

# PRACTICAL OPTIMIZATION ALGORITHMS

徐 翔

数学科学学院  
浙江大学

DEC 23, 2021

## 第十讲：线性规划 - 单纯形法 (LINEAR PROGRAMMING - THE SIMPLEX METHOD)

## 简介(INTRODUCTION)

## 简介

- 一般形式：

$$\min_x c^T x \quad (10.1)$$

$$s.t. \quad Ax = b, \quad (10.2)$$

$$x \geq 0. \quad (10.3)$$

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

- 例如  $\min c^T x$ , subject to  $Ax \leq b$  可以转化为

$$\min c^T x, \quad \text{subject to } Ax + z = b, z \geq 0.$$

- 把上面的  $x$  分裂成正部和负部:  $x = x^+ - x^-$ , 其中  $x^+ = \max(x, 0)$ ,  $x^- = \max(-x, 0)$ .

$$\min \begin{bmatrix} c \\ -c \\ 0 \end{bmatrix}^T \begin{bmatrix} x^+ \\ x^- \\ z \end{bmatrix}, \text{ subject to } [A \quad -A \quad I] \begin{bmatrix} x^+ \\ x^- \\ z \end{bmatrix} = b; \begin{bmatrix} x^+ \\ x^- \\ z \end{bmatrix} \geq 0$$

- 如果是  $Ax \geq b$ , 则可以使用  $Ax - y = b, y \geq 0$

# 最优性条件和对偶

- 构造拉格朗日函数

$$\mathcal{L}(x; \lambda, s) = c^T x - \lambda^T (Ax - b) - s^T x$$

- 由KKT条件可以得到

$$\begin{aligned} A^T \lambda + s &= c, \quad Ax = b, \\ x &\geq 0, \quad s \geq 0, \\ x_i s_i &= 0, \quad i = 1, 2, \dots, n. \end{aligned}$$

- 最后一个条件通常写成  $x^T s = 0$  (由于  $s_i > 0, x_i > 0$ ).
- 设  $(x^*; \lambda^*, s^*)$  代表解, 则

$$c^T x^* = (A^T \lambda^* + s^*)^T x^* = (Ax^*)^T \lambda^* = b^T \lambda^*$$

即所有满足KKT条件的  $(x; \lambda, s)$  使得主问题和对偶问题的目标函数值相等.

- 可以证明  $x^*$  是全局最优解. 设  $\bar{x}$  是可行点  $A\bar{x} = b, \bar{x} \geq 0$  则

$$c^T \bar{x} = (A\lambda^* + s^*)^T \bar{x} = b^T \lambda^* + \bar{x}^T s^* \geq b^T \lambda^* = c^T x^*$$

# 最优性条件和对偶

- 对偶问题为：

$$\max b^T \lambda, \text{ subject to } A^T \lambda \leq c.$$

- 引入“松弛”变量  $s$

$$\max b^T \lambda, \text{ subject to } A^T \lambda + s = c, s \geq 0$$

- 主问题-对偶问题之间的关系：

$$\min -b^T \lambda, \text{ subject to } c - A^T \lambda \geq 0.$$

- 记  $x$  是上述问题的拉格朗日乘子，

$$\bar{\mathcal{L}}(\lambda; x) = -b^T \lambda - x^T (c - A^T \lambda)$$

- KKT条件为：

$$\nabla_{\lambda} \bar{\mathcal{L}}(\lambda; x) = Ax - b = 0, A\lambda \leq c,$$

$$x \geq 0, \quad x_i(c - A\lambda)_i = 0.$$

记  $s = c - A\lambda$ ，上述KKT条件与主问题KKT条件完全一致。

该KKT条件是对偶问题的充分条件

假设  $x^*, \lambda^*$  满足KKT条件， $\bar{\lambda}$  满足对偶问题约束条件，则

$$\begin{aligned} b^T \bar{\lambda} &= (Ax^*)^T \bar{\lambda} \\ &= (x^*)^T (A\bar{\lambda} - c) + c^T x^* \\ &\leq c^T x^* = b^T \lambda^* \end{aligned}$$

因此， $\lambda^*$  是对偶问题的解。

# 最优性条件和对偶

## 定理：强对偶

以下两种情况只有一种出现：

- ① 如果主问题(或对偶问题)有一个有限解, 那么对偶问题(或主问题)也有一个有限解, 且两个问题的目标函数值相等。
- ② 如果主问题(或对偶问题)的目标函数值无界, 则对偶问题(或主问题)的可行域是空集。

## Proof.

- ① 略
- ② 如果主问题无界, 则存在一系列  $x_k$ , 满足  $c^T x_k \rightarrow -\infty$ ,  $Ax_k = b$ ,  $x_k \geq 0$ .  
如果对偶问题是可行的, 即至少存在一个  $\bar{\lambda}$ , 满足  $A\bar{\lambda} \leq c$ .  
再根据  $x_k \geq 0$ , 可以得到  $\bar{\lambda}^T Ax_k \leq c^T x_k$ .  
可以推出  $\bar{\lambda}^T b = \bar{\lambda}^T Ax_k \leq c^T x_k \rightarrow -\infty$ . 这产生了矛盾.



# 可行域的几何性质

假设 $A$ 是行满秩的，并且 $A \in R^{m \times n}$ .

定义：基本可行点(Basic feasible point)

$x$ 是基本可行点，当且仅当

- 存在某个指标子集 $\mathcal{B} \subset \{1, \dots, n\}$ ,  $|\mathcal{B}| = m$ .
- 如果 $i \notin \mathcal{B}$ ,  $x_i = 0$ . (即 $i \in \mathcal{B}$ ,  $x_i$ 不活跃).
- 如果 $B$ 是非奇异的.

这里 $B$ 是由 $\mathcal{B}$ 构成的 $m \times m$ 矩阵,  $B = [A_i]_{i \in \mathcal{B}}$ , 其中 $A_i$ 是 $A$ 的第 $i$ 列.

这里 $B$ 通常被称为**基矩阵(basis matrix)**,  $\mathcal{B}$ 被称为**基(basis)**

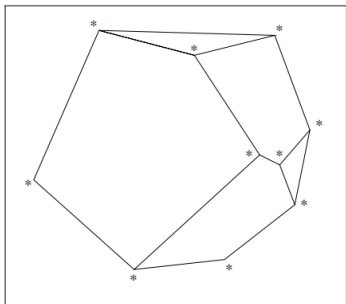
单纯形方法的基本策略：只需要检查**基本可行点**就可以收敛到最优解.



# 可行域的几何性质

定理:

- ① 如果主问题可行域非空, 则至少存在一个基本可行点.
- ② 如果主问题存在解, 则至少存在一个解是基本可行点.
- ③ 如果主问题是可行并且有界, 则至少存在一个最优解.



- 对于线性约束条件, 可行域是多面体
- 多面体的顶点就是基本可行点.(几何形式与代数表达式的对应关系)

定理

所有的基本可行点都是可行域

$\mathcal{F} = \{x | Ax = b, x \geq 0\}$  的顶点, 反之亦然.

## 可行域的几何性质

## Proof.

- " $\rightarrow$ ". 设 $x$ 是基本可行点, 即存在 $\mathcal{B}$ , s.t.  $B = [A_i]_{i \in \mathcal{B}}$  非奇异. 不妨设 $\mathcal{B} = 1, 2, \dots, m$ , 则 $x_{m+1}, \dots, x_n = 0$ . 记  $x_B = (x_1, \dots, x_m)$ . 假设 $x$ 不是可行域的顶点, 即 $x$ 可以由另外两个可行点线性组合, 即存在 $y, z \in \mathcal{F}$ 且 $y \neq x, z \neq x$ , s.t.,  $x = \alpha y + (1 - \alpha)z, \alpha > 0$ . 可以定义 $y_B = (y_1, \dots, y_m), z_B = (z_1, \dots, z_m)$ . 显然由于 $Ax = Ay = Az = b$ 能推出 $Bx_B = By_B = Bz_B = b$ . 由于 $B$ 可逆,  $x_B = y_B = z_B$ , 进而可以得到  $x = y = z$ . 产生矛盾.
- " $\leftarrow$ ". 设 $x$ 是顶点, 其中的非零分量是 $x_1, \dots, x_p$ . 假设对应的列向量 $A_1, \dots, A_p$ 是线性相关的, 即 $A_p = \sum_{j=1}^{p-1} z_j A_j$ , 可以构造一个扰动的向量  $x(\varepsilon) = x + \varepsilon z = x + \varepsilon(z_1, z_2, \dots, z_{p-1}, -1, 0, \dots, 0)$ , 当 $\varepsilon$ 很小时, 可以得出 $Ax(\varepsilon) = b, x \geq 0$ , i.e.,  $x(\varepsilon) \in \mathcal{F}$ . 这样我们可以取某个 $\hat{\varepsilon}$ 使得 $x(\hat{\varepsilon})$ 和 $x(-\hat{\varepsilon})$ 都是可行的. 显然 $x(0)$ 是落在这两点的连线上的, 即 $x$ 不是顶点. 所以, 如果 $x$ 是顶点,  $A_1, \dots, A_p$ 一定线性无关. 如果 $p = m$ , 那么就已经证明了结果. 如果 $p < m$ , 由于 $A$ 是行满秩的, 我们可以继续从剩下的 $n - p$ 个中挑选 $m - p$ 个 $A_i$ 加入到 $A_1, \dots, A_p$ , 组成 $\mathcal{B}$ , 证毕.

# 单纯形方法简介

- 单纯形方法是迭代法，从一个顶点到另一个顶点.
- 绝大多数迭代步，目标函数值都在降低（除非是无界问题）
- 迭代步中，最主要是确定如何更新 $B$ ，每一步迭代都需要加入一个新的指标 $q$ ，去除一个指标 $p$ .
- 可以从KKT条件中得到一些启发.

# 单纯形方法简介

KKT条件:

$$A^T \lambda + s = c, Ax = b,$$

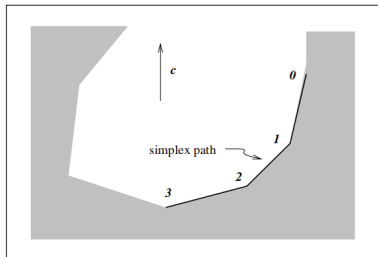
$$x \geq 0, s \geq 0,$$

$$x_i s_i = 0, i = 1, 2, \dots, n.$$

- 定义非基本指标集合  $N = \{1, \dots, n\} \setminus \mathcal{B}$
- 定义相应的  $x_B = [x_i]_{i \in \mathcal{B}}, x_N = [x_i]_{i \in N}, s_B, s_N, \lambda_B, \lambda_N, c_B, c_N$ .
- 根据  $Ax = Bx_B + Nx_N = b$ , 而  $x_N = 0$ , 所以  $x_B = B^{-1}b \geq 0$
- 根据互补性条件  $s_B = 0$ , 再根据  $[B \ N]^T \lambda + [s_B; s_N] = [c_B; c_N]$ , 得到  $\lambda = B^{-T} c_B, N^T \lambda + s_N = c_N$ . 即可以得到  $s_N = c_N - (B^{-1}N)^T c_B$
- 到目前为止, KKT条件中仅有  $s \geq 0$  没有强制满足. 如果上式计算中  $s_q < 0$ , 说明可以把相应的  $x_q$  从0变为正的并保持  $x$  仍是可行的, 对应的目标函数值  $c^T x$  可以降低.
- 假设更新  $x_q$  后的  $x$  记为  $x^+$ , 由于在  $N \setminus \{q\}$  中的分量都没发生变化, 这些  $x_i^+ = 0$ , 那么  $\mathcal{B}$  中  $x_B^+$  应该满足  $Ax^+ = Bx_B^+ + A_q x_q^+ = b = Bx_B$ , 即  $x_B^+ = x_B - B^{-1}A_q x_q^+$
- 我们来计算新的目标函数值  $c^T x^+ = c_B^T x_B^+ + c_q x_q^+ = c_B^T x_B - c_B^T B^{-1}A_q x_q^+ + c_q x_q^+, -c_B^T B^{-1}A_q x_q^+ = (c_q - s_q)x_q^+$
- 最终得到  $c^T x^+ = c_B^T x_B - (c_q - s_q)x_q^+ + c_q x_q^+ = c^T x + s_q x_q^+.$

# 单纯形法简介

- 原则上可以一直增大 $x_q$ 直到碰到下一个顶点, 即某个 $x_p^+ = 0$ ,  $p \in \mathcal{B}$ . 将 $p$ 移入 $\mathcal{N}$ 中.
- 或者如果可以一直增大 $x_q$ 到无穷大而不碰到下一个顶点, 那说明原问题无界.



## 定理

对于非退化的有界线性规划问题, 使用单纯形方法可以在有限步之内终止.

# 单纯形方法

## 单步单纯形方法

给定  $\mathcal{B}$ ,  $\mathcal{N}$ ,  $x_B = B^{-1}b \geq 0$ ,  $x_N = 0$ ;

求解  $B^T \lambda = c_B$ ,

计算  $s_N = c_N - N^T \lambda$ ;

**if**  $s_N \geq 0$

**stop**; (找到了最优点)

**else**

选择  $q \in \mathcal{N}$  with  $s_q < 0$ ,

计算  $Bd = A_q$ ;

**if**  $d \leq 0$

**stop**; (问题无界)

**else**

计算  $x_q^+ = \min_{i|d_i > 0} \frac{(x_B)_i}{d_i}$ , 记录最小的指标为  $p$ ;

更新  $x_B^+ = x_B - dx_q^+$ ,  $x_N^+ = (0, \dots, 0, x_q^+, 0, \dots, 0)^T$ ;

把  $q$  加入  $\mathcal{B}$ ,  $p$  移除  $\mathcal{B}$ .

**end(if)**

**end(if)**

## 一个例子

考虑如下问题

$$\min -4x_1 - 2x_2$$

$$s.t. \quad x_1 + x_2 + x_3 = 5;$$

$$2x_1 + \frac{1}{2}x_2 + x_4 = 8,$$

$$x \geq 0.$$

- 选取初始的  $\mathcal{B} = \{3, 4\}$ , 可以得到  $B = I$ ,

$$x_B = \begin{bmatrix} 5 \\ 8 \end{bmatrix}, \lambda = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, s_N = \begin{bmatrix} -8 \\ -2 \end{bmatrix}$$

- 计算目标函数值  $c^T x = 0$ . 选取  $q = 1$ ,  $A_q = [1, 2]^T$ , 计算  $Bd = A_q$  得到  $d = [1, 2]^T$ .
- 计算  $x_4^+ = \min_{i|d_i > 0} \frac{(x_B)_i}{d_i} = 4$ , 对应的指标是4. 即可以更新  $\mathcal{B} = \{3, 1\}$ ,  $\mathcal{N} = \{4, 2\}$

- 经计算得到

$$x_B = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \lambda = \begin{bmatrix} 0 \\ -4 \end{bmatrix}, s_N = \begin{bmatrix} \frac{3}{2} \\ -\frac{5}{4} \end{bmatrix}$$

- 计算目标函数值  $c^T x = -12$ . 选取  $q = 2$ ,  $A_2 = [1, \frac{1}{2}]^T$ , 计算  $Bd = A_q$  得到  $d = [\frac{3}{2}, -\frac{1}{2}]^T$ .
- 计算  $x_2^+ = \min_{i|d_i > 0} \frac{(x_B)_i}{d_i} = \frac{4}{3}$ , 对应的指标是2. 即可以更新  $\mathcal{B} = \{2, 1\}$ ,  $\mathcal{N} = \{4, 3\}$

- 第三次迭代计算得到

$$x_B = (x_2, x_1)^T = \left(\frac{4}{3}, \frac{11}{3}\right)^T, \\ \lambda = \left(-\frac{5}{3}, -\frac{2}{3}\right)^T, \\ s_N = \left(\frac{7}{3}, \frac{5}{3}\right)^T.$$

- 此时目标函数值  $c^T x = -\frac{41}{3}$ .