常用停止迭代规则)**

1.
$$\|\mathbf{x}^{k+1} - \mathbf{x}^k\| < \varepsilon$$
2.
$$\frac{\|\mathbf{x}^{k+1} - \mathbf{x}^k\|}{\|\mathbf{x}^k\|} < \varepsilon$$
3.
$$\|\nabla f(\mathbf{x}^k)\| < \varepsilon$$



3.2 凸函数和凸规划

3.2.1 凸函数及其性质

定义 3.2.1 (凸函数)

对于凸集 $S \subset \mathbb{R}^n$, 称 $f: S \rightarrow \mathbb{R}$ 为凸函数, 如果成立

$$\forall \mathbf{x}, \mathbf{y} \in S, \forall \lambda \in [0, 1] \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

**定义 3.2.2 (严格凸函数)**

对于凸集 $S \subset \mathbb{R}^n$, 称 $f: S \rightarrow \mathbb{R}$ 为严格凸函数, 如果成立

$$\forall \mathbf{x} \neq \mathbf{y} \in S, \forall \lambda \in [0, 1] \implies f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) < \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

**定理 3.2.1 (凸函数的运算)**

1. 对于凸集 $S \subset \mathbb{R}^n$ 上的凸函数 $f: S \rightarrow \mathbb{R}$, 如果 $\lambda \geq 0$, 那么 λf 为 S 上的凸函数。
2. 对于凸集 $S \subset \mathbb{R}^n$ 上的凸函数 $f, g: S \rightarrow \mathbb{R}$, $f + g$ 为 S 上的凸函数。



 **笔记** 凸函数的积未必为凸函数, 例如:

$$f(x) = x, \quad g(x) = -x$$

定理 3.2.2

对于凸集 $S \subset \mathbb{R}^n$ 上的凸函数 $f: S \rightarrow \mathbb{R}$, 集合

$$H_S(f, c) = \{\mathbf{x} \in S : f(\mathbf{x}) \leq c\}$$

为凸集。



证明 任取 $\mathbf{x}, \mathbf{y} \in H_S(f, c) \subset S$ 与 $\lambda \in [0, 1]$, 由于 S 为凸集, 那么

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

由于 f 为凸函数, 那么

$$f(\lambda \mathbf{x} + (1 - \lambda)\mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda)f(\mathbf{y}) \leq \lambda c + (1 - \lambda)c = c$$

因此 $\lambda \mathbf{x} + (1 - \lambda)\mathbf{y} \in H_S(f, c)$, 从而 $H_S(f, c)$ 为凸集。

定理 3.2.3

对于凸集 $S \subset \mathbb{R}^n$ 上的可微函数 $f: S \rightarrow \mathbb{R}$, f 为凸函数的充分必要条件为: 对于任意 $\mathbf{x}, \mathbf{y} \in S$, 成立

$$\nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) \leq f(\mathbf{y}) - f(\mathbf{x})$$



定理 3.2.4

对于凸集 $S \subset \mathbb{R}^n$ 上的可微函数 $f: S \rightarrow \mathbb{R}$, f 为严格凸函数的充分必要条件为: 对于任意 $\mathbf{x} \neq \mathbf{y} \in S$, 成立

$$\nabla f(\mathbf{x})(\mathbf{y} - \mathbf{x}) < f(\mathbf{y}) - f(\mathbf{x})$$



定理 3.2.5

对于凸集 $S \subset \mathbb{R}^n$ 上的二阶连续可偏导函数 $f: S \rightarrow \mathbb{R}$, f 为凸函数的充分必要条件为: f 的 Hessian 矩阵 $\nabla^2 f(\mathbf{x})$ 在 S 上是半正定矩阵, 其中

$$\nabla^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n} \end{pmatrix}$$



定理 3.2.6

对于凸集 $S \subset \mathbb{R}^n$ 上的二阶连续可偏导函数 $f: S \rightarrow \mathbb{R}$, 如果 f 的 Hessian 矩阵 $\nabla^2 f(\mathbf{x})$ 在 S 上为正定矩阵, 那么 f 为严格凸函数。



命题 3.2.1 (二次凸函数)

$f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c$ 为 \mathbb{R}^n 上的凸函数, 其中 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 为正定矩阵, $\mathbf{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$ 。



证明 令 $\mathbf{A} = (a_{ij})$ 与 $\mathbf{b} = (b_1, \dots, b_n)^T$, 那么

$$f(x_1, \dots, x_n) = \frac{1}{2} \sum_{i,j=1}^n a_{ij} x_i x_j + \sum_{k=1}^n b_k x_k + c$$

因此一阶偏导数为

$$\frac{\partial f}{\partial x_k} = \sum_{l=1}^n a_{kl} x_l + b_k, \quad 1 \leq k \leq n$$

二阶偏导数为

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = a_{ij}, \quad 1 \leq i, j \leq n$$

因此

$$\nabla f(\mathbf{x}) = \mathbf{A} \mathbf{x} + \mathbf{b}, \quad \nabla^2 f(\mathbf{x}) = \mathbf{A}$$

进而 $f(\mathbf{x})$ 为 \mathbb{R}^n 上的凸函数。

3.2.2 凸规划及其性质

定义 3.2.3 (凸规划)

称数学规划

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{aligned} \quad (*)$$

为凸规划, 如果可行域

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \right\}$$

为凸集, 且 f 为 X 上的凸函数。



定理 3.2.7

对于数学规划 (*), 如果 g_i 均为 \mathbb{R}^n 上的凸函数, h_j 均为线性函数, 且 f 为 X 上的凸函数, 那么该数学规划为凸规划。



证明 记

$$S = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, 1 \leq i \leq p\}$$

$$T = \{\mathbf{x} \in \mathbb{R}^n : h_j(\mathbf{x}) = 0, 1 \leq j \leq q\}$$

由于 g_i 均为 \mathbb{R}^n 上的凸函数, 那么由定理 3.2.2

$$H^i(g_i, 0) = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0\}, \quad 1 \leq i \leq p$$

为凸集。由引理 1.2.1

$$S = \bigcap_{i=1}^p H^i(g_i, 0)$$

为凸集。由于 h_j 均为线性函数, 从而 T 为凸集。由引理 1.2.1, $X = S \cap T$ 为凸集, 进而该数学规划为凸规划。

定理 3.2.8

凸规划的局部最优解为整体最优解。



证明 设 \mathbf{x}^* 为凸规划 (*) 的局部最优解, 那么存在 $\delta > 0$, 使得成立

$$f(\mathbf{x}^*) \leq f(\mathbf{x}), \quad \mathbf{x} \in B_\delta(\mathbf{x}^*) \cap X$$

如果 \mathbf{x}^* 不为整体最优解, 那么存在 $\bar{\mathbf{x}} \in X$, 使得成立

$$f(\bar{\mathbf{x}}) < f(\mathbf{x}^*)$$

取 $\lambda > 0$, 使得

$$\lambda \bar{\mathbf{x}} + (1 - \lambda) \mathbf{x}^* \in X \cap B_\delta(\mathbf{x}^*)$$

又因为 f 为凸函数, 所以

$$f(\lambda \bar{\mathbf{x}} + (1 - \lambda) \mathbf{x}^*) \leq \lambda f(\bar{\mathbf{x}}) + (1 - \lambda) f(\mathbf{x}^*) < f(\mathbf{x}^*) + (1 - \lambda) f(\mathbf{x}^*) = f(\mathbf{x}^*)$$

矛盾!

3.3 一维搜索方法

定义 3.3.1 (一维搜索问题)

定义目标函数为单变量的非线性规划问题为一维搜索问题，其数学模型为

$$\min_{t \geq 0} \varphi(t)$$

特别的，称数学模型

$$\min_{0 \leq t \leq T} \varphi(t)$$

对应的一维搜索问题为有效一维搜索问题。



3.3.1 0.618 法

定义 3.3.2 (单谷函数)

称 $[a, b]$ 上的函数 $\varphi(t)$ 为单谷函数，如果存在 $t^* \in [a, b]$ ，使得 $\varphi(t)$ 在 $[a, t^*]$ 上严格单调递减，在 $[t^*, b]$ 上严格单调递增。



定理 3.3.1 (0.618 法)

求解一维搜索问题

$$\min_{a \leq t \leq b} \varphi(t)$$

其中 $\varphi(t)$ 为 $[a, b]$ 上的单谷函数。

1. 确定单谷区间 $[a, b]$ ，给定精度 $\varepsilon > 0$ 。
2. 计算最初两个探索点

$$t_1 = b - 0.618(b - a), \quad t_2 = a + 0.618(b - a)$$

并计算 $\varphi_1 = \varphi(t_1)$ 与 $\varphi_2 = \varphi(t_2)$ 。

3. 若 $\varphi_1 \leq \varphi_2$ ，转至第 4 步；否则转至第 5 步。
4. 若 $t_2 - a \leq \varepsilon$ ，停止迭代，输出 t_1 ；否则令

$$b := t_2, \quad t_2 := t_1, \quad t_1 := b - 0.618(b - a), \quad \varphi_2 := \varphi_1, \quad \varphi_1 := \varphi(t_1)$$

转至第 3 步。

5. 若 $b - t_1 \leq \varepsilon$ ，停止迭代，输出 t_2 ；否则令

$$a := t_1, \quad t_1 := t_2, \quad t_2 := a + 0.618(b - a), \quad \varphi_1 := \varphi_2, \quad \varphi_2 := \varphi(t_2)$$

转至第 3 步。



3.3.2 Newton 法

定理 3.3.2 (Newton 法)

求解一维搜索问题

$$\min_{a \leq t \leq b} \varphi(t)$$

其中 $\varphi(t)$ 二阶可微，且 $\varphi''(t) \neq 0$ 。

1. 给定初始点 t_1 与精度 $\varepsilon > 0$ ，令 $n := 1$ 。

2. 如果 $|\varphi'(t_n)| < \varepsilon$, 停止迭代, 输出 t_n ; 否则, 当 $\varphi''(t_n) = 0$ 时, 停止, 解题失败; 当 $\varphi''(t_n) \neq 0$ 时, 转至第 3 步。

3. 计算

$$t_{n+1} = t_n - \frac{\varphi'(t_n)}{\varphi''(t_n)}$$

如果 $|t_{n+1} - t_n| < \varepsilon$, 停止迭代, 输出 t_{n+1} ; 否则, 令 $n := n + 1$, 转至第 2 步。



3.4 无约束最优化方法

定义 3.4.1 (无约束非线性规划问题)

定义无约束非线性规划问题的数学模型为

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x}) \quad (\text{UMP})$$



3.4.1 无约束问题的最优性条件

定理 3.4.1

对于函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 如果 f 在 $\bar{\mathbf{x}} \in \mathbb{R}^n$ 处可微, 且存在 $\mathbf{p} \in \mathbb{R}^n$, 使得成立

$$\nabla f(\bar{\mathbf{x}})^T \mathbf{p} < 0$$

那么 \mathbf{p} 为 f 在 $\bar{\mathbf{x}}$ 处的下降方法。



定理 3.4.2

对于函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 如果 f 在 $\mathbf{x}^* \in \mathbb{R}^n$ 处可微, 且 \mathbf{x}^* 为 (UMP) 的局部最优解, 那么

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$



定理 3.4.3

对于函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 如果 f 在 $\mathbf{x}^* \in \mathbb{R}^n$ 处存在 Hessian 矩阵 $\nabla^2 f(\mathbf{x}^*)$, 且

$$\nabla f(\mathbf{x}^*) = \mathbf{0}, \quad \nabla^2 f(\mathbf{x}^*) \text{ 为正定矩阵}$$

那么 \mathbf{x}^* 为 (UMP) 的严格局部最优解。



定理 3.4.4

对于可微凸函数 $f: \mathbb{R}^n \rightarrow \mathbb{R}$, 如果

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

那么 \mathbf{x}^* 为 (UMP) 的整体最优解。



3.4.2 最速下降法

定理 3.4.5 (最速下降法)

求解无约束非线性规划问题

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

1. 选出初始点 \mathbf{x}^0 , 给定终止误差 $\varepsilon > 0$, 令 $k := 0$ 。
2. 计算 $\mathbf{p}^k = -\nabla f(\mathbf{x}^k)$ 。
3. 如果 $\|\mathbf{p}^k\| \leq \varepsilon$, 停止迭代, 输出 \mathbf{x}^k ; 否则进行第 4 步。
4. 进行一维搜索, 求 t_k , 使得

$$f(\mathbf{x}^k + t_k \mathbf{p}^k) = \min_{t \geq 0} f(\mathbf{x}^k + t \mathbf{p}^k)$$

令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ 与 $k := k + 1$, 转至第 2 步。



3.4.3 共轭方向法

定义 3.4.2 (共轭方向)

对于 n 阶实对称矩阵 \mathbf{A} , 称非零向量 $\mathbf{p}, \mathbf{q} \in \mathbb{R}^n$ 相互 \mathbf{A} 共轭, 如果

$$\mathbf{p}^T \mathbf{A} \mathbf{q} = 0$$



定理 3.4.6

对于 n 阶实对称矩阵 \mathbf{A} , 如果非零向量 $\mathbf{p}^0, \dots, \mathbf{p}^{n-1} \in \mathbb{R}^n$ 相互 \mathbf{A} 共轭, 那么 $\mathbf{p}^0, \dots, \mathbf{p}^{n-1}$ 线性无关。



定义 3.4.3 (二次严格凸函数的无约束最优化问题)

$$\min f(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T \mathbf{A} \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \quad (\text{AP})$$

其中 $\mathbf{A} \in \mathbb{R}^{n \times n}$ 为正定矩阵, $\mathbf{b} \in \mathbb{R}^n$, $c \in \mathbb{R}$ 。



定理 3.4.7

对于问题 (AP), 如果非零向量 $\mathbf{p}^0, \dots, \mathbf{p}^{n-1} \in \mathbb{R}^n$ 相互 \mathbf{A} 共轭, 那么对于任意初始点 $\mathbf{x}^0 \in \mathbb{R}^n$, 依次沿 $\mathbf{p}^0, \dots, \mathbf{p}^{n-1}$ 进行精确一维搜索, 至多经过 n 轮迭代可达到 (AP) 的整体最优解。



定义 3.4.4 (共轭方向法)

称从任意初始点 $\mathbf{x}^0 \in \mathbb{R}^n$ 出发, 依次沿某组共轭方向进行一维搜索求解 (UMP) 的方法为共轭方向法。



定义 3.4.5 (具有二次终止性的方法)

称经过有限轮迭代可以达到 (AP) 的最优解的方法为具有二次终止性的方法。



定理 3.4.8 (Fletcher-Reeves 共轭梯度法)

求解无约束非线性规划问题

$$\min_{\mathbf{x} \in \mathbb{R}^n} f(\mathbf{x})$$

1. 选取初始点 \mathbf{x}^0 , 给定终止误差 $\varepsilon > 0$, 与迭代周期 n 。

2. 计算 $\nabla f(\mathbf{x}^0)$, 若 $\|\nabla f(\mathbf{x}^0)\| \leq \varepsilon$, 停止迭代, 输出 \mathbf{x}^0 ; 否则, 进行第 3 步。
3. 取 $\mathbf{p}^0 = -\nabla f(\mathbf{x}^0)$, 令 $k := 0$ 。
4. 进行一维搜索, 求 t_k , 使得

$$f(\mathbf{x}^k + t_k \mathbf{p}^k) = \min_{t \geq 0} f(\mathbf{x}^k + t \mathbf{p}^k)$$

令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$ 。

5. 计算 $\nabla f(\mathbf{x}^{k+1})$, 若 $\|\nabla f(\mathbf{x}^{k+1})\| \leq \varepsilon$, 停止迭代, 输出 \mathbf{x}^{k+1} ; 否则, 进行第 6 步。
6. 若 $k+1 = n$, 令 $\mathbf{x}^0 := \mathbf{x}^n$, 转第 3 步; 否则进行第 7 步。
7. 使用 F-R 公式取

$$\mathbf{p}^{k+1} = -\nabla f(\mathbf{x}^{k+1}) + \lambda_k \mathbf{p}^k, \quad \lambda_k = \frac{\|\nabla f(\mathbf{x}^{k+1})\|^2}{\|\nabla f(\mathbf{x}^k)\|^2}$$

令 $k := k+1$, 转第 4 步。



3.5 约束最优化方法

定义 3.5.1 (约束非线性规划问题)

约束非线性规划问题:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{aligned} \quad (\text{MP})$$

可行域:

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \mid \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \right\}$$

记

$$I = \{1, 2, \dots, p\}, \quad J = \{1, 2, \dots, q\}$$



3.5.1 约束最优化问题的最优性条件

定义 3.5.2 (积极约束)

对于 (MP) 问题的可行域中的可行点 $\mathbf{x}^* \in X$, 称使得成立 $g_i(\mathbf{x}) = 0$ 的约束 $g_i(\mathbf{x}) \leq 0$ 为 \mathbf{x}^* 的积极约束, 记 \mathbf{x}^* 的积极约束下标集为

$$I(\mathbf{x}^*) = \{i : g_i(\mathbf{x}) = 0, 1 \leq i \leq p\}$$



定理 3.5.1 (Kuhn-Tucker 条件)

对于 (MP) 问题, 设 $\mathbf{x}^* \in X$, $f, g_i, i \in I(\mathbf{x}^*)$ 在 \mathbf{x}^* 处可微, $g_i, i \in I \setminus I(\mathbf{x}^*)$ 在 \mathbf{x}^* 处连续, $h_j, j \in J$ 在 \mathbf{x}^* 处连续可微, 且各 $\nabla g_i(\mathbf{x}^*), i \in I; \nabla h_j(\mathbf{x}^*), j \in J$ 线性无关。若 \mathbf{x}^* 为 (MP) 的局部最优解, 则存在两组实数 $\lambda_i^*, i \in I(\mathbf{x}^*); \mu_j^*, j \in J$, 使得成立

$$\begin{cases} \nabla f(\mathbf{x}^*) + \sum_{i \in I(\mathbf{x}^*)} \lambda_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j \in J} \mu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \\ \lambda_i^* \geq 0, & 1 \leq i \leq p \end{cases}$$

若 $g_i, i \in I$ 在 \mathbf{x}^* 处可微, 则 K-T 条件可简化为

$$\begin{cases} \nabla f(\mathbf{x}^*) + \sum_{i=1}^p \lambda_i^* \nabla g_i(\mathbf{x}^*) + \sum_{j=1}^q \mu_j^* \nabla h_j(\mathbf{x}^*) = \mathbf{0} \\ \lambda_i^* g_i(\mathbf{x}^*) = 0, & 1 \leq i \leq p \\ \lambda_i^* \geq 0, & 1 \leq i \leq p \end{cases}$$

引入 Lagrange 函数

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^p \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^q \mu_j h_j(\mathbf{x})$$

其中 $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)^T$ 与 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_q)^T$, 则 K-T 条件为

$$\begin{cases} \nabla_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*) = \mathbf{0} \\ \lambda_i^* g_i(\mathbf{x}^*) = 0, & 1 \leq i \leq p \\ \lambda_i^* \geq 0, & 1 \leq i \leq p \end{cases}$$



定理 3.5.2 (寻找 K-T 点)

寻找约束非线性规划问题的 K-T 点。

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{aligned}$$

1. 作 Lagrange 函数

$$L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = f(\mathbf{x}) + \sum_{i=1}^p \lambda_i g_i(\mathbf{x}) + \sum_{j=1}^q \mu_j h_j(\mathbf{x})$$

其中 $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_p)^T$ 与 $\boldsymbol{\mu} = (\mu_1, \dots, \mu_q)^T$

2. 求解 K-T 条件

$$\begin{cases} L_{\mathbf{x}k}(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\mu}) = 0, & 1 \leq k \leq n \\ \lambda_i g_i(\mathbf{x}) = 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases}$$

解记为 $(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\mu}^*)$ 。

3. 验证是否成立

$$\begin{cases} g_i(\mathbf{x}^*) \leq 0, & 1 \leq i \leq p \\ \boldsymbol{\lambda}^* \geq 0 \end{cases}$$



定理 3.5.3

对于 (MP) 问题, 若 $\mathbf{x}^* \in X$ 为可行点, $f, g_i, h_j, i \in I, j \in J$ 在 \mathbf{x}^* 处连续可微, \mathbf{x}^* 成立 K-T 条件, 且 $f, g_i, i \in I(\mathbf{x}^*)$ 为凸函数, $h_j, j \in J$ 为线性函数, 则 \mathbf{x}^* 为整体最优解。



3.5.2 简约梯度法

定义 3.5.3 (可行方向法)

对于 (MP) 问题, 寻找可行点列 $\{\mathbf{x}^n\}_{n=1}^{\infty} \subset X$, 使得成立

$$f(\mathbf{x}^{n+1}) < f(\mathbf{x}), \quad n \in \mathbb{N}^*$$

使得 \mathbf{x}^n 收敛于 (MP) 的极小点或 K-T 点。



考虑问题

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} \mathbf{Ax} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases} \end{aligned} \quad (\text{MP})$$

其中 $\mathbf{A} \in \mathbb{R}^{m \times n}$, 且 $\text{rank}(\mathbf{A}) = m \leq n$ 。可行域记为

$$X_l = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$$

作约束非退化假设:

1. 对于任意 $\mathbf{x} \in X_l$, \mathbf{x} 至少存在 m 个非零分量。
2. \mathbf{A} 的任意 m 列向量线性无关。

在每一轮迭代的当前点 \mathbf{x}^n 处, 将 \mathbf{x}^n 的 m 个最大正分量称为**基变量**, 余下的 $n - m$ 个分量称为**非基变量**。

将 \mathbf{x} 分解为两部分

$$\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix}$$

其中 $\mathbf{x}_B \in (\mathbb{R}^+)^m$ 为基分量, $\mathbf{x}_N \in \mathbb{R}^{n-m}$ 为非基分量。不失一般性, 假设 \mathbf{A} 的前 m 列对应于基变量, 将 \mathbf{A} 分解为

$$\mathbf{A} = (\mathbf{B} \quad \mathbf{N})$$

其中 \mathbf{B} 为 m 阶满秩方阵。由 $\mathbf{Ax} = \mathbf{b}$, 可得

$$\mathbf{B}\mathbf{x}_B + \mathbf{N}\mathbf{x}_N = \mathbf{b}$$

那么

$$\mathbf{x}_B = \mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N$$

因此

$$F(\mathbf{x}_N) := f(\mathbf{x}) = f(\mathbf{x}_B, \mathbf{x}_N) = f(\mathbf{B}^{-1}\mathbf{b} - \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N, \mathbf{x}_N)$$

由复合函数求导法则

$$\mathbf{r}_N = \nabla F(\mathbf{x}_N) = -(\mathbf{B}^{-1}\mathbf{N})^T \nabla_B f(\mathbf{x}) + \nabla_N f(\mathbf{x})$$

称 \mathbf{r}_N 为函数 f 在 \mathbf{x} 点处关于基矩阵 \mathbf{B} 的简约梯度, 其中

$$\nabla f(\mathbf{x}) = \begin{pmatrix} \nabla_B f(\mathbf{x}) \\ \nabla_N f(\mathbf{x}) \end{pmatrix}$$

定理 3.5.4

对于约束非线性规划问题 (MP), 设 f 为可微函数, $\mathbf{x} \in X_l$, 且 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix}$, I_B 为 \mathbf{x} 的 m 个最大分量的

下标集。如果

$$\mathbf{p} = \begin{pmatrix} \mathbf{p}_B \\ \mathbf{p}_N \end{pmatrix}, \quad \begin{cases} \mathbf{p}_B = -\mathbf{B}^{-1}\mathbf{N}\mathbf{p}_N \\ \mathbf{p}_N : p_i = \begin{cases} -r_i, & r_i \leq 0 \\ -x_i r_i, & r_i > 0 \end{cases} \quad i \notin I_B \end{cases}$$

那么

1. 若 $\mathbf{p} \neq \mathbf{0}$, 则 \mathbf{p} 为 f 在 \mathbf{x} 处关于 X_l 的可行下降方向。
2. $\mathbf{p} = \mathbf{0}$ 的充分必要条件为 \mathbf{x} 为问题 (MP) 的 K-T 点。



定理 3.5.5 (Wolfe 简约梯度法)

求解约束非线性规划问题

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

1. 选取初始可行点 $\mathbf{x}^0 \in X_l$, 给定终止误差 $\varepsilon > 0$, 令 $k := 0$ 。
2. 设 I_B^k 为 \mathbf{x}^k 的 m 个最大分量的下标集, 对矩阵 \mathbf{A} 进行分解 $\mathbf{A} = (\mathbf{B}_k, \mathbf{N}_k)$ 。
3. 计算

$$\nabla f(\mathbf{x}^k) = \begin{pmatrix} \nabla_B f(\mathbf{x}^k) \\ \nabla_N f(\mathbf{x}^k) \end{pmatrix}, \quad \mathbf{r}_N^k = -(\mathbf{B}_k^{-1}\mathbf{N}_k)^T \nabla_B f(\mathbf{x}^k) + \nabla_N f(\mathbf{x}^k)$$

记 \mathbf{r}_N^k 的第 i 个分量为 r_i^k , 其中 $i \notin I_B^k$ 。

4. 构造搜索方向

$$\mathbf{p}^k = \begin{pmatrix} \mathbf{p}_B^k \\ \mathbf{p}_N^k \end{pmatrix}, \quad \begin{cases} \mathbf{p}_B^k = -\mathbf{B}_k^{-1}\mathbf{N}_k\mathbf{p}_N^k \\ \mathbf{p}_N^k : p_i^k = \begin{cases} -r_i^k, & r_i^k \leq 0 \\ -x_i^k r_i^k, & r_i^k > 0 \end{cases} \quad i \notin I_B^k \end{cases}$$

5. 进行有效一维搜索, 求解

$$\min_{0 \leq t \leq t_{\max}^k} f(\mathbf{x}^k + t\mathbf{p}^k)$$

得到最优解 t_k , 其中

$$t_{\max}^k = \begin{cases} +\infty, & \mathbf{p}^k \geq \mathbf{0} \\ \min \left\{ -\frac{x_i^k}{p_i^k} : p_i^k < 0, 1 \leq i \leq n \right\}, & \text{其他} \end{cases}$$

6. 令 $\mathbf{x}^{k+1} = \mathbf{x}^k + t_k \mathbf{p}^k$, $k := k + 1$, 转至第 2 步。



3.5.3 惩罚函数法

3.5.3.1 罚函数法

考虑约束非线性规划问题

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{aligned} \quad (\text{MP})$$

可行域为

$$X = \left\{ \mathbf{x} \in \mathbb{R}^n \left| \begin{array}{l} g_i(\mathbf{x}) \leq 0, \quad 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, \quad 1 \leq j \leq q \end{array} \right. \right\}$$

构造惩罚函数

$$p_c(\mathbf{x}) = c \sum_{i=1}^p (\max\{g_i(\mathbf{x}), 0\})^2 + \frac{c}{2} \sum_{j=1}^q h_j^2(\mathbf{x})$$

其中 c 称为罚参数或罚因子，相应构造增广目标函数

$$F_c(\mathbf{x}) = f(\mathbf{x}) + p_c(\mathbf{x})$$

容易知道，当 $f(\mathbf{x}), g_i(\mathbf{x}), h_j(\mathbf{x})$ 连续可微时， $F_c(\mathbf{x})$ 连续可微。当 c 充分大时，可使 (MP) 问题转化为无约束的极小化问题

$$\min F_c(\mathbf{x})$$

定理 3.5.6 (罚函数法)

求解约束非线性规划问题

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & \begin{cases} g_i(\mathbf{x}) \leq 0, & 1 \leq i \leq p \\ h_j(\mathbf{x}) = 0, & 1 \leq j \leq q \end{cases} \end{array}$$

1. 选取初始点 \mathbf{x}^0 ，罚参数列 $\{c_k\}_{k=1}^{\infty}$ ，终止误差 $\varepsilon > 0$ ，令 $k := 1$ 。
2. 构造罚函数

$$p_{c_k}(\mathbf{x}) = c_k \sum_{i=1}^p (\max\{g_i(\mathbf{x}), 0\})^2 + \frac{c_k}{2} \sum_{j=1}^q h_j^2(\mathbf{x})$$

构造增广目标函数

$$F_{c_k}(\mathbf{x}) = f(\mathbf{x}) + p_{c_k}(\mathbf{x})$$

3. 选用某种无约束最优化方法，以 \mathbf{x}^{k-1} 为初始点，求解

$$\min F_{c_k}(\mathbf{x})$$

得到最优解 \mathbf{x}^k 。若 \mathbf{x}^k 满足某种终止条件，停止迭代，输出 \mathbf{x}^k ；否则令 $k := k + 1$ ，转第 2 步。



3.5.3.2 障碍函数法

考虑约束非线性规划问题

$$\begin{array}{ll} \min & f(\mathbf{x}) \\ \text{s.t.} & g_i(\mathbf{x}) \leq 0, \quad 1 \leq i \leq p \end{array} \quad (\text{MP})$$

可行域为

$$X = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) \leq 0, \quad 1 \leq i \leq p\}$$

其内部为

$$X^\circ = \{\mathbf{x} \in \mathbb{R}^n : g_i(\mathbf{x}) < 0, \quad 1 \leq i \leq p\}$$

构造障碍函数

$$B_d(\mathbf{x}) = -d \sum_{i=1}^p \frac{1}{g_i(\mathbf{x})} \quad \text{或} \quad B_d(\mathbf{x}) = -d \sum_{i=1}^p \ln(-g_i(\mathbf{x}))$$

其中 d 称为罚参数或罚因子，相应构造增广目标函数

$$F_d(\mathbf{x}) = f(\mathbf{x}) + B_d(\mathbf{x})$$

当 d 充分小时，可使 (MP) 问题转化为无约束的极小化问题

$$\min F_d(\mathbf{x})$$

定理 3.5.7 (障碍函数法)

求解约束非线性规划问题

$$\min f(\mathbf{x})$$

$$\text{s.t. } g_i(\mathbf{x}) \leq 0, \quad 1 \leq i \leq p$$

1. 选取初始点 $\mathbf{x}^0 \in X^\circ$ ，罚参数列 $\{d_k\}_{k=1}^\infty$ ，终止误差 $\varepsilon > 0$ ，令 $k := 1$ 。

2. 构造障碍函数

$$B_{d_k}(\mathbf{x}) = -d_k \sum_{i=1}^p \frac{1}{g_i(\mathbf{x})} \quad \text{或} \quad B_{d_k}(\mathbf{x}) = -d_k \sum_{i=1}^p \ln(-g_i(\mathbf{x}))$$

构造增广目标函数

$$F_{d_k}(\mathbf{x}) = f(\mathbf{x}) + B_{d_k}(\mathbf{x})$$

3. 选用某种无约束最优化方法，以 \mathbf{x}^{k-1} 为初始点，求解

$$\min F_{d_k}(\mathbf{x})$$

得到最优解 \mathbf{x}^k 。若 \mathbf{x}^k 满足某种终止条件，停止迭代，输出 \mathbf{x}^k ；否则令 $k := k + 1$ ，转第 2 步。



第四章 动态规划

4.1 多阶段决策问题

有一个系统，可以分成若干个阶段，任意一个阶段 k 的系统状态可以用 x_k 表示 (x_k 可以是数量、向量、集合等)。在每一阶段 k 的每一状态 x_k 都有一个决策集合 $Q_k(x_k)$ ，在 $Q_k(x_k)$ 中选定一个决策 $q_k \in Q_k(x_k)$ ，状态 x_k 就转移到新的状态 $x_{k+1} = T_k(x_k, q_k)$ ，并且得到效益 $R_k(x_k, q_k)$ 。我们的目的就是在每一个阶段都在它的决策集合中选择一个决策，使所有阶段的总效益 $\sum_k R_k(x_k, q_k)$ 达到最优。我们称之为多阶段决策问题。

一个多阶段决策问题包括阶段数、状态变量、决策变量、状态转移方程和目标函数等基本要素，描述一个多阶段决策问题就要从以上基本要素入手，只要这些基本要素刻画清楚了，整个决策问题就明了了。

不同问题的要素不尽相同，根据要素的差异，多阶段决策问题可以分成不同类型：

1. 根据阶段数可分为：
 - (a). 有限阶段决策问题，其阶段数为有限值。
 - (b). 无限阶段决策问题，其阶段数为无穷大，决策过程可无限持续下去。
2. 根据变量取值情况可分为：
 - (a). 连续多阶段决策问题，决策变量和状态变量取连续变化的实数。
 - (b). 离散多阶段决策问题，决策变量和状态变量取有限的数值。
3. 根据阶段个数是否明确可分为：
 - (a). 定期多阶段决策问题，其阶段数是明确的，不受决策的影响。
 - (b). 不定期多阶段决策问题，其阶段数是不确定的，不同的决策下阶段数不同。
4. 根据参数取值情况可分为：
 - (a). 确定多阶段决策问题，其参数是给定的常数。
 - (b). 不确定多阶段决策问题，其参数中包含不确定因素，如随机参数，区间取值参数等。

本章着重介绍确定有限多阶段决策问题。

4.2 最优化原理

一个过程的最优策略具有这样的性质，即无论其初始状态及其初始决策如何，其后诸策略对以第一个决策所形成的状态作为初始状态而言，必须构成最优策略。

定理 4.2.1 (确定有限多阶段决策问题)

1. 明确问题，找出阶段数。
2. 确定变量，找出状态变量和决策变量。
3. 找出状态转移方程。
4. 写出递推关系式。
5. 求解递推关系式。



4.3 确定性的定期多阶段决策问题

4.3.1 递推法

定理 4.3.1 (递推法)

设 $f_{n-k+1}(x_k)$ 为第 k 个阶段的状态为 x_k , 经过 $n-k+1$ 个阶段的最优目标函数值, 则有

$$f_{n-k+1}(x_k) = \min_{q_k \in Q_k(x_k)} \{R_k(x_k, q_k) + f_{n-k}(T_k(x_k, q_k))\}$$

$$\vdots$$

$$f_1(x_n) = \min_{q_n \in Q_n(x_n)} \{R_n(x_n, q_n)\}$$



4.3.2 最短路径问题

例题 4.1 求解从 A_0 到 A_3 的最短路径。

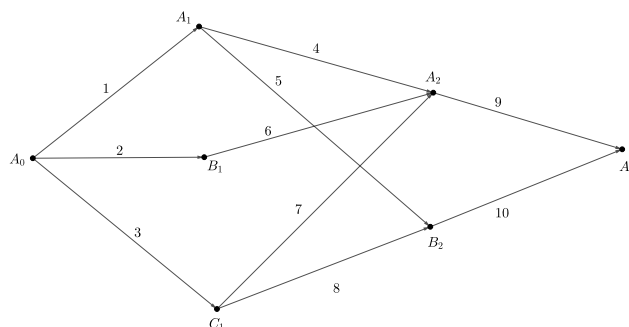


图 4.1: 最短路径问题

解 记 $f(P)$ 为点 P 到目标点 A_3 的最短路径长度, $x(P)$ 为点 P 在最短路径中的下一点, $d(P, Q)$ 为点 P 到点 Q 的距离。我们从目标点往前向出发点递推。

还有 1 步到达目标点:

$$f(A_2) = d(A_2, A_3) = 9, \quad x(A_2) = A_3$$

$$f(B_2) = d(B_2, A_3) = 10, \quad x(B_2) = A_3$$

还有 2 步到达目标点:

$$f(A_1) = \min\{d(A_1, A_2) + f(A_2), d(A_1, B_2) + f(B_2)\} = \min\{13, 15\} = 13, \quad x(A_1) = A_2$$

$$f(B_1) = d(B_1, A_2) + f(A_2) = 15, \quad x(B_1) = A_2$$

$$f(C_1) = \min\{d(C_1, A_2) + f(A_2), d(C_1, B_2) + f(B_2)\} = \min\{16, 18\} = 16, \quad x(C_1) = A_2$$

还有 3 步到达目标点:

$$f(A_0) = \min\{d(A_0, A_1) + f(A_1), d(A_0, B_1) + f(B_1), d(A_0, C_1) + f(C_1)\} = \min\{14, 17, 19\} = 14$$

$$x(A_0) = A_1$$

所以最短路径为

$$A_0 \rightarrow A_1 \rightarrow A_2 \rightarrow A_3, \quad \min = 14$$

4.3.3 旅行售货员问题

例题 4.2 求解从点 v_1 出发, 经过点 v_2, v_3, v_4 各一次, 最后回到点 v_1 的最短路径。距离矩阵如下, 其中 (i, j) 为点 v_i 到点 v_j 的距离。

$$D = \begin{pmatrix} 0 & 8 & 5 & 6 \\ 6 & 0 & 8 & 5 \\ 7 & 9 & 0 & 5 \\ 9 & 7 & 8 & 0 \end{pmatrix}$$

解 记 $f(v_i, V)$ 为从点 v_i , 经过集合 V 中的点各一次, 最后回到初始点 v_1 的最短路径长度, $x(v_i)$ 为点 v_i 在最短路径中的下一点, $d(v_i, v_j)$ 为点 v_i 到点 v_j 的距离。我们从目标点往前向出发点递推。

经过 0 个点, 回到 v_1 :

$$f(v_2, \emptyset) = d(v_2, v_1) = 6$$

$$f(v_3, \emptyset) = d(v_3, v_1) = 7$$

$$f(v_4, \emptyset) = d(v_4, v_1) = 9$$

经过 1 个点, 回到 v_1 :

$$f(v_2, \{v_3\}) = d(v_2, v_3) + f(v_3, \emptyset) = 15$$

$$f(v_2, \{v_4\}) = d(v_2, v_4) + f(v_4, \emptyset) = 14$$

$$f(v_3, \{v_2\}) = d(v_3, v_2) + f(v_2, \emptyset) = 15$$

$$f(v_3, \{v_4\}) = d(v_3, v_4) + f(v_4, \emptyset) = 14$$

$$f(v_4, \{v_2\}) = d(v_4, v_2) + f(v_2, \emptyset) = 13$$

$$f(v_4, \{v_3\}) = d(v_4, v_3) + f(v_3, \emptyset) = 15$$

经过 2 个点, 回到 v_1 :

$$f(v_2, \{v_3, v_4\}) = \min\{d(v_2, v_3) + f(v_3, \{v_4\}), d(v_2, v_4) + f(v_4, \{v_3\})\} = \min\{22, 20\} = 20, \quad x(v_2) = v_4$$

$$f(v_3, \{v_2, v_4\}) = \min\{d(v_3, v_2) + f(v_2, \{v_4\}), d(v_3, v_4) + f(v_4, \{v_2\})\} = \min\{23, 18\} = 18, \quad x(v_3) = v_4$$

$$f(v_4, \{v_2, v_3\}) = \min\{d(v_4, v_2) + f(v_2, \{v_3\}), d(v_4, v_3) + f(v_3, \{v_2\})\} = \min\{22, 23\} = 22, \quad x(v_4) = v_2$$

经过 3 个点, 回到 v_1 :

$$\begin{aligned} & f(v_1, \{v_2, v_3, v_4\}) \\ &= \min\{d(v_1, v_2) + f(v_2, \{v_3, v_4\}), d(v_1, v_3) + f(v_3, \{v_2, v_4\}), d(v_1, v_4) + f(v_4, \{v_2, v_3\})\} \\ &= \min\{28, 23, 28\} = 23 \\ & x(v_1) = v_3 \end{aligned}$$

所以最短路径为

$$v_1 \rightarrow v_3 \rightarrow v_4 \rightarrow v_2 \rightarrow v_1, \quad \min = 23$$

4.3.4 多阶段资源分配问题

引理 4.3.1

如果 $g(y), h(y)$ 为凸函数, 那么对于任意 x , $F(y) = g(y) + h(x - y)$ 为凸函数。



引理 4.3.2

如果 $f(x), g(x)$ 为凸函数, 那么 $\max\{f(x), g(x)\}$ 为凸函数。



定理 4.3.2 (多阶段资源分配问题)

多阶段资源分配问题递推公式为

$$\begin{cases} f_1(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y)\} \\ f_{k+1}(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y) + f_k(ay + b(x-y))\} \end{cases}$$

**定理 4.3.3 (多阶段资源分配问题的凸函数解)**

如果 $g(y)$ 和 $h(y)$ 为凸函数, 且 $g(0) = h(0) = 0$, 那么 n 阶段资源分配问题的最优策略 y 在每个阶段总取 $0 \leq y \leq x$ 的端点的值, 且

$$\begin{cases} f_1(x) = \max\{g(x), h(x)\} \\ f_{k+1}(x) = \max\{g(x) + f_k(ax), h(x) + f_k(bx)\} \end{cases}$$



例题 4.3 某单位有资源 100 单位, 拟分 4 个周期使用, 在每个周期有生产任务 A 和 B , 把资源用于 A 生产任务, 每单位能获利 10 元, 资源回收率为 $2/3$; 把资源用于 B 生产任务, 每单位能获利 7 元, 资源回收率为 $9/10$ 。问每个周期应如何分配资源, 使总收益最大?

解 记 A 生产任务的获利函数为 $g(x) = 10x$, B 生产任务的获利函数为 $h(x) = 7x$, 以及 $a = 2/3$, $b = 9/10$ 。

以 $f_1(x)$ 表示投入 x 单位的生产资源在 1 个周期内生产的最大总收益函数, 那么

$$f_1(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y)\} = \max_{0 \leq y \leq x} \{7x + 3y\} = 10x, \quad y = x$$

以 $f_2(x)$ 表示投入 x 单位的生产资源在 2 个周期内生产的最大总收益函数, 那么

$$f_2(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y) + f_1(ay + b(x-y))\} = \max_{0 \leq y \leq x} \left\{16x + \frac{2}{3}y\right\} = \frac{50}{3}x, \quad y = x$$

以 $f_3(x)$ 表示投入 x 单位的生产资源在 3 个周期内生产的最大总收益函数, 那么

$$f_3(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y) + f_2(ay + b(x-y))\} = \max_{0 \leq y \leq x} \left\{22x - \frac{8y}{9}\right\} = 22x, \quad y = 0$$

以 $f_4(x)$ 表示投入 x 单位的生产资源在 4 个周期内生产的最大总收益函数, 那么

$$f_4(x) = \max_{0 \leq y \leq x} \{g(y) + h(x-y) + f_3(ay + b(x-y))\} = \max_{0 \leq y \leq x} \left\{\frac{134}{5}x - \frac{32}{15}y\right\} = \frac{134}{5}x, \quad y = 0$$

因此投入 100 单位的生产资源在 4 个周期内生产的最大总收益为

$$f_4(100) = 2680$$

每个阶段的生产计划如下

表 4.1: 生产计划

	投入 A 生产任务	投入 B 生产任务	本周期收益
第一周期	0	100	700
第二周期	0	90	630
第三周期	81	0	810
第四周期	54	0	540
总收益			2680

4.3.5 可靠性问题

例题 4.4 求解规划问题

$$\begin{aligned} \max \quad & z = x_1 x_2 \cdots x_n \\ \text{s.t.} \quad & \begin{cases} x_1 + x_2 + \cdots + x_n = 1 \\ x_1, x_2, \cdots, x_n \geq 0 \end{cases} \end{aligned}$$

解 令

$$f_k(x) = \max\{x_1 x_2 \cdots x_k : x_1 + x_2 + \cdots + x_k = 1, x_1, x_2, \cdots, x_k \geq 0\}$$

那么成立递推关

$$\begin{cases} f_1(x) = x \\ f_{k+1}(x) = \max_{0 \leq x_{k+1} \leq x} x_{k+1} f_k(x - x_k) \end{cases}$$

从而

$$f_k(x) = \max_{0 \leq x_k \leq x} x_k \left(\frac{x - x_k}{k-1} \right)^{n-1} = \frac{x^k}{k^k}, \quad x_k = \frac{x}{k}$$

现在来求最优解。

第一阶段，状态变量 $x = 1$ ，最优解 $x_n = \frac{1}{n}$ 。

第二阶段，状态变量

$$x = 1 - \frac{1}{n} = \frac{n-1}{n}$$

最优解

$$x_{n-1} = \frac{x}{n-1} = \frac{1}{n}$$

第三阶段，状态变量

$$x = 1 - \frac{1}{n} - \frac{1}{n} = \frac{n-2}{n}$$

最优解

$$x_{n-2} = \frac{x}{n-2} = \frac{1}{n}$$

以此类推，最后阶段，状态变量

$$x = 1 - \frac{1}{n} - \frac{1}{n} - \cdots - \frac{1}{n} = \frac{1}{n}$$

最优解

$$x_{n-2} = \frac{x}{1} = \frac{1}{n}$$

进而最优解为 $(\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$ ，最优值为 $1/n^n$ 。

例题 4.5 求解规划问题

$$\begin{aligned} \max \quad & z = 4x_1^2 - x_2^2 + 2x_3^2 + 12 \\ \text{s.t.} \quad & \begin{cases} 3x_1 + 2x_2 + x_3 = y \\ x_1, x_2, x_3 \geq 0 \end{cases} \end{aligned}$$

解 第一阶段

$$f_1(y) = \max_{0 \leq 3x_1 \leq y} \{4x_1^2\} = \frac{4}{9}y^2, \quad x_1 = \frac{y}{3}$$

第二阶段

$$f_2(y) = \max_{0 \leq 2x_2 \leq y} \{-x_2^2 + f_1(y - 2x_2)\} = \frac{4}{9}y^2, \quad x_2 = 0$$

第三阶段

$$f_3(9) = \max_{0 \leq x_3 \leq 9} \{2x_3^2 + 12 + f_2(9 - x_3)\} = 174, \quad x_3 = 9$$

因此最优解为

$$x_1 = 0, \quad x_2 = 0, \quad x_3 = 9$$

4.4 确定性的不定期多阶段决策问题

4.4.1 最优线路问题

定义 4.4.1 (最优线路问题)

$$\begin{cases} f(i) = \min_{1 \leq j \leq n} \{c_{ij} + f(j)\}, & i = 1, 2, \dots, n-1 \\ f(n) = 0 \end{cases} \quad (*)$$

定理 4.4.1 (函数空间迭代法)

记 $f_k(i)$ 为由 v_i 点出发至多经过 k 个点到达 v_n 的最短路程, 那么

$$\begin{cases} f_1(i) = c_{in}, & i = 1, 2, \dots, n-1 \\ f_1(n) = 0 \end{cases}$$

递推关系为

$$\begin{cases} f_k(i) = \min_{1 \leq j \leq n} \{c_{ij} + f_{k-1}(j)\}, & i = 1, 2, \dots, n-1 \\ f_k(n) = 0 \end{cases}$$

得到的函数列 $\{f_k(i)\}$ 单调下降收敛于函数方程 $(*)$ 的解 $f(i)$ 。

例题 4.6 设有五个城市 1, 2, 3, 4, 5 相互的距离矩阵为

$$D = \begin{pmatrix} 0 & 3 & 4 & \infty & \infty \\ \infty & 0 & 5 & 2 & 4 \\ \infty & \infty & 0 & 3 & \infty \\ \infty & \infty & \infty & 0 & 1 \\ \infty & \infty & \infty & \infty & 0 \end{pmatrix}$$

试用函数空间迭代法求各城市到城市 5 的最短路线及最短路程。

解 记 $f_1(i)$ 为由城市 i 出发至多经过 1 个城市到达城市 5 的最短路程, 那么

$$f_1(1) = d_{15} = \infty, \quad f_1(2) = d_{25} = 4, \quad f_1(3) = d_{35} = \infty, \quad f_1(4) = d_{45} = 1, \quad f_1(5) = d_{55} = 0$$

记 $f_2(i)$ 为由城市 i 出发至多经过 2 个城市到达城市 5 的最短路程, 那么

$$f_2(1) = \min_{1 \leq j \leq 5} \{d_{1j} + f_1(j)\} = \min\{0 + \infty, 3 + 4, 4 + \infty, \infty + 1, \infty + 0\} = 7$$

$$f_2(2) = \min_{1 \leq j \leq 5} \{d_{2j} + f_1(j)\} = \min\{\infty + \infty, 0 + 4, 5 + \infty, 2 + 1, 4 + 0\} = 3$$

$$f_2(3) = \min_{1 \leq j \leq 5} \{d_{3j} + f_1(j)\} = \min\{\infty + \infty, \infty + 4, 0 + \infty, 3 + 1, \infty + 0\} = 4$$

$$f_2(4) = \min_{1 \leq j \leq 5} \{d_{4j} + f_1(j)\} = \min\{\infty + \infty, \infty + 4, \infty + \infty, 0 + 1, 1 + 0\} = 1$$

$$f_2(5) = 0$$

记 $f_3(i)$ 为由城市 i 出发至多经过 3 个城市到达城市 5 的最短路程, 那么

$$f_3(1) = \min_{1 \leq j \leq 5} \{d_{1j} + f_2(j)\} = \min\{0 + 7, 3 + 3, 4 + 4, \infty + 1, \infty + 0\} = 6$$

$$f_3(2) = \min_{1 \leq j \leq 5} \{d_{2j} + f_2(j)\} = \min\{\infty + 7, 0 + 3, 5 + 4, 2 + 1, 4 + 0\} = 3$$

$$f_3(3) = \min_{1 \leq j \leq 5} \{d_{3j} + f_2(j)\} = \min\{\infty + 7, \infty + 3, 0 + 4, 3 + 1, \infty + 0\} = 4$$

$$f_3(4) = \min_{1 \leq j \leq 5} \{d_{4j} + f_2(j)\} = \min\{\infty + 7, \infty + 3, \infty + 4, 0 + 1, 1 + 0\} = 1$$

$$f_3(5) = 0$$

记 $f_4(i)$ 为由城市 i 出发至多经过 4 个城市到达城市 5 的最短路程, 那么

$$f_4(1) = \min_{1 \leq j \leq 5} \{d_{1j} + f_3(j)\} = \min\{0 + 6, 3 + 3, 4 + 4, \infty + 1, \infty + 0\} = 6, \quad s(1) = 2$$

$$f_4(2) = \min_{1 \leq j \leq 5} \{d_{2j} + f_3(j)\} = \min\{\infty + 6, 0 + 3, 5 + 4, 2 + 1, 4 + 0\} = 3, \quad s(2) = 4$$

$$f_4(3) = \min_{1 \leq j \leq 5} \{d_{3j} + f_3(j)\} = \min\{\infty + 6, \infty + 3, 0 + 4, 3 + 1, \infty + 0\} = 4, \quad s(3) = 4$$

$$f_4(4) = \min_{1 \leq j \leq 5} \{d_{4j} + f_3(j)\} = \min\{\infty + 6, \infty + 3, \infty + 4, 0 + 1, 1 + 0\} = 1, \quad s(4) = 5$$

$$f_4(5) = 0$$

因此各城市到城市 5 的最短路线及最短路程为

城市	最短路线	最短路程
1	$1 \rightarrow 2 \rightarrow 4 \rightarrow 5$	6
2	$2 \rightarrow 4 \rightarrow 5$	3
3	$3 \rightarrow 4 \rightarrow 5$	4
4	$4 \rightarrow 5$	1

定理 4.4.2 (策略空间迭代法)

策略 $\{s(i)\}$ 就是给定在点 v_i 时, 选定下一步的位置为 $s(i)$ 。

给定无回路的初始策略 $\{s_0(i)\}$, 作方程组

$$\begin{cases} f_0(i) = c_{i,s_0(i)} + f_0(s_0(i)), & i = 1, 2, \dots, n-1 \\ f_0(n) = 0 \end{cases}$$

解得 $f_0(1), \dots, f_0(n)$ 。求解

$$\min_{1 \leq j \leq n} \{c_{ij} + f_0(j)\} = c_{i,s_1(i)} + f_0(s_1(i)), \quad i = 1, 2, \dots, n-1$$

由此得到策列 $\{s_1(i)\}$ 。重复上面过程, 得到策略 $\{s_k(i)\}$, 由此得到的函数列 $\{f_k(i)\}$ 单调下降收敛于函数方程 (*) 的解 $f(i)$ 。



例题 4.7 设有五个城市 1, 2, 3, 4, 5 相互的距离矩阵为

$$D = \begin{pmatrix} 0 & 3 & 4 & \infty & \infty \\ \infty & 0 & 5 & 2 & 4 \\ \infty & \infty & 0 & 3 & \infty \\ \infty & \infty & \infty & 0 & 1 \\ \infty & \infty & \infty & \infty & 0 \end{pmatrix}$$

试用策略空间迭代法求各城市到城市 5 的最短路线及最短路程, 其中初始策略为

$$s_0(1) = 2, \quad s_0(2) = 3, \quad s_0(3) = 4, \quad s_0(4) = 5, \quad s_0(5) = 5$$

解 作方程组

$$\begin{cases} f_0(1) = d_{1,s_0(1)} + f_0(s_0(1)) \\ f_0(2) = d_{2,s_0(2)} + f_0(s_0(2)) \\ f_0(3) = d_{3,s_0(3)} + f_0(s_0(3)) \\ f_0(4) = d_{4,s_0(4)} + f_0(s_0(4)) \\ f_0(5) = 0 \end{cases} \iff \begin{cases} f_0(1) = 3 + f_0(2) \\ f_0(2) = 5 + f_0(3) \\ f_0(3) = 3 + f_0(4) \\ f_0(4) = 1 + f_0(5) \\ f_0(5) = 0 \end{cases} \iff \begin{cases} f_0(1) = 12 \\ f_0(2) = 9 \\ f_0(3) = 4 \\ f_0(4) = 1 \\ f_0(5) = 0 \end{cases}$$

求解

$$\begin{aligned} \min_{1 \leq j \leq 5} \{d_{1j} + f_0(j)\} &= \min\{0 + 12, 3 + 9, 4 + 4, \infty + 1, \infty + 0\} = 8, & s_1(1) &= 3 \\ \min_{1 \leq j \leq 5} \{d_{2j} + f_0(j)\} &= \min\{\infty + 12, 0 + 9, 5 + 4, 2 + 1, 4 + 0\} = 3, & s_1(2) &= 4 \\ \min_{1 \leq j \leq 5} \{d_{3j} + f_0(j)\} &= \min\{\infty + 12, \infty + 9, 0 + 4, 3 + 1, \infty + 0\} = 4, & s_1(3) &= 4 \\ \min_{1 \leq j \leq 5} \{d_{4j} + f_0(j)\} &= \min\{\infty + 12, \infty + 9, \infty + 4, 0 + 1, 1 + 0\} = 1, & s_1(4) &= 5 \end{aligned}$$

由此得到策略

$$s_1(1) = 3, \quad s_1(2) = 4, \quad s_1(3) = 4, \quad s_1(4) = 5, \quad s_1(5) = 5$$

作方程组

$$\begin{cases} f_1(1) = d_{1,s_1(1)} + f_1(s_1(1)) \\ f_1(2) = d_{2,s_1(2)} + f_1(s_1(2)) \\ f_1(3) = d_{3,s_1(3)} + f_1(s_1(3)) \\ f_1(4) = d_{4,s_1(4)} + f_1(s_1(4)) \\ f_1(5) = 0 \end{cases} \iff \begin{cases} f_1(1) = 4 + f_1(3) \\ f_1(2) = 2 + f_1(4) \\ f_1(3) = 3 + f_1(4) \\ f_1(4) = 1 + f_1(5) \\ f_1(5) = 0 \end{cases} \iff \begin{cases} f_1(1) = 8 \\ f_1(2) = 3 \\ f_1(3) = 4 \\ f_1(4) = 1 \\ f_1(5) = 0 \end{cases}$$

求解

$$\begin{aligned} \min_{1 \leq j \leq 5} \{d_{1j} + f_1(j)\} &= \min\{0 + 8, 3 + 3, 4 + 4, \infty + 1, \infty + 0\} = 6, & s_2(1) &= 2 \\ \min_{1 \leq j \leq 5} \{d_{2j} + f_1(j)\} &= \min\{\infty + 8, 0 + 3, 5 + 4, 2 + 1, 4 + 0\} = 3, & s_2(2) &= 4 \\ \min_{1 \leq j \leq 5} \{d_{3j} + f_1(j)\} &= \min\{\infty + 8, \infty + 3, 0 + 4, 3 + 1, \infty + 0\} = 4, & s_2(3) &= 4 \\ \min_{1 \leq j \leq 5} \{d_{4j} + f_1(j)\} &= \min\{\infty + 8, \infty + 3, \infty + 4, 0 + 1, 1 + 0\} = 1, & s_2(4) &= 5 \end{aligned}$$

由此得到策略

$$s_2(1) = 2, \quad s_2(2) = 4, \quad s_2(3) = 4, \quad s_2(4) = 5, \quad s_2(5) = 5$$

作方程组

$$\begin{cases} f_2(1) = d_{1,s_2(1)} + f_2(s_2(1)) \\ f_2(2) = d_{2,s_2(2)} + f_2(s_2(2)) \\ f_2(3) = d_{3,s_2(3)} + f_2(s_2(3)) \\ f_2(4) = d_{4,s_2(4)} + f_2(s_2(4)) \\ f_2(5) = 0 \end{cases} \iff \begin{cases} f_2(1) = 3 + f_2(2) \\ f_2(2) = 2 + f_2(4) \\ f_2(3) = 3 + f_2(4) \\ f_2(4) = 1 + f_2(5) \\ f_2(5) = 0 \end{cases} \iff \begin{cases} f_2(1) = 6 \\ f_2(2) = 3 \\ f_2(3) = 4 \\ f_2(4) = 1 \\ f_2(5) = 0 \end{cases}$$

求解

$$\min_{1 \leq j \leq 5} \{d_{1j} + f_2(j)\} = \min\{0 + 6, 3 + 3, 4 + 4, \infty + 1, \infty + 0\} = 6, \quad s_3(1) = 2$$

$$\min_{1 \leq j \leq 5} \{d_{2j} + f_2(j)\} = \min\{\infty + 6, 0 + 3, 5 + 4, 2 + 1, 4 + 0\} = 3, \quad s_3(2) = 4$$

$$\min_{1 \leq j \leq 5} \{d_{3j} + f_2(j)\} = \min\{\infty + 6, \infty + 3, 0 + 4, 3 + 1, \infty + 0\} = 4, \quad s_3(3) = 4$$

$$\min_{1 \leq j \leq 5} \{d_{4j} + f_2(j)\} = \min\{\infty + 6, \infty + 3, \infty + 4, 0 + 1, 1 + 0\} = 1, \quad s_3(4) = 5$$

由此得到策略

$$s_3(1) = 2, \quad s_3(2) = 4, \quad s_3(3) = 4, \quad s_3(4) = 5, \quad s_3(5) = 5$$

因此各城市到城市 5 的最短路线及最短路程为

城市	最短路线	最短路程
1	$1 \rightarrow 2 \rightarrow 4 \rightarrow 5$	6
2	$2 \rightarrow 4 \rightarrow 5$	3
3	$3 \rightarrow 4 \rightarrow 5$	4
4	$4 \rightarrow 5$	1

第五章 图与网络分析

5.1 图的基本概念与矩阵表示

5.1.1 图的基本概念

定义 5.1.1 (图)

1. 称二元组 (V, E) 为无向图, 其中 V 为点集合

$$E = \{\langle u, v \rangle \text{ 为无序偶} \}$$

为无向边集合。

2. 称二元组 (V, E) 为有向图, 其中 V 为点集合

$$E = \{(u, v) \text{ 为有序偶} \}$$

为有向边集合。



定义 5.1.2 (点与边的关系)

1. 点与边的关联

(a). 称无向边 $\langle u, v \rangle$ 与点 u, v 关联。

(b). 称有向边 (u, v) 与点 u, v 关联。

2. 点与点的邻接

(a). 称点 u, v 邻接, 如果存在无向边 $\langle u, v \rangle$ 。

(b). 称点 u, v 邻接, 如果存在有向边 (u, v) 。

3. 边与边的邻接

(a). 称无向边 $\langle u, v \rangle$ 与 $\langle x, y \rangle$ 邻接, 如果 $\{u, v\} \cap \{x, y\} \neq \emptyset$ 。

(b). 称有向边 (u, v) 与 (x, y) 邻接, 如果 $\{u, v\} \cap \{x, y\} \neq \emptyset$ 。



定义 5.1.3 (特殊点)

1. 端点: 称点 u, v 为无向边 $\langle u, v \rangle$ 的端点。

2. 起点/尾: 称点 u 为有向边 (u, v) 的起点/尾。

3. 终点/头: 称点 v 为有向边 (u, v) 的终点/头。

4. 孤立点

(a). 称点 v 为孤立点, 如果不存在无向边 $\langle u, w \rangle$, 使得 $v \in \{u, w\}$ 。

(b). 称点 v 为孤立点, 如果不存在有向边 (u, w) , 使得 $v \in \{u, w\}$ 。



定义 5.1.4 (特殊边)

1. 环

(a). 称无向边 $\langle v, v \rangle$ 为环。

(b). 称有向边 (v, v) 为环。

2. 重边

(a). 称无向边 $\langle u, v \rangle$ 与 $\langle u, v \rangle$ 为重边。

(b). 称有向边 (v, v) 与 (v, v) 为重边。



定义 5.1.5 (特殊图)

1. 简单图：称图 (V, E) 为简单图，如果其既没有环，也没有重边。
2. 完全图
 - (a). 称简单无向图 (V, E) 为完全无向图 K_n ，其中 $|V| = n$ ，如果对于任意 $u, v \in V$ ，成立 $\langle u, v \rangle \in E$ 。
 - (b). 称简单有向图 (V, E) 为完全有向图 K_n ，其中 $|V| = n$ ，如果对于任意 $u, v \in V$ ，成立 $(u, v) \in E$ 且 $(v, u) \in E$ 。
3. 二分图
 - (a). 称无向图 (V, E) 为二分无向图，并记作 (S, T, E) ，如果 $V = S \sqcup T$ ，且对于任意 $\langle s, t \rangle \in E$ ，成立 $s \in S$ 与 $t \in T$ 。
 - (b). 称有向图 (V, E) 为二分有向图，并记作 (S, T, E) ，如果 $V = S \sqcup T$ ，且对于任意 $(s, t) \in E$ ，成立 $s \in S$ 与 $t \in T$ 。
4. 完全二分图
 - (a). 称简单二分无向图 (S, T, E) 为完全二分无向图 $K_{p,q}$ ，其中 $|S| = p$ 且 $|T| = q$ ，如果对于任意 $s \in S$ 与 $t \in T$ ，存在 $\langle s, t \rangle \in E$ 。
 - (b). 称简单二分有向图 (S, T, E) 为完全二分有向图 $K_{p,q}$ ，其中 $|S| = p$ 且 $|T| = q$ ，如果对于任意 $s \in S$ 与 $t \in T$ ，存在 $(s, t) \in E$ 。
5. 补图：对于简单图 $G = (V, E)$ ，令 $K_n = (V, E_K)$ 为完全图，定义 G 的补图为 $\overline{G} = (V, E_K \setminus E)$ 。
6. 基本图：称无向图 (V, E_0) 为有向图 (V, E) 的基本图，如果成立如下命题。
 - (a). 对于任意 $(u, v) \in E$ ，成立 $\langle u, v \rangle \in E_0$ 。
 - (b). 对于任意 $\langle u, v \rangle \in E_0$ ，成立 $(u, v) \in E$ 。

**定义 5.1.6 (图依大小的分类)**

1. 有限图：称图 (V, E) 为有限图，如果 $|V| < \infty$ 且 $|E| < \infty$ 。
2. 无限图：称图 (V, E) 为无限图，如 $|V| = \infty$ 或 $|E| = \infty$ 。
3. 空图：称图 (V, E) 为空图，如果 $V = \emptyset$ 。
4. 平凡图：称图 (V, E) 为平凡图，如果 $|V| = 1$ 。

**定义 5.1.7 (网络)**

对于图 (V, E) ，定义 (V, E, φ) 为网络，其中 $\varphi: E \rightarrow \mathbb{R}$ 为函数。

**定义 5.1.8 (图的运算)**

1. 并：定义图 $G_1 = (N_1, E_1)$ 与 $G_2 = (N_2, E_2)$ 的并为图

$$G_1 \cup G_2 = (N_1 \cup N_2, E_1 \cup E_2)$$
2. 交：定义图 $G_1 = (N_1, E_1)$ 与 $G_2 = (N_2, E_2)$ 的交为图

$$G_1 \cap G_2 = (N_1 \cap N_2, E_1 \cap E_2)$$
3. 不交：称图 $G_1 = (N_1, E_1)$ 与 $G_2 = (N_2, E_2)$ 不相交，如果 $N_1 \cap N_2 = \emptyset$ 。



5.1.2 图的关联矩阵与邻接矩阵

定义 5.1.9 (关联矩阵)

1. 定义简单无向图 (V, E) 的关联矩阵为 $|V| \times |E|$ 阶矩阵 $B = (b_{ij})$, 其中

$$b_{ij} = \begin{cases} 1, & \text{点 } v_i \text{ 与边 } e_j \text{ 关联} \\ 0, & \text{点 } v_i \text{ 与边 } e_j \text{ 不关联} \end{cases}$$

2. 定义简单有向图 (V, E) 的关联矩阵为 $|V| \times |E|$ 阶矩阵 $B = (b_{ij})_{|V| \times |E|}$, 其中

$$b_{ij} = \begin{cases} 1, & \text{边 } e_j \text{ 以点 } v_i \text{ 为起点} \\ -1, & \text{边 } e_j \text{ 以点 } v_i \text{ 为终点} \\ 0, & \text{其余情况} \end{cases}$$



定义 5.1.10 (邻接矩阵)

1. 定义简单无向图 (V, E) 的邻接矩阵为 $|V|$ 阶矩阵 $A = (a_{ij})$, 其中

$$a_{ij} = \begin{cases} 1, & \langle v_i, v_j \rangle \in E \\ 0, & \langle v_i, v_j \rangle \notin E \end{cases}$$

2. 定义简单有向图 (V, E) 的邻接矩阵为 $|V|$ 阶矩阵 $A = (a_{ij})$, 其中

$$a_{ij} = \begin{cases} 1, & (v_i, v_j) \in E \\ 0, & (v_i, v_j) \notin E \end{cases}$$



定义 5.1.11 (度)

1. 定义简单无向图 (V, E) 的点 v 的度 $\deg(v)$ 为图中与点 v 关联的边数; 换言之

$$\deg(v) = \sum_{e \in E \text{ 与 } v \text{ 关联}} 1$$

2. 定义简单有向图 (V, E) 的点 v 的入度 $\deg^-(v)$ 为图中以点 v 为起点的边数; 换言之

$$\deg^-(v) = \sum_{e \in E \text{ 以 } v \text{ 为起点}} 1$$

3. 定义简单有向图 (V, E) 的点 v 的出度 $\deg^+(v)$ 为图中以点 v 为终点的边数; 换言之

$$\deg^+(v) = \sum_{e \in E \text{ 以 } v \text{ 为终点}} 1$$



定理 5.1.1 (二分图的等价定义)

G 为二分图, 当且仅当 G 的邻接矩阵可表示为

$$A = \begin{pmatrix} O & M \\ M^T & O \end{pmatrix}$$



定理 5.1.2

1. 对于简单无向图 (V, E) , 成立

$$\sum_{v \in V} \deg(v) = 2|E|$$

2. 对于简单有向图 (V, E) , 成立

$$\sum_{v \in V} \deg^+(v) = \sum_{v \in V} \deg^-(v) = |E|$$



5.1.3 子图

定义 5.1.12 (子图)

1. 称无向图 (V_0, E_0) 为图 (V, E) 的子图, 如果 $V_0 \subset V$ 与 $E_0 \subset E$, 且对于任意 $\langle u, v \rangle \in E_0$, 成立 $u \in V_0$ 与 $v \in V_0$ 。
2. 称有向图 (V_0, E_0) 为图 (V, E) 的子图, 如果 $V_0 \subset V$ 与 $E_0 \subset E$, 且对于任意 $(u, v) \in E_0$, 成立 $u \in V_0$ 与 $v \in V_0$ 。



定义 5.1.13 (特殊子图)

1. 支撑子图: 称图 (V, E) 的子图 (V_0, E_0) 为支撑子图, 如果 $V_0 = V$ 。
2. 点导出子图: 称图 $G = (V, E)$ 的子图 $G_0 = (V_0, E_0)$ 为由点 V_0 导出的 G 的子图, 并记作 $G[V_0]$, 如果

$$E_0 = \bigcup_{(V_0, E') \text{ 为 } (V, E) \text{ 的子图}} E'$$

3. 边导出子图: 称图 $G = (V, E)$ 的子图 $G_0 = (V_0, E_0)$ 为由边 E_0 导出的 G 的子图, 并记作 $G[E_0]$, 如果

$$V_0 = \bigcup_{(V', E_0) \text{ 为 } (V, E) \text{ 的子图}} V'$$



5.2 图的连通性

5.2.1 图的连通

定义 5.2.1 (路)

称图中从点 a 到 z 的路为点和边的交错序列

$$(a, e_{ab}, b, \dots, y, e_{yz}, z)$$

特别的, 称简单图中从点 a 到 z 的路为点的序列

$$(a, b, \dots, y, z)$$



定义 5.2.2 (特殊路)

1. 简单路: 称边不重复的路为简单路。
2. 初级路: 称点不重复的路为初级路。
3. 回路: 称路

$$(a, e_{ab}, b, \dots, y, e_{yz}, z)$$

为回路, 如果 $a = z$ 。



定义 5.2.3 (连通性)

1. 点的连通: 对于无向图, 称点 u 与 v 连通, 如果存在从 u 到 v 的路。
2. 连通图: 称无向图为连通图, 如果其任意两点间均连通。
3. 连通分支: 无向图的极大连通子图称为连通分支。

**定义 5.2.4 (强连通性)**

1. 点的连通: 对于有向图, 称点 u 与 v 连通, 如果存在从 u 到 v 的路。
2. 强连通图: 称有向图为连通图, 如果其任意两点间均强连通。
3. 强连通分支: 有向图的极大强连通子图称为强连通分支。

**5.2.2 图的割集****定义 5.2.5 (割边)**

称图 (V, E) 的边 e 为割边, 如果成立如下命题之一。

1. 子图 $(N, E \setminus \{e\})$ 的连通分支数严格大于图 (V, E) 的连通分支数。
2. 边 e 不包含在 G 的任意简单回路中。

**定义 5.2.6 (边割)**

对于图 (V, E) , 令

$$\{S, N \setminus S\} = \{\{n_i, n_j\} \in E : n_i \in S, n_j \in N \setminus S\}$$

称子集 $T \subset \{S, N \setminus S\}$ 为边割, 如果子图 $(N, E \setminus T)$ 的连通分支数严格大于图 (V, E) 的连通分支数。

**定义 5.2.7 (割集)**

图的极小边割称为割集。

**定理 5.2.1**

边割为不交割集的并。

**定理 5.2.2**

对于图 G , 如果 C 为 G 的简单回路, Ω 为 G 的割集, 令 $E(C), E(\Omega)$ 分别表示 C, Ω 所包含的边集合, 那么或 $E(C) \cap E(\Omega) = \emptyset$, 或 $|E(C) \cap E(\Omega)| \geq 2$ 。

**5.3 树与支撑树****5.3.1 树及其基本性质****定义 5.3.1 (树)**

称无向图 $T = (V, E)$ 为树, 如果成立如下命题之一。

1. 无回路且连通。
2. 无回路且 $|E| = |V| - 1$ 。
3. 连通且 $|E| = |V| - 1$ 。

4. 连通且每一条边均为割边。
5. 任意两点间存在且存在唯一一条路。
6. 无回路，但任意增加一条新边，存在且存在唯一回路。



证明 $1 \implies 5$: 设 T 连通且无回路，任取两点 i, j ，由于 T 连通，那么 i, j 间存在路。又由于 T 无回路，那么 i, j 间仅存在唯一一条路。

$5 \implies 6$: 设 T 的任意两点间存在且存在唯一一条路，那么 T 无回路。在任意不相邻的两点 i, j 连接，并设 T 中 i, j 间的路为 P ，则 $P + e$ 为 $T + e$ 的唯一回路。

$6 \implies 4$: 设 T 无回路，但任意增加一条新边，存在且存在唯一回路。由于 T 无回路，那么 T 的每一条边均为割边。若 T 不连通，则不妨 T 存在两个连通分支 T_1, T_2 ，那么 T_1 中的点 i 与 T_2 中的点 j 间增加新边不会产生回路。

$4 \implies 3$: 设 T 连通且每一条边均为割边，下面对于 $|N|$ 进行数学归纳。

当 $|N| = 1, 2$ 时，结论平凡。

假设当 $|N| \leq k$ 时结论成立，那么当 $|N| = k + 1$ 时，任取 T 的一条边 e ，由于 e 为割边，因此 $T - e$ 存在两个连通分支 T_1, T_2 。由于 T_1, T_2 的点数 $\leq k$ ，有归纳假设

$$|E_1| = |N_1| - 1, \quad |E_2| = |N_2| - 1$$

因此

$$|E| = |E_1| + |E_2| + 1 = |N_1| - 1 + |N_2| - 1 + 1 = |N| - 1$$

$3 \implies 2$: 设 T 连通且 $|E| = |V| - 1$ ，下面对于 $|N|$ 进行数学归纳。

当 $|N| = 1, 2$ 时，结论平凡。

假设当 $|N| = k$ 时结论成立，那么当 $|N| = k + 1$ 时，由 T 的连通性， T 中点的次数 ≥ 1 。又若 T 中点的次数 ≥ 2 ，则

$$2|N| = \sum_{i \in N} d(i) = 2|E| = 2|N| - 2$$

矛盾！因此存在次数为 1 的点。令 $d(i) = 1$ ，由于 $T' = T - i$ 连通，且

$$|N'| = |N| - 1, \quad |E'| = |E| - 1$$

因此

$$|E'| = |N'| - 1$$

由归纳假设， T' 无回路，因此 T 无回路。

$2 \implies 1$: 设 T 无回路且 $|E| = |V| - 1$ 。由于 T 无回路，因此 T 的每条边均为割边，假若 T 不连通，设 T 存在 k 个连通分支，则可在这些连通分支间增加 $k - 1$ 条边得到 T' ，使得 T' 连通且每条边均为割边，由 $4 \implies 3$

$$|E'| = |N'| - 1 = |N| - 1$$

这与

$$|E'| = |E| + k - 1 = |N| + k - 2$$

矛盾！

定理 5.3.1

每棵树至少存在两个次为 1 的点。



证明 任给树 $T = (N, E)$ ，由于 T 连通，因此 T 中每个点的次数均 ≥ 1 。而

$$\sum_{i \in N} \deg(i) = 2|E| = 2|N| - 2$$

因此至少有两个点的次为 1。

定理 5.3.2

如果图 G 的点的次数均 ≥ 2 ，那么 G 存在回路。



证明 如果 G 存在环，那么其为回路。如果 G 存在重边，其为回路。下设 G 为简单图，令 P 为 G 的极长的路，其端点为 u, v 。由于 $\deg(u) \geq 2$ ，因此 u 除了路 P 中的邻点外还有还有一点 w 在 P 中，此时 $uPw + uw$ 为回路。

5.3.2 支撑树及其基本性质

定义 5.3.2 (支撑树)

称图 G 的支撑子图 T 为支撑树，如果 T 为树。



定义 5.3.3 (反树)

称图 G 的支撑树 T 的反树为 $T^* = G \setminus T$ 。



定义 5.3.4

对于图 G 的支撑树 T ，取 T 的一条边 e ， $T - e$ 的两个连通分支分别为 T_1, T_2 ，其点集合分别为 S_1, S_2 ，记 $\Omega(e)$ 为割集 $\{S_1, S_2\}$ 。



定理 5.3.3

图 G 存在支撑树当且仅当 G 为连通图。



定理 5.3.4

对于连通图 G 的支撑树 T ，如果 e 为 T 的一条边，那么存在且存在唯一割集 $\Omega(e)$ ，使得 $\Omega(e) \subset T^* + e$ 。



定理 5.3.5

对于连通图 G 的支撑树 T_1 和 T_2 ，如果 $|T_1 \cap T_2| = k$ ，那么 T_2 经过 k 次迭代后得到 T_1 。



5.4 最小树问题

5.4.1 最小树及其性质

定义 5.4.1 (最小树)

称连通网络 G 的支撑树 T 为最小树，如果成立如下命题之一。

1. T 为 G 的支撑树中权最小的支撑树。
2. 对于任意 $e \in T^*$ ，成立

$$W(e) = \max_{e' \in C(e)} W(e')$$

其中 $C(e) \subset T + e$ 为唯一回路。

3. 对于任意 $e \in T$ ，成立

$$W(e) = \min\{e' \in \Omega(e)\} W(e')$$

其中 $\Omega(e) \subset T^* + e$ 为唯一割集。



定理 5.4.1

对于连通网络 G 的支撑树 T ，如下命题等价。

1. T 为 G 的唯一最小树
2. 对于任意 $e \in G \setminus T$ ， e 为 $C(e)$ 中唯一最大边。
3. 对于任意 $e \in T$ ， e 为 $\Omega(e)$ 中唯一最小边。

**5.4.2 Kruskal 算法****定理 5.4.2 (Kruskal 算法)**

求网络 $G(V, E, W)$ 的最小树，其中 $|V| = n$ ， $|E| = m$ 。

1. 将边依权排列

$$W(e_1) \leq W(e_2) \leq \cdots \leq W(e_m)$$

令

$$S := \emptyset, \quad i := 0, \quad j := 1$$

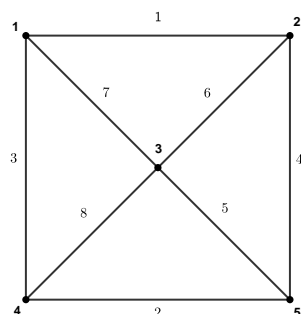
2. 若 $|S| = i = n - 1$ ，停止， $G[S] = T$ 即为最小树；否则转至第 3 步。
3. 若 $G[S \cup \{e_j\}]$ 不构成回路，则令

$$e_{i+1} := e_j, \quad S := S \cup \{e_{i+1}\}, \quad i := i + 1, \quad j := j + 1$$

转至第 2 步；否则令 $j := j + 1$ ，转至第 2 步。



例题 5.1 用 Kruskal 算法求解下图所示网络的最小树。

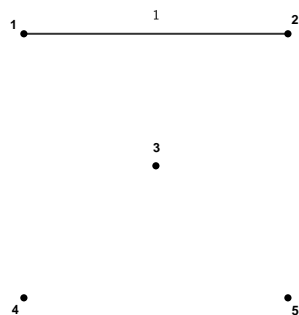


解 边依权排列如下

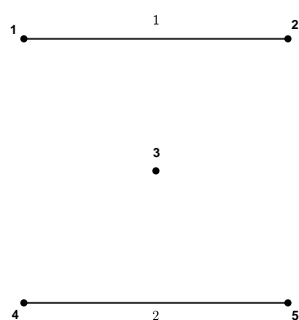
边	e_{12}	e_{45}	e_{14}	e_{25}	e_{35}	e_{23}	e_{13}	e_{34}
权	1	2	3	4	5	6	7	8

在不产生圈的前提下依次加边，直至得到最小树。

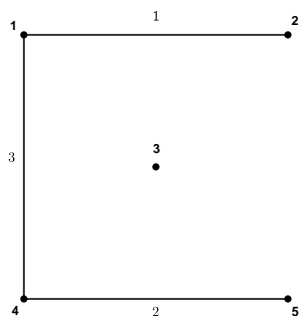
选取 e_{12} ，构成如下图



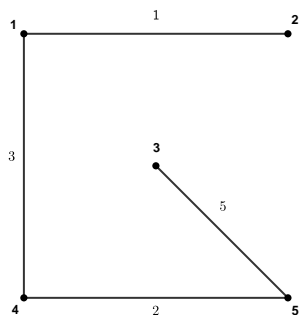
选取 e_{45} , 构成如下图



选取 e_{14} , 构成如下图



若选取 e_{25} , 则构成回路 $1 \rightarrow 2 \rightarrow 5 \rightarrow 4 \rightarrow 1$ 。选取 e_{35} , 构成如下图



此时已构成树，因此为最小树。

5.4.3 Dijkstra 算法

定理 5.4.3 (Dijkstra 算法)

求网络 $G(V, E, W)$ 的最小树，其中

$$V = \{v_1, \dots, v_n\}, \quad \langle v_i, v_j \rangle = e_{ij}, \quad W(e_{ij}) = w_{ij}$$

1. 令

$$u_j := w_{1j}, \quad T := \emptyset, \quad R := \{v_1\}, \quad S := \{v_2, \dots, v_n\}$$

2. 取

$$u_k := \min_{j \in S} \{u_j\} = w_{1k}$$

令

$$T := T \cup \{e_{1k}\}, \quad R := R \cup \{v_k\}, \quad S := S \setminus \{v_k\}$$

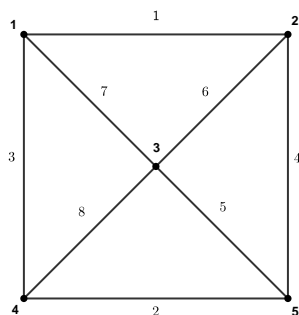
3. 若 $S = \emptyset$ ，停止；否则，令

$$u_j := \min\{u_j, w_{kj}\}, \quad j \in S$$

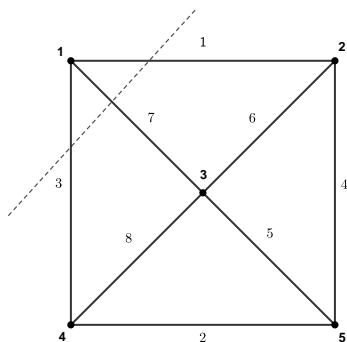
转至第 2 步。



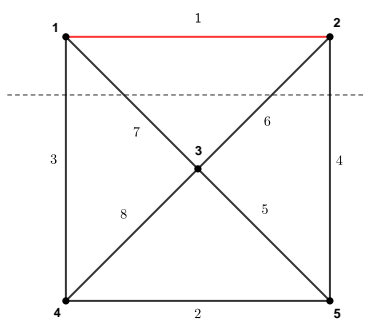
例题 5.2 用 Dijkstra 算法求解下图所示网络的最小树。



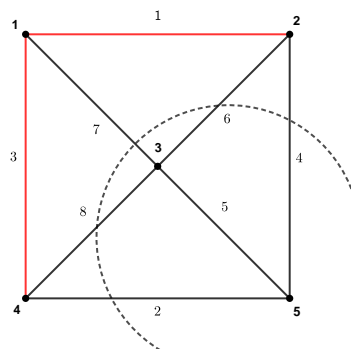
解 选取点 1 为起始点，作割集



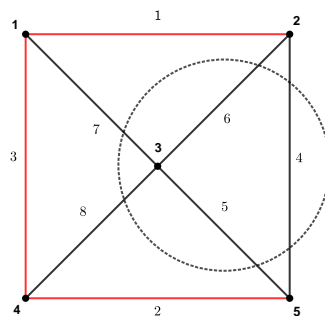
选取边 e_{12} , 作割集



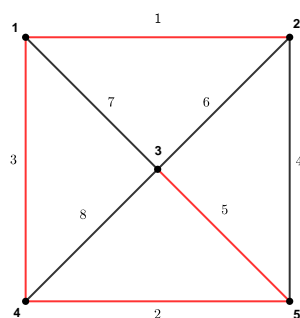
选取边 e_{14} , 作割集



选取边 e_{45} , 作割集



选取边 e_{35} , 构成最小树



5.5 最短有向路问题

定理 5.5.1 (Dijkstra 算法)

对于正权有向网络 $G = (V, E, W)$, 求解点 1 到其他各点最短有向路径。

1. 令

$$u_1 := 0, \quad u_j := w_{1j}, \quad 2 \leq j \leq n, \quad P := \{1\}, \quad T := \{2, \dots, n\}$$

2. 在 T 中寻找一点 k , 使得成立

$$u_k = \min_{j \in T} \{u_j\}$$

令

$$P := P \cup \{k\}, \quad T := T \setminus \{k\}$$

若 $T = \emptyset$, 终止; 否则, 转至第 3 步。

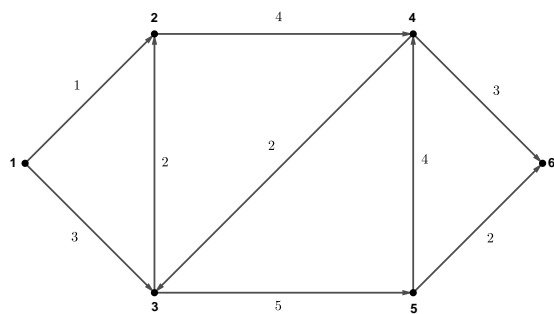
3. 对于 T 中每一点 j , 令

$$u_j = \min\{u_j, u_k + w_{kj}\}$$

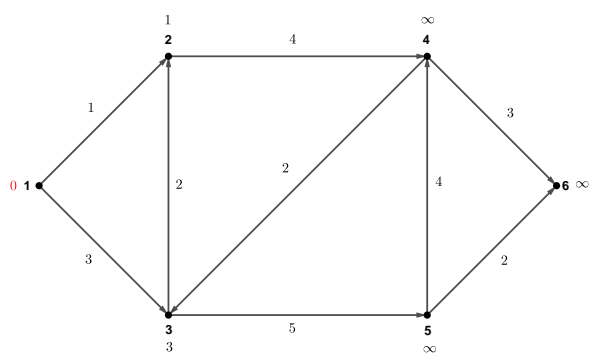
转至第 2 步。



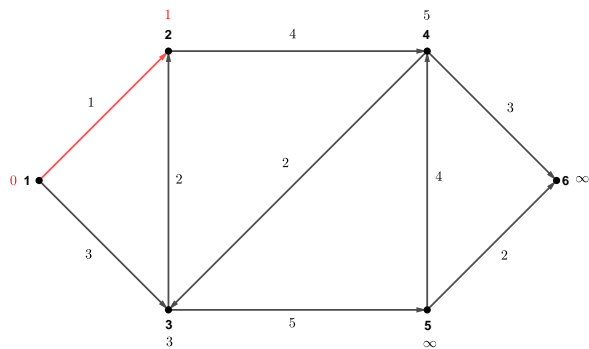
例题 5.3 用 Dijkstra 算法求解下图所示有向网络中自点 1 到其他各点的最短有向路。



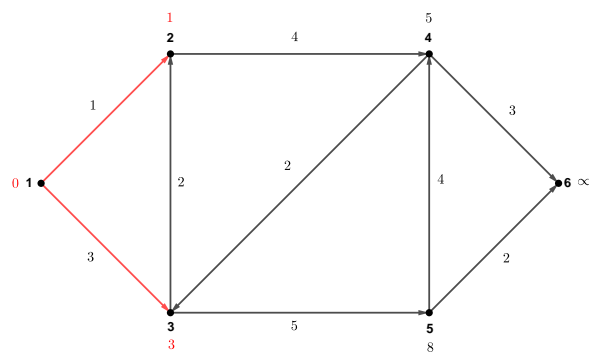
解 作临时标号



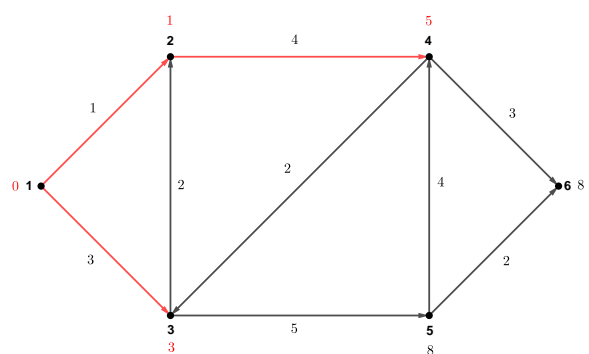
选择点 2 的临时标号为永久标号，并更新其他点的临时标号。



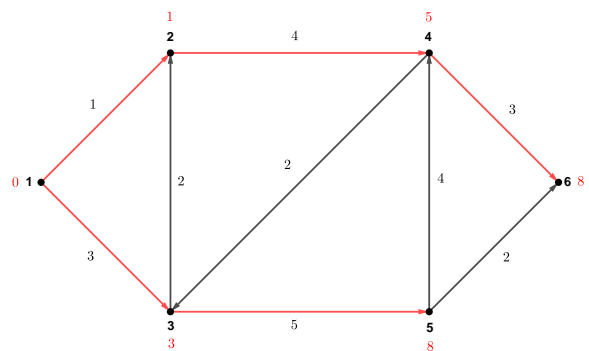
选择点 3 的临时标号为永久标号，并更新其他点的临时标号。



选择点 4 的临时标号为永久标号，并更新其他点的临时标号。



选择点 5, 8 的临时标号为永久标号，因此自点 1 到其他各点的最短有向路为



城市	最短路线	最短路程
2	$1 \rightarrow 2$	1
3	$1 \rightarrow 3$	3
4	$1 \rightarrow 2 \rightarrow 4$	5
5	$1 \rightarrow 3 \rightarrow 5$	8
6	$1 \rightarrow 2 \rightarrow 4 \rightarrow 6$	8

5.6 最大流问题

定义 5.6.1 (最大流问题)

对于有向网络 $G = (V, E, C)$, 其中 c_{ij} 表示有向边 (i, j) 的容量, 令 s 为发点, t 为收点

x_{ij} = 通过有向边 (i, j) 的流量

那么求

$$\max v = \max \sum_j x_{ij} = \max \sum_j x_{ji}$$

使得成立

$$0 \leq x_{ij} \leq c_{ij}$$

$$\sum_j x_{ij} - \sum_j x_{ji} = \begin{cases} +v, & i = s \\ -v, & i = t \\ 0, & \text{其他} \end{cases}$$



定理 5.6.1 (最大流算法)

1. 令 $x = (x_{ij})$ 为整数可行流, 可以为零流, 给 s 一个永久标号 $(-, \infty)$ 。
2. (a). 如果所有标号点的相邻点均已检查, 转至第 4 步。
(b). 找一个已标号但存在未检查的相邻点的点 i , 对于每一个有向边 (i, j) , 若 $x_{ij} < c_{ij}$ 且 j 未标号, 则给 j 一个标号 $(+i, \delta(j))$, 其中

$$\delta(j) = \min\{c_{ij} - x_{ij}, \delta(i)\}$$

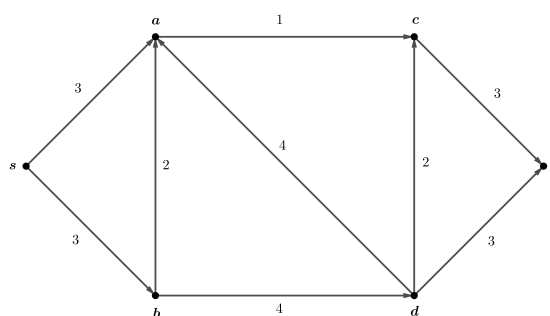
对于每一个有向边 (j, i) , 若 $x_{ji} > 0$ 且 j 未标号, 则给 j 一个标号 $(-i, \delta(j))$, 其中

$$\delta(j) = \min\{x_{ji}, \delta(i)\}$$

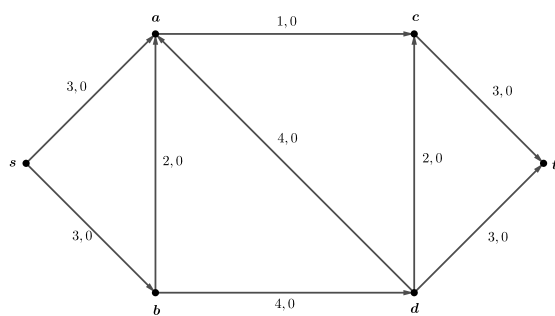
- (c). 若 t 已被标号, 转至第 3 步, 否则转至 2.a。
3. 由点 t 开始, 使用指示标号构造一个增广路, 指示标号的正负表示通过增加还是减少流量来增大流值。删去点 s 以外的所有标号, 转至第 2 步。
4. 此时现行流是最大的, 若把所有标号点的集合记为 S , 所有未标号的点的集合记为 T , 便得到最小容量割 (S, T) , 计算完成。



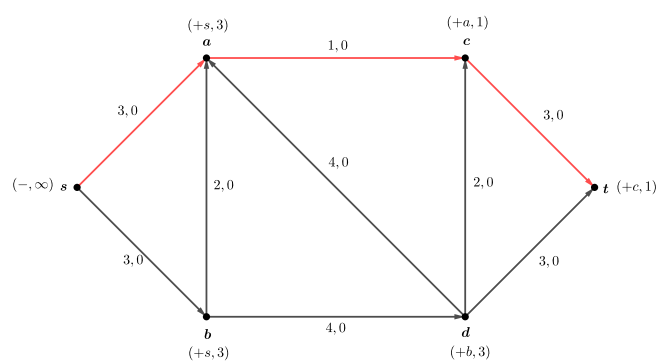
例题 5.4 使用 Ford-Fulkerson 算法求如下图所示的有向网络中从 s 到 t 的最大流。



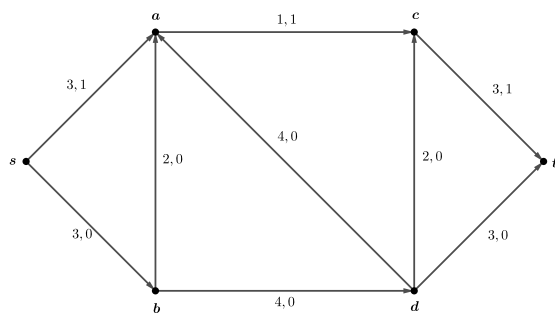
解 作零流



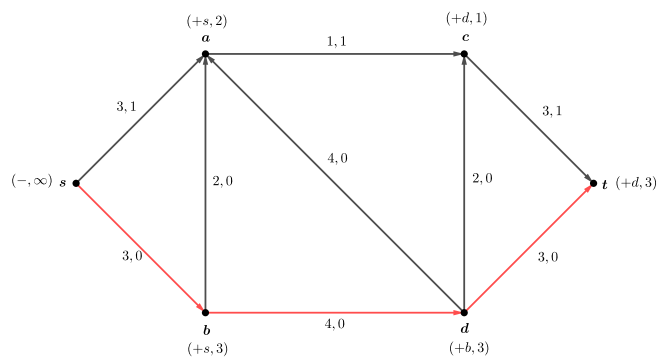
依次标号



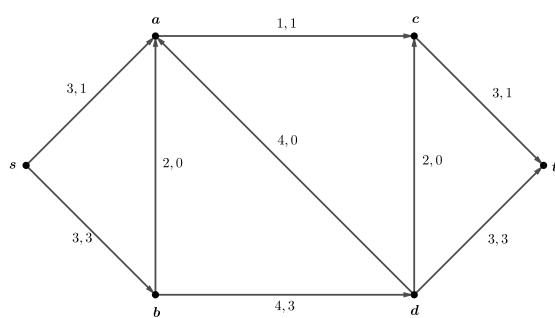
更新流量



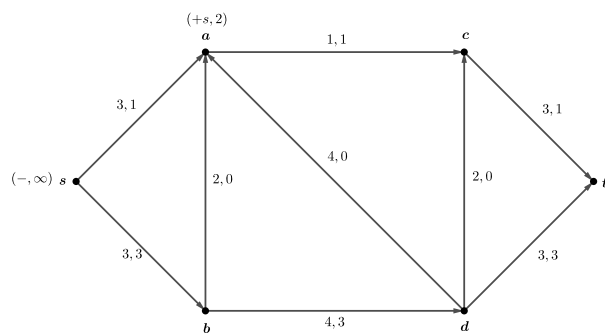
依次标号



更新流量



依次标号



此时流量最大，流量如下

边	(s, a)	(s, b)	(a, c)	(b, d)	(c, t)	(d, t)	其他
流量	1	3	1	3	1	3	0

最大流为 $v = 4$ 。

第六章 排队论

6.1 基本模型

基本组成部分：

1. 输入过程：顾客按照怎样的规律到达。
 - (a). 顾客来源
 - I. 有限
 - II. 无限
 - (b). 顾客数量
 - I. 有限
 - II. 无限
 - (c). 经常性的顾客来源
 - I. 顾客到达间隔时间：到下一个顾客到达的时间。
 - II. 服从某一频率分布（指数分布）。
 - (d). 顾客的行为假定
 - I. 在未服务之间不会离开。
 - II. 当看到队列很长的时候离开。
 - III. 从一个队列移到另一个队列。
2. 排队规则：顾客按照一定规则排队等待服务。
3. 服务机构：服务机构的设置，服务台的数列，服务的方式，服务时间分布等。

记号方案：

1. M ：指数分布
2. D ：定长分布
3. E_k ： k 级 Erlang 分布

定义 6.1.1 (负指数分布)

对于参数为 $1/\lambda$ 的指数分布 ξ ，其分布密度函数为

$$f(t) = \begin{cases} \lambda e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

分布函数为

$$F(t) = \begin{cases} 1 - e^{-\lambda t}, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

数学期望与方差为

$$E(\xi) = \frac{1}{\lambda}, \quad D(\xi) = \frac{1}{\lambda^2}$$

定理 6.1.1 (指数分布的无记忆性)

$$P\{\xi > t + x \mid \xi > t\} = P\{\xi > x\}$$

定义 6.1.2 (最简单流)

称事件流 $N(t)$ 为最简单流, 如果满足如下性质。

1. 平稳性: 对于任意时刻 t_0 , $(t_0, t_0 + t]$ 时间内出现的事件数仅与事件长度 t 有关, 而与起点 t_0 无关。
2. 无记忆性: $(t_0, t_0 + t]$ 时间内出现 k 个事件与时刻 t_0 以前出现的事件数无关。
3. 普通性: 在充分小的时间区间 Δt 内, 发生两个及两个以上事件的概率为比 Δt 高阶的无穷小量。



6.2 M/M/1/∞ 模型

λ : 单位时间内平均到达的顾客人数。

μ : 单位时间平均服务完的顾客人数。

$\rho = \lambda/\mu$: 服务强度。

$N(t)$ 表示 t 时刻在系统中的顾客数量, 其概率公式为

$$p_n = P\{N(t) = n\} = (1 - \rho)\rho^n$$

$N_q(t) = N(t) - 1$ 表示 t 时刻在系统中的排队等待的顾客数量。

在系统中的平均顾客数量, 换言之, 平均队长为

$$L = E(N) = \sum_{n=0}^{\infty} np_n = \sum_{n=0}^{\infty} n(1 - \rho)\rho^n = \frac{\rho}{1 - \rho}$$

在系统中的平均排队等待的顾客数量为

$$L_q = E(N_q) = \sum_{n=1}^{\infty} (n - 1)p_n = \sum_{n=1}^{\infty} (n - 1)(1 - \rho)\rho^n = \frac{\rho^2}{1 - \rho}$$

平均逗留时长为

$$W = W_q + \frac{1}{\mu} = \frac{1}{\mu - \lambda}$$

平均等待时长为

$$W_q = \sum_{n=0}^{\infty} \frac{n}{\mu} p_n = \frac{1}{\mu} E(N) = \frac{\rho}{\mu - \lambda} = \frac{\lambda}{\mu(\mu - \lambda)}$$

第七章 决策分析

7.1 模型建立

决策问题的基本要素：

1. 状态集 $S = \{x\}$ ：把决策的对象统称为一个系统，系统处于不同的状况称为状态。
2. 决策集 $A = \{a\}$ ：为达到预想的目标提出的每一个行动方案称为决策方案。
3. 报酬函数 $R(a, x)$ ：在状态 x 出现时，决策者采取方案 a 所得到收益值。

7.2 风险型决策分析

7.3 不确定型决策分析

乐观法：

$$\Phi : \max_{a \in A} \{ \max_{x \in S} \{ R(a, x) \} \}$$

悲观法：

$$\Phi : \max_{a \in A} \{ \min_{x \in S} \{ R(a, x) \} \}$$

乐观系数法：

$$\Phi : \max_{a \in A} \{ \alpha \max_{x \in S} \{ R(a, x) \} + (1 - \alpha) \min_{x \in S} \{ R(a, x) \} \}$$

等可能法：

$$\Phi : \max_{a \in A} \left\{ \frac{1}{n} \sum_{k=1}^n R(a, x_k) \right\}$$

后悔值法：

$$RV(a, x) = \max_{a \in A} \{ R(a, x) \} - R(a, x)$$

$$\Phi : \min_{a \in A} \{ \max_{x \in S} \{ RV(a, x) \} \}$$

例题 7.1 某公司欲购进一种新产品, 有三种可供选择的方案, 即大批量购进、中批量购进、小批量购进, 在各种市场需要下推销该产品的获利情况如下表所示, 其中负数表示亏损。

表 7.1: 获利情况

利润/万元		市场情况		
		畅销	一般	滞销
方案	大批量购进	600	200	-80
	中批量购进	400	300	-20
	小批量购进	200	100	-10

分别用乐观法、悲观法、乐观系数法（乐观系数 $\alpha = 0.4$ ）、后悔值法和等可能法进行决策分析, 找出最优方案。

解 状态集为

$$S = \{\text{畅销, 一般, 滞销}\}$$

决策集为

$$A = \{\text{大批量购进, 中批量购进, 小批量购进}\}$$

决策表为

表 7.2: 决策表

$R(a, x)/\text{万元}$		市场情况		
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)
方案	a_1 (大批量购进)	600	200	-80
	a_2 (中批量购进)	400	300	-20
	a_3 (小批量购进)	200	100	-10

1. 对于乐观法，乐观法决策表为

表 7.3: 乐观法决策表

$R(a, x)/\text{万元}$		市场情况			$\max_{x \in S}\{R(a, x)\}$
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)	
方案	a_1 (大批量购进)	600	200	-80	600
	a_2 (中批量购进)	400	300	-20	400
	a_3 (小批量购进)	200	100	-10	200

由于

$$\begin{aligned}\Phi : & \max_{a \in A}\{\max_{x \in S}\{R(a, x)\}\} \\ &= \max\{R(a_1, x_1), R(a_2, x_1), R(a_3, x_1)\} \\ &= R(a_1, x_1) \\ &= 600\end{aligned}$$

因此最优方案为大批量购进。

2. 对于悲观法，悲观法决策表为

表 7.4: 悲观法决策表

$R(a, x)/\text{万元}$		市场情况			$\min_{x \in S}\{R(a, x)\}$
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)	
方案	a_1 (大批量购进)	600	200	-80	-80
	a_2 (中批量购进)	400	300	-20	-20
	a_3 (小批量购进)	200	100	-10	-10

由于

$$\begin{aligned}\Phi : & \max_{a \in A} \{ \min_{x \in S} \{ R(a, x) \} \} \\ &= \max \{ R(a_1, x_3), R(a_2, x_3), R(a_3, x_3) \} \\ &= R(a_3, x_3) \\ &= -10\end{aligned}$$

因此最优方案为小批量购进。

3. 对于乐观系数法，乐观系数法决策表为

表 7.5: 乐观系数法决策表

$R(a, x)/\text{万元}$		市场情况			$\alpha \max_{x \in S} \{ R(a, x) \} + (1 - \alpha) \min_{x \in S} \{ R(a, x) \}$
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)	
方案	a_1 (大批量购进)	600	200	-80	192
	a_2 (中批量购进)	400	300	-20	148
	a_3 (小批量购进)	200	100	-10	74

由于

$$\begin{aligned}\Phi : & \max_{a \in A} \{ \alpha \max_{x \in S} \{ R(a, x) \} + (1 - \alpha) \min_{x \in S} \{ R(a, x) \} \} \\ &= \max_{a \in A} \{ 0.4 \max_{x \in S} \{ R(a, x) \} + 0.6 \min_{x \in S} \{ R(a, x) \} \} \\ &= \max \{ 192, 148, 74 \} \\ &= 192\end{aligned}$$

因此最优方案为大批量购进。

4. 对于后悔值法，后悔值函数为

$$RV(a, x) = \max_{a \in A} \{ R(a, x) \} - R(a, x)$$

后悔值表为

表 7.6: 后悔值表

$RV(a, x)/\text{万元}$		市场情况			$\max_{x \in S} \{ RV(a, x) \}$
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)	
方案	a_1 (大批量购进)	0	100	70	100
	a_2 (中批量购进)	200	0	10	200
	a_3 (小批量购进)	400	200	0	400

由于

$$\Phi : \min_{a \in A} \{ \max_{x \in S} \{ RV(a, x) \} \} = 100$$

因此最优方案为大批量购进。

5. 对于等可能法，等可能法决策表为

表 7.7: 等可能法决策表

$R(a, x)/\text{万元}$		市场情况			$\frac{1}{n} \sum_{k=1}^n R(a, x_k)$
		x_1 (畅销)	x_2 (一般)	x_3 (滞销)	
方案	a_1 (大批量购进)	600	200	-80	240
	a_2 (中批量购进)	400	300	-20	680/3
	a_3 (小批量购进)	200	100	-10	290/3

由于

$$\begin{aligned}
 \Phi : \quad & \max_{a \in A} \left\{ \frac{1}{n} \sum_{k=1}^n R(a, x_k) \right\} \\
 & \max_{a \in A} \left\{ \frac{1}{3} \sum_{k=1}^3 R(a, x_k) \right\} \\
 & = \max\{240, 680/3, 290/3\} \\
 & = 240
 \end{aligned}$$

因此最优方案为大批量购进。

附录 A 期末复习

A.1 填空与判断

对应于基 B 的典式: $\mathbf{x}_B + \mathbf{B}^{-1}\mathbf{N}\mathbf{x}_N = \mathbf{B}^{-1}\mathbf{b}$

检验数向量: $\boldsymbol{\zeta}^T = \mathbf{c}_B^T \mathbf{B}^{-1} \mathbf{A} - \mathbf{c}^T$

0.618 法:

$$\omega = \frac{t_2 - a}{b - a} = \frac{b - t_1}{b - a}$$

Newton 法:

$$t_{k+1} = t_k - \frac{\varphi'(t_k)}{\varphi''(t_k)}$$

共轭梯度法:

$$\lambda_k = \frac{\|\nabla f(\mathbf{x}^{k+1})\|^2}{\|\nabla f(\mathbf{x}^k)\|^2}$$

简约梯度法:

$$t_{\max}^k = \begin{cases} +\infty, & \mathbf{p}^k \geq \mathbf{0} \\ \min \left\{ -\frac{x_i^k}{p_i^k} : p_i^k < 0, 1 \leq i \leq n \right\}, & \text{其他} \end{cases}$$

简约梯度:

$$\mathbf{r}_N = \nabla F(\mathbf{x}_N) = -(\mathbf{B}^{-1}\mathbf{N})^T \nabla_B f(\mathbf{x}) + \nabla_N f(\mathbf{x})$$

树: 无回路, 连通, $|E| = |V| - 1$

排队系统的组成: 输入过程, 排队规则, 服务机构

最简单流的性质: 平稳性, 无记忆性, 普通性

决策问题的基本要素: 状态集 $S = \{x\}$, 决策集 $A = \{a\}$, 报酬函数 $R(a, x)$