第 2 章 线性规划（&单纯形法）

可行解域的顶点的代数刻画：基可行解 bfs！

第 2.4 节：LP 基可行解（basic feasible solution：bfs）

引例： $\max\ z=2x_1+3x_2$

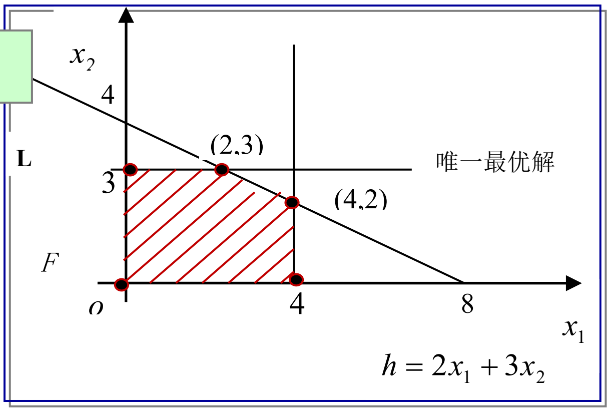
$$s.t.\begin{cases}x_1+2x_2\le 8\\ 4x_1\le 16\\ 4x_2\le 12\\ x_1,x_2\ge 0\end{cases}$$

<u>LP 可行解域及最优解的几何特征：</u>

(1) 可行解域 F 是多面体且一定是凸集；

(2) 若最优解存在，必可在多面体的**顶点**取到。

顶点对应的解是什么？



1）其标准形式(*)：

$$\max Z=2x_1+3x_2$$

$$s.t.\begin{cases}x_1+2x_2+x_3 \qquad\qquad =8\\ 4x_1 \qquad\qquad +x_4 \quad =16\\ \qquad 4x_2 \qquad\qquad +x_5=12\\ x_1,x_2,x_3,x_4,x_5\ge 0\end{cases}$$

→ 令 $x_1=x_2=0$ 得 $x_3=8,x_4=16,x_5=12$ → 点 O

2）

→ 令 $x_2=x_4=0$ 得 $x_1=4,x_3=4,x_5=12$ → 点(4,0)

将(*) 约束恒等变换得形式(**)：
$$s.t.\begin{cases}+2x_2+x_3-\frac{1}{4}x_4 \quad =4\\ x_1 \qquad\qquad +\frac{1}{4}x_4 \quad =4\\ 4x_2 \qquad\qquad\qquad +x_5=12\end{cases}$$

3）

→ 令 $x_4=x_5=0$ 得 $x_1=4,x_2=3,x_3=-2$ → 点(4,3)

将(*) 约束恒等变换得形式(***)：
$$s.t.\begin{cases}x_3-\frac{1}{4}x_4-\frac{1}{2}x_5=-2\\ x_1 \qquad +\frac{1}{4}x_4 \qquad =4\\ x_2 \qquad\qquad +\frac{1}{4}x_5=3\end{cases}$$

…… ……我们发现：这种方程组的解对应顶点！

一、LP 基本可行解（basic feasible solution—bfs） —— LP 可行解域顶点的代数刻画

1. 基可行解 (bfs) 的定义

给定标准形式(LP)：
$$\max c'x \quad s.t.\begin{cases} Ax=b, \\ x\geq 0 \end{cases}$$
A 为 $m\times n$ 矩阵且假定：$\operatorname{rank}(A)=m$。

在矩阵 A 中取 m 个线性无关列 $P_{j_1},P_{j_2},\cdots,P_{j_m}$，记它们构成的矩阵为：
$$B=\left(P_{j_1},P_{j_2},\cdots,P_{j_m}\right),$$

称：
- $P_{j_1},P_{j_2},\cdots,P_{j_m}$ 为(LP)的一个**基(basis)**；
- B 为相应的**基阵**；构成 B 的列为相应的**基列**或**基向量**；
- A 的其他列称为相应的**非基列**或**非基向量**（并将所有非基列构成的矩阵记为 $\boldsymbol{N}$）。

❑ 将基列和非基列分别排列、系数矩阵可写成 $A=(B,N)$，相应地，变量写成 $x=\begin{pmatrix}x_B\\x_N\end{pmatrix}$，

则(LP)可等价化为：
$$Ax=b \xleftrightarrow{\text{化为}} (B,N)\begin{pmatrix}x_B\\x_N\end{pmatrix}=b \xleftrightarrow{\text{矩阵分块乘法化为}} Bx_B+Nx_N=b \qquad \textbf{(1.1)}$$

❑ 在**(1.1)**式中令非基变量 $x_N=0$，得到唯一解 $x_B=B^{-1}b$.

称
- (1) $\begin{pmatrix}x_B\\x_N\end{pmatrix}=\begin{pmatrix}B^{-1}b\\0\end{pmatrix}$ 为关于基 B 的**基解**或**基本解** (Basic Solution)；
- (2) 若 $\boxed{B^{-1}b\geq 0}$，称上述解是关于基 B 的**基可行解**(Basic Feasible Solution, **bfs**)

注1° bfs 的个数有限，不超过 C_n^m。

注2° 若 x 是 bfs，则其 0 分量的个数不少于 $(n-m)$，或者说其非 0 分量的个数不超过 m。

注3° **结论 1.** x^* 是 F 的一个顶点当且仅当 x^* 是(LP)的一个 bfs。

结论 2. 线性规划(LP)最优解必能在某个 bfs 上取到（若存在）。
几何上，若(LP)存在最优解，则必在可行解域的顶点上达到。

（结论 1、2 的证明将在后面讲解）

LP 单纯形法的基本思路：

这种算法称为：迭代法

从某个 bfs（可行解域顶点）出发

↓

寻找一个相邻的“更好（目标函数值更优）”的 bfs
（两个相邻的 bfs 对应的顶点由“可行解域的**一条棱**”相连）

↓

若不存在这样的相邻 bfs，则当前的 bfs 就是最优解！End！
否则，给出这个更好的 bfs

下两部分（二、三）
初步要求：了解这里列出的定理的结论（与前面单纯形法基本框架结合起来深入理解）[2]；
高一级要求：则是掌握这些定理的证明

二、 基可行解(bfs)的性质

给定(LP)：
$$\max c'x \quad s.t.\begin{cases} Ax=b. \\ x\geq 0 \end{cases}$$

引理. (bfs 的刻画)

给定 (LP) 的可行解 $x=(x_1,x_2,\cdots,x_n)^T$，则 x 是(LP)的一个 bfs 的充分必要条件是 x 的正分量(非 0 分量)对应的系数列线性无关。

Proof: $\Rightarrow$（必要性）由 bfs 定义可知结论成立。

$\Leftarrow$（充分性）不妨设 x 的正分量为 $x_1,x_2,\cdots,x_t$，相应的系数列向量 $P_1,P_2,\cdots,P_t$ 线性无关；x 的其余分量均取值为 0。显然 $t\leq m$。

1) 若 $t=m$，则 $P_1,P_2,\cdots,P_t$ 构成了(LP)的一个基，$x=(x_1,x_2,\cdots,x_t,0,\cdots,0)^T$ 即为相应的 bfs.

2) 若 $t<m$，因为 $rank(A)=m$，故必定可以从其余列向量中找出 $(m-t)$ 个列向量与 $P_1,P_2,\cdots,P_t$ 合在一起，构成 A 的列向量的一个极大线性无关组，即(LP)一组基。

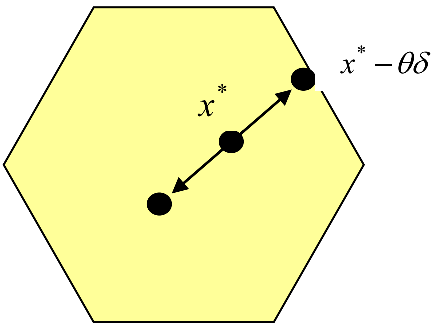
由于 x 在这些基列以外的列所对应分量上均取值为 0，故该基对应的 bfs 就是 x. ■

定理 1. (bfs 的存在性)

若线性规划(LP)可行解域非空，则它必定存在 **bfs**。

> 构造性证明思路：从(LP)的一个非 bfs 的可行解 x^* 出发，找到一个方向 δ（即满足 $A\delta=0$），使得沿该方向能得到(LP)的两个新的可行解 $x^*\pm\theta\delta$ (其中 $\theta>0$，可看作是沿所给方向 $\pm\delta$ 前进的步长)，且这两个解中至少一个的非 0 分量的个数严格减少（走到了可行解域的边界，如右图）；如此进行下去，直至得到一个 bfs（其非 0 分量对应得系数列线性无关）。

Proof:



设 x^* 是(LP)的一个可行解，
不妨设 $x^*=(x_1,x_2,\cdots,x_t,0,\cdots,0)^T$，其中 $x_1,x_2,\cdots,x_t$ 为其正分量，对应的系数列向量为 $P_1,P_2,\cdots,P_t$。

- 若 $P_1,P_2,\cdots,P_t$ 线性无关，由引理知，x^* 是(LP)的一个 bfs。
- 假设 $P_1,P_2,\cdots,P_t$ 线性相关。

1) 存在一组不全为 0 的数 $\delta_1,\delta_2,\cdots,\delta_t$，使得

$$\delta_1P_1+\delta_2P_2+\cdots+\delta_tP_t=0. \qquad \textbf{(1.1)}$$

构造 n 维向量 $\delta=(\delta_1,\delta_2,\cdots,\delta_t,0,\cdots,0)^T$，由 **(1.1)** 式和 δ 的定义知，

$$\delta_1P_1+\delta_2P_2+\cdots+\delta_nP_n=0，即 A\delta=0。$$

2) 令 $\bar{x}=x^*+\theta\delta$ 和 $\hat{x}=x^*-\theta\delta$，其中 θ 是足够小的正数。易得

$$A\bar{x}=A\left(x^*+\theta\delta\right)=Ax^*+\theta A\delta=Ax^*+0=b\text{；同理 }A\hat{x}=b\text{；}\qquad\textbf{(1.2)}$$

$$\text{因为}\begin{cases}\bar{x}=\left(x_1+\theta\delta_1,x_2+\theta\delta_2,\cdots,x_t+\theta\delta_t,0,\cdots,0\right)^T\\ \hat{x}=\left(x_1-\theta\delta_1,x_2-\theta\delta_2,\cdots,x_t-\theta\delta_t,0,\cdots,0\right)^T\end{cases},\qquad\textbf{(1.3)}$$

其中 $x_1,x_2,\cdots,x_t>0$，故 $\theta>0$ 足够小时，故能保证 $\bar{x},\hat{x}\geq 0$。
由**(1.2)**和**(1.3)**式知：当 $\theta>0$ 足够小时，$\bar{x},\hat{x}$ 均是(LP)的可行解。

3) 在保证 $\bar{x},\hat{x}\geq 0$ 的同时，我们将 θ 逐渐从 0 增大，直至 $\bar{x}$ 和 $\hat{x}$ 中某个正分量变为 0。记此时的 θ 为 θ_0。并取当 $\theta=\theta_0$ 时 $\bar{x},\hat{x}$ 中 0 分量个数增多的解为 $x^{(1)}$。

显然，$x^{(1)}$ 不仅是(LP)的可行解，且其非 0 分量个数至少减少 1。

考察解 $x^{(1)}$ 中非 0 分量对应的系数列是否线性无关：若线性无关，则由引理知该解 $x^{(1)}$ 必为 bfs；否则，重复上面的证明过程，将得到一个非 0 分量个数更少的可行解 $x^{(2)}$。

由于 A 的系数列均不是 0 向量，故重复若干次上述证明过程后，一定可找到一个可行解 $x^{(k)}$，其非 0 分量对应的列线性无关，此时我们便得到了(LP)一个的 $bfs\ x^{(k)}$。 ■

定理 2. (最优 bfs 存在性)

若线性规划(LP)存在最优解，则最优解必可在某个 bfs 上取到。

<u>Proof:</u> 设 x^* 是(LP)的一个最优解。

- 若 x^* 是一个 bfs，则问题自然得证。
- 否则，当 x^* 不是一个 bfs，可按照定理 1 的证明，存在向量 $\delta=\left(\delta_1,\delta_2,\cdots,\delta_n\right)^T$ 满足：

  （1）$A\delta=0$；

  （2）$\bar{x}=x^*+\theta_0\delta,\hat{x}=x^*-\theta_0\delta$ 都是(LP)的可行解，且其中至少一个的非 0 分量的个数比 x^* 严格减少（其中 θ_0 的取得方式与定理 2 的证明中相同）。

  因为 x^* 是(LP)的一个最优解，故有

$$c\bar{x}=c\left(x^*+\theta_0\delta\right)=cx^*+c\theta_0\delta\leq cx^*,\quad c\hat{x}=c\left(x^*-\theta_0\delta\right)=cx^*-c\theta_0\delta\leq cx^*\qquad\textbf{(1.4)}$$

同时成立。因此必有 $c\delta=0$。再根据（1.4）式知，$c\bar{x}=c\hat{x}=cx^*$，即 $\bar{x},\hat{x}$ 均为(LP)的最优解。

与定理 1 证明相同，取 $\bar{x},\hat{x}$ 中非 0 分量个数减少的一个记为 $x^{(1)}$。若 $x^{(1)}$ 的非 0 分量对应的系数列线性无关，则 $x^{(1)}$ 是最优 bfs；否则，重复以上步骤，直至找到最优 $bfs\ x^{(k)}$。 ■

三、基可行解的几何意义

1） 凸多面体 (Convex Polytopes) 相关概念

① 凸组合 (convex combination):

$\forall x,y\in R^n$，称 $\lambda x+(1-\lambda)y\ (0\leq\lambda\leq 1)$ 是 x,y 的凸组合；若 $0<\lambda<1$，则称为 x,y 的严格凸组合。凸组合的定义还可以推广到多个点的情形：

$\forall x^1,x^2,\ldots,x^k\in R^n$，

称 $\sum_{i=1}^{k}\lambda_i x^i$ 为点 $x^1, x^2, \ldots, x^k$ 的凸组合（其中 $0 \le \lambda_i \le 1\ (i = 1,2,\ldots,k)$ 且 $\sum_{i=1}^{k}\lambda_i = 1$）。

② 凸集 (convex set):

给定集合 $S \subseteq R^n$，若 $\forall x, y \in S$，它们的凸组合仍属于 S，则称集合 S 为凸集。

定理 3. (LP): $\begin{array}{ll} & \max c'x \\ s.t. & \begin{cases} Ax = b \\ x \ge 0 \end{cases}\end{array}$ 的可行解域是凸集；若最优解存在，最优解集也是凸集。

证明简单,略.

③（凸）多面体:

i) **超平面:**

$\{x \in R^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n = b\}$ 称为超平面，其中 $a_1, a_2, \cdots, a_n$ 不全为 0。

ii) **半空间:**

由超平面 $\{x \in R^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n = b\}$ 所确定的两个集合

$\{x \in R^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n \ge b\}$ 和 $\{x \in R^n : a_1x_1 + a_2x_2 + \cdots + a_nx_n \le b\}$ 称为半空间。

iii) **多面体 (polyhedra):**

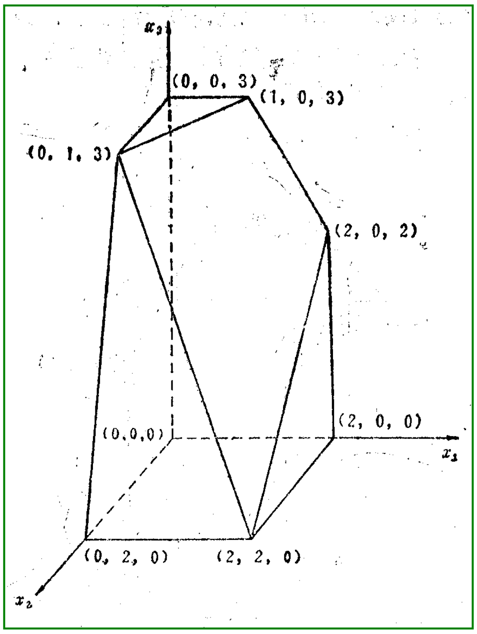
若干个半空间的交称为多面体，可表示为:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \ge b_1 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \ge b_m \end{cases};$$

结论 1: 多面体是凸集（故多面体也常被称为凸多面体）。
结论 2: 线性规划的可行解域是凸多面体。

例子: $\min z = 2x_2 + x_4 + 5x_7$

$$s.t. \begin{cases} x_1 + x_2 + x_3 + x_4 = 4 \\ x_1 + x_5 = 2 \\ x_3 + x_6 = 3 \\ 3x_2 + x_3 + x_7 = 6 \\ x_1, x_2, \cdots, x_7 \ge 0 \end{cases}$$



④ 多面体的顶点 (vertex):

点 $x \in S$ 是多面体 S 的顶点：是指 x 不能表示为 S 中两个不同点的严格凸组合；即

若存在 $x', x'' \in S$ 使得 x 表示为 $x', x'' \in S$ 的严格凸组合：$x = \lambda x' + (1-\lambda)x''$（其中 $0 < \lambda < 1$），

则必有 $x' = x'' = x$。

2）基可行解与可行解域顶点的对应

定理 4. 给定 (LP)：$\begin{array}{l}\max c'x \\ \begin{cases} Ax=b \\ x\geq 0 \end{cases}\end{array}$，其可行解域 F 非空。

则 x^* 是 F 的一个顶点当且仅当 x^* 是(LP)的一个 bfs。

Proof：不失一般性，假设 x^* 的前 t 个分量为正（其中 $t\leq m$），其他分量均为 0。

必要性： 假设 x 不是(LP)的 bfs。由定理 1 的证明知，必存在向量 $\delta=(\delta_1,\delta_2,\cdots,\delta_n)^T$ 满足：

(1) $A\delta=0$；

(2) $x^1=x^*+\theta_0\delta\in F, x^2=x^*-\theta_0\delta\in F$ (其中 $\theta_0>0$ 取值与定理 3 的证明中相同)。

→ 由此可知，$x^1\neq x^2$ 且 $x^*=x^1/2+x^2/2$。这与 $\hat{x}$ 是 F 的顶点矛盾。

充分性： 反证法. 假设 x 是(LP)的 bfs，但不是 F 的顶点。

根据顶点定义：存在两个不同点 $x',x''\in F$ $(x'\neq x'')$，使 x 可表示为 x',x'' 的严格凸组合，即 $x=\lambda x'+(1-\lambda)x''$ $(0<\lambda<1)$。

由此可知：对于下标 $j>t$ 的分量，我们有

$$x_j=x_j^{'}=x_j^{''}=0。\tag{2.1}$$

又因 x 是(LP)的 bfs，其正分量对应的系数列向量 $P_1,P_2,\cdots,P_t$ 线性无关.

故对于 $x',x''\in F$，必有 $\sum_{j=1}^{t}P_jx_j^{'}=b$，$\sum_{j=1}^{t}P_jx_j^{''}=b$. 将此两式相减即得：

$$\sum_{j=1}^{t}P_j\left(x_j^{'}-x_j^{''}\right)=0。\tag{2.2}$$

因为 $x'\neq x''$，由(2.1)、(2.2)式知，(1.6)式中 $x_j^{'}-x_j^{''}$ $(j=1,2,\ldots,t)$ 不全为 0。由此可知，向量组 $P_1,P_2,\cdots,P_t$ 线性相关，矛盾。 ■