基于一个例子的线性规划简介

刘红英

摘要：本文的素材取自 Strang 的教材[1, pp.440-446]. 我们通过一个例子，简单介绍了线性规划的核心内容，包括线性规划的图解法、基本定理、单纯形法、对偶理论和内点法. 本文可作为线性代数应用的素材，也可作为学习“线性规划”的阅读材料，帮助学生理解线性规划最基本的理论和算法思想.

关键词：线性规划，单纯形法，对偶，内点法

1 从线性方程组到线性规划

线性规划是线性代数加上两个新思想“**不等式**”和“**极小化**”. 除过可接受点是非负的外，即要求$\boldsymbol{x} \geq \mathbf{0}$（$\boldsymbol{x}$的每一个分量都是非负的），出发点仍然是矩阵方程. 设$\boldsymbol{A} \in \mathbf{R}^n, \boldsymbol{b} \in \mathbf{R}^m$，如果$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$有一个满足$\boldsymbol{x} \geq \mathbf{0}$的解，则一定有许多. 矩阵$\boldsymbol{A}$满足$n > m$，即未知数比方程多.

线性规划(linear programming) 就是找一个解$\boldsymbol{x}^*$，其极小化成本$c_1x_1 + \cdots + c_nx_n$. 最终胜出的向量$\boldsymbol{x}^*$是$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$的非负解，其具有最小的成本. 这样，问题就是极小化$\boldsymbol{c}^T\boldsymbol{x}$受约束于两组条件$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$和$\boldsymbol{x} \geq \mathbf{0}$，可以紧凑地表示为

$$\begin{aligned} &\text{minimize} \ \ \boldsymbol{c}^T\boldsymbol{x} \\ &\text{subject to} \ \ \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b} \\ &\qquad\qquad\ \ \boldsymbol{x} \geq \mathbf{0}. \end{aligned} \tag{1}$$

综上，一个线性规划问题始于一个矩阵$\boldsymbol{A}$和两个向量$\boldsymbol{b}$和$\boldsymbol{c}$：

1) $\boldsymbol{A}$满足$n > m$：例如$\boldsymbol{A} = [1\ 1\ 2]$(一个方程，三个未知数)；
2) $\boldsymbol{b}$有m个分量，其分别对应于m个方程$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$：例如$\boldsymbol{b}$=[4]；
3) 成本向量$\boldsymbol{c}$有n个分量：例如$\boldsymbol{c}$ =[5 3 8].

由这里给出的数据，得到

例 1
$$\begin{aligned} &\text{minimize} \ \ 5x_1 + 3x_2 + 8x_3 \\ &\text{subject to} \ \ x_1 + x_2 + 2x_3 = 4, \\ &\qquad\qquad\ \ x_1, x_2, x_3 \geq 0. \end{aligned} \tag{2}$$

这里我们直接给出了这个问题，却并没有解释它源自哪里. 实际上，从数学到日常管理，线性规划都有很广泛的应用. 研究快速算法和快速编码的竞争异常激烈. 因为有额外的要求：$\boldsymbol{x}^* \geq \mathbf{0}$和最小的成本$\mathbf{c}^\mathrm{T}\boldsymbol{x}^*$，大家将看到找$\boldsymbol{x}^*$比解方程组$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$更难. 在解出例 1 后，我们将解释这个例子的背景，著名的**单纯形法**，对偶理论和内点法.

2 线性规划的图解法与解的性质(线性规划基本定理)

先看约束条件：$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$和$\boldsymbol{x} \geq \mathbf{0}$. 方程$x_1 + x_2 + 2x_3 = 4$在三维空间中确定了一个平面. 非负条件$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$将平面限制成一个三角形. 解向量$\boldsymbol{x}^*$位于 ΔPQR上，如图 1.

在这个三角形内部，$\boldsymbol{x}$的所有元素都是正的. 对于ΔPQR的边上的点而言，有一个分量的取值是 0. 对于 P，Q，R 这三个顶点而言，有两个分量的取值均是 0. **最优解向量$\boldsymbol{x}^*$将是这些顶点之一员！**

这个三角形包含了所有满足条件$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$和$\boldsymbol{x} \geq \mathbf{0}$的向量$\boldsymbol{x}$. 称这些$\boldsymbol{x}$为**可行点**，称这个三角形为**可行域**. 这些点是使得$\mathbf{c} \cdot \boldsymbol{x}$取最小值的候选点，从而进入最后的环节：找到位于$\Delta PQR$上的一个$\boldsymbol{x}^*$，其使得$5x_1 + 3x_2 + 8x_3$取到最小值.

那些具有 0 成本的向量位于平面$5x_1 + 3x_2 + 8x_3 = 0$上. 这个平面和三角形不相交. 在关于$\boldsymbol{x}$的这些约束条件下，我们不能使成本为 0. 所以需要增大成本 C直到$5x_1 + 3x_2 + 8x_3 = C$与三角形相交. 当依次增大 C时，会出现向三角形移动的相互平行的平面族.

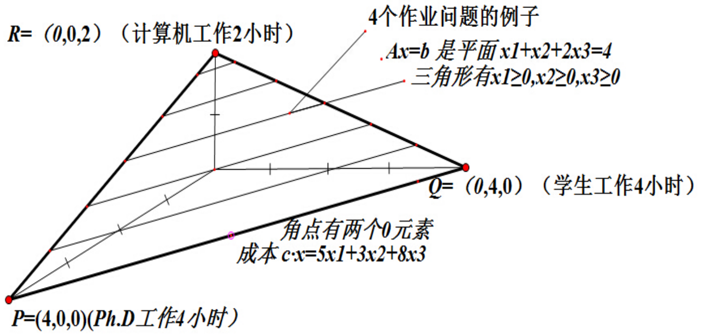

图 1 三角形包含了例 1 中方程组的所有非负解，即满足$\boldsymbol{Ax}=\boldsymbol{b}$和$\boldsymbol{x}\geq\boldsymbol{0}$的所有点. 使成本最小的解向量$\boldsymbol{x}^*$是可行域的顶点 P，Q或 R 中的一个.

第一个接触到三角形的平面$5x_1+3x_2+8x_3=C$具有最小成本 C，**这个接触点是解$\boldsymbol{x}^*$**，它一定是顶点 P，Q，R 中的一个. 一族移动的平面在接触到一个顶点之前是不可能接触到三角形内部的点. 所以，检验每个顶点处的成本$5x_1+3x_2+8x_3$的大小. 具体地，P=(4, 0, 0) 处的成本为 20，**Q**=(0, 4, 0) 处的成本为 12，R=(0, 0, 2)处的成本为 16. 最终选出的点为 Q. 所以$\boldsymbol{x}^*$=(0, 4, 0)是例 1 的解.

如果成本向量$\boldsymbol{c}$改变，平行的平面族将倾斜. 对于较小的改变来说，Q 依然是最优选择. 如果成本$\boldsymbol{c}\cdot\boldsymbol{x}=5x_1+4x_2+7x_3$，则最优解$\boldsymbol{x}^*$移到 $R=(0,0,2)$，得最小成本为$7\cdot 2=14$.

注记 1 有些线性规划**最大化利润**而非最小化成本. 所用数学原理几乎是一样的，不同之处是移动相互平行的平面族时，是从一个绝对值较大的数 C开始，而不是从绝对值较小的数开始. 他们向原点移动，而不是向无穷远处移动. 随着 C的变小，**第一个接触点还是某个顶点**.

注记 2 可能出现约束条件$\boldsymbol{Ax}=\boldsymbol{b}$和$\boldsymbol{x}\geq\boldsymbol{0}$不相容的情况，比如方程$x_1+x_2+x_3=-1$没有非负解. **这个可行域是空集**.

注记 3 也可能出现**无界的**可行域这种情况，比如如果约束条件是$x_1+x_2-2x_3=4$，则大的正向量（100, 100, 98）是一个可行点. 大的向量（1000, 1000, 998）也是. 此时平面$\boldsymbol{Ax}=\boldsymbol{b}$不再被限制成一个三角形. 两个顶点 P 和 Q 仍然是$\boldsymbol{x}^*$的候选点，但是顶点 R 移到了正无穷.

注记 4 当可行域无界时，最小的成本有可能是$-\infty$(负无穷)，比如成本是$-x_1-x_2+x_3$，则向量（100,100,98）处的成本$C=-102$. 向量（1000,1000,998）处的成本$C=-1002$. 这代表收入是 1002，而不是支出. 尽管在实际应用中不会发生这种情况，但是从理论上讲三元组$\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}$有可能产生不可预料的三角形和成本.

下面给出例 1 中的$\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}$的一种具体解释. 设未知数$x_1$，$x_2$，$x_3$分别表示由博士，学生和机器完成作业所需要的时间，他们每小时的成本分别为 5 元，3 元，8 元. 所需时间是非负的：$x_1\geq 0, x_2\geq 0, x_3\geq 0$. 博士和学生一小时分别能完成作业的一个问题，机器一小时能完成两个问题. 在原则上他们共同完成由四个问题组成的作业：$x_1+x_2+2x_3=4$.

问题：以最小的成本$\boldsymbol{c}^T\boldsymbol{x}$完成这四个问题.

如果所有的三种方式都投入工作，那么完成作业需要 1 小时：$x_1=x_2=x_3=1$，对应的成本为$5+3+8=16$. 但是可以用工作效率相同但成本更低的学生代替博士. 当学生工作 2 小时，机器工作 1 小时时，成本为 6+8=14，并且解决了所有四个问题. 三角形的边 QR 表示博士不工作：$x_1=0$. 剩下的问题是：所有的工作都由学生做还是都由机器做. 在这个例子中学生在四个小时内解决四道题的花费是 12 元，花费最少.

这个例子中的约束方程组$\boldsymbol{Ax}=\boldsymbol{b}$只有一个方程，顶点(0, 4, 0)有两个零元素. **当$\boldsymbol{Ax}=\boldsymbol{b}$有$m$个方程，顶点有$n-m$个零元素.** 选$n-m$个变量作为自由变量，并将这$n-m$个自由变量置为零后，解$\boldsymbol{Ax}=\boldsymbol{b}$可以得到$m$个非零变量(**基本解**).

可能存在的顶点数目是从n个元素中选取$n-m$个的选取方式的个数. 如果有20个未知数和8个方程，那么可行域将会有$20!/(8!\,12!)$个顶点，即$20\cdot 19\cdots\cdot 14\cdot 13=5079110400$. 逐个检查三个顶点的确可以求解最小成本问题. 但检查五百万个顶点是行不通的. 单纯形法将使加快计算速度.

3 原始问题和对偶问题(对偶理论)

在线性规划中，问题经常会以原始和对偶的方式成对出现. 原始问题和它的对偶问题中有一个最小化问题和一个最大化问题. 原始问题由矩阵 $\boldsymbol{A}$ 和两个向量$\boldsymbol{b},\boldsymbol{c}$确定. 它的对偶问题将$\boldsymbol{A}$转置，并调换$\boldsymbol{b},\boldsymbol{c}$：求$\boldsymbol{b}\cdot\boldsymbol{y}$ 的最大值.

对偶问题 在约束条件$\boldsymbol{A}^T\boldsymbol{y}\le\boldsymbol{c}$下极大化$\boldsymbol{b}\cdot\boldsymbol{y}$，即

$$\begin{aligned}&\text{maximize} && \boldsymbol{b}^T\boldsymbol{y}\\&\text{subject to} && \boldsymbol{A}^T\boldsymbol{y}\le\boldsymbol{c}.\end{aligned}\tag{3}$$

还是以例 1 为例，考虑它的对偶问题.

例 1(续，对偶) 设一个作弊者通过卖答案来提供解题服务. 他的要价是每道题y元，一共$4y$元. 这个人必须和博士，学生以及机器一样的廉价：$y\le 5$，$y\le 3$和$2y\le 8$. 这个人确定价格以最大化自己的收入$4y$元.

当$y=3$时，$b\cdot y$最大. 收益为$4y=12$. 对偶问题的最大值等于原始问题的最小值. **最大值=最小值**就是对偶性.

强对偶定理 如果两个问题中的一个有最优解（$\boldsymbol{x}^*$或$\boldsymbol{y}^*$）, 则另一个问题也有，且最小成本$\boldsymbol{c}\cdot\boldsymbol{x}^*$等于最大收益$\boldsymbol{b}\cdot\boldsymbol{y}^*$.

证明这个结论比较复杂，但是仅用一行足以证明这个定理容易的那半边：作弊者的收入$\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{y}$不可能超过诚实者的成本：

弱对偶定理 如果$\boldsymbol{Ax}=\boldsymbol{b}$，$\boldsymbol{x}\ge\boldsymbol{0}$，$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}\le\boldsymbol{c}$，则$\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{y}=(\boldsymbol{Ax})^{\boldsymbol{T}}\boldsymbol{y}=\boldsymbol{x}^{\mathbf{T}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}\le\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{c}$.

上式中最后一个≤的推导如下：

因为$\boldsymbol{x}\ge\boldsymbol{0}$和$\boldsymbol{s}=\boldsymbol{c}-\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}\ge\boldsymbol{0}$，所以二者的点积$\boldsymbol{x}^{\mathbf{T}}\boldsymbol{s}\ge 0$. 此即$\boldsymbol{x}^{\mathbf{T}}\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}\le\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{c}$.

完整的对偶定理表明：当$\boldsymbol{b}^{\boldsymbol{T}}\boldsymbol{y}$ 取到它的最大值且$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{c}$ 取到它的最小值时，它们二者相等：$\boldsymbol{b}\cdot\boldsymbol{y}^*=\boldsymbol{c}\cdot\boldsymbol{x}^*$.

互补定理 设$\boldsymbol{x}$和$\boldsymbol{y}$分别是原始问题(1)和对偶问题(3)的可行解，则$\boldsymbol{b}\cdot\boldsymbol{y}=\boldsymbol{c}\cdot\boldsymbol{x}$成立当且仅当$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{s}=0$，即对每个 j，最优解有$x_j^*=0$或者$s_j^*=0$.

假设已经得到原始问题的解$\boldsymbol{x}^*$. 求解$\boldsymbol{A}^T\boldsymbol{y}=\boldsymbol{c}$中与$\boldsymbol{x}^*$中$m$个非零元素对应的$m$元方程组得解$\boldsymbol{y}^*$. 令$\boldsymbol{s}^*=\boldsymbol{c}-\boldsymbol{A}^T\boldsymbol{y}^*$, 因为恒有$\boldsymbol{x}^{*\boldsymbol{T}}\boldsymbol{s}^*=\boldsymbol{0}$，也即对偶性：$x_j^*=0$或者$s_j^*=0$，所以$\boldsymbol{y}^*$是对偶问题的最优解.

例 1(续，互补定理) 用例 1 来说明. 那里的$\boldsymbol{x}^*=(0,4,0)$是最优顶点 Q, 此时仅有第二个分量非零，从而由互补定理，求解$\boldsymbol{A}^T\boldsymbol{y}=\boldsymbol{c}$的第二个方程$y=3$，得最优价格$y^*$=3. 这是例 1 的对偶问题的解.

4 单纯形法

消元是解线性方程组的关键. 单纯形法是解线性不等式的关键. **单纯形法从一个顶点移到相邻的成本更小的顶点. 直到找到成本最小的顶点.**

一个顶点是一个向量$\boldsymbol{x}$，它满足n元方程组$\boldsymbol{Ax}=\boldsymbol{b}$**且至多有$m$个正元素，其余的$n-m$个元素是零**(这些是自由变量. 回代即可得到$m$个基变量. 所有的变量必须是非负的，否则$\boldsymbol{x}$是

一个假顶点）. 对于一个相邻的顶点而言，$\boldsymbol{x}$的一个零元素变为正，同时一个正元素变为零.

用单纯形法必须确定：哪个元素"进基"变为正的？哪个元素"离基"变为 0？选取这样的交换以便使总成本变低. 这是单纯形法向最优解$\boldsymbol{x}^*$移动过程中的一步，称为一步单纯形迭代.

观察当前顶点的每个零元素. 如果它从 0 变到 1，那么其他非 0 元素必须改变以满足方程组$\boldsymbol{Ax}=\boldsymbol{b}$，通过回代消元过程找到新的$\boldsymbol{x}$并且观察总成本$\boldsymbol{c}\cdot\boldsymbol{x}$的改变量. 这个改变量是新变量的"**既约成本**"r(reduced cost). **最负**r对应的那个变量是"进基变量". 对于一个单位的新变量而言，这是成本的最大减少量.

例 1（**续，进基−出基**） 假定当前顶点是$P=(4,0,0)$，即由博士做所有的作业. 如果由学生工作一个小时，则$\boldsymbol{x}=(3,1,0)$处的成本下降为 18 元，既约成本r= -2. 如果由机器工作一小时，则$\boldsymbol{x}=(2,0,1)$处的成本也是 18 元，既约成本也是-2. 此时，单纯形法即可以选择学生或机器为"进基变量".

即使在很简单的例子中，也很少能在第一步就找到最优解$\boldsymbol{x}^*$. 此时，这种方法通过比较r以确定"进基变量"，几何上，即从图 1 中的顶点 P 移到顶点 Q 或 R.

尽管"进基变量"越大，成本就越低. 但当第一个正元素变为 0 时(调整$\boldsymbol{x}$以满足$\boldsymbol{Ax}=\boldsymbol{b}$)，成本就不再降低. "离基变量"指的是**第一个**变为 0 的正x_i. 当出现这种情况时，就找到了一个相邻的顶点. 然后重新开始，去找其他的进基和离基变量.

当既约成本全为正的时，当前顶点就是最优$\boldsymbol{x}^*$. 因为在不使$\boldsymbol{c}\cdot\boldsymbol{x}$增加的前提下，没有 0 元素可以变为正的，因此不存在新的进基变量. 问题得解.

注记 5 一般情况下得到$\boldsymbol{x}^*$需要$\alpha\cdot n$步，这里的α不是很大. 但是有人构造了一个例子(Klee-Minty 的例子)，求解其需要指数数量的单纯形迭代. 后来提出了一种不同的方法，能证明其用较少(但更难)步的迭代接近$\boldsymbol{x}^*$. 这种新方法在可行域的内部迭代.

例 2（**进基−出基**） 使成本$\boldsymbol{c}\cdot\boldsymbol{x}=3x_1+x_2+9x_3+x_4$最小化，限制条件是$\boldsymbol{x}\geq\boldsymbol{0}$和两个方程$\boldsymbol{Ax}=\boldsymbol{b}$：

$$x_1+2x_3+x_4=4,\qquad m=2$$
$$x_2+x_3-x_4=2,\qquad n=4.$$

开始顶点$\boldsymbol{x}=(4,2,0,0)$，它的成本$\boldsymbol{c}\cdot\boldsymbol{x}=14$. 它有$m=2$个非零元素和$n-m=2$个零元素. 零元素是$x_3$和$x_4$. 问题是：应该是$x_3$进基，还是$x_4$进基. 对它们分别进行尝试：

如果$x_3=1$ 且$x_4=0$，那么$\boldsymbol{x}=(2,1,1,0)$处的成本为16.

如果$x_3=0$ 且$x_4=1$，那么$\boldsymbol{x}=$（3,3,0,1）处的成本为 13.

与 14 相比，换入x_3时既约成本$r=2$，是正的，所以舍去. 换入x_4时既约成本$r=-1$，是负的，所以保留. 从而进基变量是x_4.

新进基的x_4应取多大？x_4增加 1 使x_1从 4 降到 3. x_4增加 4 使x_1从 4 降到 0，所以离基变量是x_1. 新的顶点$\boldsymbol{x}=(0,6,0,4)$，成本只有$\boldsymbol{c}\cdot\boldsymbol{x}=10$. 这就是最优的$\boldsymbol{x}^*$. 但是要知道此结论还需要从(0,6,0,4)开始检验. 假设让x_1或者x_3进基：

从顶点(0,6,0,4)开始：如果$x_1=1$且$x_3=0$，得$\boldsymbol{x}=(1,5,0,3)$，成本为 11. 如果$x_3=1$且$x_1=0$，得$\boldsymbol{x}=(0,3,1,2)$，成本为 14. 这些成本都比 10 大. 所以既约成本 r 都是正的，舍去. 从而当前顶点(0,6,0,4)是解$\boldsymbol{x}^*$.

可以简化这些计算. 每个单纯形需要求解与矩阵$\boldsymbol{B}$($\boldsymbol{A}$的m个基本变量对应的列组成的$m\times m$矩阵，称为**基矩阵**)有关的两个线性方程组. 当用进基的列向量替换旧的离基列向量后，可以快速更新$\boldsymbol{B}^{-1}$(**修正单纯形法**)，大多数代码是基于此实现单纯形法的.

5 内点法

单纯形法沿着可行域的边界移动，最后到达最优顶点$\boldsymbol{x}^*$. 内点法只在可行域的内部移动($\boldsymbol{x}>0$). 这种方法希望能更直接的逼近$\boldsymbol{x}^*$.

保持在内部的途径之一是在边界上设置障碍. 当任一变量x_j接近 0 时都要加上额外的成本. 最好满足$\boldsymbol{x}>\mathbf{0}$. 为此,设$\theta$是一个无限趋于 0 的正数,考虑如下的障碍罚函数的极小化,即

$$\text{障碍问题}\qquad \begin{aligned} &\text{minimize}\ \ \boldsymbol{c}^T\boldsymbol{x}-\theta(\log x_1+\cdots+\log x_n) \\ &\text{subject to}\quad \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}. \end{aligned} \tag{4}$$

这里的成本不是线性的. 因为当$x_j=0$ 时 $\log x_j$是无穷，所以不需要约束条件$x_j\geq 0$.

对每个给定的障碍参数θ，障碍问题(4)均可作为原始问题的一个近似问题. 这是一个等式约束优化问题,可以用 Langrange 乘子法求解[2]. 设m个约束条件$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$的 Lagrange 乘子分别为$y_1$，$\cdots,y_m$，则得障碍问题的 Lagrange 函数

$$L(\boldsymbol{x},\ \boldsymbol{y},\ \theta)=\boldsymbol{c}^{\boldsymbol{T}}\boldsymbol{x}-\theta\left(\sum_{i=1}^{n}\log x_i\right)-\boldsymbol{y}^{\boldsymbol{T}}(\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b})$$

由$\frac{\partial L}{\partial y}=0$可以得到$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$,而导数$\frac{\partial L}{\partial x_j}$很有意思!由这些可以得到**障碍问题(4)的最优性条件:**

$$\begin{aligned} &\frac{\partial L}{\partial x_j}=c_j-\frac{\theta}{x_j}-(\boldsymbol{A}^T\boldsymbol{y})_j=0,\qquad j=1,\cdots,n \\ &\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \end{aligned}$$

令$\boldsymbol{s}=\boldsymbol{c}-\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}$，障碍问题(4)的最优性条件可以表示为

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x}&=\boldsymbol{b} \\ \boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}+\boldsymbol{s}&=\boldsymbol{c} \\ x_js_j&=\theta,\ j=1,\cdots,n \end{aligned} \tag{5}$$

注意原始问题的解满足$x_js_j=0$，而障碍问题的解满足$x_js_j=\theta$. 解$\boldsymbol{x}^*(\theta)$在以$\boldsymbol{x}^*(0)$为端点的**中心路径**(central path)上. 上述方程组中的n个等式$x_js_j=\theta$不是线性的，用牛顿法可以求解方程组(5).

设当前向量$\boldsymbol{x}$，$\boldsymbol{y}$，$\boldsymbol{s}$满足$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$，$\boldsymbol{x}\geq\mathbf{0}$，$\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}+\boldsymbol{s}=\boldsymbol{c}$，但不满足$x_js_j=\theta$. 设牛顿法的校正步为$\Delta\boldsymbol{x}$，$\Delta\boldsymbol{y}$，$\Delta\boldsymbol{s}$. 在$(x_j+\Delta x_j)(s_j+\Delta s_j)=\theta$中忽略二次项$\Delta x_j\Delta s_j$，即得到关于校正步的线性方程组：

$$\text{牛顿步}\qquad \begin{aligned} &\boldsymbol{A}\Delta\boldsymbol{x}=\mathbf{0} \\ &\boldsymbol{A}^{\boldsymbol{T}}\Delta\boldsymbol{y}+\Delta\boldsymbol{s}=\mathbf{0} \\ &s_j\Delta x_j+x_j\Delta s_j=\theta-x_js_j,\ j=1,\cdots,n \end{aligned}$$

例 1（续，内点法） 对于例 1 中的问题(2)，障碍问题是

$$\begin{aligned} &\text{minimize}\ \ 5x_1+3x_2+8x_3-\theta(\log x_1+\log x_2+\log x_3) \\ &\text{subject to}\quad x_1+x_2+2x_3=4. \end{aligned}$$

该问题的最优性条件(5)对应的方程组共有 7 个未知数，具体为

$$\begin{aligned} &x_1+x_2+2x_3=4 \\ &y+s_1=5 \\ &y+s_2=3 \\ &2y+s_3=8 \\ &s_jx_j=\theta,\ j=1,2,3 \end{aligned}$$

给定$(\boldsymbol{x},\boldsymbol{y},\boldsymbol{s})$，其满足$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{A}^{\boldsymbol{T}}\boldsymbol{y}+\boldsymbol{s}=\boldsymbol{c},\boldsymbol{x}>0,\boldsymbol{s}>0$, 此时解如下线性方程组

$$\begin{aligned} &\Delta x_1+\Delta x_2+2\Delta x_3=0 \\ &\Delta y+\Delta s_1=0 \\ &\Delta y+\Delta s_2=0 \end{aligned}$$

$$\mathbf{2}\Delta y + \Delta s_3 = 0$$
$$s_j \Delta x_j + x_j \Delta s_j = \theta - x_j s_j, \qquad j = 1, 2, 3$$

可以得到障碍问题的**牛顿步**(Newton step). 令

$$(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{s}) + \alpha(\Delta \boldsymbol{x}, \Delta \boldsymbol{y}, \Delta \boldsymbol{s}),$$

其中$\alpha > 0$称为步长，选取其满足$(\boldsymbol{x}, \boldsymbol{s}) + \alpha(\Delta \boldsymbol{x}, \Delta \boldsymbol{s}) > \mathbf{0}$.

对每个θ，牛顿迭代是二次收敛的(解方程组的牛顿法)。然后令θ趋于 0. 迭代 20 步到 60 步后，对偶间隙$\boldsymbol{x}^{\boldsymbol{T}}\boldsymbol{s}$通常会小于$10^{-8}$. 内点法被广泛应用于商业软件，用于解决大量的线性或者非线性最优化问题.

参考文献

1. G. Strang. Introduction to Linear Algebra, 4th Edition. Wellesley-Cambridge Press: Wellesley MA 02482, USA, pp.440-446.
2. 刘红英，夏勇，周水生. 数学规划基础. 北京航空航天大学出版社：北京，2012.10.