第8讲 线性规划 (上)

罗国杰

gluo@pku.edu.cn

2024年春季学期

PKU-04833060

算法设计与分析『实验班』

本讲内容

- 线性规划模型
- 应用例子
- 单纯形法

线性规划模型

- 例 生产计划问题 用 3 种原料混合配制 2 种清洁剂

	原料1	原料2	原料3	售价(万元/吨)
清洁剂 A	0.25	0.50	0.25	12
清洁剂 B	0.50	0.50		15
存量 (吨)	120	150	50	

这 2 种清洁剂应各配制多少才能使总价值最大?

设清洁剂 A 和 B 分别配制 x 和 y 吨

$$
\begin{aligned}
&\max z = 12x + 15y \\
&\text{s.t.}\ \ 0.25x + 0.50y \le 120 \\
&\qquad 0.50x + 0.50y \le 150 \\
&\qquad 0.25x \qquad\qquad \le 50 \\
&\qquad x \ge 0,\ y \ge 0
\end{aligned}
$$

线性规划的一般形式

$$\min(\max)\ z = \sum_{j=1}^{n} c_j x_j$$ 目标函数

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \le (=, \ge) b_i, \quad i = 1, 2, \ldots, m$$ 约束条件

$$x_j \ge 0, \quad j \in J \subseteq \{1,2,\ldots,n\}$$ 非负条件

$$x_j \text{ 任意}, \ j \in \{1,2,\ldots,n\} - J$$ 自由变量

可行解　满足约束条件和非负条件的变量

可行域　全体可行解

最优解　目标函数值最小(最大)的可行解

最优值　最优解的目标函数值

线性规划求解工具

- IBM ILOG CPLEX
  - commercial
- COIN-OR Linear Programming (CLP)
  - open-source

Problem Set	CPLEX	CLP	GLPK	lp_solve[3]	MINOS[4]
Small CCO	0.0	0.1	1.3	19.0	3.1
Infeasible	0.2[1]	3.6	0.7	43.8	16.3
Netlib	9.1	29.5	52.5	14,975.1	3,198.7
Kennington	12.9	16.1	624.3	19,417.5	10,123.8
Large CCO	13.0	19.0	108.4	3,175.8	41,976.1
FOME	54.5	182.7	6,061.4	33,544.5	59,301.9
Rail	152.5	212.9	N/A[2]	29,012.2	28,899.9
PDS	179.6	224.5	34,118.3	115,200.0	115,200.0
Grand Total	**421.8**	**688.4**	**40,966.9**	**215,387.9**	**258,719.8**

[1]Infeasible problem set solution time for CPLEX does not include CPLEX2.mps in summation.
[2]None of the Rail problems for GLPK could be solved due to read error.
[3]lp_solve included 1 Netlib time out, 1 Kennington time out, 1 FOME time out, 2 Rail time outs, and 8 PDS time outs.
[4]MINOS included 3 FOME time outs, 2 Rail time outs, and 8 PDS time outs, 8 solve Small COO solve errors and 8 Large OCC solve errors.

source: Gearhart, “Comparison of Open-Source Linear Programming Solvers,” SANDIA REPORT 2013

求解工具：CPLEX 例子

```
IloNumVarArray x(env);
x.add(IloNumVar(env, 0.0, IloInfinity));
x[0].setName("x");
x.add(IloNumVar(env, 0.0, IloInfinity));
x[1].setName("y");

IloObjective obj = IloMinimize(env);
obj.setLinearCoef(x[0], 12);
obj.setLinearCoef(x[1], 15);

IloRangeArray c(env);
c.add(IloRange(env, -IloInfinity, 120));
c[0].setLinearCoef(x[0], 0.25);
c[0].setLinearCoef(x[1], 0.50);
c.add(IloRange(env, -IloInfinity, 150));
c[1].setLinearCoef(x[0], 0.50);
c[1].setLinearCoef(x[1], 0.50);
c.add(IloRange(env, -IloInfinitty, 50));
c[2].setLinearCoef(x[0], 0.25);
```

```
IloEnv env;

…

IloModel model(env);
model.add(obj);
model.add(c);

IloCplex cplex(model);
cplex.solve();

IloNumArray vals(env);
cplex.getValues(vals, x);
```

求解工具：CVXOPT 例子

```
Python 3.7.6 (default, Jan  8 2020, 19:59:22)
[GCC 7.3.0] :: Anaconda, Inc. on linux
Type "help", "copyright", "credits" or "license" for more information.
>>> from cvxopt import matrix, solvers
>>> c = matrix([ -12.0, -15.0 ])
>>> A = matrix([[ 0.25, 0.50, 0.25, -1.0, 0.0 ], [ 0.50, 0.50, 0.0, 0.0, -1.0 ]])
>>> b = matrix([ 120.0, 150.0, 50.0, 0.0, 0.0 ])
>>> sol = solvers.lp(c, A, b)
     pcost       dcost       gap    pres   dres   k/t
 0: -1.8893e+03 -4.8941e+03  3e+03  0e+00  9e-17  1e+00
 1: -3.9740e+03 -4.9444e+03  1e+03  2e-16  2e-16  8e+01
 2: -4.1350e+03 -4.2540e+03  1e+02  3e-16  4e-16  1e+01
 3: -4.1399e+03 -4.1413e+03  1e+00  2e-16  8e-16  1e-01
 4: -4.1400e+03 -4.1400e+03  1e-02  6e-16  7e-16  1e-03
 5: -4.1400e+03 -4.1400e+03  1e-04  2e-16  4e-16  1e-05
Optimal solution found.
>>> print(sol['x'])
[ 1.20e+02]
[ 1.80e+02]

>>>
```

线性规划简史

- 1939（Kantorovich）：《组织和计划的数学方法》最早提出线性规划
- 1940s（Dantzig, Kantorovich, Koopmans, von Neumann, …）：奠定线性规划的基础，起源于经济问题与物流问题的研究
- 1947（Dantzig）：单纯形法
- 1950s-60s：线性规划应用于其他领域
- 1979（Khachiyan）：椭球算法，最坏情况复杂度（多项式时间）优于单纯形法，但求解实际问题远慢于单纯形法
- 1984（Karmarkar）：投影算法（内点法），最坏情况多项式时间，实际也高效
- 1984后：内点法变种（降低复杂度、或提高实际效率），求解大规模问题的软件

应用：无穷范数（切比雪夫范数）拟合

$$\text{minimize} \quad \|Ax - b\|_\infty$$

with $A \in \mathbf{R}^{m\times n}$, $b \in \mathbf{R}^m$

- ℓ_∞**-norm (Chebyshev norm)** of m-vector y is

$$\|y\|_\infty = \max_{i=1,\ldots,m} |y_i|$$

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

应用：L1范数拟合

$$\text{minimize} \quad \|Ax - b\|_1$$

- ℓ_1-**norm** of m-vector y is

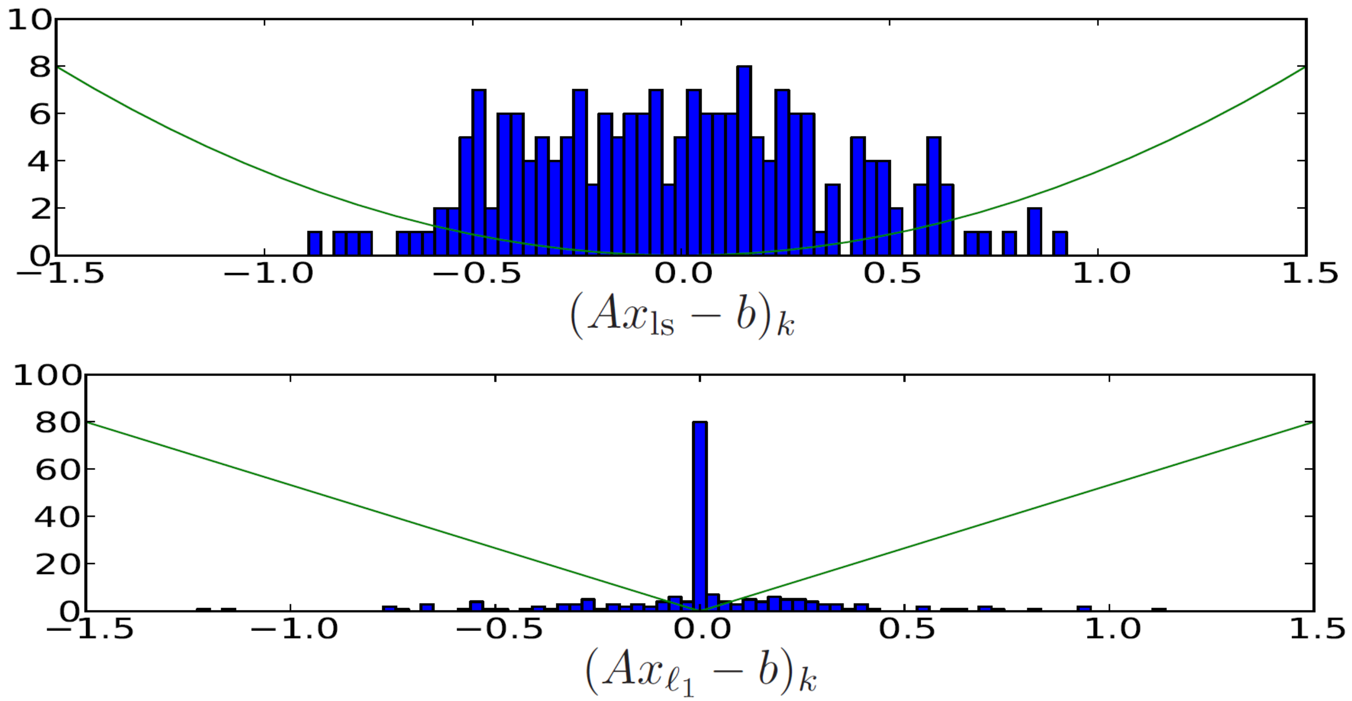
$$\|y\|_1 = \sum_{i=1}^{m} |y_i|$$

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

L1范数拟合 vs L2范数拟合（最小二乘）

histograms of residuals $Ax - b$, with randomly generated $A \in \mathbf{R}^{200\times 80}$, for

$$x_{\mathrm{ls}} = \operatorname{argmin} \|Ax - b\|, \qquad x_{\ell_1} = \operatorname{argmin} \|Ax - b\|_1$$



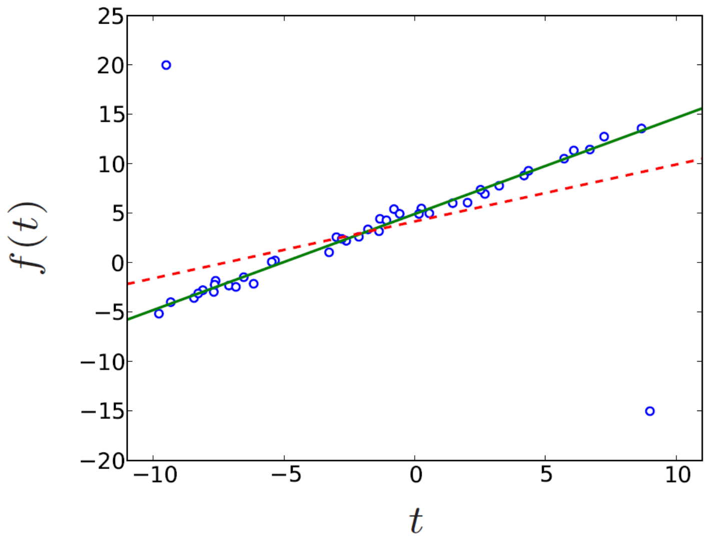
ℓ_1-norm distribution is **wider** with a **high peak at zero**

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

应用：鲁棒的线性拟合

- fit affine function $f(t) = \alpha + \beta t$ to m points (t_i, y_i)
- an approximation problem $Ax \approx b$ with

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix}, \qquad x = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \qquad b = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix}$$



- dashed: minimize $\|Ax - b\|$
- solid: minimize $\|Ax - b\|_1$

ℓ_1-norm approximation is more robust against outliers

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

应用：稀疏信号重建

- $\hat{x} \in \mathbf{R}^n$ is unknown signal, known to be very sparse
- we make linear measurements $y = A\hat{x}$ with $A \in \mathbf{R}^{m \times n}$, $m < n$

estimation by ℓ_1-norm minimization: compute estimate by solving

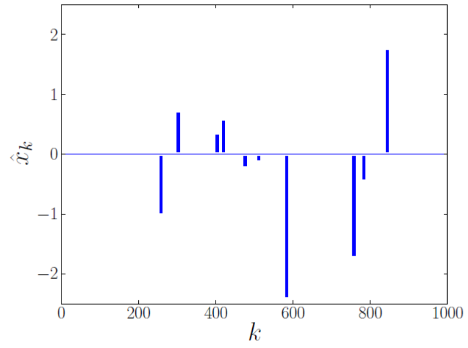
$$\begin{array}{ll} \text{minimize} & \|x\|_1 \\ \text{subject to} & Ax = y \end{array}$$

estimate is signal with smallest ℓ_1-norm, consistent with measurements

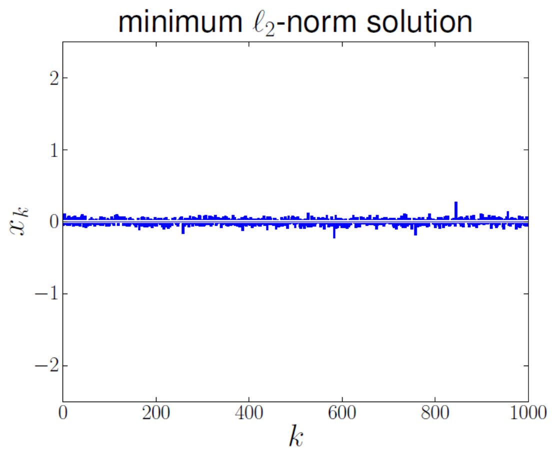
http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

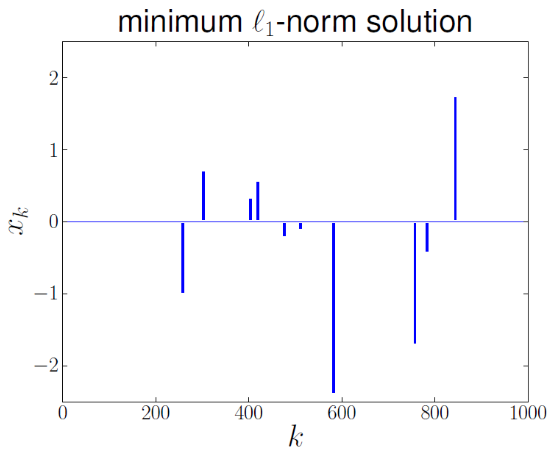
应用：稀疏信号重建例子

- exact signal $\hat{x} \in \mathbf{R}^{1000}$
- 10 nonzero components



least-norm solutions (randomly generated $A \in \mathbf{R}^{100\times 1000}$)





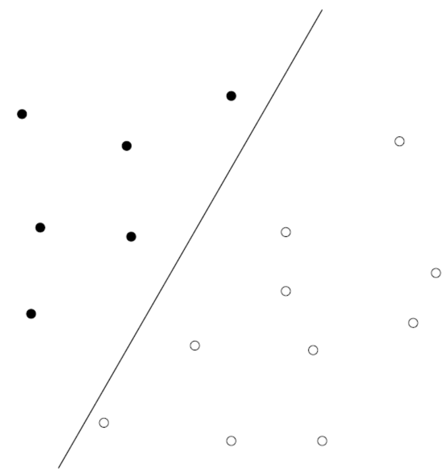
ℓ_1-norm estimate is **exact**

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

应用：线性分类

- given a set of points $\{v_1, \ldots, v_N\}$ with binary labels $s_i \in \{-1, 1\}$
- find hyperplane that strictly separates the two classes

$$
\begin{aligned}
a^T v_i + b > 0 &\quad \text{if } s_i = 1 \\
a^T v_i + b < 0 &\quad \text{if } s_i = -1
\end{aligned}
$$



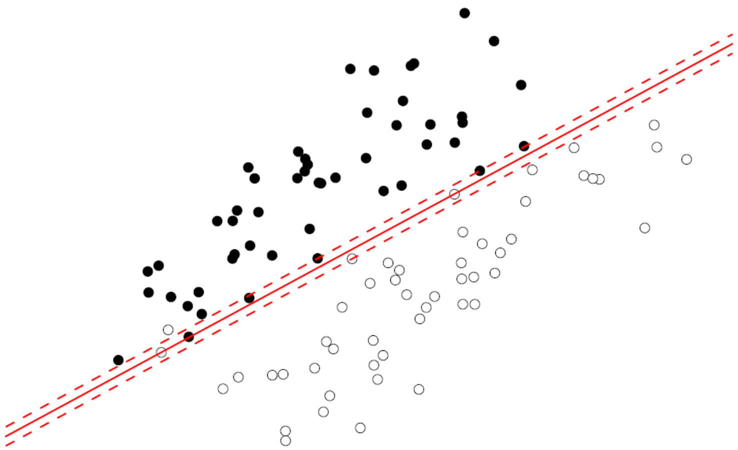
homogeneous in a, b, hence equivalent to the linear inequalities (in a, b)

$$
s_i(a^T v_i + b) \geq 1, \quad i = 1, \ldots, N
$$

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

应用：线性不可分集合的近似线性分类

$$\text{minimize} \quad \sum_{i=1}^{N} \max\{0, 1 - s_i(a^T v_i + b)\}$$



- penalty $1 - s_i(a_i^T v_i + b)$ for misclassifying point v_i
- can be interpreted as a heuristic for minimizing #misclassified points
- a piecewise-linear minimization problem with variables a, b

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

二维线性规划图解法（1/2）

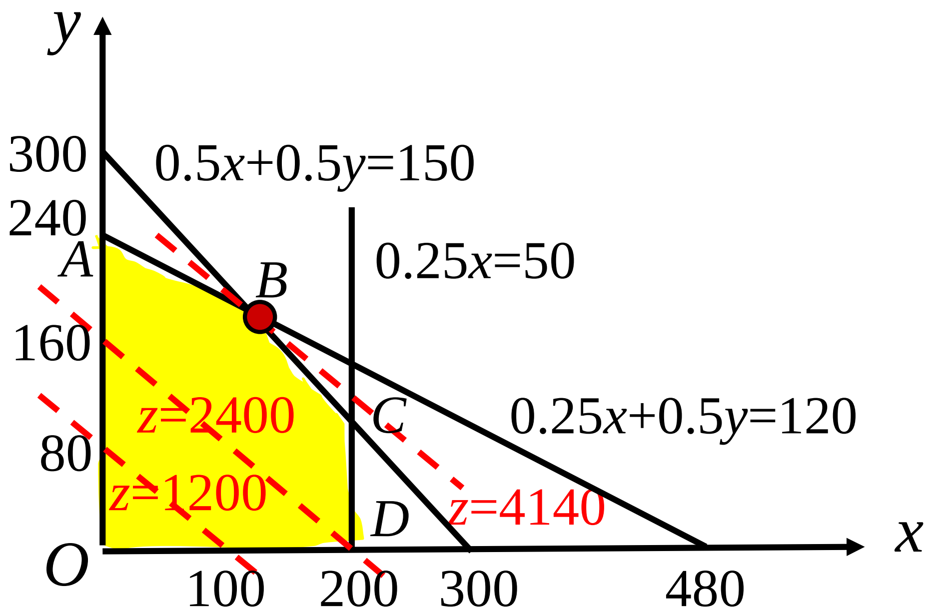
例 $\max z = 12x + 15y$

s.t. $0.25x+0.50y \le 120$

$0.50x+0.50y \le 150$

$0.25x \le 50$

$x \ge 0, y \ge 0$



$O(0,0)$, $A(0,240)$, $B(120,180)$, $C(200,100)$, $D(200,0)$

最优解 $x^*=120$, $y^*=180$ (点B) **最优值** $z^* = 4140$.

目标函数改为 $\max z = 12x + 12y$

最优解 $x^*= 120t + 200(1- t) = 200-80t$

$y^*=180t +100(1- t) = 100 + 80t$,

$z^*=3600$ 最优值

($0\le t \le 1$，线段BC)

二维线性规划图解法（2/2）

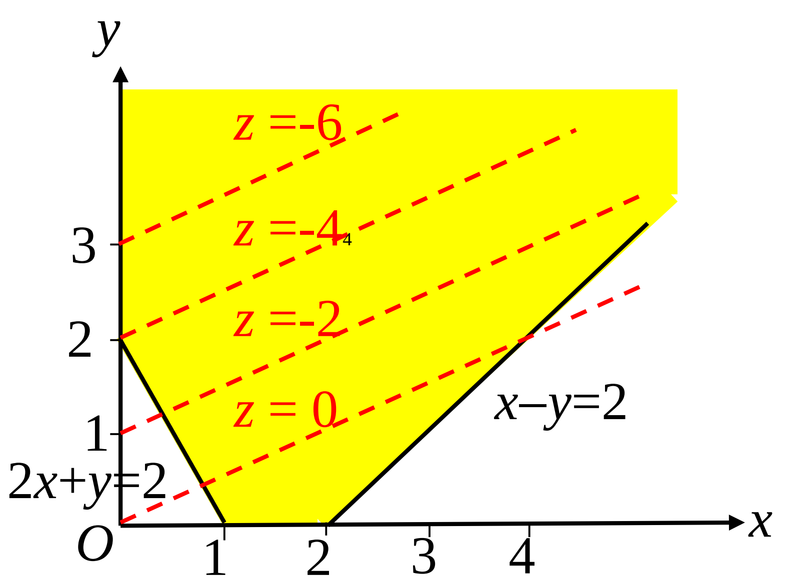
例 $\min z = x - 2y$

$$\text{s.t. } 2x + y \geq 2$$
$$x - y \leq 2$$
$$x \geq 0, y \geq 0$$

有可行解

目标函数值可以任意小

无最优解.



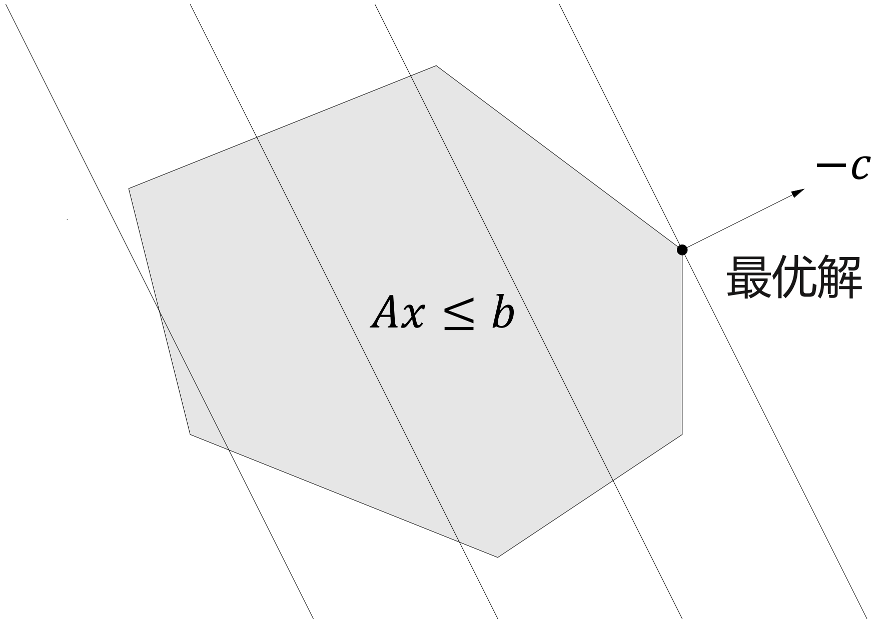
$2x + y \geq 2$ 改为 $2x + y \leq 2$,

$x - y \leq 2$ 改为 $x - y \geq 2$

则可行域为空集, 无可行解

线性规划的几何解释

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \le b \end{array}$$

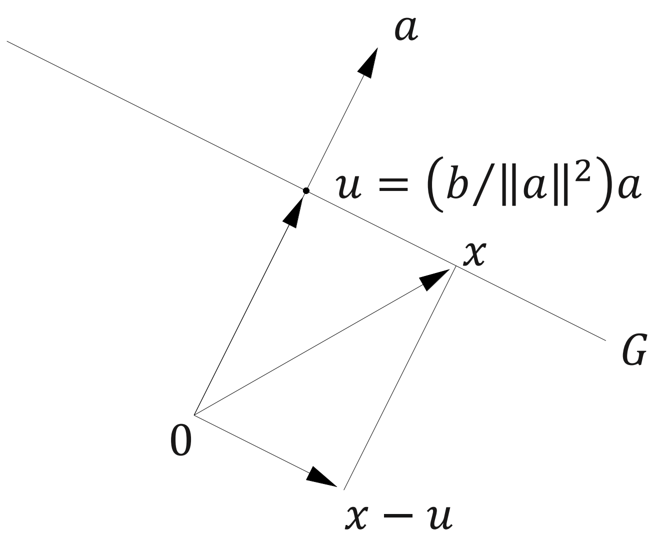


- 虚线（超平面）对应 $c^T x = \alpha$ 在不同 α 值的水平集（level set）

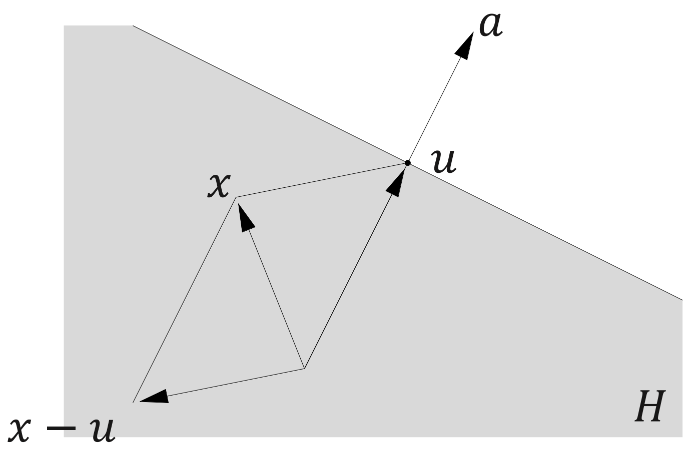
http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释：超平面和半空间

- 超平面 $G = \{x \mid a^T x = b\}$



- 半空间 $H = \{x \mid a^T x \leq b\}$



http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

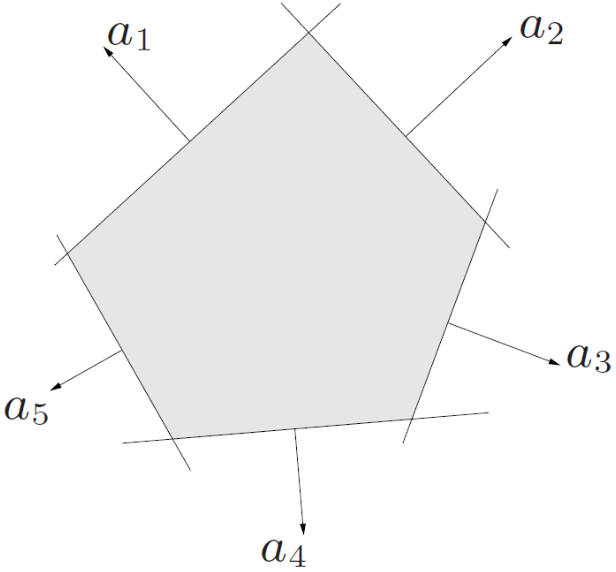
线性规划的几何解释：超平面和半空间例子

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释：多面体

- 可行域是多面体（有限个半空间的交集）

$$a_1^T x \le b_1, \qquad a_2^T x \le b_2, \qquad \ldots, \qquad a_m^T x \le b_m$$



– 超平面 $Fx = g$ 等价于两个半空间 $Fx \le g$ 和 $-Fx \le -g$ 的交集

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释：多面体例子

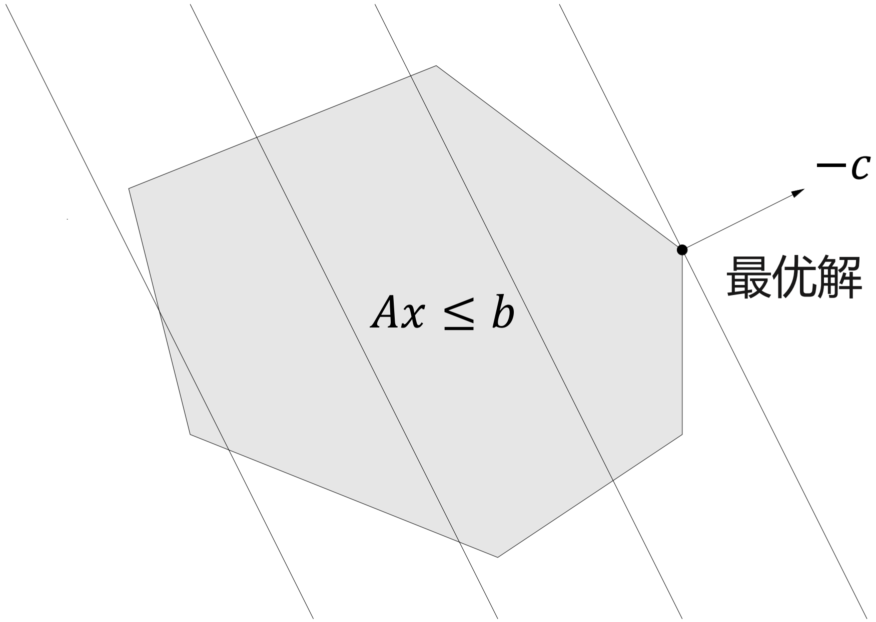
http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释：多面体例子

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \le b \end{array}$$

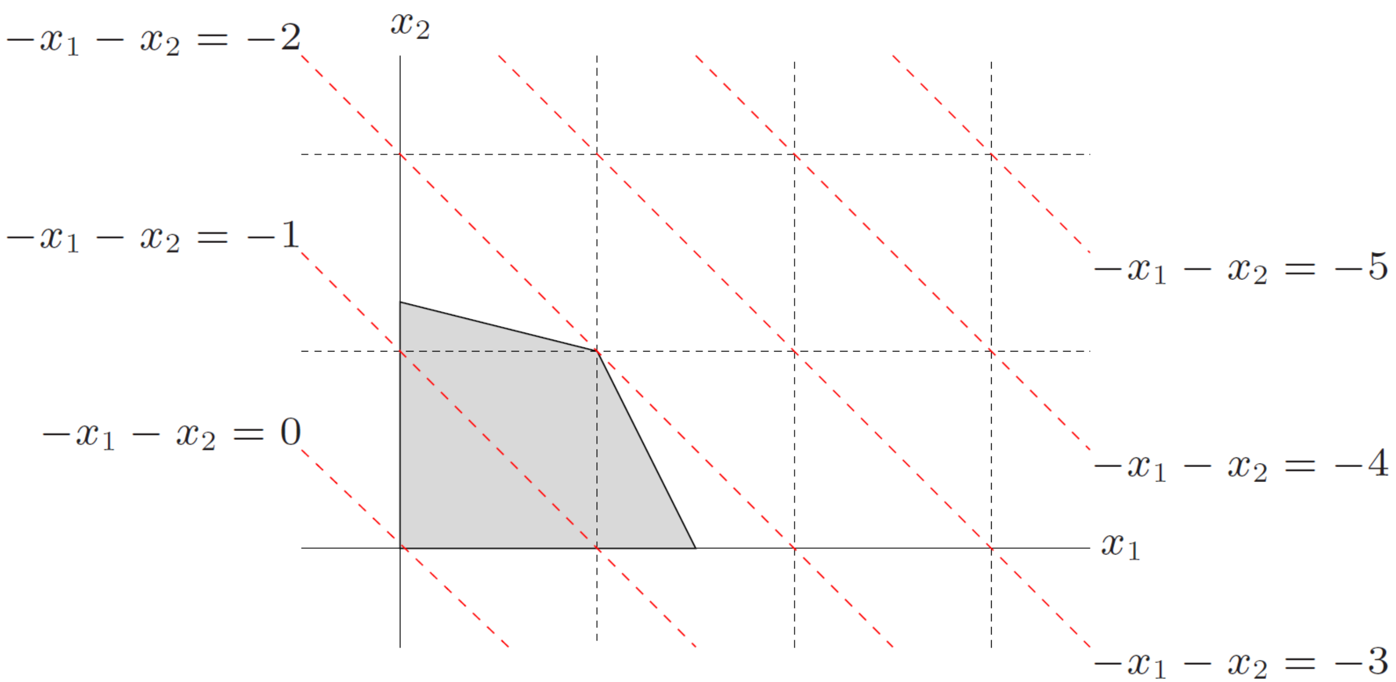


- 虚线（超平面）对应 $c^T x = \alpha$ 在不同 α 值的水平集（level set）

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

线性规划的几何解释：例子

$$
\begin{array}{ll}
\text{minimize} & -x_1 - x_2 \\
\text{subject to} & 2x_1 + x_2 \leq 3 \\
& x_1 + 4x_2 \leq 5 \\
& x_1 \geq 0, \quad x_2 \geq 0
\end{array}
$$



optimal solution is $(1,1)$

http://www.seas.ucla.edu/~vandenbe/ee236a/ee236a.html

几种解的情况

(1) 解有4种可能

(a) 有唯一的最优解.

(b) 有无穷多个最优解.

(c) 有可行解, 但无最优解 (目标函数值无界).

(d) 无可行解, 更无最优解.

(2) 可行域是一个凸多边形 (可能无界, 也可能是空集).

如果有最优解, 则一定可以在凸多边形的顶点取到.

一般的 n 维线性规划也是如此

单纯形法 (Simplex)

- 标准形
  - 目标最小 min $c^T x$
  - 等式约束 $Ax=b$，且 $b\geq 0$
  - 变量非负 $x\geq 0$
- 基本可行解
  - 构造 $Ax=b$ 和 $x\geq 0$ 且满足一定性质的可行解
- 单纯形法基础
  - 定理1：如果标准形有可行解，则必有基本可行解。
  - 定理2：如果标准形有最优解，则必存在一个基本可行解是最优解。
- 单纯形法步骤
  - 确定初始基本可行解
  - 从一组基本可行解变换到另一组“相邻的”基本可行解，且使目标函数下降
  - 重复上述步骤直至找到最优解

单纯形法的线性规划标准形

标准形

$$\min z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i \geq 0, \quad i = 1,2,\ldots,m$$

$$x_j \geq 0, \quad j = 1,2,\ldots,n$$

特点

目标函数：最小化

约束条件：等式约束，常数非负；变量非负

化成标准形

(1) 把 max z 替换成 min $z' = -z$, 即取 $c_j' = -c_j$.

(2) $b_i < 0$. 两边同时变号, $\leq$ 改变成 $\geq$, $\geq$ 改变成 $\leq$.

(3) $\sum_{j=1}^{n} a_{ij} x_j \leq b_i$. 引入松弛变量 $y_i \geq 0$, 替换成

$$\sum_{j=1}^{n} a_{ij} x_j + y_i = b_i$$

(4) $\sum_{j=1}^{n} a_{ij} x_j \geq b_i$. 引入剩余变量 $y_i \geq 0$, 替换成

$$\sum_{j=1}^{n} a_{ij} x_j - y_i = b_i$$

(5) 自由变量 x_j 替换成 $x_j' - x_j''$, $x_j' \geq 0$, $x_j'' \geq 0$

化成标准形：例子

写出下述线性规划的标准形

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

$$\text{s.t.}\quad x_1 + 3x_2 - 3x_3 \le 10$$

$$4x_1 - x_2 - 5x_3 \le -30$$

$$x_1 \ge 0,\ x_2 \ge 0,\ x_3 \text{ 任意}$$

解

$$\min z' = -3x_1 + 2x_2 - x_3' + x_3''$$

$$\text{s.t.}\quad x_1 + 3x_2 - 3x_3' + 3x_3'' + x_4 = 10$$

$$-4x_1 + x_2 + 5x_3' - 5x_3'' - x_5 = 30$$

$$x_1 \ge 0,\ x_2 \ge 0,\ x_3' \ge 0,\ x_3'' \ge 0,\ x_4 \ge 0,\ x_5 \ge 0$$

标准形的其他形式

$$\boldsymbol{A}=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad \boldsymbol{b}=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad \boldsymbol{c}=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad \boldsymbol{x}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

矩阵形式

$$\min z = c^T x$$

$$\text{s.t.}\ \ Ax = b$$

$$x \geq 0$$

向量形式

$$\min z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} P_j x_j = b$$

$$x_j \geq 0,\ \ j = 1,2,\ldots,n$$

$$P_j = \begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

标准形的可行解的性质

定义 设 A 的秩为 m,

A的 m 个线性无关的列向量称作标准形的基.

给定基 $B = (\boldsymbol{P}_{i_1}, \boldsymbol{P}_{i_2}, \cdots, \boldsymbol{P}_{i_m})$,

对应基中列向量的变量 $\boldsymbol{x}_{i_1}, \boldsymbol{x}_{i_2}, \cdots, \boldsymbol{x}_{i_m}$ 称作基变量,

其余的变量称作非基变量.

基变量构成的向量记作 x_B, 非基变量构成的向量记作 x_N. 令 $x_N = 0$, 等式约束变成

$$B\, x_B = b$$

解得 $x_B = B^{-1}b$. 向量 x 满足约束 $Ax = b$且非基变量全为 0, 称作关于基 B 的基本解 .

x是一个基本解且 $x \geq 0$, 则称 x是基本可行解, 对应的基 B为可行基.

基变量/基本可行解/可行基：例子

$$
\begin{aligned}
&\min z = -12x_1 - 15x_2 \\
&\text{s.t.}\ \ 0.25x_1+0.50x_2+x_3 \qquad\qquad =120 \\
&\qquad 0.50x_1+0.50x_2 \qquad +x_4 \qquad =150 \\
&\qquad 0.25x_1 \qquad\qquad\qquad\qquad\qquad +x_5=50 \\
&\qquad x_i\geq 0, \qquad i=1,2,\ldots,5
\end{aligned}
$$

$$
\boldsymbol{A}=\begin{bmatrix} \mathbf{0.25} & \mathbf{0.50} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0.50} & \mathbf{0.50} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{0.25} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}
$$

基 $B_1=(P_1,P_2,P_3)$. 基变量 x_1, x_2, x_3, 非基变量 x_4, x_5.

令 $x_4= 0$, $x_5 = 0$, 得

$$
\begin{aligned}
0.25x_1 + 0.50x_2 + x_3 &= 120 \\
0.50x_1+ 0.50x_2 \qquad &= 150 \\
0.25x_1 \qquad\qquad\qquad &= 50
\end{aligned}
$$

解得 $x_1 = 200$, $x_2 = 100$, $x_3 = 20$.

$x^{(1)} = (200,100,20,0,0)^T$ 是基本可行解, B_1是可行基.

基变量/基本可行解/可行基：例子（续）

取基 $B_2=(P_1,P_2,P_4)$. 基变量 x_1, x_2, x_4, 非基变量 x_3, x_5.

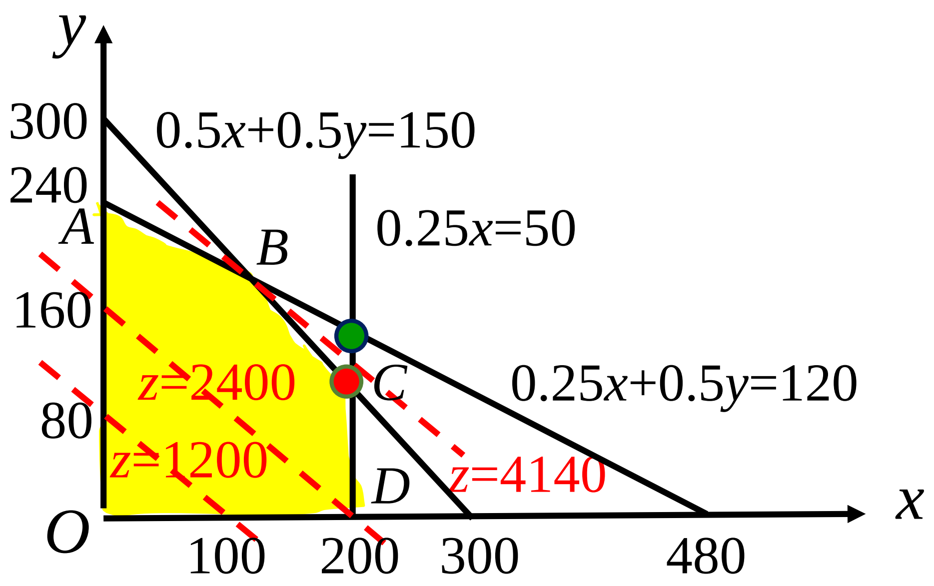
令$x_3=0,\ x_5=0$, 由

$$\begin{aligned} 0.25x_1 + 0.50x_2 \quad\quad &= 120 \\ 0.50x_1 + 0.50x_2 + x_4 &= 150 \\ 0.25x_1 \quad\quad\quad\quad\quad &= 50 \end{aligned}$$

解得 $x_1=200,\ x_2=140,\ x_4=-20.$

$x^{(2)}=(200, 140, 0, -20, 0)^T$是基本解,

不是基本可行解.



基本可行解的性质

引理1 α 是 $Ax=b$ 的解，即 $A\alpha=b$

α 是基本解 $\Leftrightarrow$ α 中非零分量对应的列向量线性无关.

证 $\Rightarrow$ 必要性 根据基本解的定义, 这是显然的.

$\Leftarrow$ 充分性 设α 的非零分量为 $\boldsymbol{\alpha}_{j_1}, \boldsymbol{\alpha}_{j_2}, \cdots, \boldsymbol{\alpha}_{j_r}$, 对应的列向量 $\boldsymbol{P}_{j_1}, \boldsymbol{P}_{j_2}, \cdots, \boldsymbol{P}_{j_r}$ 线性无关. A的秩为 m, 必存在 $\boldsymbol{P}_{j_r+1}, \cdots, \boldsymbol{P}_{j_m}$

使得 $\boldsymbol{P}_{j_1}, \boldsymbol{P}_{j_2}, \cdots, \boldsymbol{P}_{j_m}$ 线性无关, 构成一个基, 记作B. α 是方程 $Bx_B = b$ 的解, 而这个方程的解是惟一的, 故α 是关于B 的基本解.

基本可行解的性质

定理1 如果标准形有可行解, 则必有基本可行解.

证 设α是一个可行解, 从α开始, 构造出一个基本可行解.

设α的非零分量为 $\alpha_1, \alpha_2, \ldots, \alpha_r$, $r \le n$. 如果对应的列向量 $P_1, P_2, \ldots, P_r$ 线性无关, 则α是一个基本可行解.

否则, 存在不全为 0 的 $\lambda_1, \lambda_2, \ldots, \lambda_r$ 使

$$\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_r P_r = 0$$

取 $\lambda_{r+1} = \ldots = \lambda_n = 0$, 有

$$\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_n P_n = 0$$

于是, 对任意的δ,

$$\sum_{j=1}^{n} (\alpha_j + \delta\lambda_j) P_j = \sum_{j=1}^{n} \alpha_j P_j + \delta \sum_{j=1}^{n} \lambda_j P_j = b$$

定理1：如果标准形有可行解, 则必有基本可行解（续）

记 $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)^T$, 为使 $\alpha+\delta\lambda$ 成为一个可行解, 要求所有 $\alpha_j+\delta\lambda_j\geq 0$. 当 $\lambda_j=0$ 时, 不等式自然成立.

当 $\lambda_j>0$时, 要求 $\delta\geq -\alpha_j/\lambda_j$
当 $\lambda_j<0$时, 要求 $\delta\leq -\alpha_j/\lambda_j$ $\Rightarrow$ 当 $\lambda_j\neq 0$ 时, $\delta\leq|\alpha_j/\lambda_j|$

设 $$\left|\boldsymbol{\alpha}_{j_0}/\boldsymbol{\lambda}_{j_0}\right|=\min\left\{|\boldsymbol{\alpha}_j/\boldsymbol{\lambda}_j|:\boldsymbol{\lambda}_j\neq\mathbf{0}\right\},\quad 1\leq j_0\leq r$$

取 $\delta^*=-\boldsymbol{\alpha}_{j_0}/\boldsymbol{\lambda}_{j_0}$, 令 $\beta_j=\alpha_j+\delta^*\lambda j$, $(j=1,2,\ldots,n)$

$$\beta_1P_1+\ldots+\beta_nP_n=b$$

$\beta_j\geq 0\ (j=1,2,\ldots,n)$, $\boldsymbol{\beta}_{j_0}=0$, $\beta_{r+1}=\ldots=\beta_n=0$.
$\beta=(\beta_1,\beta_2,\ldots,\beta_n)$是可行解且比 α 至少少一个非零分量.

上述过程至多进行 r-1 次一定可以得到一个基本可行解.

基本可行解的性质

定理2 如果标准形有最优解, 则必存在一个基本可行解是最优解.

证　当α是最优解时, β也是最优解. （定理1的 α 和 β）

由$\alpha_j = 0 \Rightarrow \lambda_j = 0$, 对足够小的$\delta > 0$, $\alpha + \delta\lambda$ 和 $\alpha - \delta\lambda$ 都是可行解.

$$\left.\begin{array}{l}\sum_{j=1}^{n} c_j\alpha_j \le \sum_{j=1}^{n} c_j(\alpha_j + \delta\lambda_j) = \sum_{j=1}^{n} c_j\alpha_j + \delta\sum_{j=1}^{n} c_j\lambda_j \\ \sum_{j=1}^{n} c_j\alpha_j \le \sum_{j=1}^{n} c_j(\alpha_j - \delta\lambda_j) = \sum_{j=1}^{n} c_j\alpha_j - \delta\sum_{j=1}^{n} c_j\lambda_j\end{array}\right\} \Rightarrow \sum_{j=1}^{n} c_j\lambda_j = 0$$

$$\sum_{j=1}^{n} c_j\beta_j = \sum_{j=1}^{n} c_j(\alpha_j + \delta^*\lambda_j) = \sum_{j=1}^{n} c_j\alpha_j + \delta^*\sum_{j=1}^{n} c_j\lambda_j = \sum_{j=1}^{n} c_j\alpha_j$$

$\beta = \alpha + \delta^*\lambda$ 也是最优解.

A有m行n列, 至多有 C_n^m个基, 故至多有 C_n^m 个基本解.

单纯形法

基本步骤

(1) 确定初始基本可行解.

(2) 检查当前的基本可行解.

若是最优解或无最优解, 计算结束;

否则作基变换, 用一个非基变量替换一个基变量, 得到一个新的可行基和对应的基本可行解, 且使目标函数值下降 (至少不升).

(3) 重复 (2).

确定初始基本可行解

先考虑最简单的情况，设约束条件为

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \le b_i \text{ , } b_i \ge 0, \ \ i = 1,2,\dots,m$$

引入 m 个松弛变量 $x_{n+i} \ge 0$ （ $i = 1, 2, \dots, m$ ），

$$a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n + x_{n+i} = b_i \text{ , } \quad i=1,2,\dots,m$$

取 x_{n+i} ($i = 1, 2, \dots, m$) 作为基变量，初始基本可行解为

$$x^{(0)} = (0, 0, \dots, 0, b_1, b_2, \dots, b_m)^T$$

$$\begin{aligned}
&\min z' = -12x_1 - 15x_2 \\
&\text{s.t.} \ \ 0.25x_1 + 0.50x_2 + x_3 = 120 \\
&\qquad 0.50x_1 + 0.50x_2 + x_4 = 150 \\
&\qquad 0.25x_1 + x_5 = 50 \\
&\qquad x_j \ge 0, \quad i = 1,2,\dots,5
\end{aligned}$$

例 $\max z = 12x + 15y$

$$\begin{aligned}
&\text{s.t.} \ \ 0.25x + 0.50y \le 120 \\
&\qquad 0.50x + 0.50y \le 150 \\
&\qquad 0.25x \le 50 \\
&\qquad x \ge 0, y \ge 0
\end{aligned}$$

取 x_3, x_4, x_5作为基变量，$x^{(0)} = (0,0,120,150,50)^T$

最优性检验（1/2）

$$\min z = c^T x$$
$$\text{s.t.}\ \ Ax = b$$
$$x \geq 0$$

给定可行基 $B=(P_{\pi(1)}, P_{\pi(2)},\cdots, P_{\pi(m)})$

非基变量取 0，基变量由约束 $Ax=b$ 确定唯一取值

目标函数

$$\begin{aligned} z &= c^T x \\ &= c^T x + c_B^T B^{-1}(b - Ax) \\ &= c_B^T B^{-1} b + (c^T - c_B^T B^{-1} A)x \\ &= z_0 + (c^T - c_B^T B^{-1} A)x \end{aligned}$$

以 B 为可行基时，目标函数取值 z_0

记 $\lambda^T = c^T - c_B^T B^{-1} A$ （**检验数**）

目标函数可改写成 $z = z_0 + \lambda^T x$

最优性检验（2/2）

$$z = z_0 + \lambda^T x$$
$$x_B = B^{-1}b - B^{-1}Nx_N$$

记 $B^{-1}A = (\alpha_{ij})_{m\times n}$, $P_j'=B^{-1}P_j$ $(1\le j\le n)$, $\beta=B^{-1}b$.

定理 3　给定基本可行解 $x^{(0)}$,
(1) 若所有检验数大于等于0, 则 $x^{(0)}$是最优解.
(2) 若存在检验数 $\lambda_k< 0$且所有$\alpha_{ik}\le 0$ $(1\le i\le m)$, 则无最优解.

证 (1) 如果 $\lambda\ge 0$, 则对任意可行解, $x\ge 0$, $z\ge z_0$, 故 $x^{(0)}$是最优解.

(2) 若存在$\lambda_k< 0$ (λ_k必对应非基变量) 且所有$\alpha_{ik}\le 0$ $(1\le i\le m)$,
取 $x_k = M > 0$, 其余非基变量 $x_j = 0$, 解得

$$x_{\pi(i)} = \beta_i - \alpha_{ik}M \ge 0, \qquad 1\le i\le m$$

这是一个可行解, 其目标函数值为 $z = z_0 + \lambda_k M$
当 $M\to +\infty$ 时, $z\to -\infty$. 得证无最优解.

问题 存在检验数 $\lambda_k< 0$且有$\alpha_{lk} > 0$的情况呢？基变换.

基变换（1/3）

给定可行基 $B=(P_{\pi(1)},P_{\pi(2)},\ldots,P_{\pi(m)})$, 设 $\lambda_k< 0$且$\alpha_{lk} > 0$, x_k必是非基变量.

基变换: 用非基变量 x_k 替换基变量 $x_{\pi(l)}$, 用P_k替换 B 中的$P_{\pi(l)}$, 新的基为 $B'=(P_{\pi(1)}, \ldots, P_{\pi(l-1)}, P_k, P_{\pi(l+1)}, \ldots, P_{\pi(m)})$.

称 x_k 为换入变量, $x_{\pi(l)}$ 为换出变量.

(1) 证 B' 是基, 即 $P_{\pi(1)}, \ldots, P_{\pi(l-1)}, P_k, P_{\pi(l+1)}, \ldots, P_{\pi(m)}$ 线性无关. 由于$P_{\pi(1)}, P_{\pi(2)}, \ldots, P_{\pi(m)}$ 线性无关, 只需证 $P_{\pi(l)}$可表成 $P_{\pi(1)}, \ldots, P_{\pi(l-1)}, P_k, P_{\pi(l+1)}, \ldots, P_{\pi(m)}$的线性组合.

由于 $$(B^{-1}P_{\pi(1)}, B^{-1}P_{\pi(2)}, \ldots, B^{-1}P_{\pi(m)}) = B^{-1}B = E,$$

$$(B^{-1}P_k) = P'_k = \sum_{i=1}^{m} \alpha_{ik} P'_{\pi(i)} = \sum_{i=1}^{m} \alpha_{ik} (B^{-1}P_{\pi(i)}) \Longrightarrow \boldsymbol{P_k = \sum_{i=1}^{m} \alpha_{ik} P_{\pi(i)}}$$

基变换（2/3）

解得

$$\boldsymbol{P}_{\pi(l)} = \frac{1}{\alpha_{lk}}\boldsymbol{P}_k - \sum_{\substack{i=1\\ 且\neq l}}^{m}\frac{\alpha_{ik}}{\alpha_{lk}}\boldsymbol{P}_{\pi(i)}$$

得证 B' 是一个基.

(2) 要保证 B' 是可行基.

$$Ax = b$$

$$B^{-1}Ax = B^{-1}b = \beta$$

$$x_B + B^{-1}Nx_N = \beta$$

$B^{-1}A=(P_1',P_2',\ldots,P_m')$ 中对应 x_B 的列 (第$\pi(1),\ldots,\pi(m)$列) 构成单位矩阵.

用P_k替换 $P_{\pi(l)}$得到 B', 将 x_B中的 $x_{\pi(l)}$替换成 x_k , 即解出第 l 个方程中的 x_k. 这只需用α_{lk}除第l个方程, 再用第 l 个方程消去其它方程中的 x_k

基变换（3/3）

计算公式

$$\boxed{\begin{array}{l} \alpha_{lj}'=\alpha_{lj}/\alpha_{lk}, \quad 1\le j\le n \\ \alpha_{ij}'=\alpha_{ij}-\alpha_{ik}\alpha_{lj}/\alpha_{lk}, \ 1\le i\le m \text{且 } i\ne l,\ 1\le j\le n \\ \beta_l' = \beta_l/\alpha_{lk} \\ \beta_i' = \beta_i-\alpha_{ik}\beta_l/\alpha_{lk}, \quad 1\le i\le m \text{ 且 } i\ne l \end{array}}$$

为保证 B'是可行的, 只需

$$\beta_i' = \beta_i - \alpha_{ik}\beta_l/\alpha_{lk} \ge 0, \quad 1\le i\le m\text{且 } i\ne l$$

$\beta_i \ge 0$, $\beta_l \ge 0$, $\alpha_{lk} > 0$. $\alpha_{ik} \le 0$时不等式成立； $\alpha_{ik} > 0$时$\beta_l/\alpha_{lk} \le \beta_i/\alpha_{ik}$

取 l 使得 $\beta_l/\alpha_{lk} = \min\{ \beta_i/\alpha_{ik} \mid \alpha_{ik} > 0,\ 1\le i\le m \}$

用第 l 个方程消去简化的目标函数中的 x_k,

$$\lambda_j' = \lambda_j-\lambda_k\alpha_{lj}/\alpha_{lk}, \quad 1\le j\le m$$

$$z_0' = z_0+\lambda_k\beta_l/\alpha_{lk}$$

单纯形法

算法 单纯形法 (针对最小化)

1. 设初始可行基 $B=(P_{\pi(1)}, P_{\pi(2)},\ldots, P_{\pi(m)})$, $\alpha = B^{-1}A$, $\beta = B^{-1}b$,

   $\lambda^T = c^T - c_B^T B^{-1}A$, $z_0 = B^{-1}b$.

2. 若所有$\lambda_j \geq 0$ $(1\leq j\leq n)$, 则 $x_B=\beta$, $x_N=0$是最优解, 计算结束.

3. 取$\lambda_k < 0$. 若所有$\alpha_{ik} \leq 0$ $(1\leq i \leq m)$, 则无最优解, 计算结束.

4. 取 l 使得

$$\beta_l / \alpha_{lk} = \min\{ \beta_i / \alpha_{ik} \mid \alpha_{ik} > 0,\ 1\leq i \leq m \}$$

5. 以 x_k为换入变量、$x_{\pi(l)}$ 为换出变量做基变换.

6. 转 2.

对最大化, 2中$\lambda_j \geq 0$ 改为 $\lambda_j \leq 0$, 3中 $\lambda_k < 0$ 改为 $\lambda_k > 0$.

单纯形表

			c_1	c_2	$\dots$	c_n	
c_B	x_B	b	x_1	x_2	$\dots$	x_n	θ
$c_{\pi(1)}$	$x_{\pi(1)}$	β_1	α_{11}	α_{12}	$\dots$	α_{1n}	
$c_{\pi(2)}$	$x_{\pi(2)}$	β_2	α_{21}	α_{22}	$\dots$	α_{2n}	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$	
$c_{\pi(m)}$	$x_{\pi(m)}$	β_m	α_{m1}	α_{m2}	$\dots$	α_{mn}	
	-z	-z_0	λ_1	λ_2	$\dots$	λ_n	

$$-z + \lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n = -z_0$$

$$\min z = c^T x$$
$$\text{s.t. } Ax = b$$
$$x \geq 0$$

基变量 B

$(\alpha_{i,j}) = B^{-1}A$
$(\beta_i) = B^{-1}b$

$\lambda^T = c^T - c_B^T B^{-1} A$

选择 $\lambda_k < 0$:
$\theta_i = \beta_i / \alpha_{i,k}$
for $\alpha_{i,k} > 0$

令 $\theta_l = min\{\theta_i\}$:
基变量换出 $x_{\pi(l)}$
基变量换入 x_k

重复上述过程

单纯形表：例1

$$\min z = -12x_1 - 15x_2$$
$$\text{s.t.}\quad 0.25x_1+0.50x_2+x_3 = 120$$
$$0.50x_1+0.50x_2+x_4 = 150$$
$$0.25x_1+x_5 = 50$$

			-12	-15	0	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	120	0.25	0.50	1	0	0	240
0	x_4	150	0.50	0.50	0	1	0	300
0	x_5	50	0.25	0	0	0	1	
	-z	0	-12	-15	0	0	0	
-15	x_2	240	0.50	1	2	0	0	480
0	x_4	30	0.25	0	-1	1	0	120
0	x_5	50	0.25	0	0	0	1	200
	-z	3600	-4.5	0	30	0	0	
-15	x_2	180	0	1	4	-2	0	
-12	x_1	120	1	0	-4	4	0	
0	x_5	20	0	0	1	- 1	1	
	-z	4140	0	0	12	18	0	

单纯形法：例2

用单纯形法解下述线性规划

$$\min z = x_1 - 2x_2$$

$$\text{s.t.}\ \ x_1 - x_2 \le 1$$

$$-2x_1 + x_2 \le 4$$

$$x_1 \ge 0,\ x_2 \ge 0$$

解 引入2个松弛变量 x_3, x_4, 得到标准形

$$\min z = x_1 - 2x_2$$

$$\text{s.t.}\ \ x_1 - x_2 + x_3 = 1$$

$$-2x_1 + x_2 + x_4 = 4$$

$$x_j \ge 0,\quad j = 1,2,3,4$$

单纯形表：例2

			1	-2	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	θ
0	x_3	1	1	-1	1	0	
0	x_4	4	-2	1	0	1	4
	-z	0	1	-2	0	0	
0	x_3	5	-1	0	1	1	
-2	x_2	4	-2	1	0	1	
	-z	8	-3	0	0	2	

目标函数值没有下界, 无最优解

构造初始基本可行解（其他约束条件）

- 不等式约束 $\sum_{j=1}^{n} a_{ij}x_j \leq b_i$
  - 按前述方法，松弛变量取为不等式右值，构造初始基本可行解
- 不等式条件 $\sum_{j=1}^{n} a_{ij}x_j \geq b_i$
  - 添加剩余变量，转化为等式约束
- 等式约束 $\sum_{j=1}^{n} a_{ij}x_j = b_i$
  - 引入人工变量和辅助问题（两阶段法）
  - 人工变量：$\sum_{j=1}^{n} a_{ij}x_j + y_i = b_i$ 且 $y_i \geq 0$
  - 辅助问题：最小化人工变量之和，满足所有约束
    - 若最优值非零，原问题不存在可行解
    - 若最优解的基变量不含人工变量，作为原问题的初始基本可行解
    - 若最优解的基变量含有人工变量，可推断存在线性相关的等式约束，将等式删掉后，继续求最优解，直到基变量不含人工变量

避免换出换入变量循环的Bland规则

- 永不终止的单纯形法 [Beal 1955]
  - 规定：当有多个 $\lambda_j < 0$ 时，设 $|\lambda_k| = \max\{|\lambda_j|: \lambda_j < 0\}$，取 x_k 作为换入变量；当有多个 θ_i 同时取到最小值时，取对应的下标最小的基变量作为换出变量。
  - 如下例子，取x_5, x_6, x_7作为初始基变量，计算经过6次基变换回到初始可行基

$$
\begin{aligned}
\min z &= -\frac{3}{4}x_1 + 20x_2 - \frac{1}{2}x_3 + 6x_4 \\
s.t.\ \ & \frac{1}{4}x_1 - 8x_2 - x_3 + 9x_4 + x_5 = 0 \\
& \frac{1}{2}x_1 - 12x_2 - \frac{1}{2}x_3 + 3x_4 + x_6 = 0 \\
& x_3 + x_7 = 1 \\
& x_i \geq 0,\ 1 \leq i \leq 7
\end{aligned}
$$

- Bland规则
  - 规则1：当有多个 $\lambda_i < 0$ 时，取对应的非基变量中下标最小的作为换入变量
  - 规则2：当有多个 $\theta_i = \beta_i/\alpha_{ik}$ （$\alpha_{ik} > 0$) 同时取到最小值时，取对应的基变量中下标最小的作为换出变量

（可用反证法证明Bland规则不会出现循环）

本讲小结

- 线性规划模型
- 应用例子
- 单纯形法

to-be continued...