最优化方法

Methods of Optimization

目录

第一章　基础知识

1.1　数学规划

规划论，又称数学规划，是运筹学的分支，研究在约束条件下通过分配资源寻求目标函数极值的数学方法，应用于经济管理、工程设计和过程控制等领域. 其核心是建立约束条件与目标函数的数学模型，这种数学模型的一般形式为

$$\begin{aligned} \text{Opt.} \quad & f(x_i, y_j, \xi_k) \\ \text{s.t.} \quad & g_h(x_i, y_j, \xi_k) \leq (=, \geq) 0 \\ & h = 1, 2, \cdots, m \end{aligned}$$

其中，x_i 是控制变量，y_j 是已知参数，ξ_k 是随机因素，f 是目标函数，g_h 是约束函数.

按照目标函数和自变量取值范围的不同，数学规划可分为：

- 线性规划：f, g 都是线性函数；
- 非线性规划：f, g 中有非线性函数；
- 多目标规划：f 是向量函数；
- 整数规划：决策变量 x_i 为整数；
- 动态规划：含多阶段决策过程；
- 随机规划：含有随机因子.

若记

$$S = \{\boldsymbol{x} \mid g_h(x_i, y_j, \xi_k) \leq (=, \geq) 0\} \subseteq \mathbb{R}^n$$

为约束集合，则这种数学模型可进一步简记为

$$(f\ S) \begin{cases} \min & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{x} \in S \end{cases}$$

其中 $\boldsymbol{x} \in S$ 称模型 $(f\ S)$ 的可行解.

定义 1.1　（全局最优解、最优值）

若对于 $\boldsymbol{x}^* \in S$，满足 $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$，$\forall \boldsymbol{x} \in S$，则称 $\boldsymbol{x}^*$ 为 $(f\ S)$ 的全局最优解（最优解），记作 g.opt.，简记为 opt.. 称此时的 $f(\boldsymbol{x}^*)$ 为 $(f\ S)$ 的最优值（最优目标函数值）.

定义 1.2 （局部最优解）

若对于 $\boldsymbol{x}^* \in S$，$\exists \boldsymbol{x}^*$ 的某邻域 $U(\boldsymbol{x}^*)$，使得 $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$，$\forall \boldsymbol{x} \in S$，则称 $\boldsymbol{x}^*$ 为 $(f\ S)$ 的局部最优解，记作 l.opt.，

在上述定义中，若当 $\boldsymbol{x} \neq \boldsymbol{x}^*$ 时有严格不等式成立，则分别称 $\boldsymbol{x}^*$ 为 $(f\ S)$ 的严格全局最优解和严格局部最优解.

1.2　基本概念和符号

接下来将约定一些术语的符号表示：

- $\mathbb{R}^n$——n 维欧式空间；
- $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^\top\in\mathbb{R}^n$——表示 $\mathbb{R}^n$ 中的一个点或一个向量，其中分量 $x_i\in\mathbb{R}$；
- $\boldsymbol{d}=(d_1,d_2,\cdots,d_n)^\top\in\mathbb{R}^n$，其中 $\boldsymbol{d}\neq 0$——表示从 0 指向 $\boldsymbol{d}$ 的方向. 此外，设 $\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)}\in\mathbb{R}^n$，若 $\exists\lambda>0$，使得 $\boldsymbol{d}^{(1)}=\lambda\boldsymbol{d}^{(2)}$，则称 $\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)}$ 是**同方向**的.
- $\boldsymbol{x}+\lambda\boldsymbol{d}$——表示从 $\boldsymbol{x}$ 点出发，沿 d 方向移动 $\lambda\boldsymbol{d}$ 长度得到的点.

下面是对欧式空间中向量的运算的定义.

定义 1.3 （向量的运算）

设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^\top\in\mathbb{R}^n$，$\boldsymbol{y}=(y_1,y_2,\cdots,y_n)^\top\in\mathbb{R}^n$，则

（1）**向量的内积运算**：$\boldsymbol{x}^\top\boldsymbol{y}=\boldsymbol{y}^\top\boldsymbol{x}=\sum_{i=1}^{n}x_iy_i$；

（2）**向量长度**：$\|\boldsymbol{x}\|=\sqrt{\boldsymbol{x}^\top\boldsymbol{x}}=\sqrt{\sum_{i=1}^{n}x_i^2}$；

（3）**两点间距离**：$\|\boldsymbol{x}-\boldsymbol{y}\|$.

定理 1.1 （三角不等式）

$\forall\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$，有

$$\|\boldsymbol{x}+\boldsymbol{y}\|\leq\|\boldsymbol{x}\|+\|\boldsymbol{y}\| \tag{1.1}$$

定义 1.4 （点列的收敛）

设 $\{\boldsymbol{x}^{(n)}\}$ 是定义在 $\mathbb{R}^n$ 上的点列，若 $\exists\boldsymbol{x}\in\mathbb{R}^n$，使得 $\lim_{n\to\infty}\|\boldsymbol{x}^{(n)}-\boldsymbol{x}\|=0$，则称点列 $\{\boldsymbol{x}^{(n)}\}$ 收敛到 $\boldsymbol{x}$，记作

$$\lim_{n\to\infty}\boldsymbol{x}^{(n)}=\boldsymbol{x}$$

定义 1.5 （向量的大小关系）

设 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$，若 $\forall i\in\{1,2,\cdots,n\}$，都有 $x_i\leq y_i$，则称 $\boldsymbol{x}\leq\boldsymbol{y}$. 类似可规定 $\boldsymbol{x}\geq\boldsymbol{y}$，$\boldsymbol{x}<\boldsymbol{y}$，$\boldsymbol{x}>\boldsymbol{y}$.

显然，在如上定义中，$(\mathbb{R}^n,\leq)$ 并不是全序集（证明略）. 特别地，$\boldsymbol{x}\leq 0\Leftrightarrow\forall i\in\{1,2,\cdots,n\},x_i\leq 0$.

定理 1.2

设 $\boldsymbol{x},\boldsymbol{y}\in\mathbb{R}^n$，若 $\forall\boldsymbol{y}\geq 0$，都有 $\boldsymbol{x}^\top\boldsymbol{y}\leq\alpha$，则有 $\boldsymbol{x}\leq 0$ 且 $\alpha\geq 0$.

该命题还有以下等价表述：

- 若 $\forall\boldsymbol{y}\leq 0$，都有 $\boldsymbol{x}^\top\boldsymbol{y}\leq\alpha$，则有 $\boldsymbol{x}\geq 0$ 且 $\alpha\geq 0$；

- 若 $\forall \boldsymbol{y} \geq 0$，都有 $\boldsymbol{x}^\top \boldsymbol{y} \geq \alpha$，则有 $\boldsymbol{x} \geq 0$ 且 $\alpha \leq 0$；
- 若 $\forall \boldsymbol{y} \leq 0$，都有 $\boldsymbol{x}^\top \boldsymbol{y} \geq \alpha$，则有 $\boldsymbol{x} \leq 0$ 且 $\alpha \leq 0$.

1.3　子空间、正交子空间

定义 1.6 （子空间）

设 $\boldsymbol{d}^{(i)} \in \mathbb{R}^n, i = 1,2,\cdots,m$ 是线性无关的向量组，称集合 $\{\boldsymbol{x} \mid \boldsymbol{x} = \sum_{j=1}^{m} \alpha_j \boldsymbol{d}^{(j)}, \alpha_J \in \mathbb{R}\}$ 是由向量 $\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)},\cdots,\boldsymbol{d}^{(m)}$ 生成的**子空间**，记作 $L(\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)},\cdots,\boldsymbol{d}^{(m)})$，简记作 L.

可以看到，$\forall L \subseteq \mathbb{R}^n$，都有 $0 \in L$. 若令 $\boldsymbol{A} = (\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)},\cdots,\boldsymbol{d}^{(m)}) \in \mathbb{R}^{n\times m}$，$\boldsymbol{\alpha} = (\alpha_1,\alpha_2,\cdots,\alpha_m)^\top \in \mathbb{R}^m$，则

$$\boldsymbol{x} \in L \Leftrightarrow \exists \boldsymbol{\alpha} \in \mathbb{R}^m, \boldsymbol{x} = \boldsymbol{A}\boldsymbol{\alpha}$$

并且 $\mathrm{Rank}(\boldsymbol{A}) = m$，也就是说 $\boldsymbol{A}$ 是**列满秩**的. 我们将会在接下来的证明中用到这个性质.

定义 1.7 （正交子空间）

称集合 $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{x}^\top \boldsymbol{y} = 0, \forall \boldsymbol{y} \in L\}$ 为子空间 L 的**正交子空间**，记作 $L^\perp$.

不难看出 $\forall L \subseteq \mathbb{R}^n$，都有 $0 \in L^\top$. 且有

$$\begin{aligned} \boldsymbol{x} \in L^\perp &\Leftrightarrow \forall \boldsymbol{y} \in L, \boldsymbol{x}^\top \boldsymbol{y} = 0 \\ &\Leftrightarrow \forall \boldsymbol{\alpha} \in \mathbb{R}^m, \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{\alpha} = 0 \end{aligned}$$

定理 1.3 （子空间投影定理）

设 L 为 $\mathbb{R}^n$ 的子空间，则 $\forall \boldsymbol{z} \in \mathbb{R}^n$，则必有以下两个命题成立：

（1）$\exists! \boldsymbol{x} \in L, \boldsymbol{y} \in L^\perp$，使得 $\boldsymbol{z} = \boldsymbol{x} + \boldsymbol{y}$；

（2）在（1）中的 $\boldsymbol{x}$ 为问题

$$\begin{aligned} \min_{\boldsymbol{u}} \quad & \|\boldsymbol{z} - \boldsymbol{u}\| \\ \text{s.t.} \quad & \boldsymbol{u} \in L \end{aligned}$$

的唯一解，且该问题的最优值为（1）中的 $\|\boldsymbol{y}\|$.

定理 1.4

设 $\boldsymbol{x},\boldsymbol{y} \in \mathbb{R}^n$，若 $\forall \boldsymbol{y} \in L$，都有 $\boldsymbol{x}^\top \boldsymbol{y} \le \alpha$，则有 $\boldsymbol{x} \in L^\perp$ 且 $\alpha \ge 0$.

该命题还有以下等价表述：

- 若 $\forall \boldsymbol{y} \in L$，都有 $\boldsymbol{x}^\top \boldsymbol{y} \ge \alpha$，则有 $\boldsymbol{x} \in L^\perp$ 且 $\alpha \le 0$.

1.4 多元函数及其导数

多元函数 $f:\mathbb{R}^n\to\mathbb{R}$ 可看作一个 n 维向量 $\boldsymbol{x}$ 的函数 $f(\boldsymbol{x})$. 受篇幅所限，我们并不给出矩阵导数的定义，具体请参见《矩阵分析》课程相关内容. 在此我们只给出梯度（一阶偏导数向量）和 Hesse 阵（二阶偏导数矩阵）.

定义 1.8 （梯度、Hesse 阵）

设函数 $f:S\subseteq\mathbb{R}^n\to\mathbb{R}$ 在其定义域上分别一阶可微和二阶可微，则

（1）梯度 $\nabla f(\boldsymbol{x})=\left(\dfrac{\partial f}{\partial x_1},\dfrac{\partial f}{\partial x_2},\cdots,\dfrac{\partial f}{\partial x_n}\right)^\top$；

（2）Hesse 阵 $\nabla^2 f(x)=\begin{pmatrix}\dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_2\partial x_1} & \cdots & \dfrac{\partial^2 f}{\partial x_n\partial x_1}\\ \dfrac{\partial^2 f}{\partial x_1\partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f}{\partial x_n\partial x_2}\\ \vdots & \vdots & \ddots & \vdots\\ \dfrac{\partial^2 f}{\partial x_1\partial x_n} & \dfrac{\partial^2 f}{\partial x_2\partial x_n} & \cdots & \dfrac{\partial^2 f}{\partial x_n^2}\end{pmatrix}$

定理 1.5 （一些常见函数的梯度）

- 线性函数：$f(\boldsymbol{x})=\boldsymbol{c}^\top\boldsymbol{x}+\boldsymbol{d}$，$\nabla f(\boldsymbol{x})=\boldsymbol{c}$，$\nabla^2 f(x)=0$；
- 二次函数：$f(\boldsymbol{x})=\dfrac{1}{2}\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{c}^\top\boldsymbol{x}+\boldsymbol{d}$，$\nabla f(\boldsymbol{x})=\boldsymbol{Q}\boldsymbol{x}+\boldsymbol{c}$，$\nabla^2 f(x)=\boldsymbol{Q}$，其中 $\boldsymbol{Q}$ 为对称阵.
- 向量值线性函数：$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}+\boldsymbol{d}\in\mathbb{R}^m$，$\dfrac{\mathrm{d}\boldsymbol{f}}{\mathrm{d}\boldsymbol{x}}=\boldsymbol{A}^\top$.

例 1.1 求下列函数的梯度和 Hesse 矩阵.

（1）$f_1(\boldsymbol{x})=x_2^1-x_1+x_2^2+x_1x_2+9$；

（2）$f_2(\boldsymbol{x})=2x_2^1+x_2^2+5x_3^2+3x_1x_2-6x_1x_3+4x_1x_2x_3+17$；

（3）$f_3(\boldsymbol{x})=10-(x_2-x_1^2)^2$.

接下来依旧不加证明地介绍多元函数的 Taylor 展开式及中值定理. 此处不加证明是因为此内容可以被读者在任何一本《数学分析》教材中找到.

定理 1.6 （Taylor 展开式）

设 $f:\mathbb{R}^n\to\mathbb{R}$ 在 $\boldsymbol{x}^*$ 的某邻域内分别一阶、二阶可微，则分别有：

- 一阶 Taylor 展开式：$f(\boldsymbol{x})=f(\boldsymbol{x}^*)+\nabla f^\top(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)+o(\|\boldsymbol{x}-\boldsymbol{x}^*\|)$；
- 二阶 Taylor 展开式：$f(\boldsymbol{x})=f(\boldsymbol{x}^*)+\nabla f^\top(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)+\dfrac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^\top\nabla^2 f(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)+o(\|\boldsymbol{x}-\boldsymbol{x}^*\|^2)$.

定理 1.7 （中值定理）

设 $f:\mathbb{R}^n\to\mathbb{R}$ 在 $\boldsymbol{x}^*$ 的某邻域 $U(\boldsymbol{x}^*)$ 内二阶可导，则 $\forall x\in U(\boldsymbol{x}^*)$，都有：

- Lagrange 中值定理：$\exists\lambda\in(0,1)$，记 $\boldsymbol{x}_\lambda=\boldsymbol{x}^*+\lambda(\boldsymbol{x}-\boldsymbol{x}^*)$，则有

$$f(\boldsymbol{x})=f(\boldsymbol{x}^*)+\nabla f^\top(\boldsymbol{x}_\lambda)(\boldsymbol{x}-\boldsymbol{x}^*)$$

- Taylor 中值定理：$\exists\mu\in(0,1)$，记 $\boldsymbol{x}_\mu=\boldsymbol{x}^*+\mu(\boldsymbol{x}-\boldsymbol{x}^*)$，则有

$$f(\boldsymbol{x})=f(\boldsymbol{x}^*)+\nabla f^\top(\boldsymbol{x}^*)(\boldsymbol{x}-\boldsymbol{x}^*)+\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^*)^\top\nabla^2 f(\boldsymbol{x}_\mu)(\boldsymbol{x}-\boldsymbol{x}^*)$$

定义 1.9 （方向导数）

设 $S\subseteq\mathbb{R}^n$ 为非空凸集，函数 $f:S\to\mathbb{R}$，$\boldsymbol{x}^*\in S$，方向 $\boldsymbol{d}\in\mathbb{R}^n$，若 $\exists\lambda>0$，使得 $\boldsymbol{x}^*+\lambda\boldsymbol{d}\in S$（通常令 λ 充分小，此时称 $\boldsymbol{d}$ 为**可行方向**），且此时极限

$$\lim_{\lambda\to0^+}\frac{f(\boldsymbol{x}^*+\lambda\boldsymbol{d})-f(\boldsymbol{x}^*)}{\lambda}$$

存在，则称 $f(\boldsymbol{x})$ 在点 $\boldsymbol{x}^*$ 沿方向 $\boldsymbol{d}$ 的**方向导数**存在，记作 $f'(\boldsymbol{x}^*;\boldsymbol{d})$.

定理 1.8 （方向导数和梯度的关系）

若 $f(\boldsymbol{x})$ 在 $\boldsymbol{x}^*$ 处可微，则

$$f'(\boldsymbol{x}^*;\boldsymbol{d})=\nabla f^\top(\boldsymbol{x}^*)\boldsymbol{d}$$

1.5　凸集、凸函数和凸规划

1.5.1　凸集

1. 凸集、凸组合

定义 1.10 （凸集）

设 $S \subseteq \mathbb{R}^n$，若 $\forall \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S$，$\lambda \in [0,1]$，必有

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

则称 S 为**凸集**.

规定单点集 $\{\boldsymbol{x}\}$ 和空集 $\emptyset$ 都是空集. 需要注意的是，方程 $\boldsymbol{x} = \lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)} = \boldsymbol{x}^{(2)} + \lambda(\boldsymbol{x}^{(1)} - \boldsymbol{x}^{(2)})$，$\lambda \in [0,1]$ 表示连接 $\boldsymbol{x}^{(1)}$ 和 $\boldsymbol{x}^{(2)}$ 的线段.

例 1.2 设 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$，$\boldsymbol{b} \in \mathbb{R}^n$，证明以下命题成立：

（1）集合 $S = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}\}$ 是凸集；

（2）集合 $S = \{\boldsymbol{x} \mid \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}\}$ 是凸集.

在介绍凸集的性质之前，我们先对一些将要用到的名词进行必要的说明.

定义 1.11 （内点、内点集、开集）

设 $S \subseteq \mathbb{R}^n$，$\boldsymbol{x} \in S$，如果存在一个开球 $B(\boldsymbol{x}, \varepsilon) = \{\boldsymbol{y} \in \mathbb{R}^n \mid \|\boldsymbol{y} - \boldsymbol{x}\| < \varepsilon\}$，使得 $B(\boldsymbol{x}, \varepsilon) \subseteq S$，则称 $\boldsymbol{x}$ 是 S 的**内点**，S 的内点的全体称为 S 的**内点集**，记作 $\mathrm{int}(S)$.

进一步地，如果 $\mathrm{int}(S) = S$，则称 S 是**开集**

所以 $\boldsymbol{x} \in \mathrm{int}(S) \Leftrightarrow \exists \varepsilon > 0, (\|\boldsymbol{y} - \boldsymbol{x}\| < \varepsilon \Rightarrow \boldsymbol{y} \in S)$.

定义 1.12 （聚点、闭包）

设 $S \subseteq \mathbb{R}^n$，$\boldsymbol{x} \in \mathbb{R}^n$，如果 $\forall \varepsilon > 0$ 都有 $B(\boldsymbol{x}, \varepsilon) \cap S \neq \emptyset$，则称 $\boldsymbol{x}$ 是 S 的**聚点**，S 的聚点的全体称为 S 的**闭包**，记作 $\overline{S}$.

定义 1.13 （超平面）

设 $\boldsymbol{\alpha} \in \mathbb{R}^n \backslash \{0\}$，$\beta \in \mathbb{R}$，称集合 $H = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{\alpha}^\top \boldsymbol{x} = \beta\}$ 是**超平面**，简记作 $H : \boldsymbol{\alpha}^\top \boldsymbol{x} = \beta$.

定义 1.14 （支撑超平面）

设 $S \subseteq \mathbb{R}^n$，$\boldsymbol{x}_0 \in S$，若超平面 $H : \boldsymbol{\alpha}^\top \boldsymbol{x} = \beta$ 满足：

（i）$\exists \boldsymbol{x}^{(0)} \in S$，使得 $\boldsymbol{x}^{(0)} \in H$；

（ii）$\forall \boldsymbol{x} \in S$，$\boldsymbol{\alpha}^\top \boldsymbol{x} \le \beta$

则称 H 是 S 的**支撑超平面**.

定义 1.15 （分离超平面）

设 $S_1, S_2 \subseteq \mathbb{R}^n$，且 $S_1 \cap S_2 = \emptyset$. 若存在超平面 $H: \boldsymbol{\alpha}^\top \boldsymbol{x} = \beta$，使得 $\forall \boldsymbol{x}^{(1)} \in S_1$，$\boldsymbol{\alpha}^\top \boldsymbol{x}^{(1)} \geq \beta$，且 $\forall \boldsymbol{x}^{(2)} \in S_2$，$\boldsymbol{\alpha}^\top \boldsymbol{x}^{(2)} \leq \beta$，则称 H 是 S_1, S_2 的**分离超平面**.

定理 1.9 （凸集的性质）

凸集有以下性质：

（1）凸集的交是凸集；

（2）凸集的内点集是凸集；

（3）凸集的闭包是凸集；

（4）凸集边界上任意点存在支撑超平面；

（5）两个互相不交的凸集之间存在分离超平面.

定义 1.16 （凸组合）

设 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)} \in \mathbb{R}^n$. 若 $\forall j \in \{1, 2, \cdots, m\}$，$\lambda_j \geq 0$ 且 $\sum\limits_{j=1}^{m} \lambda_j = 1$，则称 $\sum\limits_{j=1}^{m} \lambda_j \boldsymbol{x}^{(j)}$ 为 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)}$ 的**凸组合**.

定理 1.10 （凸集和凸组合的关系）

S 是凸集 $\Leftrightarrow$ S 中任意有限点的凸组合属于 S.

定义 1.17 （凸包）

设非空集合 $S \subseteq \mathbb{R}^n$，由 S 中所有有限点的凸组合所构成的集合，被称为 S 的**凸包**，记作 $\operatorname{cov}(S)$.

定理 1.11 （凸集和凸包的关系）

如果 S 是凸集，那么 $\operatorname{cov}(S) = S$.

定义 1.18 （多胞形）

设 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)} \in \mathbb{R}^n$. 由 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)}$ 的所有凸组合构成的集合称为多胞形，记作 $H(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)})$.

定义 1.19 （单纯形）

设 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)} \in \mathbb{R}^n$. 若多胞形 $H(\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)})$ 满足 $\boldsymbol{x}^{(2)} - \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(3)} - \boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(m)} - \boldsymbol{x}^{(1)}$ 线性无关，则称该多胞形是单纯形.

2. 凸锥、半正组合

定义 1.20 （锥、凸锥）

设 $S \subseteq \mathbb{R}^n$，若 $\forall \boldsymbol{x} \in S$，$\lambda > 0$，必有 $\lambda \boldsymbol{x} \in S$，则称 S 为以 0 为顶点的**锥**. 若 S 还是凸集，则称 S 为**凸锥**.

规定 $\{0\}, \mathbb{R}^n$ 都是凸锥. 锥不一定包含 0.

定义 1.21 （半正组合）

设 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)} \in \mathbb{R}^n$. 若 $\forall j \in \{1,2,\cdots,m\}$，$\lambda_j \geq 0$ 且 $\sum\limits_{j=1}^{m} \lambda_j > 0$，则称 $\sum\limits_{j=1}^{m} \lambda_j \boldsymbol{x}^{(j)}$ 为 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)}$ 的**半正组合**.

定理 1.12 （凸锥和半正组合的关系）

S 是凸锥 $\Leftrightarrow$ S 中任意有限点的半正组合属于 S.

1.5.2 凸函数

1. 凸函数

定义 1.22 （凸函数）

设集合 $S \subseteq \mathbb{R}^n$ 为凸集，函数 $f: S \to R$. 若 $\forall \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S$，$\lambda \in (0,1)$，均有

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

则称 $f(\boldsymbol{x})$ 为凸集 S 上的**凸函数**.

若进一步有上面不等式以严格不等式成立，则称 $f(\boldsymbol{x})$ 为凸集 S 上的**严格凸函数**. 当 $-f(\boldsymbol{x})$ 为凸函数（严格凸函数时），则称 $f(\boldsymbol{x})$ 为**凹函数（严格凹函数）**.

定理 1.13 （Jensen 不等式）

$f(\boldsymbol{x})$ 为凸集 S 上的凸函数的充要条件是 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，即 $\forall \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(m)} \in S$，对于任意的序列 $\{\lambda_j\}_{j=1}^{m}$ 满足 $\forall j \in \{1,2,\cdots,m\}$，$\lambda_j \geq 0$ 且 $\sum\limits_{j=1}^{m} \lambda_j = 1$，都有

$$f\left(\sum_{j=1}^{m} \lambda_j \boldsymbol{x}^{(j)}\right) \leq \sum_{j=1}^{m} \lambda_j f(\boldsymbol{x}^{(j)})$$

例 1.3 设 f_1, f_2 是凸函数，判断：

（1）设 $\lambda_1, \lambda_2 > 0$，$\lambda_1 f_1 + \lambda_2 f_2$ 是否为凸函数？$\lambda_1 f_1 - \lambda_2 f_2$ 是否为凸函数？

（2）$f(\boldsymbol{x}) = \max\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$ 是否为凸函数？$g(\boldsymbol{x}) = \min\{f_1(\boldsymbol{x}), f_2(\boldsymbol{x})\}$ 是否为凸函数？

定理 1.14 （凸函数的性质）

设 $S \subseteq \mathbb{R}^n$ 为非空凸集，函数 $f: S \to \mathbb{R}$，则有：

（1）若 f 是凸函数，则 f 在 $\mathrm{int}(S)$ 上连续；

（2）若 f 是凸函数，则其对任意方向的方向导数（若方向可行）存在；

（3）设 S 是开集，f 在 S 上可微，则 f 是凸函数的充要条件是 $\forall \boldsymbol{x}^* \in S$，都有

$$f(\boldsymbol{x}) \geq f(\boldsymbol{x}^*) + \nabla f^{\top}(\boldsymbol{x}^*)(\boldsymbol{x} - \boldsymbol{x}^*)$$

$\forall \boldsymbol{x} \in S$ 都成立. 进一步地，如果 f 是严格凸函数，则其充要条件是上述不等式的严格不等式成立；

（4）设 S 是开集，f 在 S 上二次可微，则：

（i）f 是凸函数的充要条件是 $\forall \boldsymbol{x} \in S$，$\nabla^2 f(\boldsymbol{x})$ 是半正定的；

（ii）若 $\forall \boldsymbol{x} \in S$，$\nabla^2 f(\boldsymbol{x})$ 是正定的，则 f 是严格凸函数.

例 1.4 若 $f(x) = \begin{cases} 2, & x = \pm 1, \\ x^2, & -1 < x < 1 \end{cases}$，容易验证 f 是凸函数，但是 f 在边界点上不连续.

2. 水平集

定义 1.23 （水平集）

设集合 $S \subseteq \mathbb{R}^n$，函数 $f: S \to \mathbb{R}$，$\alpha \in \mathbb{R}$，称 $S_\alpha = \{\boldsymbol{x} \in S \mid f(\boldsymbol{x}) \leq \alpha\}$ 为 $f(\boldsymbol{x})$ 在 S 上的 α **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.

定理 1.15 （凸函数的水平集是凸集）

设集合 $S \subseteq \mathbb{R}^n$ 是凸集，函数 $f: S \to \mathbb{R}$ 是凸函数，则对 $\forall \alpha \in \mathbb{R}$，$S_\alpha$ 是凸集.

上述定理的逆不真.

例 1.5 考虑分段函数 $f(x) = \begin{cases} 1, & x \geq 0, \\ 0, & x < 0 \end{cases}$，函数非凸，但其任意水平集是凸集.

1.5.3　凸规划

定义 1.24 （凸规划）

（1）若 $(f\ S)$ 中的 S 为凸集，f 是 S 上的凸函数，优化目标为 min 时，则称 $(f\ S)$ 为凸规划；

（2）若 (fgh) 中的 f, g_i 为凸函数，h_j 为线性函数，则称 (fgh) 为凸规划.

定理 1.16 （凸优化的最优解）

设 $(f\ S)$ 为凸优化. 若 $\boldsymbol{x}^*$ 为问题 $(f\ S)$ 的 l.opt.，则 $\boldsymbol{x}^*$ 为 g.opt. 进一步地，若 f 是严格凸函数，则 $\boldsymbol{x}^*$ 是 $(f\ S)$ 的唯一 g.opt.

1.6 多面体、极点、极方向

定义 1.25 （半空间）

设超平面 $H: \boldsymbol{\alpha}^\top \boldsymbol{x} = \beta \subseteq \mathbb{R}^n$，则这个超平面把 $\mathbb{R}^n$ 分成两部分 $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{\alpha}^\top \boldsymbol{x} \geq \beta\}$ 和 $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{\alpha}^\top \boldsymbol{x} \leq \beta\}$，这样的两部分称为**闭半空间**，若有严格不等式成立，则称其为**开半空间**.

定义 1.26 （多面体）

有限个闭半空间的交称为**多面体**.

由定义看出，多面体实际上由一系列线性不等式组定义，故一般可以将其写作 $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}\}$，其中 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$，$\boldsymbol{b} \in \mathbb{R}^m$. 其中我们选取一种特殊的多面体 $\{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$，称其为**标准型多面体**.

定义 1.27 （极点）

设 $S \subseteq \mathbb{R}^n$，$\boldsymbol{x} \in S$，若不存在 $\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in S$，$\boldsymbol{x}^{(1)} \neq \boldsymbol{x}^{(2)}$，使得 $\boldsymbol{x} = \lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}$，则称 x 是 S 的**极点**（或**顶点**）.

定义 1.28 （极方向）

设 $S \subseteq \mathbb{R}^n$，$\boldsymbol{d} \in \mathbb{R}^n$，若 $\boldsymbol{d}$ 是 S 的可行方向（见**方向导数**的定义），且不能被表示为 S 的两个不同可行方向的非负组合，则称 $\boldsymbol{d}$ 是 S 的**极方向**.

定理 1.17 （极点特征）

设多面体 $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$，其中 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$ 行满秩，则 $\boldsymbol{x}$ 是多面体 S 的极点的充要条件是存在分解 $\boldsymbol{A} = (\boldsymbol{B}, \boldsymbol{N})$，使得

(i) B 为 m 阶可逆阵；

(ii) $\boldsymbol{x} = (\boldsymbol{x}_B^\top, \boldsymbol{x}_N^\top)^\top$（分解方式与 $\boldsymbol{A}$ 相同），其中 $\boldsymbol{x}_B = \boldsymbol{B}^{-1}\boldsymbol{b} \geq 0$，$\boldsymbol{x}_N = 0$.

例 1.6 证明：多面体 $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$ 中必存在有限多的极点.

例 1.7 设多面体 $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0\}$，在以下条件下，求该多面体的极点：

(1) $A = \begin{pmatrix} 3 & 1 & 2 & 0 \\ 0 & 4 & 5 & 1 \end{pmatrix}, b = \begin{pmatrix} 3 \\ 6 \end{pmatrix}$;

(2) $A = \begin{pmatrix} 3 & 2 & 1 & 0 & 0 \\ 2 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 1 \end{pmatrix}, b = \begin{pmatrix} 65 \\ 40 \\ 75 \end{pmatrix}$.

定理 1.18 （极方向特征）

设多面体 $S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq 0\}$，其中 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 行满秩，则

（1）$\boldsymbol{d}$ 是 S 的方向的充要条件是 $\boldsymbol{A}\boldsymbol{d}=0$ 且 $\boldsymbol{d}\geq 0$;

（2）$\boldsymbol{d}$ 是 S 的极方向的充要条件是存在分解 $\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N})$，使得

（i）$\boldsymbol{B}$ 为 m 阶可逆阵;

（ii）$\boldsymbol{d}=(\boldsymbol{d}_B^\top,\boldsymbol{d}_N^\top)^\top$（分解方式与 $\boldsymbol{A}$ 相同），其中 $\boldsymbol{d}_B=-\boldsymbol{B}^{-1}\boldsymbol{a}_j\geq 0$，$\boldsymbol{d}_N=\boldsymbol{e}_j$. 这里 $\boldsymbol{e}_j$ 是第 j 个元素为 1，其余元素为 0 的单位向量. j 是使得 $\boldsymbol{N}$ 中的列向量 $\boldsymbol{a}_j$ 满足 $\boldsymbol{B}^{-1}\boldsymbol{a}_j\leq 0$ 的下标.

例 1.8 证明：多面体 $S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq 0\}$ 中必有极点，且个数不超过 $(n-m)C_n^m$.

例 1.9 设多面体 $S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq 0\}$，其中 $A=\begin{pmatrix}3&1&2&-6&0\\0&4&5&-5&1\end{pmatrix}$, $b=\begin{pmatrix}3\\6\end{pmatrix}$，求该多面体的极方向.

定理 1.19 （表示定理，Minkowski-Weyl）

设多面体 $S=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq 0\}$，其中 $\boldsymbol{A}\in\mathbb{R}^{m\times n}$ 行满秩，$\boldsymbol{x}^{(1)},\boldsymbol{x}^{(2)},\cdots,\boldsymbol{x}^{(k)}$ 为所有极点，$\boldsymbol{d}^{(1)},\boldsymbol{d}^{(2)},\cdots,\boldsymbol{d}^{(l)}$ 为所有极方向，则 $\forall\boldsymbol{x}\in S$，存在序列 $\{\lambda_i\}_{i=1}^k,\{\mu_j\}_{j=1}^l$，其中：

（1）$\lambda_i>0$，$i=1,2,\cdots,k$，且 $\sum\limits_{i=1}^k\lambda_i=1$;

（2）$\mu_j\geq 0$，$j=1,2,\cdots,l$，使得

$$\begin{aligned}\boldsymbol{x}=&\lambda_1\boldsymbol{x}^{(1)}+\lambda_2\boldsymbol{x}^{(2)}+\cdots+\lambda_k\boldsymbol{x}^{(k)} && \text{（极点的凸组合）}\\ &+\mu_1\boldsymbol{d}^{(1)}+\mu_2\boldsymbol{d}^{(2)}+\cdots+\mu_l\boldsymbol{d}^{(l)} && \text{（极方向的半正组合）}\end{aligned}$$

特别地，如果上述多面体 S 是有界的，则有 $\mu_j=0$.

第二章　线性规划

2.1　线性规划模型

2.1.1　线性规划的定义

我们先来介绍线性规划的一般形式.

定义 2.1 （线性规划的一般形式）

线性规划的一般形式如下:

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

其中，称 c_i 为**目标函数系数**或**价值系数**或**费用系数**；称 b_j 为**右端项**或**资源常数**；称 a_{ij} 为**约束系数**或**技术系数**.

我们为了方便解决问题，常把上述一般形式转化为标准形式:

定义 2.2 （线性规划的标准形式）

线性规划的标准形式如下:

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

其中右端项 $b_1, b_2, \cdots, b_m \geq 0$.

为了研究问题的方便，我们将其写成矩阵形式

$$\begin{aligned} \max \quad & z = \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned}$$

其中，$\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{b} \geq 0$，$\boldsymbol{A} \in \mathbb{R}^{m \times n}$. 线性规划的标准形式记作 (LP).

定义 2.3 （线性规划的规范形式）

线性规划的规划形式如下：

$$\begin{aligned} \max \quad & z = \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b} \\ & \boldsymbol{x} \geq 0 \end{aligned}$$

其中，$\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$，$\boldsymbol{b} \geq 0$,，$\boldsymbol{A} \in \mathbb{R}^{m \times n}$.

线性规划的规范形式记作 (P).

不难发现，标准形式有区别于一般形式的以下四个特点：

- 目标最大化；
- 约束为等式；
- 决策变量均非负；
- 右端项非负.

接下来我们将会讨论如何将一般形式一步步转化为标准形式.

2.1.2 将一般形式转化为标准形式

1. 极小目标极大化

在标准形式中，要求优化目标为极大化目标函数值. 若目标函数为

$$\min f = c_1 x_1 + c_2 x_2 + \cdots + c_n x_n$$

则令 $z = -f$，从而上述极小化问题与下面的极大化问题

$$\max z = -c_1 x_1 - c_2 x_2 - \cdots - c_n x_n$$

有相同的最优解，但它们的最优解的目标函数值相差一个符号，即

$$\min f = -\max z$$

2. 约束条件不为等式

在标准形式中，要求每约束条件均为等式. 若约束条件为小于等于号，即设第 i 个约束条件为

$$a_{i1} x_1 + a_{i2} x_2 + \cdots + a_{in} x_n \leq b_i$$

则可以引入一个松弛变量 s，使得

$$s = b_i - (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n)$$

从而 $s \geq 0$，也具有非负约束，这时新的约束条件成为

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n + s = b_i$$

若约束条件为大于等于号，即设第 i 个约束条件为

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i$$

相似地，也可以引入一个松弛变量 s，使得

$$s = (a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n) - b_i$$

从而 $s \geq 0$，也具有非负约束，这时新的约束条件成为

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - s = b_i$$

3. 变量无符号限制

在标准形式中，要求每一个变量均有非负约束. 当某一个变量 x_j 没有非负约束时，可以令

$$x_j = x_j' - x_j''$$

其中 $x_j', x_j'' \geq 0$，即用两个非负数之差表示一个无符号限制的变量.

4. 右端项有负值

在标准形式中，要求每一个约束条件右端均为正值. 当某一个右端项系数为负时，假设是第 i 个，即

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \leq 0$$

则把该等式两端同时乘以 -1，得到

$$-a_{i1}x_1 - a_{i2}x_2 - \cdots - a_{in}x_n = -b_i \geq 0$$

例 2.1 将以下线性规划问题转化为标准形式.

（1）
$$\begin{aligned}
\min \quad & f = 3.6x_1 - 5.2x_2 + 1.8x_3 \\
\text{s.t.} \quad & 2.3x_1 + 5.2x_2 - 6.1x_3 \leq 15.7 \\
& 4.1x_1 + 3.3x_3 \geq 8.9 \\
& x_1 + x_2 + x_3 = 38 \\
& x_1, x_2, x_3 \geq 0
\end{aligned}$$

（2）
$$
\begin{aligned}
\min\quad & f = -3x_1 + 5x_2 + 8x_3 - 7x_4 \\
\text{s.t.}\quad & 2x_1 - 3x_2 + 5x_3 + 6x_4 \le 28 \\
& 4x_1 + 2x_2 + 3x_3 - 9x_4 \ge 39 \\
& 6x_2 + 2x_3 + 3x_4 \le -58 \\
& x_1, x_2, x_3 \ge 0
\end{aligned}
$$

2.2 线性规划的单纯形法

2.2.1 线性规划理论

线性规划的单纯形法依托以下一个定理:

定理 2.1 (解的存在性)

考虑 (LP) 和多面体 $S = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, x \geq 0\}$,设 $\boldsymbol{A}$ 满秩,$\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \cdots, \boldsymbol{x}^{(k)}$ 为所有极点,$\boldsymbol{d}^{(1)}, \boldsymbol{d}^{(2)}, \cdots, \boldsymbol{d}^{(l)}$ 为所有极方向,则

(1)(LP) 存在有界最优解 $\Leftrightarrow \boldsymbol{c}^\top \boldsymbol{d}^{(j)} \leq 0$,$\forall j = 1, 2, \cdots, l$;

(2)若 (LP) 存在有界最优解,则最优解可以在某个极点得到.

由此,我们可以得到,我们要求 (LP) 的最优解,只需要求多面体有限个极点的函数值即可. 看上去问题似乎是解决了,但是极点的个数大致随着多面体的维数成指数级增加(参考例题 1.8)! 此外,求取极点的过程伴随着矩阵的求逆,而这是非常昂贵的. 所以我们需要一个有选择的迭代算法代替枚举算法.

在进行相关讨论之前,由于原 ppt 和书中的概念过于杂乱,笔者认为有必要将线性规划相关的基础概念总结如下:

定义 2.4 (线性规划相关概念)

(1)**基**:对于线性规划的约束条件 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq 0$,设 $\boldsymbol{B}$ 是矩阵 $\boldsymbol{A}$ 中的一个可逆的 $m \times m$ 的子矩阵,则称 $\boldsymbol{B}$ 为线性规划的一个**基矩阵**,简称**基**,称 $\boldsymbol{N}$ 为基矩阵 $\boldsymbol{B}$ 对应的**非基矩阵**;

(2)**基变量、非基变量**:对应于基 $\boldsymbol{B}$ 中列向量 $\boldsymbol{a}_{j_1}, \boldsymbol{a}_{j_2}, \cdots, \boldsymbol{a}_{j_m}$ 的变量 $x_{j_1}, x_{j_2}, \cdots, x_{j_m}$ 称为**基变量**,其他变量(即对应于非基矩阵 $\boldsymbol{N}$ 中列向量的变量)称为**非基变量**;

(3)**基本解、基本可行解、可行基**:令所有非基变量为 0,可以得到 m 个关于基变量 $x_{j_1}, x_{j_2}, \cdots, x_{j_m}$ 的线性方程组,解这个线形方程组得到基变量的值,称如上得到的解为一个**基本解**;若得到的基变量的值均非负,则称其为**基本可行解**,称这个基 $\boldsymbol{B}$ 为**可行基**.

写成矩阵形式,

$$\boldsymbol{x} = \begin{pmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{pmatrix} = \begin{pmatrix} \boldsymbol{B}^{-1}\boldsymbol{b} \\ 0 \end{pmatrix}$$

称为线性规划问题的一个**基本可行解**.

显然必有以下定理成立:

定理 2.2 (线性规划基本定理)

线性规划的基本可行解就是可行域的极点.

下面的定理表明了我们不需要将所有的极点都代入目标函数,也可以精确地求出最优解

opt..

定理 2.3 （最优性定理）

考虑 (LP)，条件同上，设 $\boldsymbol{x}^*$ 为极点，存在基分解 $\boldsymbol{A} = (\boldsymbol{B}, \boldsymbol{N})$，其中 $\boldsymbol{B}$ 为 m 阶可逆矩阵，使 $\boldsymbol{x}^* = (\boldsymbol{x}_B^{*\top}, \boldsymbol{x}_N^{*\top})^\top$ ，此处分解与 $\boldsymbol{A}$ 相同，且 $\boldsymbol{x}_B^* = \boldsymbol{B}^{-1}\boldsymbol{b} \geq 0$ ，$\boldsymbol{x}_N^* = 0$. 相应 $\boldsymbol{c}^\top = (\boldsymbol{c}_B^\top, \boldsymbol{c}_N^\top)^\top$ ，则

（1）若 $\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{N} \leq 0$，则 $\boldsymbol{x}^*$ 为 opt.;

（2）若存在属于 $\boldsymbol{N}$ 的分量 j ，使得 $\boldsymbol{c}_j - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{a}_j > 0$ ，且 $\boldsymbol{B}^{-1}\boldsymbol{a}_j \leq 0$ ，则 (LP) 无有界解.

2.2.2 表格单纯形法原理及算法原理

让我们回忆线性规划的标准形式如下

$$
\text{(LP)} \quad
\begin{aligned}
\max \quad & z = \boldsymbol{c}^\top \boldsymbol{x} \\
\text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\
& \boldsymbol{x} \geq 0
\end{aligned}
$$

其中，$\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n$ ，$\boldsymbol{b} \in \mathbb{R}^m$ ，$\boldsymbol{A}$ 为 $m \times n$ 行满秩矩阵，即 $\operatorname{rank}(\boldsymbol{A}) = m$.

1. 单纯形法原理及算法过程

下边的算法介绍了单纯形法求线性规划最优解的过程.

算法 1: 单纯形法迭代算法 (Simplex Method)

输入　：系数矩阵 $\boldsymbol{A} \in \mathbb{R}^{m\times n}$，向量 $\boldsymbol{b} \in \mathbb{R}^m$，价值向量 $\boldsymbol{c} \in \mathbb{R}^n$

输出　：最优解 $\boldsymbol{x}^*$ 或问题无界结论

初始化: 已知极点 $\boldsymbol{x}^{(k)}$，基分解 $\boldsymbol{A} = (\boldsymbol{B}, \boldsymbol{N})$，满足 $\boldsymbol{x}_B^{(k)} = \boldsymbol{B}^{-1}\boldsymbol{b} \geq 0$，$\boldsymbol{x}_N^{(k)} = 0$

计算检验向量：$\boldsymbol{\sigma}_N^\top = \boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{N}$;

if $\boldsymbol{\sigma}_N \leq 0$ **then**

　　return $\boldsymbol{x}^{(k)}$ 为 opt.;

选择进基变量索引 j 使得 $\sigma_j > 0$;

if $\boldsymbol{B}^{-1}\boldsymbol{a}_j \leq 0$ **then**

　　return 问题无有界解;

else

　　计算步长 $\alpha = \min\left\{ \dfrac{(\boldsymbol{B}^{-1}\boldsymbol{b})_i}{(\boldsymbol{B}^{-1}\boldsymbol{a}_j)_i} \mid (\boldsymbol{B}^{-1}\boldsymbol{a}_j)_i > 0 \right\} = \dfrac{(\boldsymbol{B}^{-1}\boldsymbol{b})_r}{(\boldsymbol{B}^{-1}\boldsymbol{a}_j)_r}$;

　　更新解向量：$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha \boldsymbol{d}$;

　　　其中 $\boldsymbol{d}^\top = (\boldsymbol{d}_B^\top, \boldsymbol{d}_N^\top)$，且 $\boldsymbol{d}_B = -\boldsymbol{B}^{-1}\boldsymbol{a}_j$，$\boldsymbol{d}_N = \boldsymbol{e}_j$;

　　更新基矩阵 $\boldsymbol{B}$ 为 $\boldsymbol{B}'$，令 $k \leftarrow k+1$ 并返回步骤 1;

为了说明这个算法的合理性，我们需要说明以下两点成立：

定理 2.4

上述算法中，更新点 $\boldsymbol{x}^{(k+1)}$

（1）是更优的；

（2）也是 S 的一个极点.

2. 单纯形表

设 $\boldsymbol{x}$ 为初始极点，相应分解 $\boldsymbol{A}=(\boldsymbol{B},\boldsymbol{N})$，则列单纯形表如下：

	f	$\boldsymbol{x}_B^\top$	$\boldsymbol{x}_N^\top$	RHS	
目标行	1	$\boldsymbol{c}_B^\top$	$\boldsymbol{c}_N^\top$	0	1 行
约束行	0	$\boldsymbol{B}$	$\boldsymbol{N}$	$\boldsymbol{b}$	m 行
	1 列	m 行	$n-m$ 行	1 列	

作变换，使前 $m+1$ 列对应的 $m+1$ 阶矩阵变为单位矩阵，相当于该表左乘 $\begin{pmatrix}1 & \boldsymbol{c}_B^\top\\ 0 & \boldsymbol{B}\end{pmatrix}^{-1}=\begin{pmatrix}1 & -\boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\\ 0 & \boldsymbol{B}^{-1}\end{pmatrix}$. 消去方法为 Gauss 主元消去法. 于是得到

	f	$\boldsymbol{x}_B^\top$	$\boldsymbol{x}_N^\top$	RHS	
目标行	1	$0^\top$	$\boldsymbol{c}_N^\top-\boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\boldsymbol{N}$	$-\boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\boldsymbol{b}$	1 行
约束行	0	$\boldsymbol{I}$	$\boldsymbol{B}^{-1}\boldsymbol{N}$	$\boldsymbol{B}^{-1}\boldsymbol{b}$	m 行
	1 列	m 行	$n-m$ 行	1 列	

根据算法内容，令**检验向量** $\boldsymbol{\sigma}_N^\top=\boldsymbol{c}_N^\top-\boldsymbol{c}_B^\top\boldsymbol{B}^{-1}\boldsymbol{N}$，则单纯形表转化为

	f	$\boldsymbol{x}_B^\top$	$\boldsymbol{x}_N^\top$	RHS	
目标行	1	$0^\top$	$\boldsymbol{\sigma}^\top$	$-z=-\boldsymbol{c}^\top\boldsymbol{x}^*$	1 行
约束行	0	$\boldsymbol{I}$	$\boldsymbol{B}^{-1}\boldsymbol{N}$	$\boldsymbol{x}_B^*$	m 行
	1 列	m 行	$n-m$ 行	1 列	

可是我们知道，对矩阵求逆的代价是十分昂贵的（对一个 n 阶矩阵求逆的时间复杂度为 $O(n^3)$）. 所以如果我们能使 $\boldsymbol{B}=\boldsymbol{I}$，则 $\boldsymbol{B}^{-1}=\boldsymbol{I}$，此时求矩阵 $\boldsymbol{B}$ 的逆将会很简单，此时

	f	$\boldsymbol{x}_B^\top$	$\boldsymbol{x}_N^\top$	RHS	
目标行	1	$0^\top$	$\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{N}$	$-\boldsymbol{c}_B^\top \boldsymbol{b}$	1 行
约束行	0	$\boldsymbol{I}$	$\boldsymbol{N}$	$\boldsymbol{b}$	m 行
	1 列	m 行	$n-m$ 行	1 列	

综上所述，我们将单次迭代所需要用到的所有元素集中在一张表中，就得到了如下经过改造的单纯形表：（其中，表??对矩阵 $\boldsymbol{B}$ 不作除可逆外的任何要求，表??中令 $\boldsymbol{B}=\boldsymbol{I}$ ）

表 2.1: 经过改造后的单纯形表

<table>
<tr><td rowspan="2">$\boldsymbol{c}_B$</td><td rowspan="2">$\boldsymbol{x}_B$</td><td rowspan="2">$\boldsymbol{b}$</td><td>$\boldsymbol{c}_B^\top$</td><td>$\boldsymbol{c}_N^\top$</td><td rowspan="2">$\boldsymbol{\alpha}$</td></tr>
<tr><td>$\boldsymbol{x}_B^\top$</td><td>$\boldsymbol{x}_N^\top$</td></tr>
<tr><td>*</td><td>*</td><td>$\boldsymbol{B}^{-1}\boldsymbol{b}$</td><td>$\boldsymbol{I}$</td><td>$\boldsymbol{B}^{-1}\boldsymbol{N}$</td><td>*</td></tr>
<tr><td colspan="2">$-z$</td><td>$-\boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{b}$</td><td>$0^\top$</td><td>$\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{N}$</td><td></td></tr>
</table>

表 2.2: 经过改造后的单纯形表（令 $\boldsymbol{B}=\boldsymbol{I}$）

<table>
<tr><td rowspan="2">$\boldsymbol{c}_B$</td><td rowspan="2">$\boldsymbol{x}_B$</td><td rowspan="2">$\boldsymbol{b}$</td><td>$\boldsymbol{c}_B^\top$</td><td>$\boldsymbol{c}_N^\top$</td><td rowspan="2">$\boldsymbol{\alpha}$</td></tr>
<tr><td>$\boldsymbol{x}_B^\top$</td><td>$\boldsymbol{x}_N^\top$</td></tr>
<tr><td>*</td><td>*</td><td>$\boldsymbol{b}$</td><td>$\boldsymbol{I}$</td><td>$\boldsymbol{N}$</td><td>*</td></tr>
<tr><td colspan="2">$-z$</td><td>$-\boldsymbol{c}_B^\top \boldsymbol{b}$</td><td>$0^\top$</td><td>$\boldsymbol{c}_N^\top - \boldsymbol{c}_B^\top \boldsymbol{N}$</td><td></td></tr>
</table>

该表的填写规则如下：（合并的单元格按照合并之前最左上角的单元格编号）

- 表格第 3 行第 1-2 列（即 $\boldsymbol{c}_B$ 和 $\boldsymbol{x}_B$ 所示单元格下边的单元格），将列向量 $\boldsymbol{c}_B,\boldsymbol{x}_B$ 分别展开后写入；
- 表格第 3 行第 3 列（即 $\boldsymbol{B}^{-1}\boldsymbol{b}$ 或 $\boldsymbol{b}$ 所示单元格），将列向量 $\boldsymbol{B}^{-1}\boldsymbol{b}$ 或 $\boldsymbol{b}$ 展开后写入；
- 表格第 4 行第 3 列（即 $-\boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{b}$ 或 $-\boldsymbol{c}_B^\top \boldsymbol{b}$ 所示单元格），将标量 $-\boldsymbol{c}_B^\top \boldsymbol{B}^{-1}\boldsymbol{b}$ 或 $-\boldsymbol{c}_B^\top \boldsymbol{b}$ 计算出来后写入；
- 表格第 4-5 列，将所示行向量或矩阵展开后写入；
- 表格第 3 行第 6 列（即 $\boldsymbol{\alpha}$ 所示单元格下边的单元格），将列向量 $\boldsymbol{\alpha}$ 展开后写入. 其中，步长 $\alpha_i=\begin{cases}\dfrac{b_i}{a_{ik}}, & a_{ik}>0\\ \infty, & a_{ik}\le 0\end{cases}$ ，k 为进基变量索引.

3. 引入单位阵 I

这一小节我们分为两种情况来讨论：对**规范形式**求解线性规划问题和**一般情况下**的求解线性规划问题.

考虑线性规划的规范形式 (P)

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

加入松弛变量 $x_{n+1}, x_{n+2}, \cdots, x_{n+m}$ 后，转化为标准形式

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

我们可以得到，$x_j = 0$，$j = 1, 2, \cdots, n$；$x_{n+1} = b_i$，$i = 1, 2, \cdots, m$ 是 S 的一个基本可行解，对应的 $\boldsymbol{B}$ 是单位矩阵.

应用单纯形法解线性规划问题有如下注意事项：

- 每一步运算只能用矩阵初等行变换；
- 表中第 3 列的数总应保持非负；
- 当所有检验数均非正时，得到最优单纯形表；
- 检验数的更新也相当于是主元消去.

例 2.2 用单纯形法解如下线性规划问题：

$$
\begin{aligned}
(1)\ \max \quad & z = 1500x_1 + 2500x_2 \\
\text{s.t.} \quad & 3x_1 + 2x_2 + x_3 = 65 \\
& 2x_1 + x_2 + x_4 = 40 \\
& 3x_2 + x_5 = 75 \\
& x_1, x_2, \cdots, x_5 \ge 0
\end{aligned}
$$

（2）$$\begin{aligned} \max \quad & z = 1500x_1 + 1000x_2 \\ \text{s.t.} \quad & 3x_1 + 2x_2 + x_3 = 65 \\ & 2x_1 + x_2 + x_4 = 40 \\ & 3x_2 + x_5 = 75 \\ & x_1, x_2, \cdots, x_5 \geq 0 \end{aligned}$$

接下来我们讨论一般情况，即初始基本可行解不明显时的情况.

考虑一般问题

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

其中右端项 $b_1, b_2, \cdots, b_m \geq 0$. 接下来介绍两种方法.

大 M 法：

引入人工变量 $x_{n+i} \geq 0$ ，$i = 1, 2, \cdots, m$ 及充分大正数 M，得到

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

显然，$x_j = 0$，$j = 1, 2, \cdots, n$；$x_{n+1} = b_i$，$i = 1, 2, \cdots, m$ 是 S 的一个基本可行解，对应的 $\boldsymbol{B}$ 是单位矩阵. 若得到的最优解满足 $x_{n+i} = 0$，$i = 1, 2, \cdots, m$，则其为原问题的**最优解**；否则，原问题**无可行解**.

例 2.3 应用大 M 法解如下线性规划问题

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

两阶段法：

两阶段法的核心思想是：先求出原问题的一个基本可行解（第一阶段），再求解原问题（第二阶段）.

第一阶段中，引入人工变量 $x_{n+i} \geq 0$ ，$i = 1,2,\cdots,m$ ，构造：

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

之后求解上述问题. 显然，$x_j = 0$，$j = 1,2,\cdots,n$；$x_{n+1} = b_i$，$i = 1,2,\cdots,m$ 是 S 的一个基本可行解，对应的 $\boldsymbol{B}$ 是单位矩阵. 若得到的最优解满足 $x_{n+i} = 0$，$i = 1,2,\cdots,m$，则其为原问题的**基本可行解**；否则，原问题**无可行解**. 得到原问题的基本可行解之后，求解原问题.

例 2.4 应用两阶段法解如下线性规划问题

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

2.3　线性规划的对偶

2.3.1　对偶问题

我们用下面的例子引出本节的讨论.

设某工厂有 A、B、C 三种类型的设备，生产甲、乙两种设备. 每件产品在生产中需要占用的设备机时数、每件产品可以获得的利润以及三种设备可利用的机时数如下表所示：

	甲产品	乙产品	设备能力（h）
设备 A	3	2	65
设备 B	2	1	40
设备 C	0	3	75
利润（元/件）	1500	2500	/

试问：

（1）工厂应如何安排生产可获得最大的总利润?

（2）若工厂的设备都用于外协加工，工厂收取加工费. 试问：设备 A、B、C 每工时各如何收费才最有竞争力?

对于第（1）个问题，设工厂生产 x_1 个甲产品，x_2 个乙产品，则可列线性规划问题如下：

$$\begin{aligned} \max \quad & z = 1500x_1 + 2500x_2 \\ \text{s.t.} \quad & 3x_1 + 2x_2 \le 65 \\ & 2x_1 + x_2 \le 40 \\ & 3x_2 \le 75 \\ & x_1, x_2 \ge 0 \end{aligned}$$

对于第（2）个问题，设设备 A、B、C 每工时各收费 y_1, y_2, y_3 元，则可列线性规划问题如下：

$$\begin{aligned} \min \quad & f = 65y_1 + 40y_2 + 75y_3 \\ \text{s.t.} \quad & 3y_1 + 2y_2 \ge 1500 \\ & 2y_1 + y_2 + 3y_3 \ge 2500 \\ & y_1, y_2, y_3 \ge 0 \end{aligned}$$

定义 2.5（对偶规划）

称具有以下对称形式的两个线性规划问题互为对偶规划：

$$\begin{aligned}(\mathrm{LP})\max \quad & z = \boldsymbol{c}^\top \boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \\ & \boldsymbol{x} \ge 0\end{aligned} \qquad\qquad \begin{aligned}(\mathrm{DP})\min \quad & f = \boldsymbol{b}^\top \boldsymbol{y} \\ \text{s.t.} \quad & \boldsymbol{A}^\top \boldsymbol{y} \ge \boldsymbol{c} \\ & \boldsymbol{y} \ge 0\end{aligned}$$

一对对称形式的对偶规划之间具有下面的对应关系[①]：

- 若一个模型为目标求“极大”，约束为“小于等于”的不等式，则它的对偶模型为目标求“极小”，约束是“大于等于”的不等式. 即“$\max, \le$”和“$\min, \ge$”相对应；
- 从约束系数矩阵看：一个模型中为 $\boldsymbol{A}$，则另一个模型中为 $\boldsymbol{A}^\top$. 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 $\boldsymbol{b}, \boldsymbol{c}$ 的位置看：在两个规划模型中，$\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的位置对换；
- 两个规划模型中的变量皆非负.

对于不满足上述定义形式的线性规划问题，其对偶规划可按如下规则直接给出：

- 将模型统一为“$\max, \le$”或“$\min, \ge$”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约束为等式.

例 2.5 写出下面线性规划的对偶规划模型.

$$\begin{aligned}\max \quad & z = x_1 - x_2 + 5x_3 - 7x_4 \\ \text{s.t.} \quad & x_1 + 3x_2 - 2x_3 + x_4 = 25 \\ & 2x_1 + 7x_2 + 2x_4 \ge -60 \\ & 2x_1 + 2x_2 - 4x_3 \le 30 \\ & x_1, x_2 \ge 0, -5 \le x_4 \le 10\end{aligned}$$

2.3.2 对偶定理

定理 2.5 （弱对偶定理）

若 $\boldsymbol{x}, \boldsymbol{y}$ 分别为 (LP) 和 (DP) 的可行解，那么 $\boldsymbol{c}^\top \boldsymbol{x} \le \boldsymbol{b}^\top \boldsymbol{y}$

推论

设 (LP) 有可行解，那么若 (LP) 无有界最优解，则 (DP) 无可行解.

定理 2.6 （最优性准则）

若 $\boldsymbol{x}, \boldsymbol{y}$ 分别为 (LP) 和 (DP) 的可行解，且 $\boldsymbol{c}^\top \boldsymbol{x} = \boldsymbol{b}^\top \boldsymbol{y}$，那么 $\boldsymbol{x}, \boldsymbol{y}$ 分别为 (LP) 和 (DP) 的最优解.

[①] $\boldsymbol{c}$ 换 $\boldsymbol{b}$，$\boldsymbol{x}$ 换 $\boldsymbol{y}$，约束矩阵变转置.

定理 2.7 （主对偶定理）

若 (LP) 有最优解，则 (DP) 也有最优解. 反之也成立，且最优值相等.