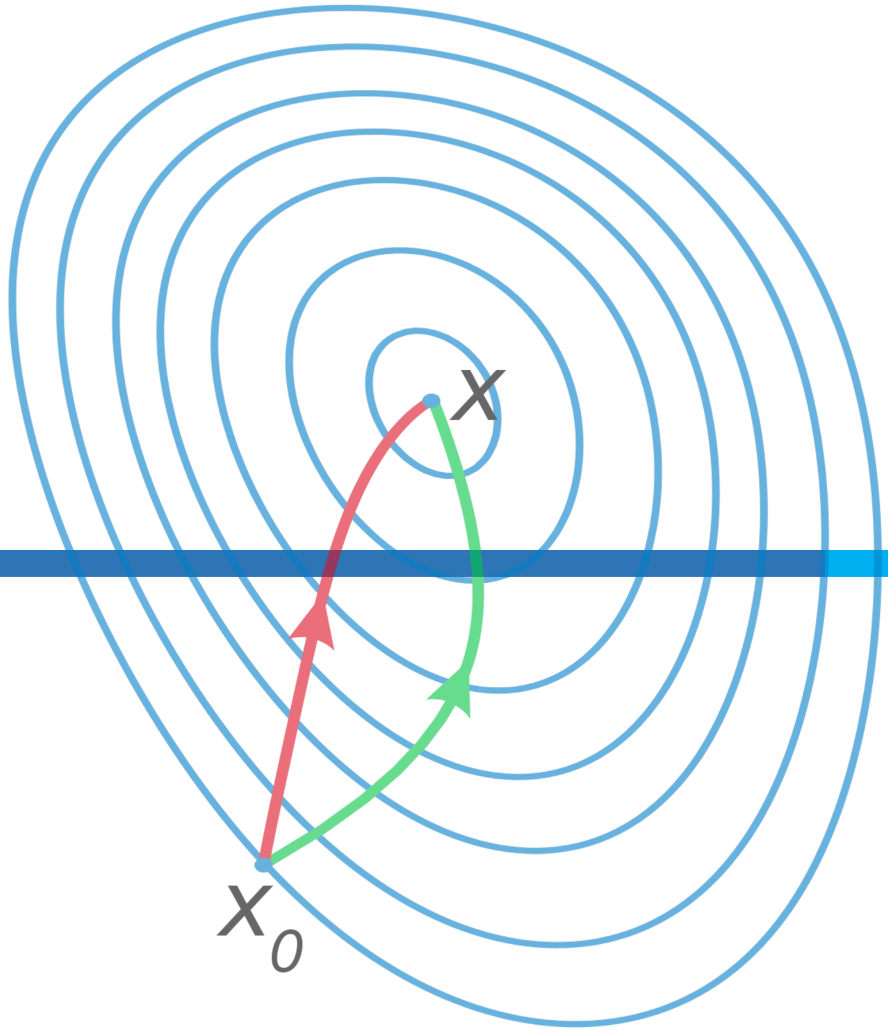
最优化方法



第三周

计算机学院
余皓然

2024/3/11

任意线性规划的求解（单纯形法）

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

例子

$$
\begin{aligned}
\text{min} \quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.} \quad & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.} \quad & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{aligned}
$$

例子修改自Wen Shen <Math 484 - Linear Programming>

例子

$$
\begin{aligned}
\min \quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.} \quad & -6x_1 \qquad + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.} \quad & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{aligned}
$$

第0步 检查是否已整理成标准形式

例子

$$\begin{aligned}
\min \quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.} \quad & -6x_1 \qquad + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 \ = 9, \\
\text{var.} \quad & x_i \geq 0, \forall i = 1,2,\dots,5.
\end{aligned}$$

第1步 找到一个基可行解

由于有5个变量和2个独立等式，基可行解一定有2个变量非零、3个变量为零

注意: 在教材中，非零变量称为基变量、零变量称为非基变量

例子

$$
\begin{array}{ll}
\min & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.} & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
 & 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.} & x_i \geq 0, \forall i = 1,2,\dots,5.
\end{array}
$$

第1步 找到一个基可行解

由于有5个变量和2个独立等式，基可行解一定有2个变量非零、3个变量为零

任意选择2个变量作为非零变量，围绕它们做旋转，整理等式

例如，选择x_2和x_3

$$
\begin{array}{l}
-6x_1 + x_3 - 2x_4 + 2x_5 = 6 \\
-3x_1 + x_2 + 6x_4 + 3x_5 = 15
\end{array}
$$

旋转: 每个等式对应一个非零变量、不包含其它非零变量，在该等式中该非零变量系数为1

本质是便于用其它零变量（x_1, x_4, x_5）分别表示非零变量（x_2和x_3）

例子

$$\begin{array}{ll} \min & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\ \text{s.t.} & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\ & 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\ \text{var.} & x_i \geq 0, \forall i = 1,2,\ldots,5. \end{array}$$

第1步 找到一个基可行解

$$\begin{aligned} -6x_1 \quad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

将零变量取零值，即$x_1 = x_4 = x_5 = 0$，易得$x_2 = 15, x_3 = 6$

由于$x_2 \geq 0, x_3 \geq 0$，这是一个基可行解

如果不是可行解，需重新选取2个变量作为非零变量，做旋转，测试是否是可行解

例子

$$\begin{aligned} \min \quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\ \text{s.t.} \quad & -6x_1 \qquad + x_3 - 2x_4 + 2x_5 = 6, \\ & 3x_1 + x_2 - x_3 + 8x_4 + x_5 \quad = 9, \\ \text{var.} \quad & x_i \geq 0, \forall i = 1,2,\ldots,5. \end{aligned}$$

第2步 计算该基可行解对应的目标函数值

利用

$$\begin{aligned} -6x_1 \qquad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

将目标函数改写成关于零变量的函数

即目标函数等于 $\boxed{5x_1 + 3x_4 - 2x_5 + 21}$

由于在该基可行解下，$x_1 = x_4 = x_5 = 0$，目标函数值为21

例子

$$
\begin{array}{ll}
\min & -4x_1+x_2+x_3+7x_4+3x_5 \\
\text{s.t.} & -6x_1 \qquad\; +x_3-2x_4+2x_5=6, \\
& 3x_1+x_2-x_3+8x_4+x_5 \quad =9, \\
\text{var.} & x_i \geq 0, \forall i=1,2,\ldots,5.
\end{array}
$$

第3步 找到进一步减小目标函数值的相邻基可行解

当前基可行解: 非零变量$\{x_2, x_3\}$ ，零变量$\{x_1, x_4, x_5\}$

与它相邻的基可行解一定与它有一个相同的非零变量

所以找相邻基可行解: 从零变量x_1, x_4, x_5中选取一个把非零变量x_2和x_3中的一个替换掉

问题1: 从零变量x_1, x_4, x_5中选择哪一个加入 ?

问题2: 从非零变量x_2, x_3中选择哪一个去掉 ?

例子

$$
\begin{aligned}
\text{min}\quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.}\quad & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.}\quad & x_i \geq 0, \forall i = 1,2,\dots,5.
\end{aligned}
$$

第3步 找到进一步减小目标函数值的相邻基可行解

当前基可行解: 非零变量$\{x_2, x_3\}$ ，零变量$\{x_1, x_4, x_5\}$

问题1: 从零变量x_1, x_4, x_5中选择哪一个加入？

经整理、仅包含零变量x_1, x_4, x_5的目标函数等于 $\boxed{5x_1 + 3x_4 - 2x_5 + 21}$

因为在当前基可行解下， $x_1 = x_4 = x_5 = 0$，而在可行域内恒有$x_1, x_4, x_5 \geq 0$，

只有当把零变量x_5（在目标函数中系数为负）加入成非零变量才有利于进一步降低目标函数

答案: 选择零变量 x_5加入为非零变量

例子

$$\begin{array}{ll} \min & -4x_1+x_2+x_3+7x_4+3x_5 \\ \text{s.t.} & -6x_1 \qquad\quad +x_3-2x_4+2x_5=6, \\ & \ 3x_1+x_2-x_3+8x_4+x_5 \ \ =9, \\ \text{var.} & x_i \geq 0, \forall i=1,2,\ldots,5. \end{array}$$

第3步 找到进一步减小目标函数值的相邻基可行解

当前基可行解: 非零变量$\{x_2, x_3\}$ ，零变量$\{x_1, x_4, x_5\}$

问题1: 从零变量x_1, x_4, x_5中选择哪一个加入？ 答案: 选择零变量 x_5加入为非零变量

问题2: 从非零变量x_2, x_3中选择哪一个去掉？

现在确定$x_1 = x_4 = 0$（已经将x_5加入非零变量），

当把非零变量x_2变为零变量时： $x_1 = x_2 = x_4 = 0, x_3 = -4, x_5 = 5$，不是可行解

当把非零变量x_3变为零变量时：$x_1 = x_3 = x_4 = 0, x_2 = 6, x_5 = 3$，是可行解

答案: 选择非零变量 x_3变为零变量

例子

$$\begin{array}{ll} \min & -4x_1+x_2+x_3+7x_4+3x_5 \\ \text{s.t.} & -6x_1 \qquad\quad +x_3-2x_4+2x_5=6, \\ & \ \ 3x_1+x_2-x_3+8x_4+x_5 \ \ =9, \\ \text{var.} & \ \ x_i \geq 0, \forall i=1,2,\ldots,5. \end{array}$$

第3步 找到进一步减小目标函数值的相邻基可行解

当前基可行解: 非零变量$\{x_2, x_3\}$ ，零变量$\{x_1, x_4, x_5\}$

问题1: 从零变量x_1, x_4, x_5中选择哪一个加入 ？ 答案: 选择零变量 x_5加入为非零变量

问题2: 从非零变量x_2, x_3中选择哪一个去掉 ？答案: 选择非零变量 x_3变为零变量

新基可行解: 非零变量$\{x_2, x_5\}$ ，零变量$\{x_1, x_3, x_4\}$

原基可行解 $x_1=x_4=x_5=0, x_2=15, x_3=6$，目标函数值21

新基可行解 $x_1=x_3=x_4=0, x_2=6, x_5=3$，目标函数值15，小于21

例子

$$
\begin{aligned}
\min\quad & -4x_1+x_2+x_3+7x_4+3x_5\\
\text{s.t.}\quad & -6x_1 + x_3 - 2x_4 + 2x_5 = 6,\\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9,\\
\text{var.}\quad & x_i \ge 0, \forall i = 1,2,\dots,5.
\end{aligned}
$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

例子

$$\begin{aligned}
\min\quad & -4x_1+x_2+x_3+7x_4+3x_5 \\
\text{s.t.}\quad & -6x_1 \qquad +x_3-2x_4+2x_5=6, \\
& 3x_1+x_2-x_3+8x_4+x_5 \quad =9, \\
\text{var.}\quad & x_i\geq 0,\forall i=1,2,\dots,5.
\end{aligned}$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_2,x_5\}$ ，零变量$\{x_1,x_3,x_4\}$

围绕它们做旋转，整理等式

$$\begin{aligned}
-3x_1 \qquad +\frac{1}{2}x_3-x_4+x_5 &= 3 \\
6x_1+x_2-\frac{3}{2}x_3+9x_4 \qquad &= 6
\end{aligned}$$

例子

$$
\begin{aligned}
\min\quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.}\quad & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.}\quad & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{aligned}
$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_2, x_5\}$ ，零变量$\{x_1, x_3, x_4\}$

围绕它们做旋转，整理等式

$$
\begin{aligned}
-3x_1 + \frac{1}{2}x_3 - x_4 + x_5 &= 3 \\
6x_1 + x_2 - \frac{3}{2}x_3 + 9x_4 &= 6
\end{aligned}
$$

将目标函数改写成关于零变量的函数 $-x_1 + x_3 + x_4 + 15$

例子

$$
\begin{aligned}
\min \quad & -4x_1 + x_2 + x_3 + 7x_4 + 3x_5 \\
\text{s.t.} \quad & -6x_1 \qquad + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.} \quad & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{aligned}
$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_2, x_5\}$ ，零变量$\{x_1, x_3, x_4\}$

根据 $\boxed{-x_1 + x_3 + x_4 + 15}$ 可知应该把x_1零变量转为非零变量

例子

$$\begin{aligned}
\min\quad & -4x_1+x_2+x_3+7x_4+3x_5 \\
\text{s.t.}\quad & -6x_1 \qquad\;\; +x_3-2x_4+2x_5=6, \\
& \;\; 3x_1+x_2-x_3+8x_4+x_5 \;\; =9, \\
\text{var.}\quad & x_i\geq 0, \forall i=1,2,\dots,5.
\end{aligned}$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_2, x_5\}$ ，零变量$\{x_1, x_3, x_4\}$

根据 $\boxed{-x_1+x_3+x_4+15}$ 可知应该把x_1零变量转为非零变量

现在确定$x_3=x_4=0$（已经将x_1加入非零变量），

当把非零变量x_2变为零变量时：$x_2=x_3=x_4=0, x_1=1, x_5=6$，是可行解

当把非零变量x_5变为零变量时：$x_3=x_4=x_5=0, x_1=-1, x_2=12$，不是可行解

新基可行解: 非零变量$\{x_1, x_5\}$ ，零变量$\{x_2, x_3, x_4\}$

例子

$$\begin{array}{ll} \min & -4x_1+x_2+x_3+7x_4+3x_5 \\ \text{s.t.} & -6x_1 \qquad\quad +x_3-2x_4+2x_5=6, \\ & \;\; 3x_1+x_2-x_3+8x_4+x_5 \;\; =9, \\ \text{var.} & x_i \geq 0, \forall i=1,2,\dots,5. \end{array}$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

原基可行解 $x_1 = x_3 = x_4 = 0, x_2 = 6, x_5 = 3$，目标函数值15

新基可行解 $x_2 = x_3 = x_4 = 0, x_1 = 1, x_5 = 6$，目标函数值14，小于15

例子

$$\begin{aligned}\min\quad & -4x_1+x_2+x_3+7x_4+3x_5\\ \text{s.t.}\quad & -6x_1 \qquad\quad +x_3-2x_4+2x_5=6,\\ & 3x_1+x_2-x_3+8x_4+x_5 \quad =9,\\ \text{var.}\quad & x_i\ge 0,\forall i=1,2,\dots,5.\end{aligned}$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_1,x_5\}$ ，零变量$\{x_2,x_3,x_4\}$

围绕它们做旋转，整理等式

$$\begin{aligned} \frac{1}{2}x_2-\frac{1}{4}x_3+\frac{7}{2}x_4+x_5&=6\\ x_1+\frac{1}{6}x_2-\frac{1}{4}x_3+\frac{3}{2}x_4 &=1\end{aligned}$$

将目标函数改写成关于零变量的函数

$$\frac{1}{6}x_2+\frac{3}{4}x_3+\frac{5}{2}x_4+14$$

例子

$$\begin{aligned}
\min\quad & -4x_1+x_2+x_3+7x_4+3x_5 \\
\text{s.t.}\quad & -6x_1 \qquad\quad +x_3-2x_4+2x_5=6, \\
& \ 3x_1+x_2-x_3+8x_4+x_5 \ \ =9, \\
\text{var.}\quad & x_i \ge 0, \forall i=1,2,\ldots,5.
\end{aligned}$$

第4步 重复搜索相邻基可行解的过程，直至找到最优解

基可行解: 非零变量$\{x_1, x_5\}$ ，零变量$\{x_2, x_3, x_4\}$

在目标函数 $\frac{1}{6}x_2+\frac{3}{4}x_3+\frac{5}{2}x_4+14$ 中，所有零变量的系数都为正

说明无法进一步通过调整非零变量/零变量降低目标函数值

当前基可行解 $x_2=x_3=x_4=0, x_1=1, x_5=6$ 就是全局最优解，最优目标函数值14

单纯形法的一般步骤

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、$\boxed{\text{单纯形法、}}$内点法； 有约束凸优化问题、KKT条件
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

单纯形法的一般步骤

单纯形法

0 检查是否已整理成标准形式

1 找到初始基可行解

2 围绕非零变量旋转，改写目标函数

3.1 通过改写后的目标函数，在（函数中系数为负的）当前零变量中选择一个，变为非零变量

3.2 在当前非零变量中选择一个，变为零变量

3.3 围绕非零变量旋转，改写目标函数

3.4 若改写后的目标函数中，所有变量系数非负，结束算法、输出最优解；否则，返回3.1

单纯形法的一般步骤

单纯形法

0 检查是否已整理成标准形式

1 找到初始基可行解

2 围绕非零变量旋转，改写目标函数

3.1 通过改写后的目标函数，在（函数中系数为负的）当前零变量中选择一个，变为非零变量

3.2 在当前非零变量中选择一个，变为零变量

3.3 围绕非零变量旋转，改写目标函数

3.4 若改写后的目标函数中，所有变量系数非负，结束算法、输出最优解；否则，返回3.1

如果有多个系数为负的零变量，该选择哪个？

零变量变为非零变量

如果有多个系数为负的零变量，该选择哪个？

$$
\begin{array}{ll}
\min & -4x_1+x_2+x_3-7x_4+3x_5 \\
\text{s.t.} & -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 = 6, \\
 & \ \ 3x_1 + x_2 - x_3 + 8x_4 + x_5 \ \ = 9, \\
\text{var.} & \ x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{array}
$$

零变量变为非零变量

如果有多个系数为负的零变量，该选择哪个？

$$
\begin{aligned}
\text{min} \quad & -4x_1 + x_2 + x_3 - 7x_4 + 3x_5 \\
\text{s.t.} \quad & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
& 3x_1 + x_2 - x_3 + 8x_4 + x_5 = 9, \\
\text{var.} \quad & x_i \geq 0, \forall i = 1,2,\dots,5.
\end{aligned}
$$

假设当前基可行解是：$x_2 = 15, x_3 = 6, x_1 = x_4 = x_5 = 0$

经旋转，得到

$$
\begin{aligned}
-6x_1 + x_3 - 2x_4 + 2x_5 &= 6 \\
-3x_1 + x_2 + 6x_4 + 3x_5 &= 15
\end{aligned}
$$

改写后的目标函数等于 $5x_1 - 11x_4 - 2x_5 + 21$

有两个零变量在目标函数中的系数为负，把哪一个变成非零变量？

零变量变为非零变量

方法不唯一，偏工程的问题

- （1）根据最大系数绝对值规则
- （2）根据最小化目标函数规则
- （3）根据随机规则
- （4）根据最陡峭边（steepest edge）规则
- （5）根据布兰德规则
- …

本页内容取自Carl Kingsford <CMU-15-451/651: Algorithms>

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （1）根据最大系数绝对值规则

将x_4转为非零变量

零变量变为非零变量

改写后的目标函数

$$5x_1 - 11x_4 - 2x_5 + 21$$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （2）根据最小化目标函数规则（与最大系数绝对值规则不同）

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$
\begin{aligned}
-6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\
-3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15
\end{aligned}
$$

- （2）根据最小化目标函数规则（与最大系数绝对值规则不同）

如果将x_4转为非零变量：$x_1 = x_5 = 0$，由两个等式约束可得

$$
\begin{aligned}
x_3 &= 6 + 2x_4 \\
x_2 &= 15 - 6x_4
\end{aligned}
$$

如果将x_5转为非零变量：$x_1 = x_4 = 0$，由两个等式约束可得

$$
\begin{aligned}
x_3 &= 6 - 2x_5 \\
x_2 &= 15 - 3x_5
\end{aligned}
$$

零变量变为非零变量

改写后的目标函数

$$5x_1 - 11x_4 - 2x_5 + 21$$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \quad\quad\;\; + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad\;\; + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （2）根据最小化目标函数规则（与最大系数绝对值规则不同）

如果将x_4转为非零变量：$x_1 = x_5 = 0$，由两个等式约束可得

$$\begin{aligned} x_3 &= 6 + 2x_4 \geq 0 \Rightarrow x_4 \geq -3 \\ x_2 &= 15 - 6x_4 \geq 0 \Rightarrow x_4 \leq 2.5 \end{aligned}$$

如果将x_5转为非零变量：$x_1 = x_4 = 0$，由两个等式约束可得

$$\begin{aligned} x_3 &= 6 - 2x_5 \geq 0 \Rightarrow x_5 \leq 3 \\ x_2 &= 15 - 3x_5 \geq 0 \Rightarrow x_5 \leq 5 \end{aligned}$$

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$
\begin{aligned}
-6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\
-3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15
\end{aligned}
$$

- （2）根据最小化目标函数规则（与最大系数绝对值规则不同）

如果将x_4转为非零变量：$x_1 = x_5 = 0$，由两个等式约束可得

$$
\begin{aligned}
x_3 &= 6 + 2x_4 \geq 0 \Rightarrow x_4 \geq -3 \\
x_2 &= 15 - 6x_4 \geq 0 \Rightarrow x_4 \leq 2.5
\end{aligned}
$$

说明x_4最多可以由0升至2.5，即相应目标函数降低27.5

所以，选择将x_4转为非零变量

如果将x_5转为非零变量：$x_1 = x_4 = 0$，由两个等式约束可得

$$
\begin{aligned}
x_3 &= 6 - 2x_5 \geq 0 \Rightarrow x_5 \leq 3 \\
x_2 &= 15 - 3x_5 \geq 0 \Rightarrow x_5 \leq 5
\end{aligned}
$$

说明x_5最多可以由0升至3 ，即相应目标函数降低6

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （2）根据最小化目标函数规则（与最大系数绝对值规则不同）

如果将x_4转为非零变量：$x_1 = x_5 = 0$，由两个等式约束可得

$$\begin{aligned} x_3 &= 6 + 2x_4 \geq 0 \Rightarrow x_4 \geq -3 \\ x_2 &= 15 - 6x_4 \geq 0 \Rightarrow x_4 \leq 2.5 \end{aligned}$$

说明x_4最多可以由0升至2.5，即相应目标函数降低27.5

所以，选择将x_4转为非零变量

如果将x_5转为非零变量：$x_1 = x_4 = 0$，由两个等式约束可得

$$\begin{aligned} x_3 &= 6 - 2x_5 \geq 0 \Rightarrow x_5 \leq 3 \\ x_2 &= 15 - 3x_5 \geq 0 \Rightarrow x_5 \leq 5 \end{aligned}$$

说明x_5最多可以由0升至3，即相应目标函数降低6

相比最大系数绝对值，最小化目标函数规则需更多计算

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （3）根据**随机**规则

随机从$\{x_4, x_5\}$中选择一个转为非零变量

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \quad\quad\; + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad\; + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （4）根据**最陡峭边（steepest edge）**规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new}-\boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new}-\boldsymbol{x}_{old}\|}$

$\boldsymbol{x}_{new}$是新基可行解， $\boldsymbol{x}_{old}$是当前基可行解，$\boldsymbol{c}$是当前目标函数中$\boldsymbol{x}$的系数向量

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束
$$\begin{aligned} -6x_1 \quad\quad\;\; + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad\;\; + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （4）根据最陡峭边（steepest edge）规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

$\boldsymbol{x}_{new}$是新基可行解， $\boldsymbol{x}_{old}$是当前基可行解，$\boldsymbol{c}$是当前目标函数中$\boldsymbol{x}$的系数向量

分子化简: $(5,0,0,-11,-2)^T(x_{new,1} - x_{old,1}, x_{new,2} - x_{old,2}, x_{new,3} - x_{old,3}, x_{new,4} - x_{old,4}, x_{new,5} - x_{old,5})$

如果将x_4转为非零变量： $x_1 = x_5 = 0$，即$x_{new,1} - x_{old,1} = 0, x_{new,5} - x_{old,5}=0$

分子简化为$-11(x_{new,4} - x_{old,4}) = -27.5$

分母$\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|$是将$x_4$转为非零变量时新的基可行解和当前基可行解在解空间中的距离

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束
$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

➢ （4）根据最陡峭边（steepest edge）规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new}-\boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new}-\boldsymbol{x}_{old}\|}$

$\boldsymbol{x}_{new}$是新基可行解，$\boldsymbol{x}_{old}$是当前基可行解，$\boldsymbol{c}$是当前目标函数中$\boldsymbol{x}$的系数向量

分子化简: $(5,0,0,-11,-2)^T\left(x_{new,1}-x_{old,1}, x_{new,2}-x_{old,2}, x_{new,3}-x_{old,3}, x_{new,4}-x_{old,4}, x_{new,5}-x_{old,5}\right)$

如果将x_5转为非零变量：$x_1 = x_4 = 0$，即$x_{new,1} - x_{old,1} = 0, x_{new,4} - x_{old,4}=0$

分子简化为$-2\left(x_{new,5} - x_{old,5}\right) = -6$

分母$\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|$是将$x_5$转为非零变量时新的基可行解和当前基可行解在解空间中的距离

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束
$$
\begin{aligned}
-6x_1 \quad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\
-3x_1 + x_2 \quad\quad + 6x_4 + 3x_5 &= 15
\end{aligned}
$$

- （4）根据**最陡峭边（steepest edge）**规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

如果将x_4转为非零变量 $\dfrac{-27.5}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

VS

如果将x_5转为非零变量 $\dfrac{-6}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

注意在上面两个式子中$\boldsymbol{x}_{new}$的取值不同，对应不同的新基可行解

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束
$$\begin{aligned} -6x_1 \quad\quad\;\; + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad\;\; + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （4）根据最陡峭边（steepest edge）规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

如果将x_4转为非零变量 $\dfrac{-27.5}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

VS

如果将x_5转为非零变量 $\dfrac{-6}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

相比于（2）最小化目标函数规则，区别是这里多了分母

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束
$$\begin{aligned} -6x_1 \quad\quad\; + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \quad\quad\; + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （4）根据最陡峭边（steepest edge）规则

最小化 $\dfrac{\boldsymbol{c}^T(\boldsymbol{x}_{new} - \boldsymbol{x}_{old})}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

如果将x_4转为非零变量 $\dfrac{-27.5}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

VS

如果将x_5转为非零变量 $\dfrac{-6}{\|\boldsymbol{x}_{new} - \boldsymbol{x}_{old}\|}$

相比于（2）最小化目标函数规则，区别是这里多了分母

含义：确保在解空间中每移动一个单位，目标函数下降的幅度最大

零变量变为非零变量

改写后的目标函数

$$5x_1 - 11x_4 - 2x_5 + 21$$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （5）根据布兰德规则

按序号选择零变量，即选择x_4 （因为4小于5）

Robert G. Bland

零变量变为非零变量

改写后的目标函数

$$5x_1 - 11x_4 - 2x_5 + 21$$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \qquad\quad + x_3 - 2x_4 + 2x_5 &= 6 \\ -3x_1 + x_2 \qquad\quad + 6x_4 + 3x_5 &= 15 \end{aligned}$$

- （5）根据**布兰德**规则

按序号选择零变量，即选择x_4（因为4小于5）

有什么优势？

零变量变为非零变量

改写后的目标函数 $5x_1 - 11x_4 - 2x_5 + 21$

经旋转后得到的约束

$$\begin{aligned} -6x_1 \quad\quad &+ x_3 - 2x_4 + 2x_5 = 6 \\ -3x_1 + x_2 \quad &\quad\;\, + 6x_4 + 3x_5 = 15 \end{aligned}$$

- （5）根据布兰德规则

按序号选择零变量，即选择x_4 （因为4小于5）

将布兰德规则同时用于零变量变为非零变量、非零变量变为零变量时，可以避免循环（可理论证明），确保单纯形法一定能收敛到最优解

循环

使用单纯形法时，可能遇到循环，即找到的一系列基可行解形成循环 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \dots, \boldsymbol{x}_K, \boldsymbol{x}_1, \boldsymbol{x}_2, \dots$

目标函数一直不下降，算法无法收敛到全局最优解

与找相邻基可行解的规则有关，即将哪个零变量变为非零变量、非零变量变为零变量有关

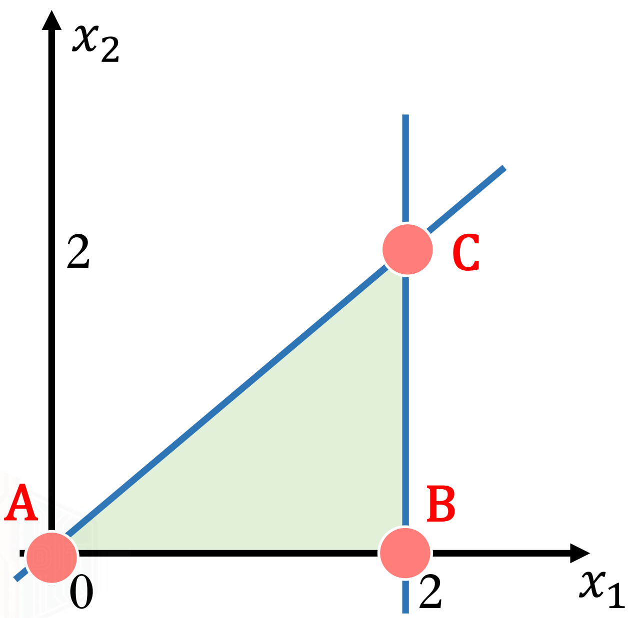
循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2 \le 0, \\ & x_1 \le 2, \\ \text{var.} & x_1 \ge 0, x_2 \ge 0. \end{array}$$

即

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2+x_3 \qquad = 0, \\ & x_1 \qquad\qquad + x_4 = 2, \\ \text{var.} & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{array}$$



可以构造出更复杂的例子，看到循环的出现
上面这个例子并没有出现循环

循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

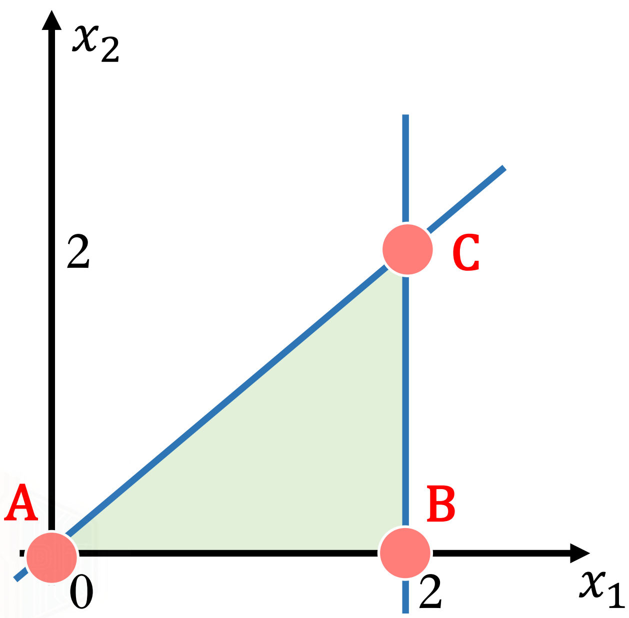
$$\begin{aligned} \min \quad & -x_2 \\ \text{s.t.} \quad & -x_1+x_2 \le 0, \\ & x_1 \le 2, \\ \text{var.} \quad & x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

即

$$\begin{aligned} \min \quad & -x_2 \\ \text{s.t.} \quad & -x_1+x_2+x_3 \qquad\quad = 0, \\ & x_1 \qquad\qquad\quad + x_4 = 2, \\ \text{var.} \quad & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{aligned}$$

当前基可行解，取x_3, x_4作非零变量（基变量）

有$x_1 = x_2 = 0$，可计算得$x_3 = 0, x_4 = 2$，对应图中顶点A



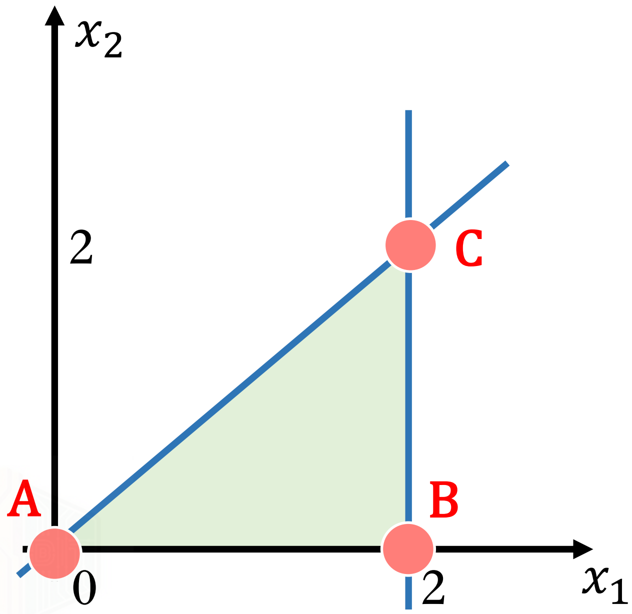
循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2 \leq 0, \\ & x_1 \leq 2, \\ \text{var.} & x_1 \geq 0, x_2 \geq 0. \end{array}$$

即

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2+x_3 \qquad = 0, \\ & x_1 \qquad\qquad + x_4 = 2, \\ \text{var.} & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0. \end{array}$$



当前基可行解，取x_3, x_4作非零变量（基变量）

有$x_1 = x_2 = 0$，可计算得$x_3 = 0, x_4 = 2$，对应图中顶点A

改写后的目标函数 $\boxed{-x_2}$

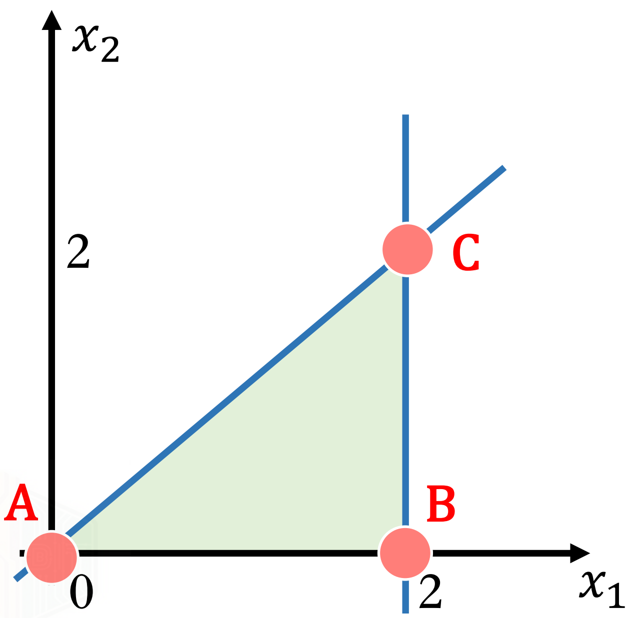
循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

$$\begin{aligned} \min\quad & -x_2 \\ \text{s.t.}\quad & -x_1+x_2 \le 0, \\ & x_1 \le 2, \\ \text{var.}\quad & x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

即

$$\begin{aligned} \min\quad & -x_2 \\ \text{s.t.}\quad & -x_1+x_2+x_3 = 0, \\ & x_1 + x_4 = 2, \\ \text{var.}\quad & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{aligned}$$



当前基可行解，取x_3, x_4作非零变量（基变量）

有$x_1 = x_2 = 0$，可计算得$x_3 = 0, x_4 = 2$，对应图中顶点A

改写后的目标函数 $\boxed{-x_2}$

把零变量x_2变成非零变量（基变量），x_1依旧是零变量

$$\begin{aligned} & x_2 + x_3 = 0, \\ & x_4 = 2, \end{aligned}$$

此时，只能把非零变量x_3变成零变量

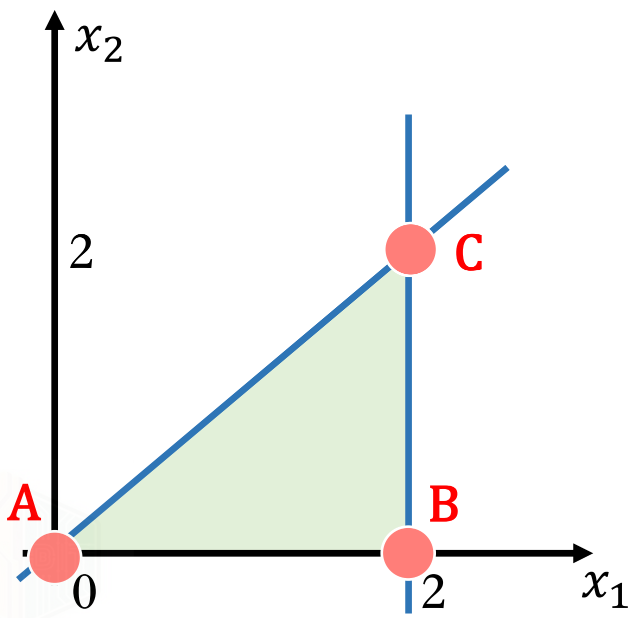
循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2 \le 0, \\ & x_1 \le 2, \\ \text{var.} & x_1 \ge 0, x_2 \ge 0. \end{array}$$

即

$$\begin{array}{ll} \min & -x_2 \\ \text{s.t.} & -x_1+x_2+x_3 = 0, \\ & x_1 + x_4 = 2, \\ \text{var.} & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0. \end{array}$$



当前基可行解，取x_3, x_4作非零变量（基变量）

有$x_1 = x_2 = 0$，可计算得$x_3 = 0, x_4 = 2$，对应图中顶点A

改写后的目标函数 $\boxed{-x_2}$

把零变量x_2变成非零变量（基变量），x_1依旧是零变量

$$\begin{array}{l} x_2 + x_3 = 0, \\ x_4 = 2, \end{array}$$

此时，只能把非零变量x_3变成零变量，得到新的基可行解

有$x_1 = x_3 = 0$，可计算得$x_2 = 0, x_4 = 2$，对应图中顶点A

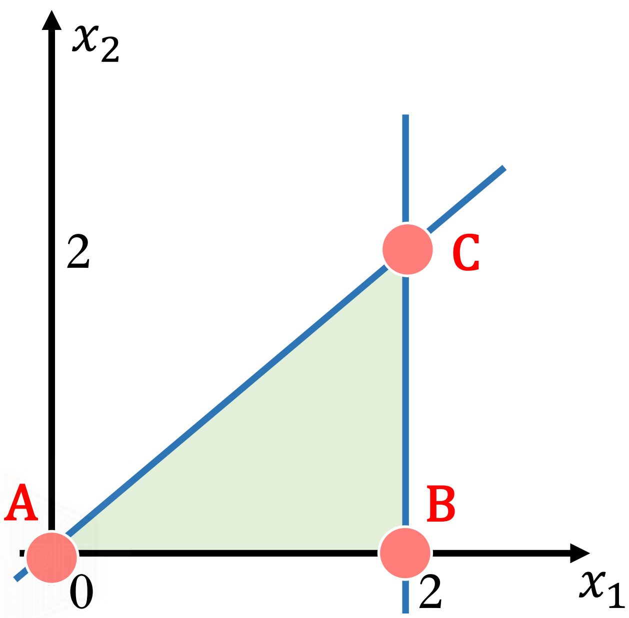
循环

以下例子说明，由一个基可行解到另一个基可行解，目标函数值可能不会下降（即不变）

$$
\begin{array}{ll}
\min & -x_2 \\
\text{s.t.} & -x_1+x_2 \le 0, \\
 & x_1 \le 2, \\
\text{var.} & x_1 \ge 0, x_2 \ge 0.
\end{array}
$$

即

$$
\begin{array}{ll}
\min & -x_2 \\
\text{s.t.} & -x_1+x_2+x_3 \qquad = 0, \\
 & x_1 \qquad\qquad + x_4 = 2, \\
\text{var.} & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0.
\end{array}
$$



当前基可行解

有$x_1 = x_2 = 0$，可计算得$x_3 = 0, x_4 = 2$，对应图中顶点A

新的基可行解

有$x_1 = x_3 = 0$，可计算得$x_2 = 0, x_4 = 2$，对应图中顶点A

两个不同的基可行解，对应图形中相同的顶点，函数值不变

循环

循环：找到的一系列基可行解持续出现类似情况且形成循环 $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \dots, \boldsymbol{x}_K, \boldsymbol{x}_1, \boldsymbol{x}_2, \dots$

- （5）根据**布兰德**规则

按序号选择零变量变为非零变量

将布兰德规则同时用于零变量变为非零变量、非零变量变为零变量时，可以避免循环（可理论证明），确保单纯形法一定能收敛到最优解

零变量变为非零变量

方法不唯一，偏工程的问题

- （1）根据最大系数绝对值规则 计算量小
- （2）根据最小化目标函数规则
- （3）根据随机规则
- （4）根据最陡峭边（steepest edge）规则 实际效果最好
- （5）根据布兰德规则 理论保证
- …

零变量变为非零变量

方法不唯一，偏工程的问题

- （1）根据最大系数绝对值规则　计算量小
- （2）根据最小化目标函数规则
- （3）根据随机规则
- （4）根据最陡峭边（steepest edge）规则　实际效果最好
- （5）根据布兰德规则　理论保证
- …

工程

将一个算法在实际中运用需要解决很多细节问题，细节影响效率

有时候需要深入实际应用场景，才能捕捉到需要解决的关键细节问题

研究

一个经典算法背后有许多研究工作的支撑

很多工作可能只是解决了一个看似非常细小的问题

零变量变为非零变量

Robert G. Bland于1977年发表关于布兰德规则及单纯形法在该规则下收敛性证明的论文

New finite pivoting **rules** for the simplex **method**

RG **Bland** - Mathematics of operations Research, 1977 - pubsonline.informs.org

… A **simple** proof of finiteness is given for the simplex **method** under an easily described pivoting **rule**. A second new finite version of the simplex **method** is also presented. … The …

☆ Save 99 Cite Cited by 769 Related articles All 17 versions

单纯形法的一般步骤

单纯形法

0 检查是否已整理成标准形式

1 找到初始基可行解

2 围绕非零变量旋转，改写目标函数

3.1 通过改写后的目标函数，在（函数中系数为负的）当前零变量中选择一个，变为非零变量

3.2 在当前非零变量中选择一个，变为零变量

3.3 围绕非零变量旋转，改写目标函数

3.4 若改写后的目标函数中，所有变量系数非负，结束算法、输出最优解；否则，返回3.1

该选择哪个非零变量变为零变量？

非零变量变为零变量

该选择哪个非零变量变为零变量？

回顾之前的例子

$$
\begin{aligned}
\text{min} \quad & 5x_1+3x_4-2x_5+21 \\
\text{s.t.} \quad & -6x_1+x_3-2x_4+2x_5=6, \\
& -3x_1+x_2+6x_4+3x_5=15, \\
\text{var.} \quad & x_i \geq 0, \forall i=1,2,\dots,5.
\end{aligned}
$$

当前基可行解为$x_1=x_4=x_5=0, x_2=15, x_3=6$

已知要把x_5变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

非零变量变为零变量

该选择哪个非零变量变为零变量？

回顾之前的例子

$$\begin{array}{ll} \min & 5x_1+3x_4-2x_5+21 \\ \text{s.t.} & -6x_1+x_3-2x_4+2x_5=6, \\ & -3x_1+x_2+6x_4+3x_5=15, \\ \text{var.} & x_i \geq 0, \forall i=1,2,\ldots,5. \end{array}$$

当前基可行解为$x_1=x_4=x_5=0, x_2=15, x_3=6$

已知要把x_5变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_1=x_4=0$（已经将x_5加入非零变量），

当把非零变量x_2变为零变量时：$x_1=x_2=x_4=0, x_3=-4, x_5=5$，不是可行解

当把非零变量x_3变为零变量时：$x_1=x_3=x_4=0, x_2=6, x_5=3$，是可行解

选择原则是确保新的解是可行解，可以不通过穷举所有可能的替换后的解来保证

非零变量变为零变量

该选择哪个非零变量变为零变量？

回顾之前的例子

$$
\begin{array}{ll}
\min & 5x_1+3x_4-2x_5+21 \\
\text{s.t.} & -6x_1+{\color{red}x_3}-2x_4+2x_5=6, \\
 & -3x_1+{\color{red}x_2}+6x_4+3x_5=15, \\
\text{var.} & x_i \geq 0, \forall i=1,2,\dots,5.
\end{array}
$$

当前基可行解为$x_1=x_4=x_5=0, x_2=15, x_3=6$

已知要把x_5变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_1=x_4=0$（已经将x_5加入非零变量），可以将关于x_2, x_3, x_5的两个等式化为

$$
\begin{aligned}
{\color{red}x_3} &= 6-2x_5, \\
{\color{red}x_2} &= 15-3x_5,
\end{aligned}
$$

非零变量变为零变量

该选择哪个非零变量变为零变量？

回顾之前的例子

$$
\begin{array}{ll}
\text{min} & 5x_1+3x_4-2x_5+21 \\
\text{s.t.} & -6x_1+x_3-2x_4+2x_5=6, \\
 & -3x_1+x_2+6x_4+3x_5=15, \\
\text{var.} & x_i \geq 0, \forall i=1,2,\ldots,5.
\end{array}
$$

当前基可行解为$x_1 = x_4 = x_5 = 0, x_2 = 15, x_3 = 6$

已知要把x_5变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_1 = x_4 = 0$（已经将x_5加入非零变量），可以将关于x_2, x_3, x_5的两个等式化为

$$
\begin{array}{l}
x_3 = 6-2x_5 \geq 0, \\
x_2 = 15-3x_5 \geq 0,
\end{array}
\Rightarrow
\begin{array}{l}
x_5 \leq 3, \\
x_5 \leq 5,
\end{array}
$$

令x_5取让不等式都成立的最大值，有$x_5 = 3, x_3=0$

非零变量变为零变量: 特殊情况

该选择哪个非零变量变为零变量？

特殊情况的例子

$$
\begin{array}{ll}
\text{min} & -7x_1 - x_4 + 2x_5 + 9 \\
\text{s.t.} & -6x_1 + {\color{red}x_3} - 2x_4 + 2x_5 = 6, \\
 & -3x_1 + {\color{red}x_2} + 6x_4 + 3x_5 = 15 \\
\text{var.} & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{array}
$$

当前基可行解为$x_1 = x_4 = x_5 = 0, x_2 = 15, x_3 = 6$

已知要把x_1变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

非零变量变为零变量: 特殊情况

该选择哪个非零变量变为零变量？

特殊情况的例子

$$
\begin{array}{ll}
\min & -7x_1 - x_4 + 2x_5 + 9 \\
\text{s.t.} & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
& -3x_1 + x_2 + 6x_4 + 3x_5 = 15 \\
\text{var.} & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{array}
$$

当前基可行解为$x_1 = x_4 = x_5 = 0, x_2 = 15, x_3 = 6$

已知要把x_1变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_4 = x_5 = 0$（已经将x_1加入非零变量），可以将关于x_1, x_2, x_3的两个等式化为

$$
\begin{aligned}
x_3 &= 6 + 6x_1 \geq 0, \\
x_2 &= 15 + 3x_1 \geq 0,
\end{aligned}
$$

非零变量变为零变量: 特殊情况

该选择哪个非零变量变为零变量？

特殊情况的例子

$$
\begin{array}{ll}
\min & -7x_1 - x_4 + 2x_5 + 9 \\
\text{s.t.} & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
 & -3x_1 + x_2 + 6x_4 + 3x_5 = 15 \\
\text{var.} & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{array}
$$

当前基可行解为$x_1 = x_4 = x_5 = 0, x_2 = 15, x_3 = 6$

已知要把x_1变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_4 = x_5 = 0$（已经将x_1加入非零变量），可以将关于x_1, x_2, x_3的两个等式化为

$$
\begin{array}{l}
x_3 = 6 + 6x_1 \geq 0, \\
x_2 = 15 + 3x_1 \geq 0,
\end{array}
$$

令x_1取让不等式都成立的最大值?

非零变量变为零变量: 特殊情况

该选择哪个非零变量变为零变量？

特殊情况的例子

$$
\begin{array}{ll}
\min & -7x_1 - x_4 + 2x_5 + 9 \\
\text{s.t.} & -6x_1 + x_3 - 2x_4 + 2x_5 = 6, \\
 & -3x_1 + x_2 + 6x_4 + 3x_5 = 15 \\
\text{var.} & x_i \geq 0, \forall i = 1,2,\ldots,5.
\end{array}
$$

当前基可行解为$x_1 = x_4 = x_5 = 0, x_2 = 15, x_3 = 6$

已知要把x_1变为非零变量，从当前非零变量x_2, x_3中选择哪一个变为零变量？

现在确定$x_4 = x_5 = 0$（已经将x_1加入非零变量），可以将关于x_1, x_2, x_3的两个等式化为

$$
\begin{array}{l}
x_3 = 6 + 6x_1 \geq 0, \\
x_2 = 15 + 3x_1 \geq 0,
\end{array}
$$

令x_1取让不等式都成立的最大值? $x_1 \to \infty$，

说明最优解不存在，目标函数趋近于-∞

单纯形法的一般步骤

单纯形法

0 检查是否已整理成标准形式

1 找到初始基可行解

2 围绕非零变量旋转，改写目标函数

3.1 通过改写后的目标函数，在（函数中系数为负的）当前零变量中选择一个，变为非零变量

3.2 在当前非零变量中选择一个，变为零变量

3.3 围绕非零变量旋转，改写目标函数

3.4 若改写后的目标函数中，所有变量系数非负，结束算法、输出最优解；否则，返回3.1

单纯形法的一般步骤

单纯形法

0 检查是否已整理成标准形式

1 找到初始基可行解

2 围绕非零变量旋转，改写目标函数

3.1 通过改写后的目标函数，在（函数中系数为负的）当前零变量中选择一个，变为非零变量

3.2 在当前非零变量中选择一个，变为零变量

3.3 围绕非零变量旋转，改写目标函数

3.4 若改写后的目标函数中，所有变量系数非负，结束算法、输出最优解；否则，返回3.1

应用中还有很多细节，比如如何迅速找到一个初始基可行解？

现在方法是在n个变量中随机取m个作为非零变量、其余是零变量，检查是否可行解

这种方法太耗时 ⇒ 大M法（本课程略过）

单纯形法的复杂度？

平均复杂度（average complexity）不高

最差复杂度（worst-case complexity）是指数时间复杂度

Victor Klee and George J. Minty构造出一种例子说明最初的单纯形法以及很多解决线性规划的算法的最差复杂度都很高

[PDF] How **good** is the **simplex algorithm**

V Klee, GJ Minty - Inequalities, 1972 - cgm.cs.mcgill.ca

… convex polytopes, it is shown that the **simplex algorithm** for linear programs (at least with its most commonly used pivot rule) is not a '**good algorithm**' in the sense of J. Edmonds. That is, …

Save Cite Cited by 1518 Related articles All 4 versions

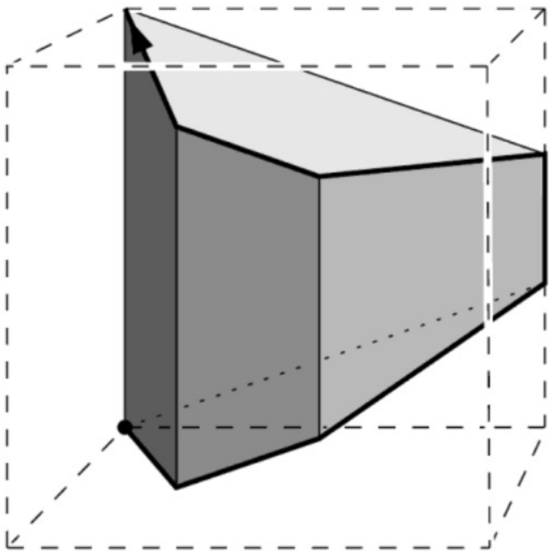
单纯形法的复杂度？

Victor Klee and George J. Minty提出一种结构（Klee – Minty cube），

在该结构下（最初的）单纯形法需要遍历该结构（线性规划问题可行域）所有的顶点

$$\max \quad 2^{D-1}x_1 + 2^{D-2}x_2 + \cdots + 2x_{D-1} + x_D,$$

$$\begin{aligned} x_1 &\le 5 \\ 4x_1 + x_2 &\le 25 \\ 8x_1 + 4x_2 + x_3 &\le 125 \\ &\vdots \\ 2^D x_1 + 2^{D-1}x_2 + \cdots + 4x_{D-1} + x_D &\le 5^D \\ x_1 \ge 0, \ \ldots, \ x_D &\ge 0. \end{aligned}$$



D=3时可行域的示意图

按最初的单纯形法，从零点出发 (即$x_1 = x_2 = x_3 = \cdots = x_D = 0$)

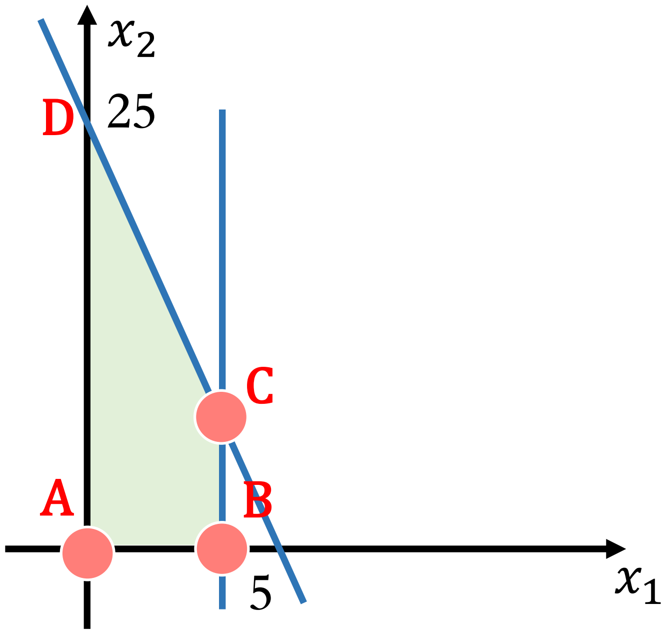
需要遍历所有2^D个顶点，

最后抵达最优解 ($x_1 = x_2 = x_3 = \cdots x_{D-1} = 0, x_D = 2^D$)

公式取自Wikipedia Klee – Minty cube词条

单纯形法的复杂度？

$$\begin{aligned}
\text{min} \quad & -2x_1 - x_2 \\
\text{s.t.} \quad & x_1 + {\color{red}x_3} = 5, \\
& 4x_1 + x_2 + {\color{red}x_4} = 25, \\
\text{var.} \quad & x_i \ge 0, \forall i = 1,2,\dots,4.
\end{aligned}$$



D=2时可以自行计算，顺序是不是 A→B → C → D

单纯形法、线性规划求解器

全球顶级数学难题——“预知未来”的求解器

https://www.bilibili.com/video/BV1AU4y1p74t?spm_id_from=333.337.search-card.all.click

内点法求解线性规划的思路

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； 线性规划问题、单纯形法、内点法； 有约束凸优化问题、KKT条件
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

单纯形法 VS 内点法

- 单纯形法的缺憾：没有利用可行域内部各点对应目标函数值的信息

- 内点法（Interior-Point Method）的思想：能不能利用可行域内部各点的信息?

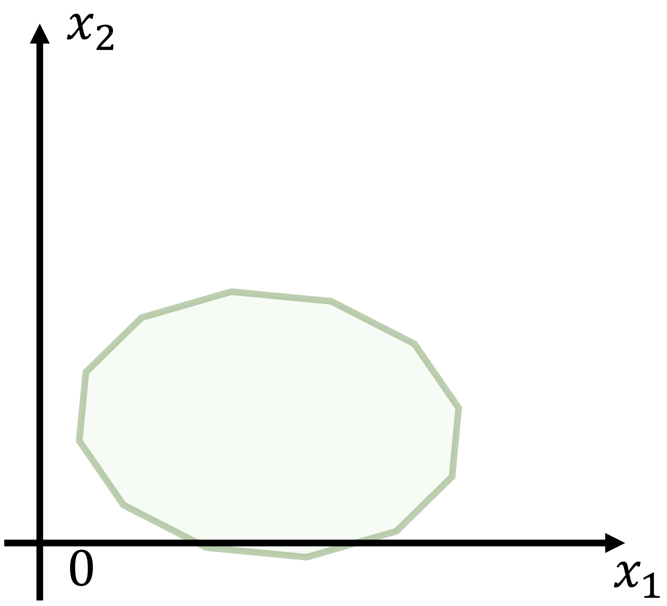
单纯形法 VS 内点法

假设有一个线性规划问题

除了松弛变量，有两个变量x_1, x_2

可行域如x_1, x_2平面上浅色区域所示

目标函数为$c_1x_1 + c_2x_2$



二维视图

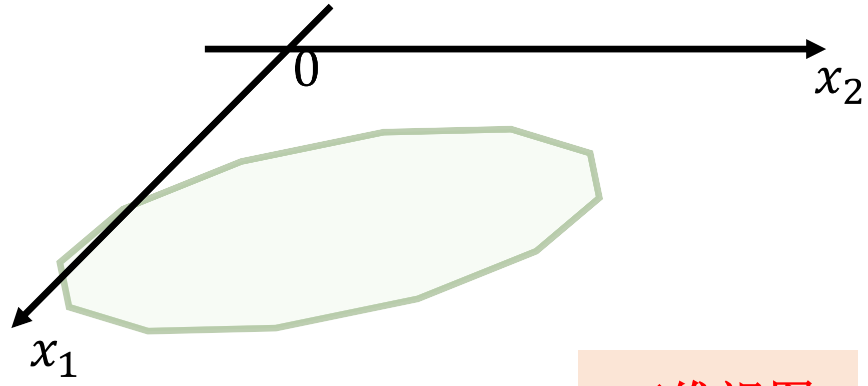
单纯形法 VS 内点法

假设有一个线性规划问题

除了松弛变量，有两个变量x_1, x_2

可行域如x_1, x_2平面上浅色区域所示

目标函数为$c_1x_1 + c_2x_2$



三维视图

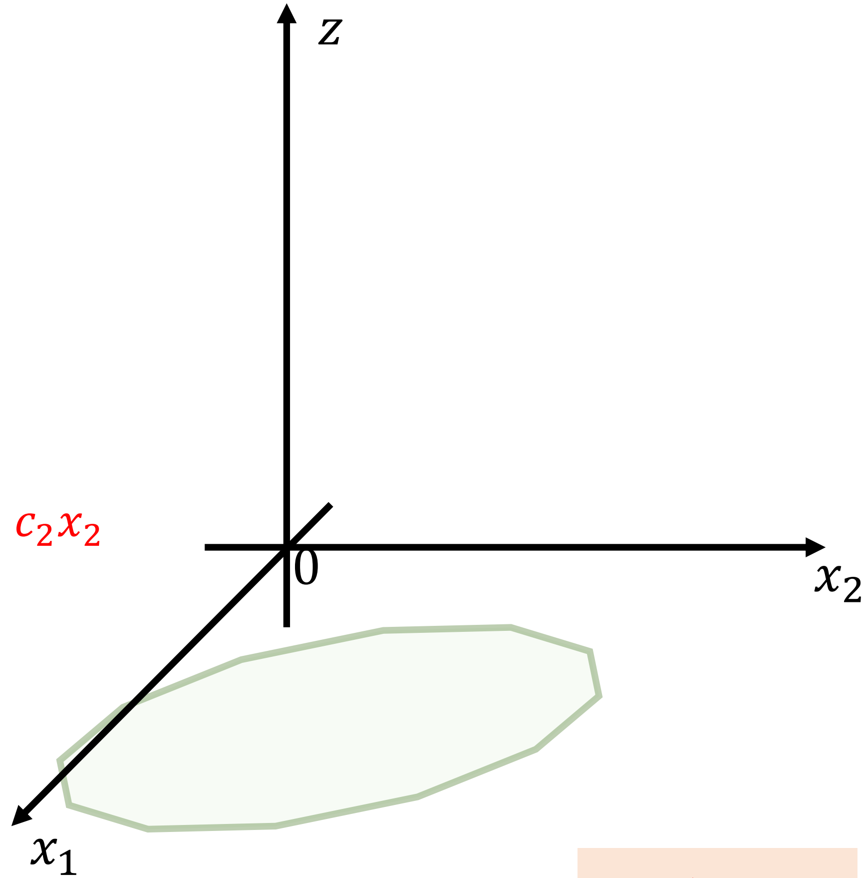
单纯形法 VS 内点法

假设有一个线性规划问题

除了松弛变量，有两个变量x_1, x_2

可行域如x_1, x_2平面上浅色区域所示

目标函数为$c_1x_1 + c_2x_2$，不妨令$z = c_1x_1 + c_2x_2$



三维视图

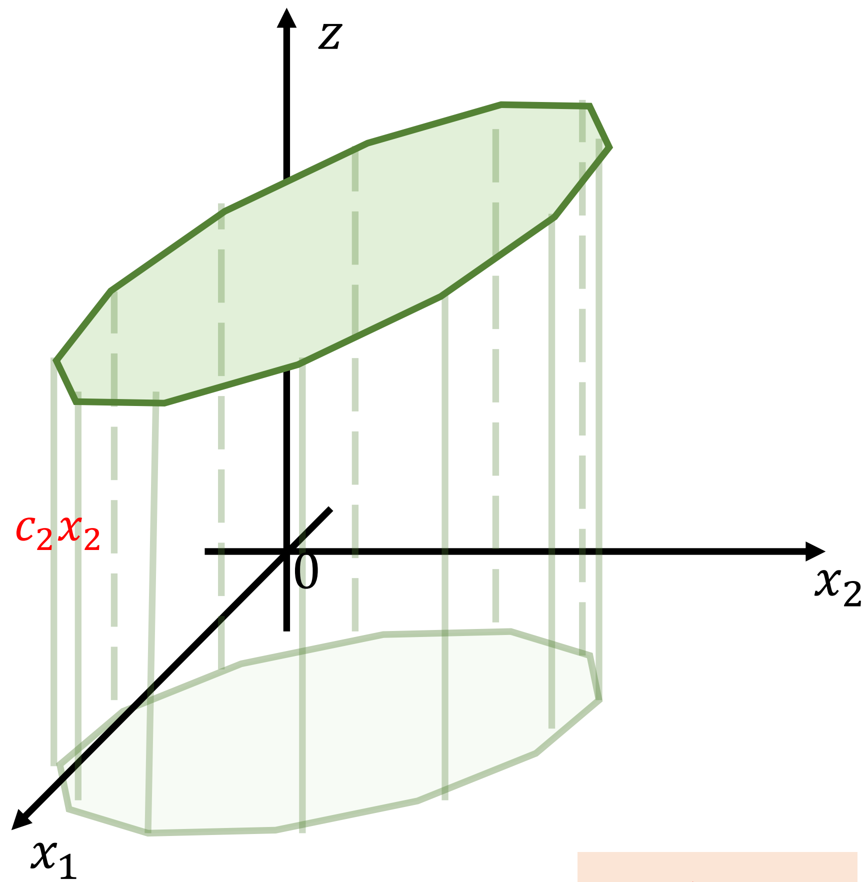
单纯形法 VS 内点法

假设有一个线性规划问题

除了松弛变量，有两个变量x_1, x_2

可行域如x_1, x_2平面上浅色区域所示

目标函数为$c_1x_1 + c_2x_2$，不妨令$z = c_1x_1 + c_2x_2$



三维视图

深绿色平面表示在可行域的各个可行解对应的目标函数值

为什么一定是一个平面？

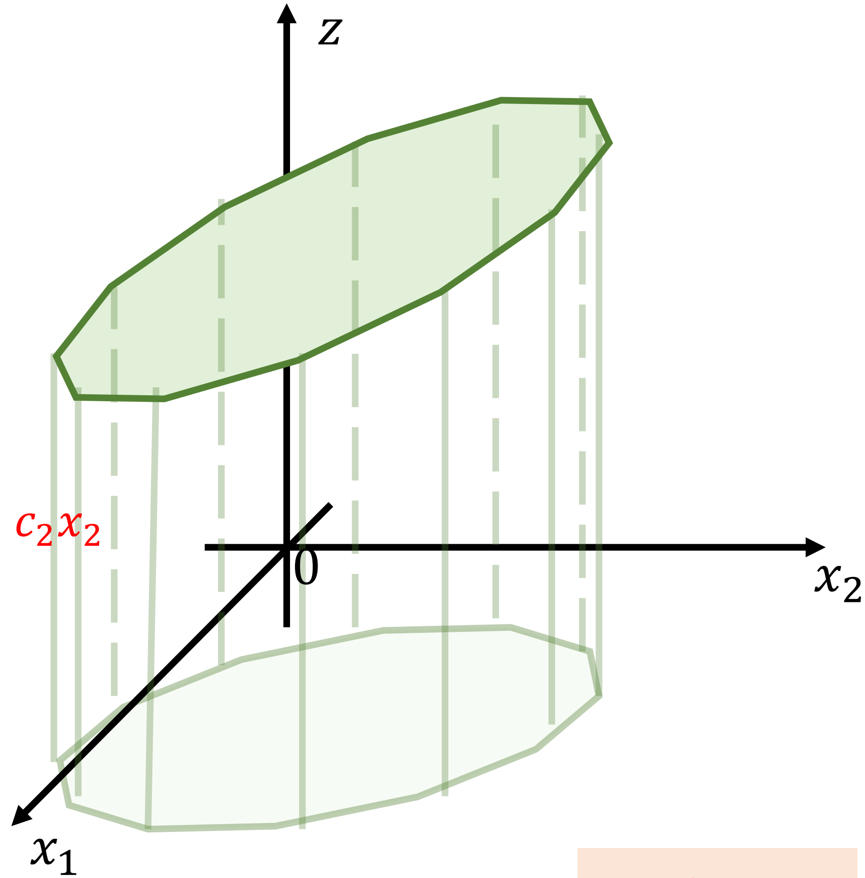
单纯形法 VS 内点法

假设有一个线性规划问题

除了松弛变量，有两个变量x_1, x_2

可行域如x_1, x_2平面上浅色区域所示

目标函数为$c_1x_1 + c_2x_2$，不妨令$z = c_1x_1 + c_2x_2$



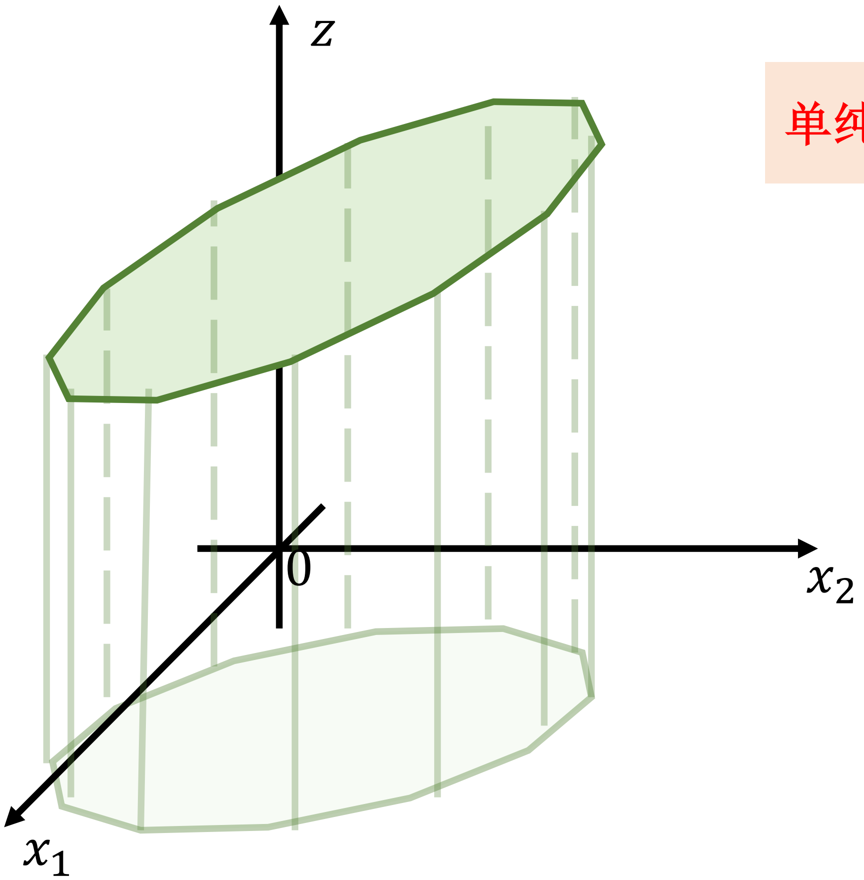
三维视图

深绿色平面表示在可行域的各个可行解对应的目标函数值

为什么一定是一个平面 ?

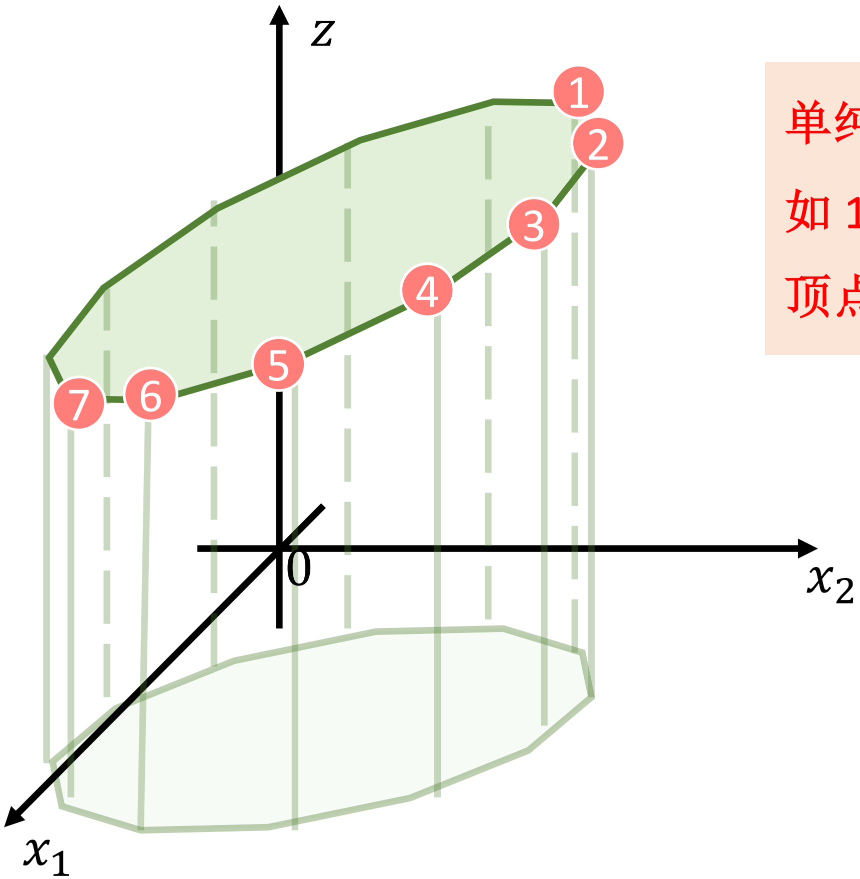
因为$c_1x_1 + c_2x_2 - z = 0$在(x_1, x_2, z)坐标系内就是一个平面的表达式

单纯形法 VS 内点法



单纯形法思路：沿着边按顺序搜索顶点

单纯形法 VS 内点法

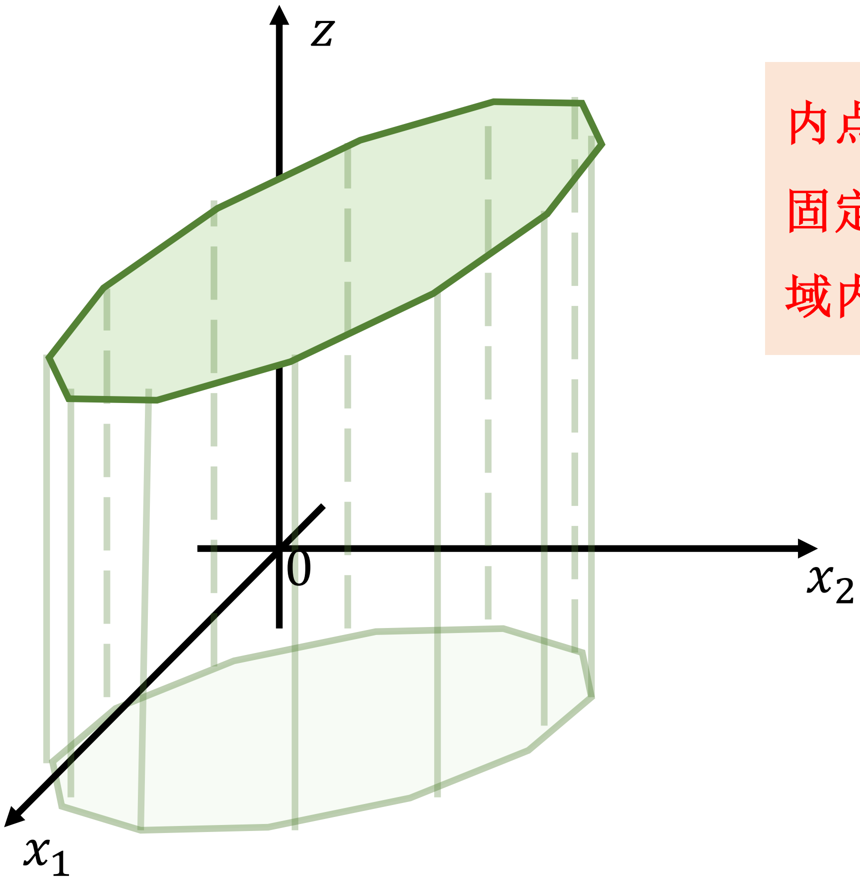


单纯形法思路：沿着边按顺序搜索顶点

如 1 → 2 → 3 → 4 → 5 → 6 → 7

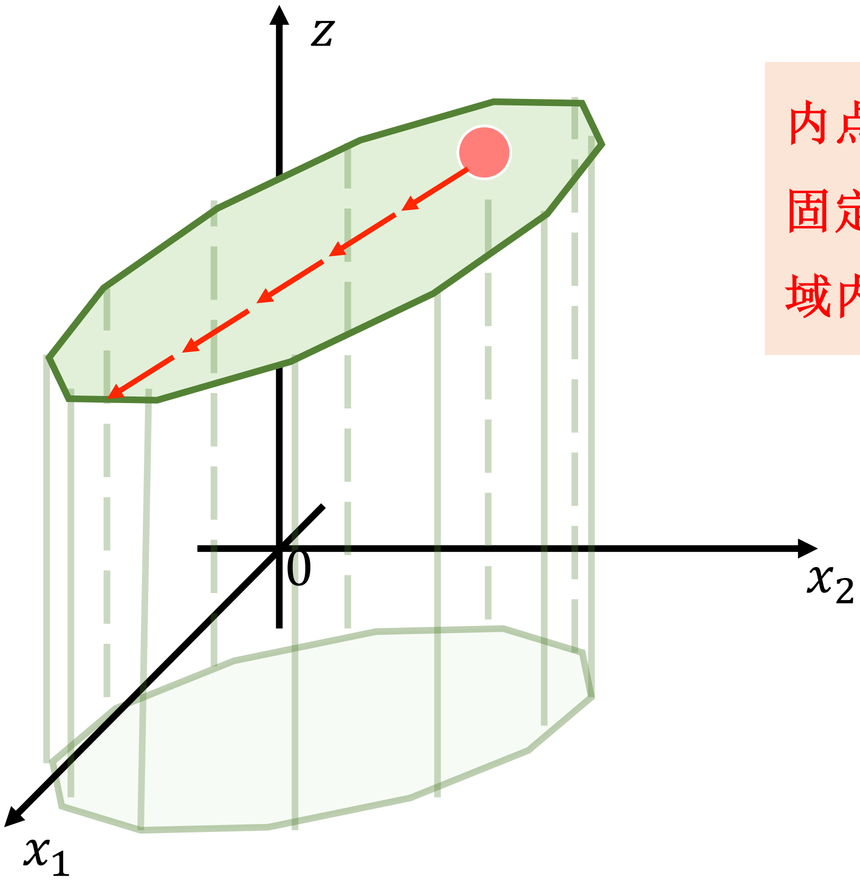
顶点越多计算步数越多

单纯形法 VS 内点法



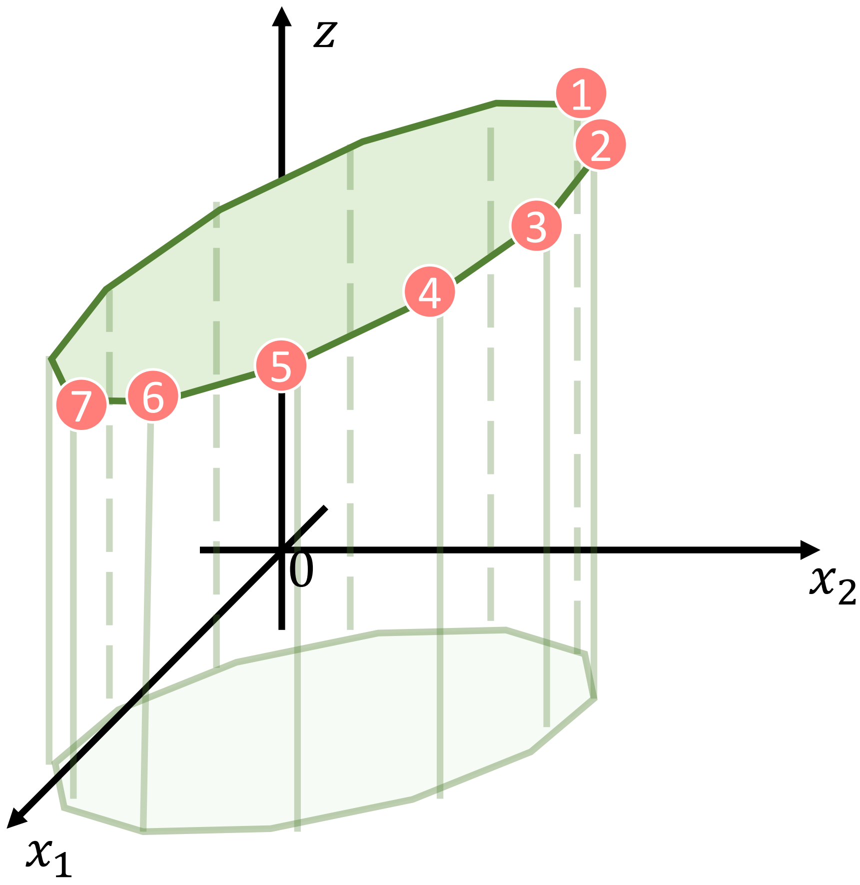
内点法思路：利用深绿色区域各点梯度是固定值、不随x_1, x_2变化的特点，从可行域内部产生路径，沿着平面倾斜方向搜索

单纯形法 VS 内点法

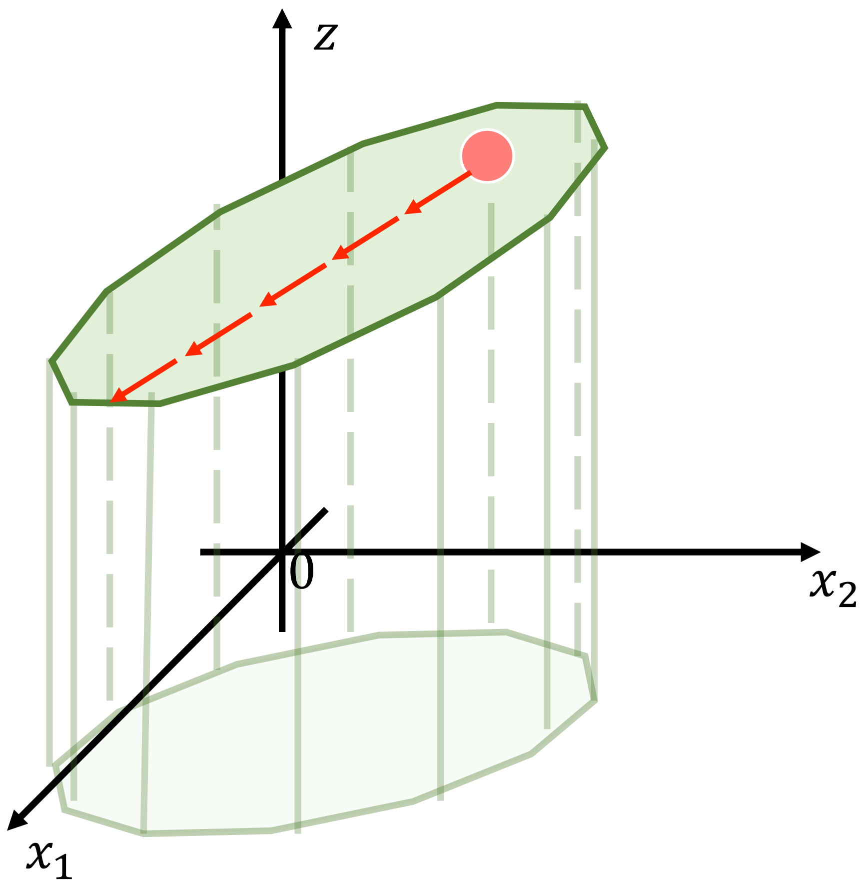


内点法思路：利用深绿色区域各点梯度是固定值、不随x_1, x_2变化的特点，从可行域内部产生路径，沿着平面倾斜方向搜索

单纯形法 VS 内点法



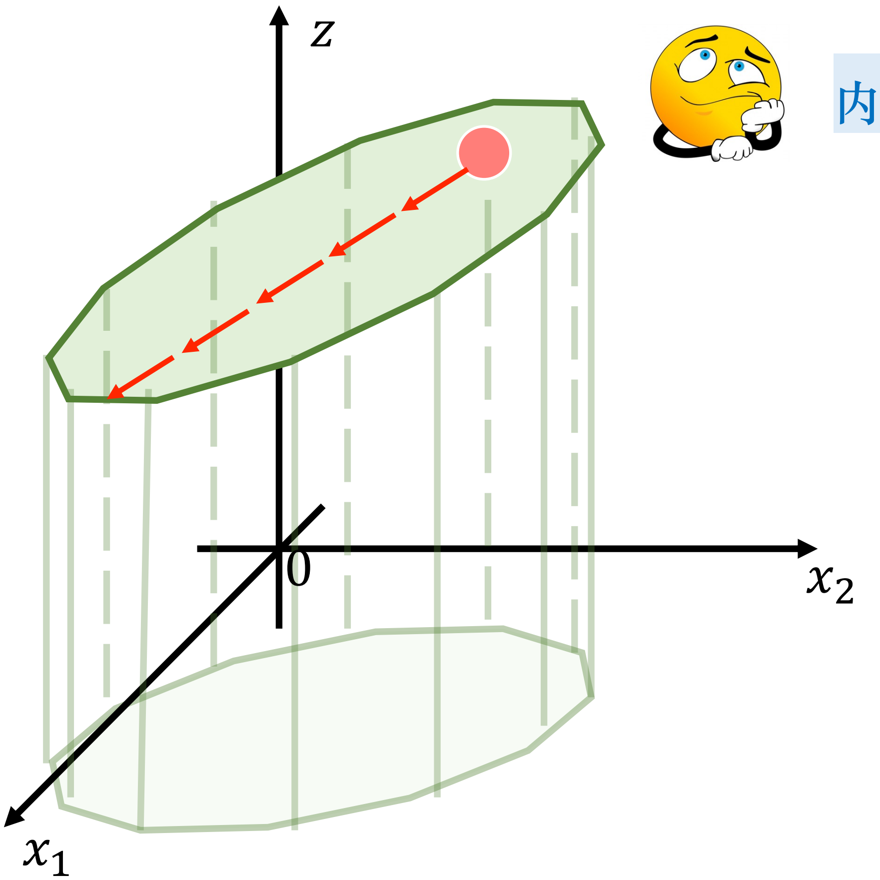
单纯形法



内点法

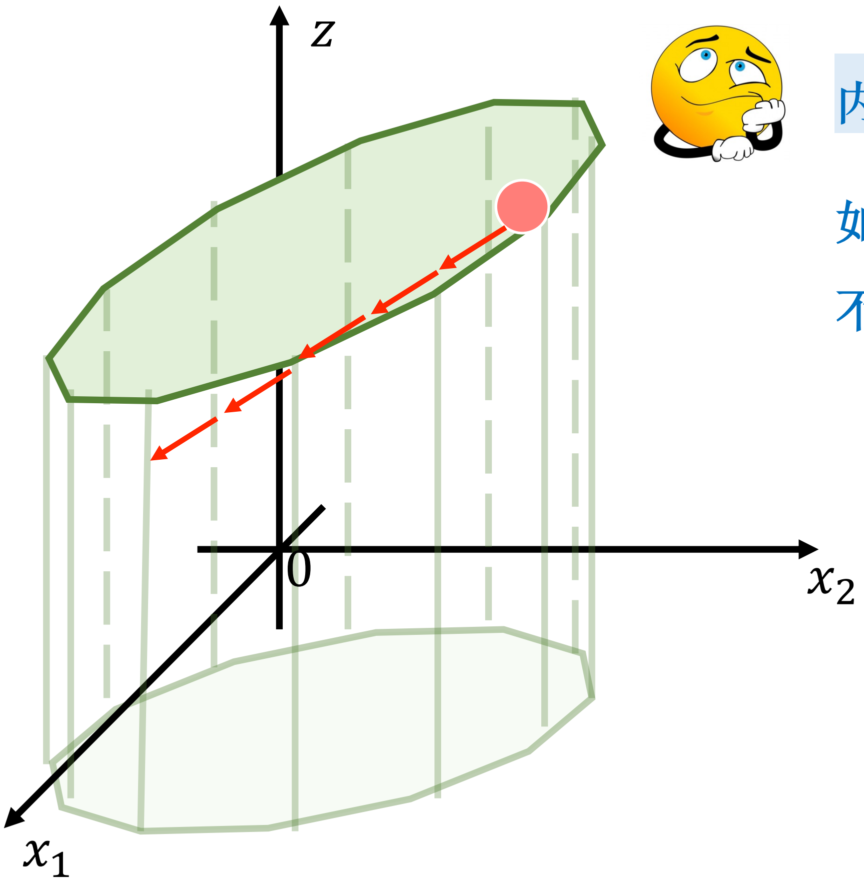
可以利用可行域内部的梯度信息向最优解快速移动

内点法的设计难点



内点法设计的难点在哪里？

内点法的设计难点

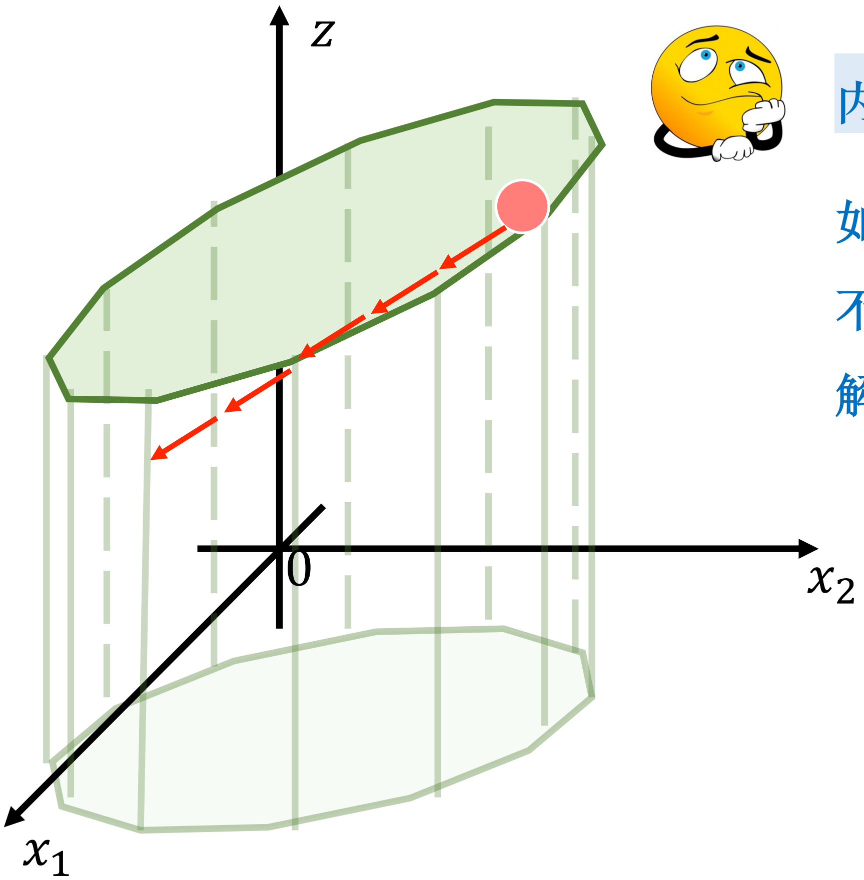


内点法设计的难点在哪里？

如何保障搜索路径一直在可行域内？

不能只看梯度、不看可行域

内点法的设计难点



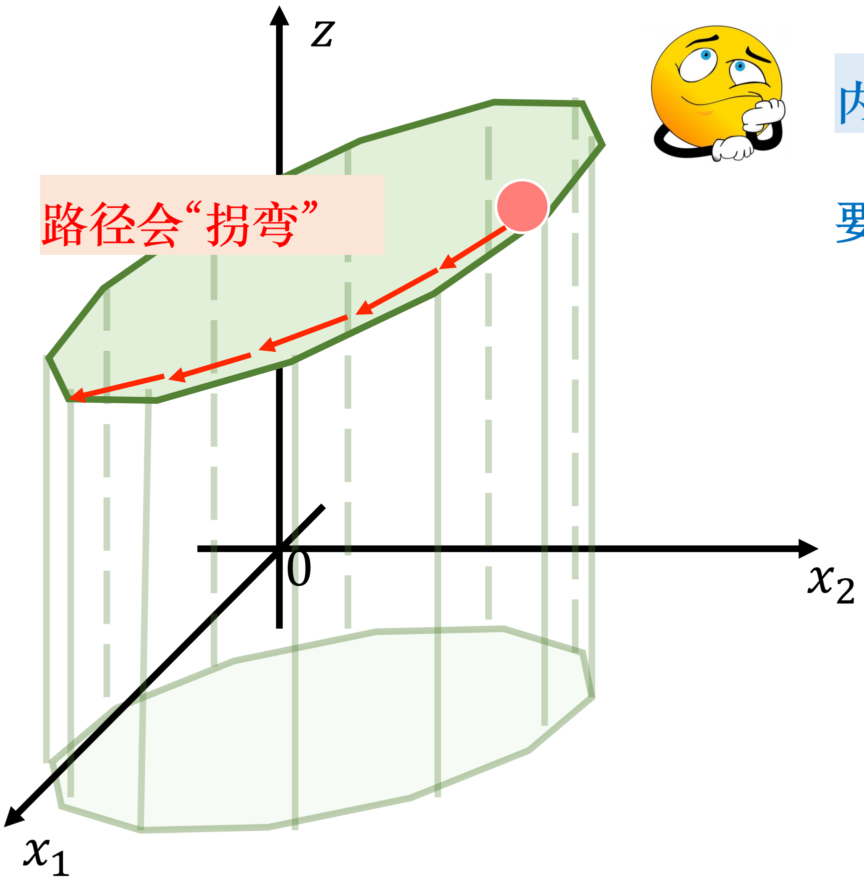
内点法设计的难点在哪里？

如何保障搜索路径一直在可行域内？

不能只看梯度、不看可行域

解决方案？每搜索一步都检查可行性？

内点法的设计难点



内点法设计的难点在哪里？

要保证搜索路径在可行域内且收敛到最优解

一种简单的内点法

Part1 如何从复杂度的角度衡量算法好坏？
Part2 针对一些有特定形式的最优化问题，如何找到最优解？
无约束凸优化问题、梯度下降法、牛顿法； **线性规划问题、单纯形法、内点法；** **有约束凸优化问题、KKT条件**
Part3 针对较复杂的最优化问题（如NP难问题），如何找到较好解？
Part4 针对较复杂的多目标最优化问题，如何找到较好解？

内点法 (Interior-Point Method)

- 内点法是一类方法，包含许多具体的算法
- 大约在20世纪80年代，不同学者相继提出在可行域内生成路径搜索最优解的算法
  - Leonid Khachiyan于1979年提出椭球算法，求解线性规划问题的最差时间复杂度是$O(n^6)$，n为变量个数。第一个多项式时间复杂度求解线性规划的算法

Leonid Khachiyan (1952-2005)
数学家、计算机科学家

内点法

- 内点法是一类方法，包含许多具体的算法

- 大约在20世纪80年代，不同学者相继提出在可行域内生成路径搜索最优解的算法
  - Leonid Khachiyan于1979年提出椭球算法，求解线性规划问题的最差时间复杂度是$O(n^6)$，n为变量个数。第一个多项式时间复杂度求解线性规划的算法

本课介绍一种简化的内点法

- 经典的求解线性规划的内点法
  - 原始牛顿屏障法（Primal Newton Barrier Method）
  - 原始-对偶内点法（Primal-Dual Interior Point Method）

完整介绍这些方法需要一些最优化理论知识基础（对偶问题、KKT条件等）

内点法

- 内点法是一类方法，包含许多具体的算法

- 大约在20世纪80年代，不同学者相继提出在可行域内生成路径搜索最优解的算法
  - Leonid Khachiyan于1979年提出椭球算法，求解线性规划问题的最差时间复杂度是$O(n^6)$，n为变量个数。第一个多项式时间复杂度求解线性规划的算法

- 经典的求解线性规划的内点法
  - 原始牛顿屏障法（Primal Newton Barrier Method）
  - 原始-对偶内点法（Primal-Dual Interior Point Method）

- 内点法可以用于解决多种优化问题（不限于线性规划）

内点法

内点法设计的难点：要保证搜索路径在可行域内且收敛到最优解

内点法

内点法设计的难点：要保证搜索路径在可行域内且收敛到最优解

线性规划问题

$$
\begin{aligned}
\min\quad & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\
\text{s.t.}\quad & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.}\quad & \boldsymbol{x}.
\end{aligned}
$$

为了用单纯形法

线性规划问题的标准形式

$$
\begin{aligned}
\min\quad & \boldsymbol{c}^T\boldsymbol{x} \\
\text{s.t.}\quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.}\quad & \boldsymbol{x} \ge \mathbf{0}.
\end{aligned}
$$

方便找基可行解

内点法

内点法设计的难点：要保证搜索路径在可行域内且收敛到最优解

线性规划问题

$$
\begin{aligned}
\min \quad & \boldsymbol{a}_0^T \boldsymbol{x} + b_0 \\
\text{s.t.} \quad & \boldsymbol{a}_i^T \boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
& \boldsymbol{d}_j^T \boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} \quad & \boldsymbol{x}.
\end{aligned}
$$

为了用单纯形法

线性规划问题的标准形式

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} = \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x} \ge \mathbf{0}.
\end{aligned}
$$

方便找基可行解

为了用内点法

线性规划问题

$$
\begin{aligned}
\min \quad & \boldsymbol{c}^T \boldsymbol{x} \\
\text{s.t.} \quad & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\
\text{var.} \quad & \boldsymbol{x}.
\end{aligned}
$$

方便看是否在可行域内

内点法

内点法设计的难点：要保证搜索路径在可行域内且收敛到最优解

$$
\begin{array}{ll}
\min & \boldsymbol{a}_0^T\boldsymbol{x} + b_0 \\
\text{s.t.} & \boldsymbol{a}_i^T\boldsymbol{x} + b_i \le 0, i = 1, \dots, m, \\
 & \boldsymbol{d}_j^T\boldsymbol{x} + c_j = 0, j = 1, \dots, p, \\
\text{var.} & \boldsymbol{x}.
\end{array}
\longrightarrow
\begin{array}{ll}
\min & \boldsymbol{c}^T\boldsymbol{x} \\
\text{s.t.} & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \\
\text{var.} & \boldsymbol{x}.
\end{array}
$$

注意：决策变量和参数等都已经重新定义

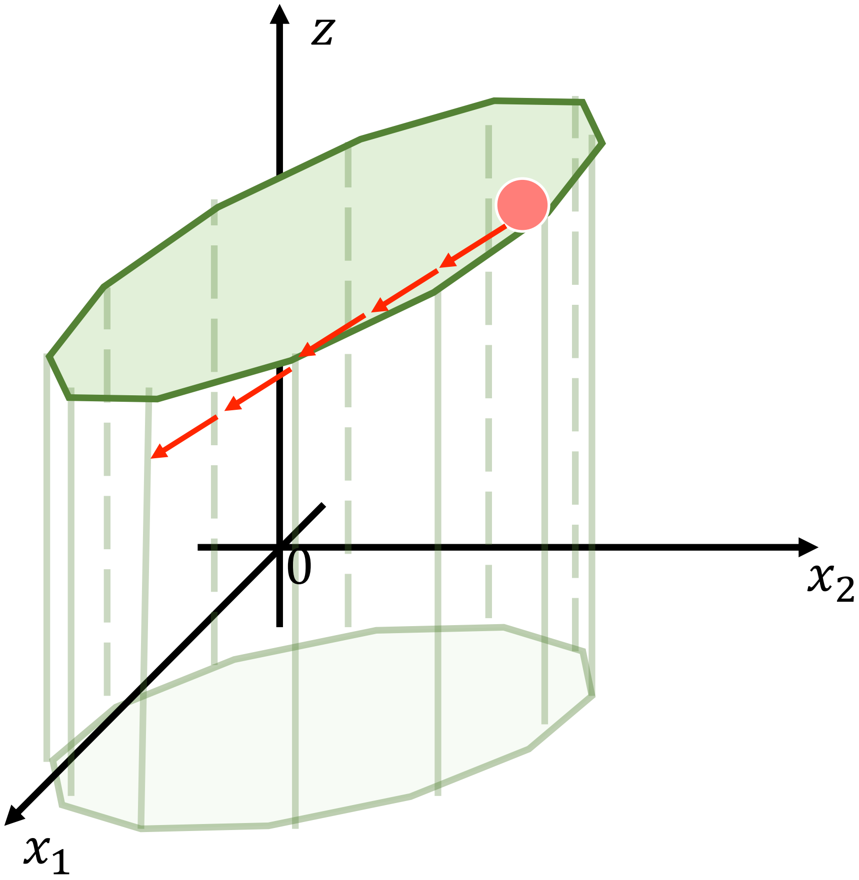
为什么一定能改写？
怎么把等式约束改写成不等式约束？

内点法

内点法设计的难点：要保证搜索路径在可行域内且收敛到最优解

$$\begin{array}{ll} \min & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array}$$

如何避免出现右图中“出界”情况
把不等式约束考虑到搜索路径中



障碍函数

$$\begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^{m} I\left(\sum_{j=1}^{n} A_{ij}x_j - b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

不妨设系数矩阵$\boldsymbol{A}$有m行n列，即m个不等式约束、n个变量

引入了障碍函数 $I(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$

本部分参考自李恩志 <凸优化算法I: 内点法求解线性规划问题>

Mung Chiang < Princeton ELE539A-Lecture 18: Interior Point Algorithm >

障碍函数

$$\begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^{m} I\left(\sum_{j=1}^{n} A_{ij}x_j - b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

不妨设系数矩阵$\boldsymbol{A}$有m行n列，即m个不等式约束、n个变量

引入了障碍函数 $I(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$

该障碍函数的问题：导致新的目标函数不是处处可微

对于不可微的函数一般难以搜索最小值

例如深度学习中损失函数的构造，一般情况需要对网络参数可微（即存在梯度）

障碍函数

$$\begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^{m} I\left(\sum_{j=1}^{n} A_{ij}x_j - b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

不妨设系数矩阵$\boldsymbol{A}$有m行n列，即m个不等式约束、n个变量

引入了障碍函数 $I(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$

该障碍函数的问题：导致新的目标函数不是处处可微

解决方案：换一个可微的、同时能近似 $I(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$ 效果的新障碍函数

障碍函数

$$\begin{aligned} \min \quad & \boldsymbol{c}^T \boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x} \leq \boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \min \quad & \boldsymbol{c}^T \boldsymbol{x} + \textstyle\sum_{i=1}^{m} I_t\left(\sum_{j=1}^{n} A_{ij} x_j - b_i\right) \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

不妨设系数矩阵$\boldsymbol{A}$有m行n列，即m个不等式约束、n个变量

引入了障碍函数 $I(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases}$

引入了新障碍函数 $I_t(u) = -\dfrac{1}{t}\log(-u)$

$t > 0$控制新障碍函数的形状

可以试手画一下函数的形状

障碍函数

$$\begin{aligned} \min \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned} \quad \Leftrightarrow \quad \begin{aligned} \min \quad & \boldsymbol{c}^T\boldsymbol{x} + \textstyle\sum_{i=1}^m I_t\left(\sum_{j=1}^n A_{ij}x_j - b_i\right) \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

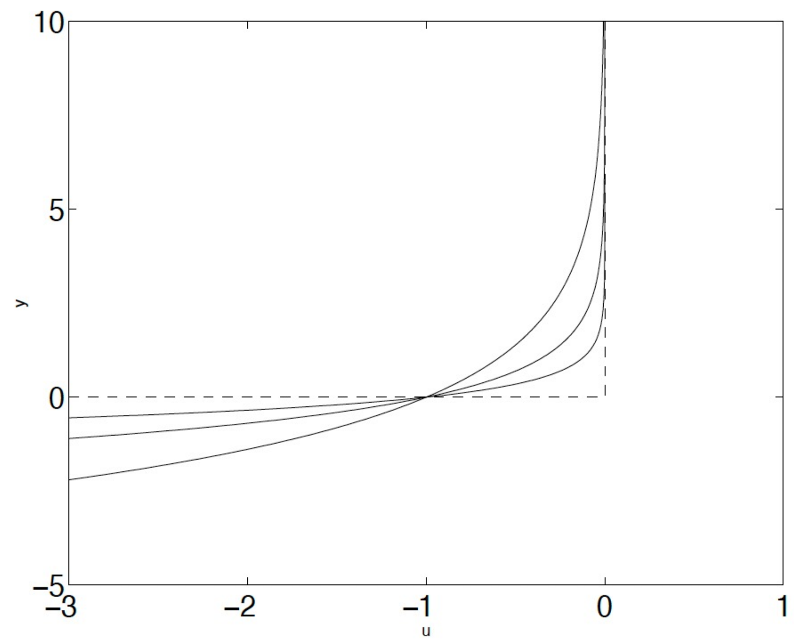
不妨设系数矩阵$\boldsymbol{A}$有m行n列，即m个不等式约束、n个变量

引入了障碍函数 $I(u) = \begin{cases} 0 & u \le 0 \\ \infty & u > 0 \end{cases}$

引入了新障碍函数 $I_t(u) = -\dfrac{1}{t}\log(-u)$

$t > 0$控制新障碍函数的形状

t越大，形状越逼近原障碍函数



障碍函数

$$\begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array} \quad \Leftrightarrow \quad \begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} + \sum_{i=1}^{m} I_t\left(\sum_{j=1}^{n} A_{ij}x_j - b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

引入了新障碍函数 $I_t(u) = -\frac{1}{t}\log(-u)$

$$\begin{array}{ll} \min & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

障碍函数

$$\begin{array}{ll} \min & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} & \mathbf{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} & \boldsymbol{x}. \end{array}$$

$$\begin{array}{ll} \min & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

$t > 0$是控制变量，对每一个t都可以求到一个$\boldsymbol{x}^*(t)$

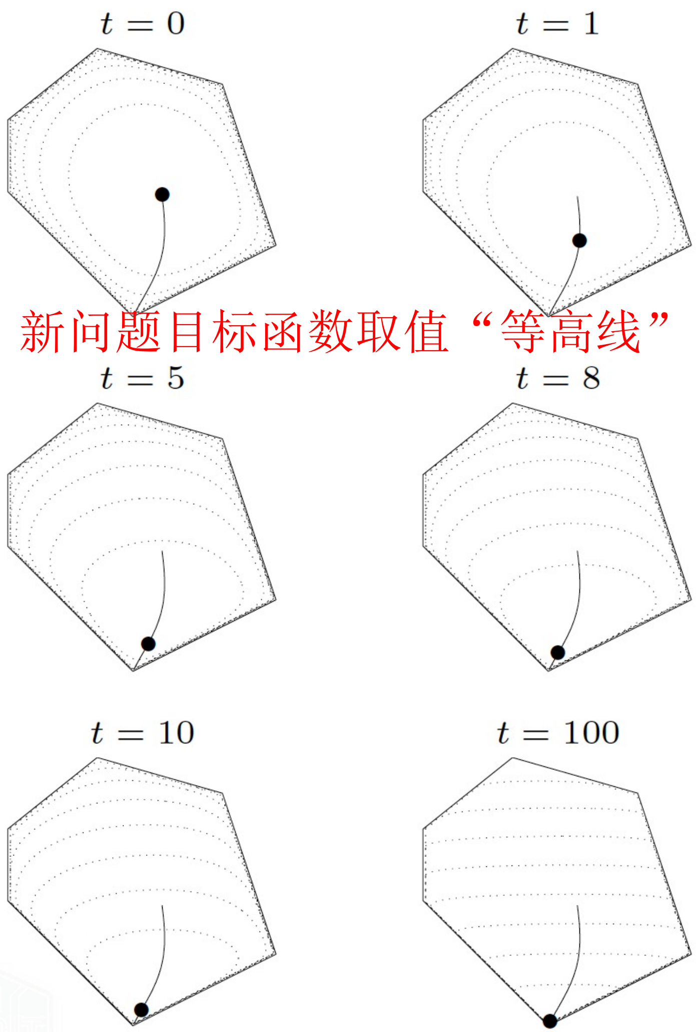
随着$t \to \infty$, 对数障碍函数将逼近于原不可微的障碍函数，

$\boldsymbol{x}^*(t)$趋近于原线性规划问题的全局最优解

$$\begin{array}{ll}\min & \color{red}{t}\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m}\log\left(-\sum_{j=1}^{n}A_{ij}x_j + b_i\right)\\ \text{var.} & \boldsymbol{x}.\end{array}$$

障碍函数

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解

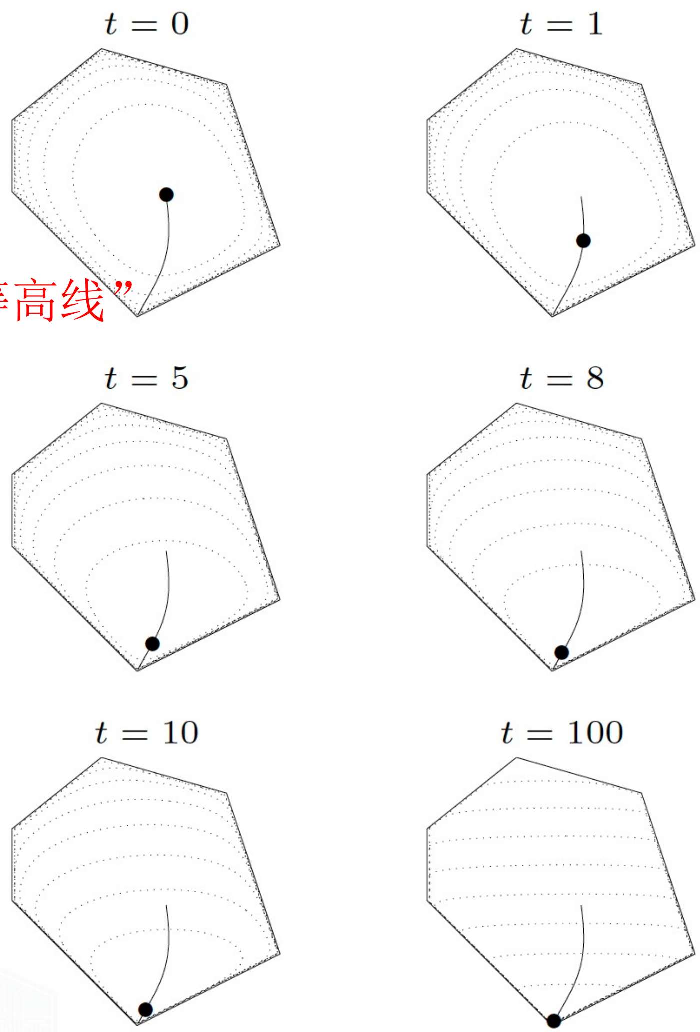


图片取自Stephen Boyd <Stanford-EE364 Sequential unconstrained minimization>

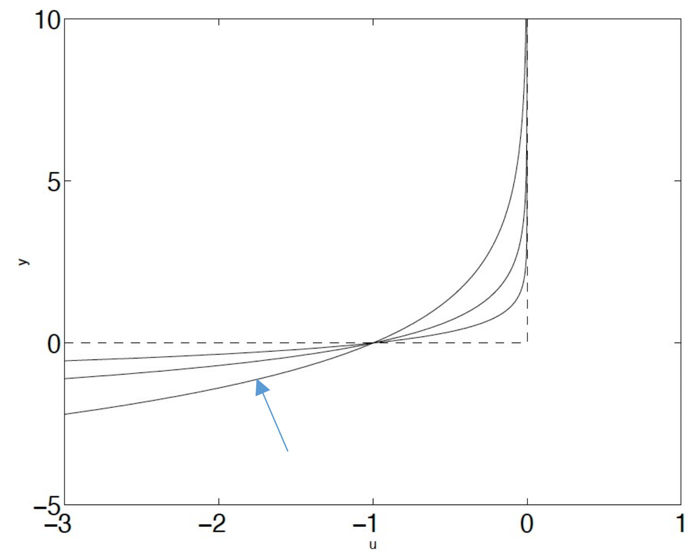
障碍函数

$$\begin{array}{ll} \min & {\color{red}t}\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



$t = 0$时，可行域的中心地带（离所有边界都远）取得最小值，即$\boldsymbol{x}^*(t)$在中心地带



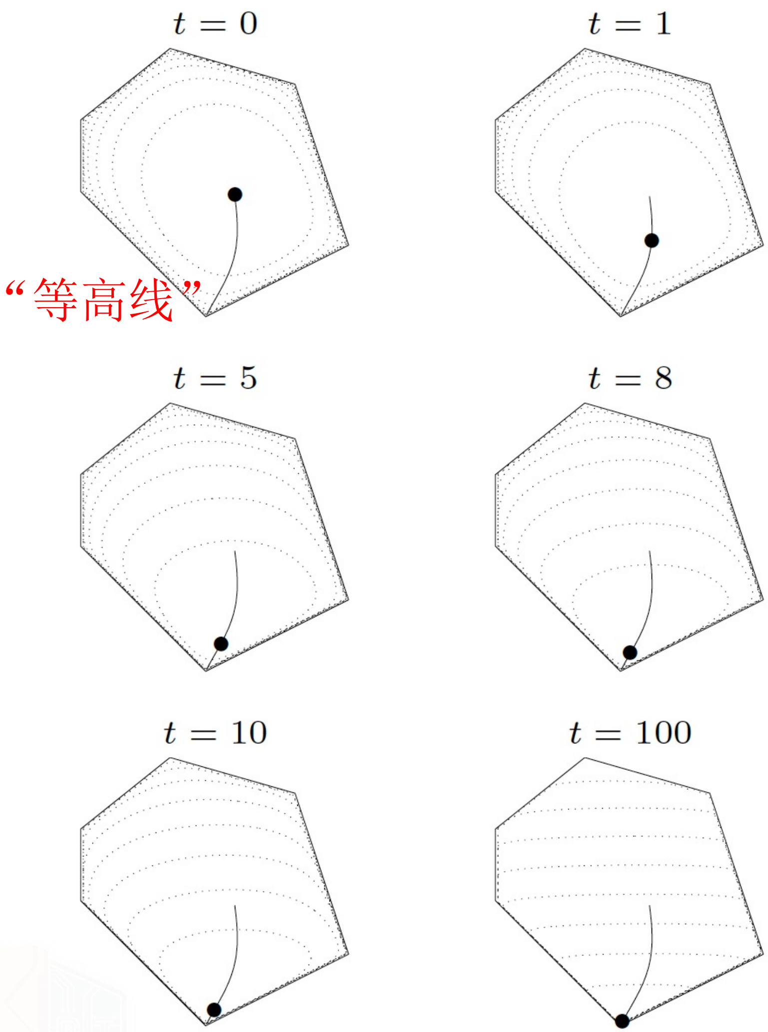
$$u = A_{ij}x_j - b_i$$

图片取自Stephen Boyd < Stanford-EE364 Sequential unconstrained minimization>

$$
\begin{aligned}
\min\quad & t\boldsymbol{c}^T\boldsymbol{x} - \textstyle\sum_{i=1}^m \log\left(-\sum_{j=1}^n A_{ij}x_j + b_i\right)\\
\text{var.}\quad & \boldsymbol{x}.
\end{aligned}
$$

障碍函数

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解

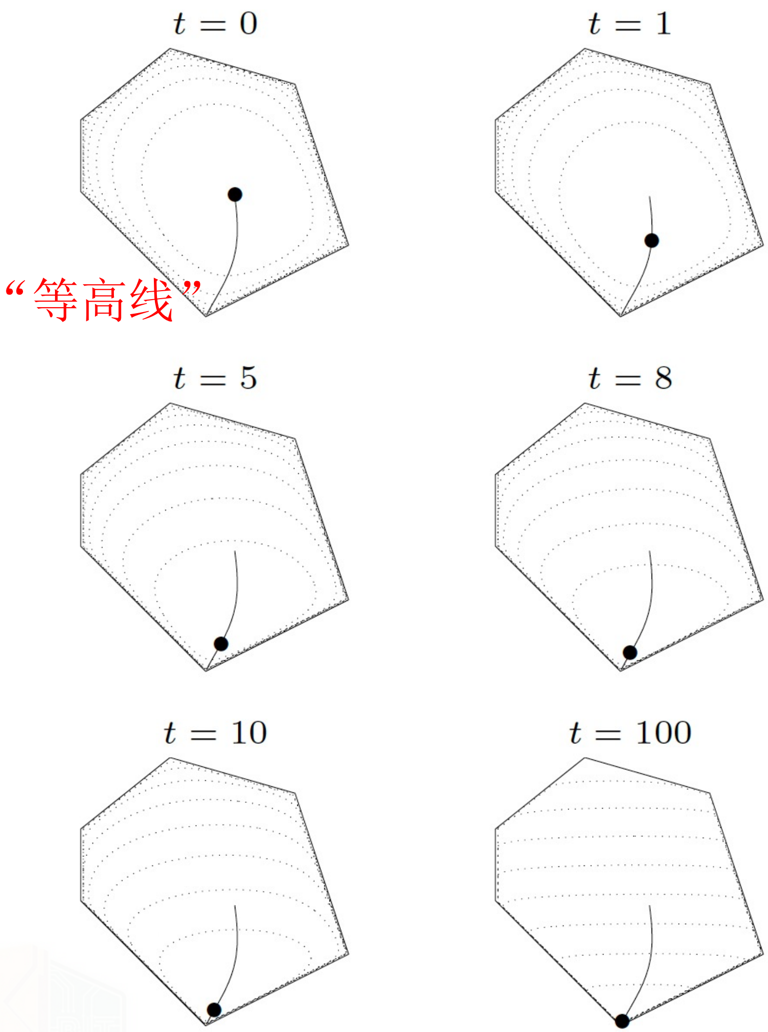


$t > 0$时，原目标函数$\boldsymbol{c}^T\boldsymbol{x}$的权重增加，$\boldsymbol{x}^*(t)$向着原线性规划问题的最优解方向移动

障碍函数

$$\begin{array}{ll} \min & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



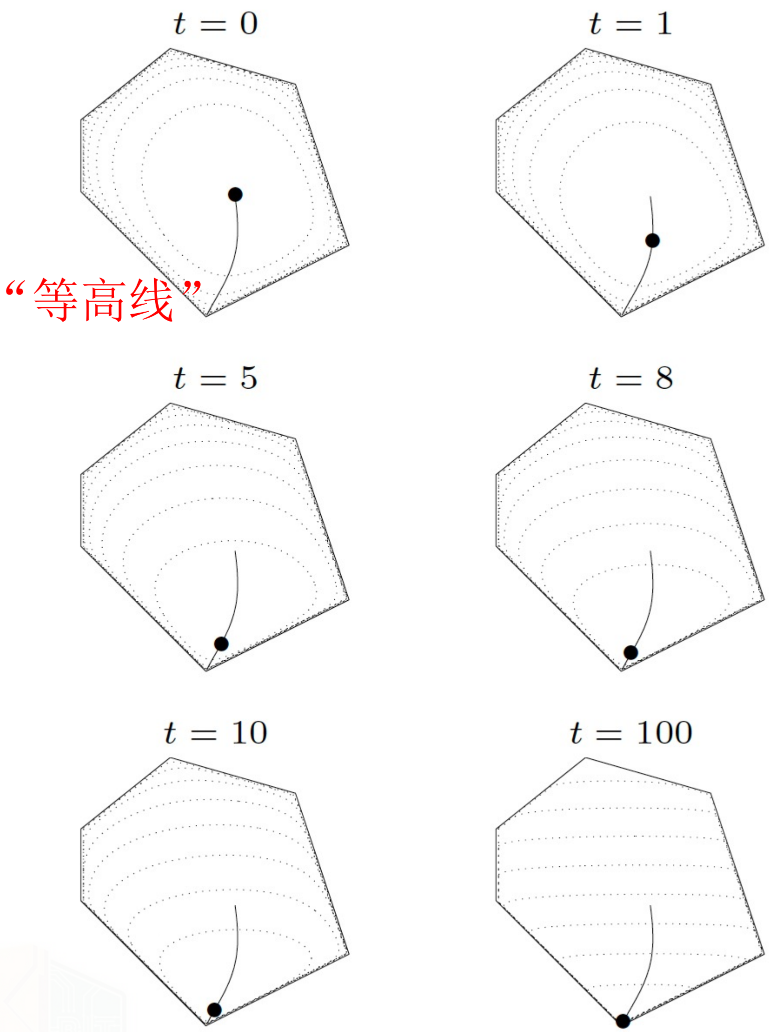
t很大时，$\boldsymbol{x}^*(t)$趋近于原问题最优解

注意等高线的变化

障碍函数

$$\begin{array}{ll} \min & \color{red}{t}\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



（本讲介绍的）内点法思路

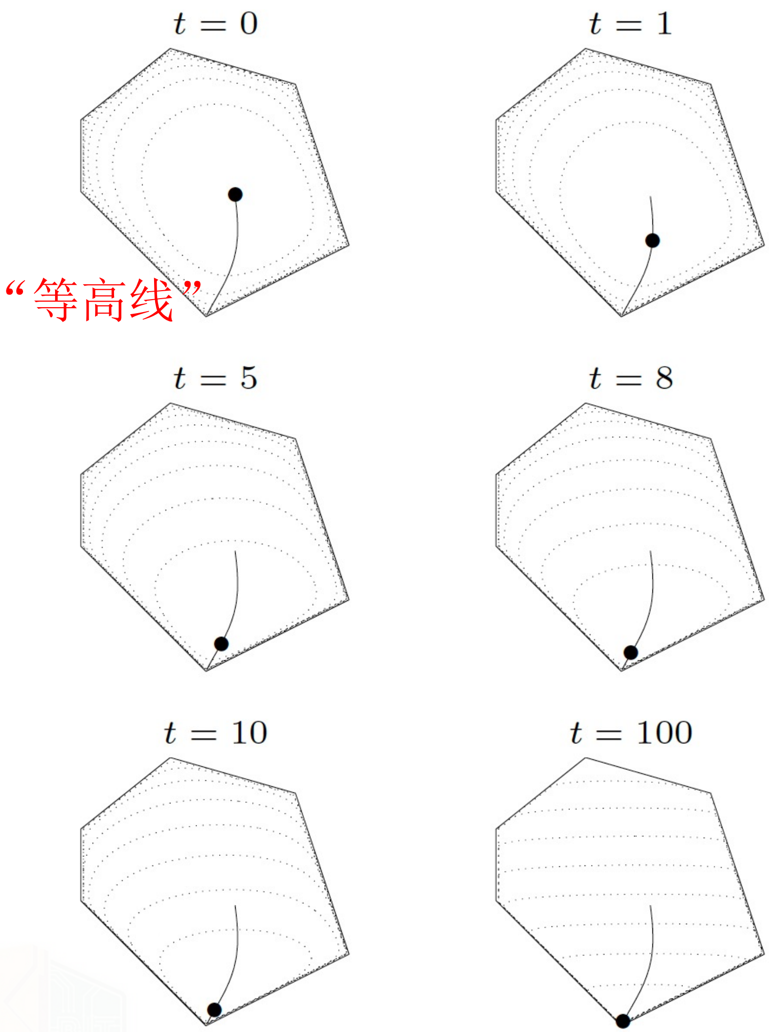
由一个较小的t_0出发，计算一个$\boldsymbol{x}^*(t)$；
逐渐增大t，用前一个$\boldsymbol{x}^*(t)$帮助计算当前t下的$\boldsymbol{x}^*(t)$；
当t较大时，趋近到原线性规划问题最优解

相邻的两个$\boldsymbol{x}^*(t)$一般距离很近，便于搜索

$$\begin{array}{ll}\min & {\color{red}t}\boldsymbol{c}^T\boldsymbol{x}-\sum_{i=1}^m\log\left(-\sum_{j=1}^n A_{ij}x_j+b_i\right)\\ \text{var.} & \boldsymbol{x}.\end{array}$$

障碍函数

$t\to\infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



（本讲介绍的）内点法思路

由一个较小的t_0出发，计算一个$\boldsymbol{x}^*(t)$；

逐渐增大t，用前一个$\boldsymbol{x}^*(t)$帮助计算当前t下的$\boldsymbol{x}^*(t)$；

当t较大时，趋近到原线性规划问题最优解

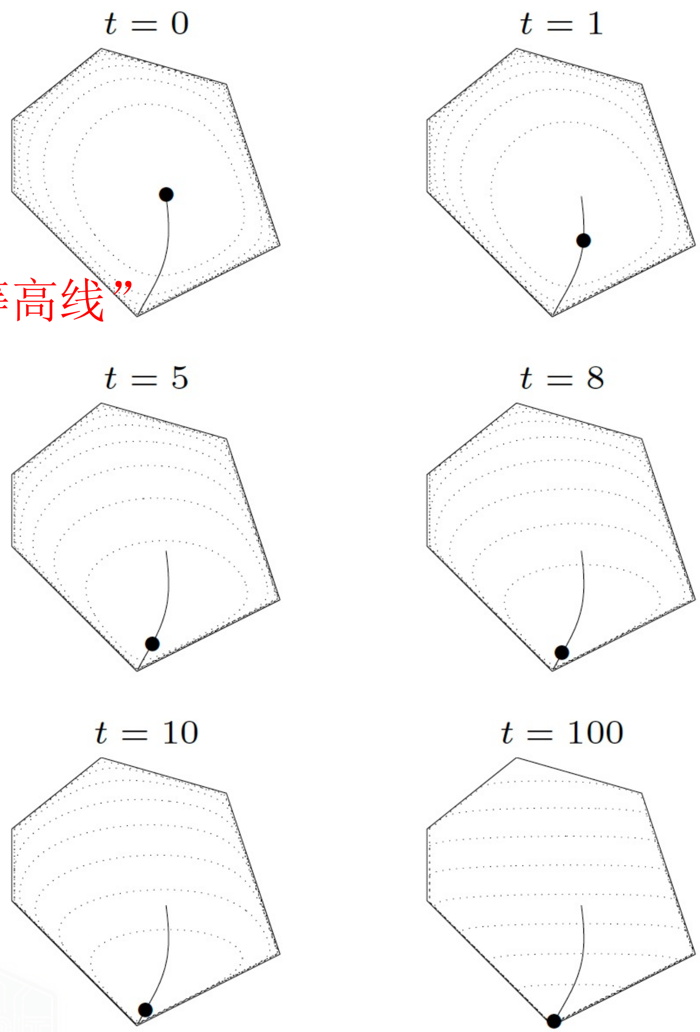
相邻的两个$\boldsymbol{x}^*(t)$一般距离很近，便于搜索

为什么不直接把t_0设成一个非常大的数，直接得到原问题最优解？

$$\begin{array}{ll} \min & {\color{red}t}\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m}\log\left(-\sum_{j=1}^{n}A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

障碍函数

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



“等高线”

（本讲介绍的）内点法思路

由一个较小的t_0出发，计算一个$\boldsymbol{x}^*(t)$；
逐渐增大t，用前一个$\boldsymbol{x}^*(t)$帮助计算当前t下的$\boldsymbol{x}^*(t)$；
当t较大时，趋近到原线性规划问题最优解

相邻的两个$\boldsymbol{x}^*(t)$一般距离很近，便于搜索

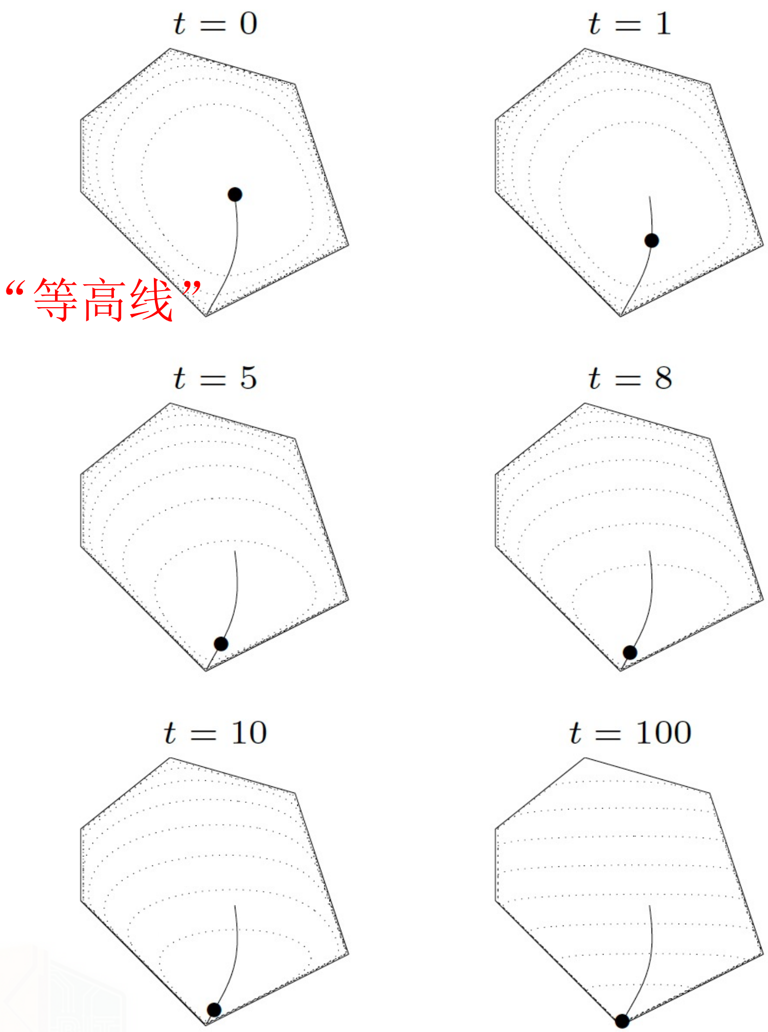
为什么不直接把t_0设成一个非常大的数，直接得到原问题最优解？

在缺少前序$\boldsymbol{x}^*(t)$位置信息的情况下，直接精确求解大t对应的最优解计算量非常大

$$\begin{aligned} &\min \quad {\color{red}t}\boldsymbol{c}^T\boldsymbol{x} - \textstyle\sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ &\text{var.} \quad \boldsymbol{x}. \end{aligned}$$

障碍函数

$t \to \infty$, $\boldsymbol{x}^*(t)$趋近于原线性规划问题全局最优解



（本讲介绍的）内点法思路

由一个较小的t_0出发，计算一个$\boldsymbol{x}^*(t)$；
逐渐增大t，用前一个$\boldsymbol{x}^*(t)$帮助计算当前t下的$\boldsymbol{x}^*(t)$；
当t较大时，趋近到原线性规划问题最优解

还未解决的问题？

给定t时怎么计算$\boldsymbol{x}^*(t)$呢？
新的目标函数已经不是线性函数了

内点法

如何在给定t时计算$\boldsymbol{x}^*(t)$

$$
\begin{aligned}
&\min \quad t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\
&\text{var.} \quad \boldsymbol{x}.
\end{aligned}
$$

内点法

如何在给定t时计算$\boldsymbol{x}^*(t)$

$$
\begin{aligned}
&\min \quad t\boldsymbol{c}^T\boldsymbol{x} - \textstyle\sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\
&\text{var.} \quad \boldsymbol{x}.
\end{aligned}
$$

证明目标函数是凸函数：

（1）凸函数之和依然是凸函数

（2）证明$g_i(\boldsymbol{x}) \triangleq -\log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right)$是凸函数

内点法

如何在给定t时计算$\boldsymbol{x}^*(t)$

$$\begin{aligned}\min\quad & t\boldsymbol{c}^T\boldsymbol{x}-\textstyle\sum_{i=1}^{m}\log\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)\\ \text{var.}\quad & \boldsymbol{x}.\end{aligned}$$

证明目标函数是凸函数：

（1）凸函数之和依然是凸函数

（2）证明$g_i(\boldsymbol{x})\triangleq-\log\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)$是凸函数

$$\nabla g_i(\boldsymbol{x})=\left(\frac{A_{i1}}{-\sum_{j=1}^{n}A_{ij}x_j+b_i},\frac{A_{i2}}{-\sum_{j=1}^{n}A_{ij}x_j+b_i},\dots,\frac{A_{in}}{-\sum_{j=1}^{n}A_{ij}x_j+b_i}\right)^T$$

$$\nabla^2 g_i(\boldsymbol{x})=\begin{bmatrix}\frac{A_{i1}^2}{\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)^2} & \cdots & \frac{A_{i1}A_{in}}{\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)^2}\\ \vdots & \ddots & \vdots\\ \frac{A_{i1}A_{in}}{\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)^2} & \cdots & \frac{A_{in}^2}{\left(-\sum_{j=1}^{n}A_{ij}x_j+b_i\right)^2}\end{bmatrix}^T=\nabla g_i(\boldsymbol{x})\left(\nabla g_i(\boldsymbol{x})\right)^T$$

由定义，易证
是半正定矩阵

内点法

如何在给定t时计算$\boldsymbol{x}^*(t)$

$$\begin{array}{ll} \min & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} & \boldsymbol{x}. \end{array}$$

梯度下降法：

$$x^{(k)} = x^{(k-1)} - \alpha_k \nabla f(x^{(k-1)}), \quad k = 1, 2, 3, \ldots$$

牛顿法：

$$x^{(k)} = x^{(k-1)} - \left(\nabla^2 f(x^{(k-1)})\right)^{-1} \nabla f(x^{(k-1)}), \quad k = 1, 2, 3, \ldots$$

注意k与t不同，k是在给定的一个t下调用梯度下降法/牛顿法时迭代次数的计数

内点法

- 利用障碍函数改写问题，把约束条件“放入”目标函数

$$\begin{aligned} \min \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{s.t.} \quad & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned} \Rightarrow \begin{aligned} \min \quad & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

- 对每一个给定的t，求解新的优化问题，得到$\boldsymbol{x}^*(t)$
  - 证明新的目标函数是凸函数
  - 用梯度下降法或牛顿法（更快）搜索使梯度为零的$\boldsymbol{x}^*(t)$

内点法

内点法

0 确定一个初始可行解$\tilde{\boldsymbol{x}}$，初始参数值$t > 0$

1 由$\tilde{\boldsymbol{x}}$出发，在当前t取值下，计算$\boldsymbol{x}^*(t)$

$$\begin{aligned} \min \quad & t\boldsymbol{c}^T\boldsymbol{x} - \sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ \text{var.} \quad & \boldsymbol{x}. \end{aligned}$$

调用牛顿法等算法求解$\boldsymbol{x}^*(t)$时，需要内迭代

2 将$\boldsymbol{x}^*(t)$赋值给$\tilde{\boldsymbol{x}}$

3 如果循环结束条件（如t大于一定阈值）满足，则结束；
否则，增加t取值（如令$t \leftarrow \mu t$，μ一般在$10 \sim 100$中取值），返回1

内容修改自Stephen Boyd < Stanford-EE364 Sequential unconstrained minimization>

内点法

内点法

0 确定一个初始可行解$\tilde{\boldsymbol{x}}$，初始参数值$t > 0$

1 由$\tilde{\boldsymbol{x}}$出发，在当前t取值下，计算$\boldsymbol{x}^*(t)$

$$\begin{aligned} &\min \quad t\boldsymbol{c}^T\boldsymbol{x} - \textstyle\sum_{i=1}^{m} \log\left(-\sum_{j=1}^{n} A_{ij}x_j + b_i\right) \\ &\text{var.} \quad \boldsymbol{x}. \end{aligned}$$

调用牛顿法等算法求解$\boldsymbol{x}^*(t)$时，需要内迭代

2 将$\boldsymbol{x}^*(t)$赋值给$\tilde{\boldsymbol{x}}$

3 如果循环结束条件（如t大于一定阈值）满足，则结束；

否则，增加t取值（如令$t \leftarrow \mu t$，μ一般在10 ~ 100中取值），返回1

控制参数μ如何影响算法效率？

内容修改自Stephen Boyd < Stanford-EE364 Sequential unconstrained minimization>

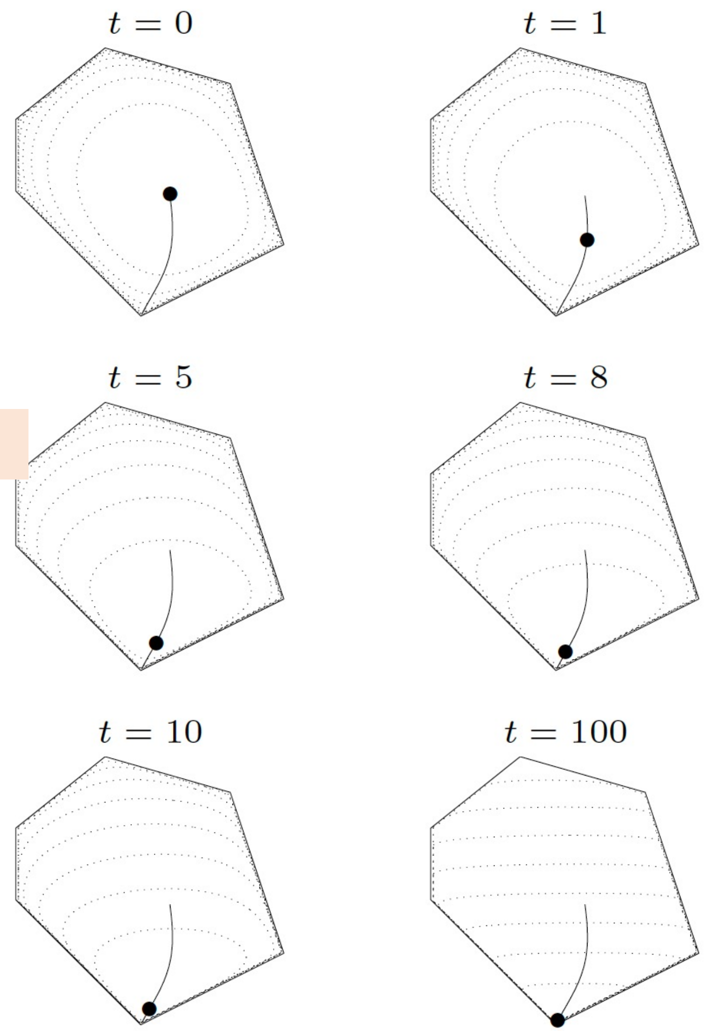
内点法

控制参数μ如何影响算法效率？

$$t \leftarrow \mu t$$

大μ令t增长很快，牛顿法的负担增大（内迭代次数增多）

小μ令t增长变慢，外迭代次数增多



内点法

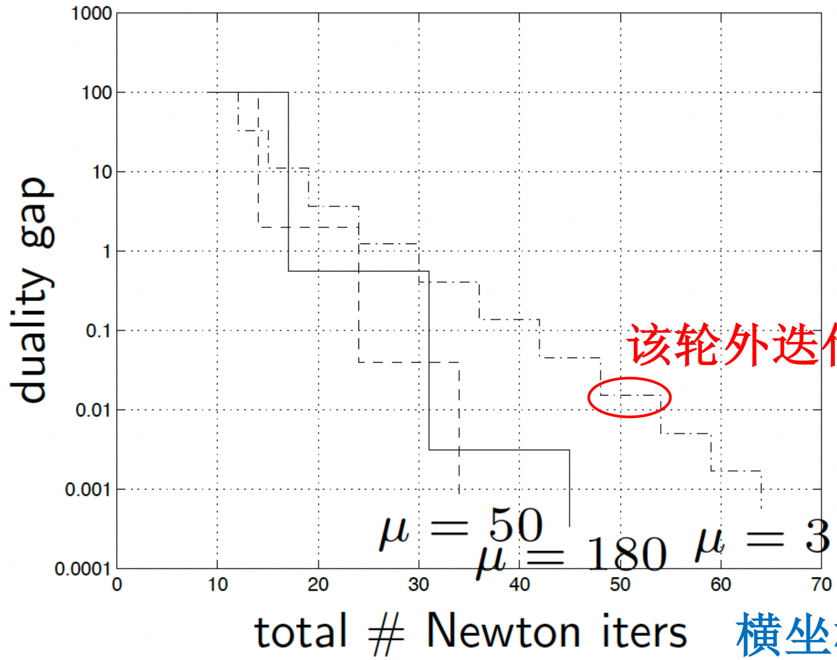
控制参数μ如何影响算法效率？

$t \leftarrow \mu t$

大μ令t增长很快，牛顿法的负担增大（内迭代次数增多）

小μ令t增长变慢，外迭代次数增多

纵坐标：原线性规划问题求解情况



该轮外迭代中内迭代的次数

说明μ需要取中间值

横坐标：内迭代总次数

内容修改自Stephen Boyd < Stanford-EE364 Sequential unconstrained minimization>

本讲小结

- 单纯形法的一般步骤
- 内点法的思想、一种简单的内点法

主要参考资料

Carl Kingsford <CMU-15-451/651: Algorithms> Slides

Wen Shen <Math 484 Linear Programming> Video

Wikipedia <Klee – Minty cube> Page

李恩志 <凸优化算法I: 内点法求解线性规划问题> 文章

Mung Chiang <Princeton ELE539A-Lecture 18: Interior Point Algorithm> Slides

Stephen Boyd <Stanford-EE364 Sequential unconstrained minimization> Slides

Ryan Tibshirani <Convex Optimization 10-725/36-725> Slides

谢谢！