数 学 规 划 基 础

部 分 习 题 参 考 解 答

刘红英　编

北京航空航天大学数学与系统科学学院

2015 年 5 月

内容简介

本书是《数学规划基础》(刘红英，夏勇，周水生，北京航空航天大学出版社，2012.10)的配套教学辅导材料，较详细地给出了该教材各章后部分习题的参考解答.

前　言

本习题解答自 2008 年春季开始编写，当时由硕士研究生阎凤玉提供部分习题解答，经讨论和确认后，由作者首次录入排版. 后来陆续参加习题解答修订的硕士研究生包括王浩、欧林鑫、朱丽媛、易彩霞和杨茜，其中的数值结果由欧林鑫提供. 作者在此向他们的辛勤劳动表示衷心的感谢.

本解答得到了？项目的资助，在此表示感谢.

由于这些参考解答尚未经过特别严格的校对，仅供参考. 任何意见、建议或其它反馈都可以发送至liuhongying@buaa.edu.cn，在此深表感谢.

刘红英

2015.5 于北京

目 录

第一章　引言

1.2 (该练习的目的是提高你的建模技巧，同时熟悉利用计算机求解线性优化问题) 一个原油精练场有 8 百万桶原油 A 和 5 百万桶原油 B 用以安排下个月的生产. 可用这些资源来生产售价为 38 元/桶的汽油，或者生产售价为 33 元/桶的民用燃料油. 有三种生产过程可供选择，各自的生产参数如下：

	过程1	过程2	过程3
输入原油A	3	1	5
输入原油B	5	1	3
输出汽油	4	1	3
输出燃料油	3	1	4
成本(单位：元)	51	11	40

除成本外，所有的量均以桶为单位. 例如，对于第一个过程而言，利用 3 桶原油 A 和 5 桶原油 B 可以生产 4 桶汽油和 3 桶民用燃料油，成本为 51 元. 表格中的成本指总的成本(即原油成本和生产过程的成本). 将此问题建模成线性规划，其能使管理者极大化下个月的净利润. 请利用Lingo,Cplex或Matlab在计算机上求解此问题.

解: 设下个月利用第一个过程生产x次, 第二个过程生产y次, 第三个过程生产z次. 则利润为

$$\begin{aligned} f(x,y,z) &= (38\times 4+33\times 3-51)x+(38+33-11)y+(38\times 3+33\times 4-40)z \\ &= 200x+60y+206z \end{aligned}$$

其数学模型为

$$\begin{aligned} \text{maximize} \quad & 200x+60y+206z \\ \text{subject to} \quad & 3x+y+5z\leq 8000000 \\ & 5x+y+3z\leq 5000000 \\ & x,y,z\geq 0,\ \text{且}\ x,y,z\ \text{是整数}. \end{aligned}$$

忽略掉整性要求后，调用 Matlab 中的 linprog.m 函数求解，得最优解 $x=0, y=500000, z=1500000$，自动满足整性要求.

1.3 利用图解法和优化软件两种方法求解下列问题

$$\begin{aligned} \text{minimize} \quad & (x_1-2)^2+(x_2-1)^2 \\ \text{subject to} \quad & x_1^2-x_2\leq 0, \\ & x_1+x_2\leq 2. \end{aligned}$$

1.4 确定下列 n 元函数的梯度向量和 Hessian 阵：

(a) $\boldsymbol{a}^T\boldsymbol{x}$: $\boldsymbol{a}$ 是常向量；

(b) $\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}$: $\boldsymbol{A}$ 是非对称的常矩阵；

(c) $\frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{x}$: $\boldsymbol{A}$ 是对称的常矩阵，$\boldsymbol{b}$ 是常向量；

(d) $\boldsymbol{r}(\boldsymbol{x})^T\boldsymbol{r}(\boldsymbol{x})$: $\boldsymbol{r}(\boldsymbol{x})=(r_1(\boldsymbol{x}),\cdots,r_m(\boldsymbol{x}))^T$ 是依赖于 $\boldsymbol{x}$ 的 m 维向量，记 $\nabla\boldsymbol{r}^T$ 为 $\boldsymbol{A}^T$，它一般不是常量.

解:

(a) $\nabla f(\boldsymbol{x})=\boldsymbol{a},\quad \nabla^2 f(\boldsymbol{x})=\boldsymbol{0}_{n\times n}$;

(b) $\nabla f(\boldsymbol{x})=(\boldsymbol{A}+\boldsymbol{A}^T)\boldsymbol{x},\quad \nabla^2 f(\boldsymbol{x})=\boldsymbol{A}+\boldsymbol{A}^T$;

(c) $\nabla f(\boldsymbol{x})=\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b},\quad \nabla^2 f(\boldsymbol{x})=\boldsymbol{A}$;

(d) $f(\boldsymbol{x})=\sum_{i=1}^m r_i^2(\boldsymbol{x}), \nabla f(\boldsymbol{x})=2\sum_{i=1}^m r_i(\boldsymbol{x})\nabla r_i(\boldsymbol{x})=2\boldsymbol{A}(\boldsymbol{x})^T\boldsymbol{r}(\boldsymbol{x})$,

$$\begin{aligned}\nabla^2 f(\boldsymbol{x}) &= 2\sum_{i=1}^m r_i(\boldsymbol{x})\nabla^2 r_i(\boldsymbol{x})+2\sum_{i=1}^n \nabla r_i(\boldsymbol{x})(\nabla r_i(\boldsymbol{x}))^T\\ &= 2\sum_{i=1}^m r_i(\boldsymbol{x})\nabla^2 r_i(\boldsymbol{x})+2\boldsymbol{A}(\boldsymbol{x})^T\boldsymbol{A}(\boldsymbol{x}).\end{aligned}$$

1.6 考虑向量值函数 $\boldsymbol{f}(\boldsymbol{x}):\mathbb{R}^n\to\mathbb{R}^m$ ，设 $\boldsymbol{f}$ 的每个分量函数 $f_i(\boldsymbol{x})$ 在 $\boldsymbol{x}'$ 都可微. 写出 $\boldsymbol{f}$ 在 $\boldsymbol{x}'$ 的Taylor展式，请用 $\boldsymbol{A}(x)^T$ 表示 $\nabla\boldsymbol{f}(x)^T(=[\nabla f_1(\boldsymbol{x}),\cdots,\nabla f_m(\boldsymbol{x})])$.

解: 为了具体，考虑 $m=2,n=3$ 给出，再表示成一般形式. 此时

$$\boldsymbol{f}(\boldsymbol{x})=\begin{pmatrix} f_1(\boldsymbol{x})\\ f_2(\boldsymbol{x})\end{pmatrix}=\begin{pmatrix} f_1(x_1,x_2,x_3)\\ f_2(x_1,x_2,x_3)\end{pmatrix}.$$

因为函数 $f_1(\boldsymbol{x})$ 和 $f_2(\boldsymbol{x})$ 可微，则由多元函数的Taylor展式，有

$$f_i(\boldsymbol{x})=f_i(\boldsymbol{x}')+\nabla f_i(\boldsymbol{x}')^T(\boldsymbol{x}-\boldsymbol{x}')+o(\|\boldsymbol{x}-\boldsymbol{x}'\|), i=1,2.$$

写成向量形式，即

$$\boldsymbol{f}(\boldsymbol{x})=\boldsymbol{f}(\boldsymbol{x}')+\boldsymbol{A}(\boldsymbol{x}')(\boldsymbol{x}-\boldsymbol{x}')+o(\|\boldsymbol{x}-\boldsymbol{x}'\|), \tag{1.1}$$

这里 $o(\|\boldsymbol{x}-\boldsymbol{x}'\|)$ 表示

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}'}\frac{\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{f}(\boldsymbol{x}')-\boldsymbol{A}(\boldsymbol{x}')(\boldsymbol{x}-\boldsymbol{x}')}{\|\boldsymbol{x}-\boldsymbol{x}'\|}=\boldsymbol{0}.$$

这里的式(1.1)即为 $\boldsymbol{f}$ 在 $\boldsymbol{x}'$ 的Taylor展式，其中的矩阵 $\boldsymbol{A}(\boldsymbol{x})$ 称为雅可比(Jacob)矩阵，它的第 i 行为 $f_i(\boldsymbol{x})$ 在 $\boldsymbol{x}$ 处的梯度向量的转置.

1.7 假设在点 $\boldsymbol{x}'$ 有 $\boldsymbol{g}'\neq\boldsymbol{0}$，证明在所有单位向量 $\boldsymbol{p}^T\boldsymbol{p}=1$ 中，函数沿方向向量 $\boldsymbol{p}=\boldsymbol{g}'/\|\boldsymbol{g}'\|_2$ 的斜率最大. 称该方向是函数的**最速上升**(steepest ascent)方向.

证：记 $\boldsymbol{g}'=\nabla f(\boldsymbol{x}')$. 因为函数可微，由方向导数与梯度的关系知函数沿方向 $\boldsymbol{p}$ 的方向导数，即斜率为 $\boldsymbol{p}^T\boldsymbol{g}'$. 设 θ 为方向向量 $\boldsymbol{p}$ 与梯度向量 $\boldsymbol{g}'$ 的夹角，则由向量夹角的定义和 $\|\boldsymbol{p}\|_2=1$ ，有

$$\boldsymbol{p}^T\boldsymbol{g}'=\|\boldsymbol{p}^T\|_2\|\boldsymbol{g}'\|_2\cos\theta\le\|\boldsymbol{p}^T\|_2\|\boldsymbol{g}'\|_2=\|\boldsymbol{g}'\|_2,$$

其中等式成立当且仅当 $\theta=0$ ，即 $\boldsymbol{p}$ 与梯度向量 $\boldsymbol{g}'$ 同方向. 又因为 $\boldsymbol{p}$ 为单位向量，所以当 $\boldsymbol{p}=\boldsymbol{g}'/\|\boldsymbol{g}'\|_2$ 时，函数沿该方向的斜率(也即方向导数)最大.

第二章　线性规划: 基本理论与方法

习题设计说明：

1. 化标准形练习：习题2.1-习题2.3，其中习题2.2和习题2.3是最优化中常用的两种重新表述技巧，这两种技巧的应用和进一步说明分别见习题2.27和习题7.19.

2. 基本解、基本可行解、退化基本可行解的练习：习题2.4, 习题2.5，习题2.6，习题2.7，教材第25页的例2.2.3.

3. 习题2.8, 习题2.9、习题2.12(b)是为了理解使用单纯形法时，如何根据单纯形表的数据判断线性规划何时有惟一解，何时有多解. 如果有多解时，如何得到多个解. 结论是：最优解不惟一时，某基本可行解的非基变量的相对费用系数非负，并且至少有一个非基变量的相对费用系数是零. 此外，习题2.30 说明，当原始问题的最优解是对偶非退化的(即非基变量的既约费用系数严格大于零)，对偶问题有惟一解；否则，对偶问题有多个极点解，进而有无穷多个解(这些极点解的凸组合都是原始问题的解).

4. 单纯形法的练习：习题2.10，习题2.11，习题2.12，习题2.13，习题2.20(说明单纯形法的效率的一般性例子中，自变量为三个时所得问题)，习题2.21(说明单纯形法采用最小相对费用系数进基原则确定进基变量时，如果所求解问题是退化的，则单纯形法会出现循环！)，习题2.31.

5. 两阶段法的练习：习题2.14-习题2.16；大 M 法的练习：习题2.18.

6. 修正单纯形法的练习：习题2.17，习题；单纯形法的矩阵表示：2.19.

7. 习题2.11，习题2.12(c)，2.32是关于灵敏度分析的练习，这也可以看成是单纯形法的应用，是难点.

8. 关于对偶性的练习：习题2.22－习题2.36.

2.1 将下面的线性规划问题化成标准形，并求解第 3 个问题(c)：

(c)

$$\begin{aligned} \text{minimize} \quad & x_1 + 4x_2 + x_3 \\ \text{subject to} \quad & x_1 - 2x_2 + x_3 = 4 \\ & x_1 - x_3 = 1 \\ & x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

解:

(c) 由于变量 x_1 无限制，可利用约束 $x_1 = x_3 + 1$ 对其消去. 因此，得其标准形

$$\begin{aligned} \text{minimize} \quad & 4x_2 + 2x_3 \\ \text{subject to} \quad & -2x_2 + 2x_3 = 3 \\ & x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

再把约束 $2x_3 = 3 + 2x_2$ 代入目标函数，得$6x_2 + 4$, 又因为 $x_2 \geq 0$ ，所以其最小值为 4，最优解为 $x_1 = 2.5, x_2 = 0, x_3 = 1.5$.

2.2 将下面的问题化成线性规划

$$\begin{aligned} \text{minimize} \quad & |x|+|y|+|z| \\ \text{subject to} \quad & x+y\leq 1 \\ & 2x+z=3. \end{aligned}$$

方法1: 令 $x=u_1-v_1, |x|=u_1+v_1, u_1\geq 0, v_1\geq 0$, 类似地表示 y 和 z ，则可将原问题重新编述为

$$\begin{aligned} \text{minimize} \quad & u_1+u_2+u_3+v_1+v_2+v_3 \\ \text{subject to} \quad & u_1-v_1+u_2-v_2+s=1, \\ & 2u_1-2v_1+u_3-v_3=3, \\ & u_i, v_i, s\geq 0, i=1,2,3. \end{aligned}$$

方法2: 引入非负变量 t_1, t_2, t_3 ，将原问题转化成等价问题

$$\begin{aligned} \text{minimize} \quad & t_1+t_2+t_3 \\ \text{subject to} \quad & x+y\leq 1, \\ & 2x+z=3, \\ & |x|=t_1, |y|=t_2, |z|=t_3. \end{aligned}$$

该问题的最优值与

$$\begin{aligned} \text{minimize} \quad & t_1+t_2+t_3 \\ \text{subject to} \quad & x+y\leq 1, \\ & 2x+z=3, \\ & |x|\leq t_1, |y|\leq t_2, |z|\leq t_3. \end{aligned}$$

的最优值相同，将这个问题的最优解投影到 (x,y,z) 所在的空间可以得到原问题的解. 这个问题可以写成线性规划问题：

$$\begin{aligned} \text{minimize} \quad & t_1+t_2+t_3 \\ \text{subject to} \quad & x+y\leq 1, \\ & 2x+z=3, \\ & -t_1\leq x\leq t_1, \\ & -t_2\leq y\leq t_2, \\ & -t_3\leq z\leq t_3. \end{aligned}$$

2.3 一类逐段线性函数$f(\boldsymbol{x}) = \max\{\boldsymbol{c}_1^T\boldsymbol{x}+d_1, \boldsymbol{c}_2^T\boldsymbol{x}+d_2, \cdots, \boldsymbol{c}_p^T\boldsymbol{x}+d_p\}$，其中$\boldsymbol{c}_i\in\mathbb{R}^n, d_i\in\mathbb{R}, i=1,\cdots,p$. 针对这样的函数，考虑问题

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b} \\ & \boldsymbol{x}\geq \boldsymbol{0}. \end{aligned}$$

将此问题化成线性规划.

解: 引入变量 t，所给问题等价于

$$\begin{aligned}\text{minimize}\quad & t\\ \text{subject to}\quad & f(\boldsymbol{x}) = t,\\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\\ & \boldsymbol{x} \geq \boldsymbol{0}.\end{aligned}$$

考虑问题

$$\begin{aligned}\text{minimize}\quad & t\\ \text{subject to}\quad & f(\boldsymbol{x}) \leq t,\\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\\ & \boldsymbol{x} \geq \boldsymbol{0},\end{aligned}$$

因为该问题关于 t 最小化，故将最优解代入第一个不等式，必有等号成立，即问题的最优解和最优值与上一个问题的相同. 从而所给问题等价于线性规划问题

$$\begin{aligned}\text{minimize}\quad & t\\ \text{subject to}\quad & \boldsymbol{c}_i^T\boldsymbol{x} + d_i \leq t, \quad i = 1, \cdots, p,\\ & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b},\\ & \boldsymbol{x} \geq \boldsymbol{0}.\end{aligned}$$

2.5 考虑问题

$$\begin{aligned}\text{minimize}\quad & c_1x_1 + c_2x_2 + c_3x_3\\ \text{subject to}\quad & x_1 + x_2 + x_3 \leq 4\\ & x_1 \leq 2\\ & x_3 \leq 3\\ & 3x_2 + x_3 \leq 6\\ & x_1, x_2, x_3 \geq 0.\end{aligned}$$

注意系数 c_1, c_2, c_3 尚未确定. 表示成标准形 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ 后, 其中

$$\boldsymbol{A} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 1 & 0\\ 0 & 3 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}, \ \boldsymbol{x} = \begin{pmatrix} x_1\\ x_2\\ x_3\\ x_4\\ x_5\\ x_6\\ x_7 \end{pmatrix}, \ \boldsymbol{b} = \begin{pmatrix} 4\\ 2\\ 3\\ 6 \end{pmatrix};$$

记 $\boldsymbol{A}$ 的第 i 列为 $\boldsymbol{a}_i$.

(a) 画出所给问题的可行域(三维空间中).

(b) 点 $(0,1,3,0,2,0,0)^T$ 是基本可行解吗？

(c) 点 $(0,1,3,0,2,0,0)^T$ 是退化基本可行解吗？如果是的话，找出可能的与其对应的基.

解:

(a) 略.

(b) 是基本可行解，因为满足约束条件，且非零元素对应列 $\boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_5$ 线性无关.

(c) 是退化的基本可行解，因为非零元素个数是3，小于系数矩阵 $\boldsymbol{A}$ 的秩4; 共有四个基与该基本可行解对应，他们是 $\boldsymbol{B} = [\boldsymbol{a}_2 \ \ \boldsymbol{a}_3 \ \ \boldsymbol{a}_5 \ \ \boldsymbol{a}_j]$ ，其中 $j = 1, 4, 6, 7$.

2.8 如果与每个非基变量 x_j 对应的既约费用系数 $r_j > 0$ ，证明与其对应的基本可行解是唯一的最优解.

证明: 不妨设满足条件的基本可行解 $\boldsymbol{x}^*$ 对应的基 $\boldsymbol{B}$ 为系数矩阵 $\boldsymbol{A}$ 的前 m 列，即 $\boldsymbol{x}^* = (x_1^*, \ldots, x_m^*, 0, 0, \ldots 0)^T$ ，且 $r_j > 0, j = m+1, \cdots, n$. 则对所有可行的 $\boldsymbol{x}$ ，有

$$\boldsymbol{c}^T\boldsymbol{x} - \boldsymbol{c}^T\boldsymbol{x}^* = \sum_{j=m+1}^{n} r_j x_j \geq 0.$$

设 $\bar{\boldsymbol{x}}$ 是另一个最优解，则必有 $\boldsymbol{c}^T\bar{\boldsymbol{x}} - \boldsymbol{c}^T\boldsymbol{x}^* = 0$. 由于诸 $r_j > 0$ ，而 $\bar{x}_j \geq 0$ ，由上式，对 $j = m+1, \ldots, n$ 有 $\bar{x}_j = 0$ ；

此外，因为 $\bar{\boldsymbol{x}}$ 满足 $\sum_{j=1}^{n} \boldsymbol{a}_j \bar{x}_j = \boldsymbol{b}$，将 $\bar{x}_j = 0, j = m+1, \ldots, n$ 代入，得 $\sum_{j=1}^{m} \boldsymbol{a}_j \bar{x}_j = \boldsymbol{b}$. 再由 $\boldsymbol{x}^*$ 的可行性，也有 $\sum_{j=1}^{m} \boldsymbol{a}_j x_j^* = \boldsymbol{b}$. 而 $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_m$ 线性无关，从而有 $\bar{x}_j = x_j^*, j = 1, \cdots, m$.

综上，问题的任一最优解均和所给解 $\boldsymbol{x}^*$ 相同，从而满足条件的基本可行解是问题的惟一最优解.

2.9 举例说明退化基本可行解不用满足所有 $r_j \geq 0$ 也可以是最优的.

解: 考虑问题

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^3}{\text{minimize}} \quad & -x_1 - x_2 \\ \text{subject to} \quad & x_1 + x_2 + x_3 = 0 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \end{aligned}$$

显然可行解只有 $\boldsymbol{x} = (0,0,0)^T$ ，这也是最优解。但若用单纯形法来求解，例如选取 x_3 为基变量,则表格为

	$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{b}$
	1	1	1	0
$\boldsymbol{r}^T$	-1	-1	0	0

此时非基变量所对应的相对费用系数都是负的，但可见这已经达到最优.

2.10 将下面的问题转化成标准形，用单纯形法求解，然后画出问题在 x_1, x_2 空间的可行域，并标明单纯形法的迭代路径.

(b)

$$\begin{aligned} \text{maximize} \quad & x_1 + x_2 \\ \text{subject to} \quad & -2x_1 + x_2 \leq 1 \\ & x_1 - x_2 \leq 1 \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

解:

(b) 引入松驰变量 x_3, x_4，化为标准形

$$\begin{aligned} \text{minimize} \quad & -x_1 - x_2 \\ \text{subject to} \quad & -2x_1 + x_2 + x_3 = 1, \\ & x_1 - x_2 + x_4 = 1, \\ & x_i \geq 0, i = 1,2,3,4. \end{aligned}$$

写出初始表，其已是第一张单纯形表

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_3	-2	1	1	0	1
x_4	$\boxed{1}$	-1	0	1	1
$\boldsymbol{c}^T(\boldsymbol{r}^T)$	-1	-1	0	0	1

$\rightarrow$

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_3	0	-1	1	2	3
x_1	1	-1	0	1	1
$\boldsymbol{r}^T$	0	-2	0	1	2

因为 x_2 对应列无正元素, 所以原问题是无界的.

2.11 (a) 利用单纯形法求解

$$\begin{aligned} \text{maximize} \quad & 2x_1 + 4x_2 + x_3 + x_4 \\ \text{subject to} \quad & x_1 + 3x_2 + x_4 \leq 4 \\ & 2x_1 + x_2 \leq 3 \\ & x_2 + 4x_3 + x_4 \leq 3 \\ & x_i \geq 0, i = 1,2,3,4. \end{aligned}$$

利用(a)中的求解结果回答以下问题：

(b) 为使最优基保持不变，给出 $\boldsymbol{b} = (4,3,3)^T$ 中第一个元素的可变范围(其它的保持不变)；

(c) 为使最优基保持不变，给出 $\boldsymbol{c} = (2,4,1,1)^T$ 中第一个元素的可变范围(其它的保持不变)；第四个的？

(d) 对于 $\boldsymbol{b}$ 微小的改变，最优解将发生怎样的改变？

(e) 对于 $\boldsymbol{c}$ 微小的改变，最优值将发生怎样的改变？

解: 将原问题化为标准形，得初始表，并依次演算，得

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	$\boldsymbol{x_B}$
x_5	1	3	0	1	1	0	0	4
x_6	$\boxed{2}$	1	0	0	0	1	0	3
x_7	0	1	4	1	0	0	1	3
$\boldsymbol{c}^T(\boldsymbol{r}^T)$	-2	-4	-1	-1	0	0	0	0

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	$\boldsymbol{x_B}$
x_4	0	$\boxed{5/2}$	0	1	1	$-1/2$	0	$5/2$
x_1	1	$1/2$	0	0	0	$1/2$	0	$3/2$
x_6	0	1	4	1	0	0	1	3
$\boldsymbol{r}^T$	0	-3	-1	-1	0	1	0	3

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_B
	x_2	0	1	0	2/5	2/5	$-1/5$	0	1
$\rightarrow$	x_1	1	0	0	$-1/5$	$-1/5$	3/5	0	1
	x_6	0	0	$\boxed{4}$	3/5	$-2/5$	1/5	1	2
	$\boldsymbol{r}^T$	0	0	-1	1/5	6/5	2/5	0	6

		x_1	x_2	x_3	x_4	x_5	x_6	x_7	$\boldsymbol{x_B}$
	x_2	0	1	0	2/5	2/5	$-1/5$	0	1
$\rightarrow$	x_1	1	0	0	$-1/5$	$-1/5$	3/5	0	1
	x_3	0	0	1	3/20	$-1/10$	1/20	1/4	1/2
	$\boldsymbol{r}^T$	0	0	0	7/20	11/10	9/20	1/4	13/2

至此，所有 $r_j \geq 0$, 得到最优解 $\boldsymbol{x}^* = (1,1,1/2,0)^T$. 最优基 $\boldsymbol{B} = [\boldsymbol{a}_2\ \boldsymbol{a}_1\ \boldsymbol{a}_3]$. 因为初始单纯形表的最后三列是单位矩阵，故

$$\boldsymbol{B}^{-1} = \begin{bmatrix} 2/5 & -1/5 & 0 \\ -1/5 & 3/5 & 0 \\ -1/10 & 1/20 & 1/4 \end{bmatrix}.$$

(a) 设 $\boldsymbol{b}$ 的改变量为 $\Delta\boldsymbol{b} = (\Delta b_1, \Delta b_2, \Delta b_3)^T$. 要使最优基不变, 此时相对费用系数保持不变，故仅需要 $\boldsymbol{B}^{-1}(\boldsymbol{b}+\Delta\boldsymbol{b}) = \boldsymbol{x_B} + \boldsymbol{B}^{-1}\Delta\boldsymbol{b} \geq \boldsymbol{0}$, 即

$$\begin{bmatrix} 1 \\ 1 \\ \frac{1}{2} \end{bmatrix} + \begin{bmatrix} 2/5 & -1/5 & 0 \\ -1/5 & 3/5 & 0 \\ -1/10 & 1/20 & 1/4 \end{bmatrix} \begin{bmatrix} \Delta b_1 \\ \Delta b_2 \\ \Delta b_3 \end{bmatrix} \geq \boldsymbol{0}. \tag{2.1}$$

当 $\boldsymbol{b}$ 的第一个分量发生变化时，$\Delta b_2 = \Delta b_3 = 0$. 将此代入(2.1)，可得第一个分量的变化范围为 $\Delta b_1 \in [-5/2, 5]$, 相应的，分量 b_1 的取值范围为 $b_1 \in [3/2, 9]$.

(b) 设 $\boldsymbol{c}$ 的改变量为 $\Delta\boldsymbol{c} = (\Delta c_1, \Delta c_2, \Delta c_3, \Delta c_4)^T$. 要使最优基不变, 基变量的取值不变，仅需要 $(-\boldsymbol{c_N} - \Delta\boldsymbol{c_N})^T - (-\boldsymbol{c_B} - \Delta\boldsymbol{c_B})^T\boldsymbol{B}^{-1}\boldsymbol{N} \geq \boldsymbol{0}$ (请注意该问题是 max). 而

$$\begin{aligned} & (-\boldsymbol{c_N} - \Delta\boldsymbol{c_N})^T + (\boldsymbol{c_B} + \Delta\boldsymbol{c_B})^T\boldsymbol{B}^{-1}\boldsymbol{N} \\ = \ & (-\boldsymbol{c}_{\boldsymbol{N}}^T + \boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{B}^{-1}\boldsymbol{N}) - \Delta\boldsymbol{c}_{\boldsymbol{N}}^T + \Delta\boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{B}^{-1}\boldsymbol{N} \\ = \ & \boldsymbol{r}_{\boldsymbol{N}}^T - \Delta\boldsymbol{c}_{\boldsymbol{N}}^T + \Delta\boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{B}^{-1}\boldsymbol{N}. \end{aligned}$$

将所需条件写成分量形式，再由 $\boldsymbol{y}_j = \boldsymbol{B}^{-1}\boldsymbol{a}_j$ ，得

$$r_j - \Delta c_j + \Delta\boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{y}_j \geq 0, \quad j = 4,5,6,7. \tag{2.2}$$

由(a)中的最优单纯形表读出 $r_j, \boldsymbol{y}_j, j = 4,5,6,7$ ，代入不等式组(2.2). 当 $\boldsymbol{c}$ 只有一个分量改变时, 令其他分量的改变量为零，解不等式组(2.2)，可得各分量的变化范围. 具体地，对第一和第四个分量分别有

$$\Delta c_1 \in [-3/4, 7/4],\ \Delta c_4 \in (-\infty, 7/20]\ c_1 \in [5/4, 15/4],\ c_4 \in (-\infty, 27/20]..$$

(c) 当 $\boldsymbol{b}$ 改变时，新的基本解是 $\boldsymbol{B}^{-1}(\boldsymbol{b}+\Delta\boldsymbol{b}) = \boldsymbol{B}^{-1}\boldsymbol{b}+\boldsymbol{B}^{-1}\Delta\boldsymbol{b}$，因为 $\boldsymbol{B}^{-1}\boldsymbol{b} > 0$，故当 $\boldsymbol{b}$ 变化不大时，新的基本解仍然是可行的，从而也是最优的. 这时，最优解的改变量

$$\boldsymbol{B}^{-1}\Delta\boldsymbol{b} = \left(\frac{2}{5}\Delta b_1 - \frac{1}{5}\Delta b_2, -\frac{1}{5}\Delta b_1 + \frac{3}{5}\Delta b_2, -\frac{1}{10}\Delta b_1 + \frac{1}{20}\Delta b_2 + \frac{1}{4}\Delta b_3\right)^T.$$

(d) 当 $\boldsymbol{c}$ 改变时，新的相对费用系数见(2.2). 因为 $r_j > 0, j = 4,5,6,7$，故当 $\boldsymbol{c}$ 的改变量不大时，相对费用系数仍然是非负的，从而原来的最优解也是最优的. 此时，新的最优值是 $(-\boldsymbol{c_B} - \Delta\boldsymbol{c_B})^T\boldsymbol{B}^{-1}\boldsymbol{b}$，最优值的改变量

$$-\Delta\boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{B}^{-1}\boldsymbol{b} = -\Delta\boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{x}_{\boldsymbol{B}}^* = -\Delta c_1 - \Delta c_2 - \frac{1}{2}\Delta c_3.$$

2.12 考虑问题

$$\begin{aligned} \text{minimize} \quad & x_1 - 3x_2 - 0.4x_3 \\ \text{subject to} \quad & 3x_1 - x_2 + 2x_3 \le 7 \\ & -2x_1 + 4x_2 \le 12 \\ & -4x_1 + 3x_2 + 3x_3 \le 14 \\ & x_1 \ge 0, x_2 \ge 0, x_3 \ge 0. \end{aligned}$$

(a) 找出一个最优解.

(b) 存在多少个最优基本可行解？

(c) 证明：如果 $c_4 + \frac{1}{5}a_{14} + \frac{4}{5}a_{24} \ge 0$，则以费用系数 c_4 和系数向量 $(a_{14}, a_{24}, a_{34})^T$ 引入另一个变量 x_4 后，最优解仍保持不变.

解: (a) 引入松驰变量 x_4, x_5, x_6 化为标准形后，得初始表，其也是第一张单纯形表，然后依次演算，

	x_1	x_2	x_3	x_4	x_5	x_6	$x_{\boldsymbol{B}}$
x_4	3	−1	2	1	0	0	7
x_5	−2	$\boxed{4}$	0	0	1	0	12
x_6	−4	3	3	0	0	1	14
$\boldsymbol{c}^T(\boldsymbol{r}^T)$	1	−3	−0.4	0	0	0	0

→

	x_1	x_2	x_3	x_4	x_5	x_6	$x_{\boldsymbol{B}}$
x_4	$\boxed{5/2}$	0	2	1	1/4	0	10
x_2	−1/2	1	0	0	1/4	0	3
x_6	−5/4	0	3	0	−3/4	1	5
$\boldsymbol{r}^T$	−1/2	0	−0.4	0	3/4	0	9

→

	x_1	x_2	x_3	x_4	x_5	x_6	$x_{\boldsymbol{B}}$
x_1	1	0	4/5	2/5	1/10	0	4
x_2	0	1	2/5	1/5	3/10	0	5
x_6	0	0	5	1	−1/2	1	15
$\boldsymbol{r}^T$	0	0	0	1/5	4/5	0	11

因为 $r_j \geq 0$, 得到一个最优解 $(4,5,0)^T$, 且最优值 $f^* = -11$.

(b) 在 (a) 的最后一张单纯形表中，因为 $r_3 = 0$ ，令 x_3 进基, 由最小正比率法则，知 x_6 出基, 即以 5 为转轴元转轴后，得

	x_1	x_2	x_3	x_4	x_5	x_6	$x_{\boldsymbol{B}}$
x_1	1	0	0	6/25	9/50	−4/25	8/5
x_2	0	1	0	3/25	8/25	−2/25	19/5
x_3	0	0	1	1/5	−1/10	1/5	3
$\boldsymbol{r}^T$	0	0	0	1/5	4/5	0	11

因为 $r_j \geq 0$, 所以又得到一个最优基本可行解 $(8/5, 19/5, 3)^T$, 且目标值仍为 $f^* = -11$.

(c) 将 c_4 和 (a_{14}, a_{24}, a_{34}) 代入单纯形表，得最后的单纯形表

	$\boldsymbol{a}_1$	$\boldsymbol{a}_2$	$\boldsymbol{a}_3$	$\boldsymbol{a}_4$	$\boldsymbol{a}_5$	$\boldsymbol{a}_6$	$\boldsymbol{a}_7$	$\boldsymbol{B}^{-1}b$
	1	0	4/5	$(2/5)a_{14} + (1/10)a_{24}$	2/5	1/10	0	4
	0	1	2/5	$(1/5)a_{14} + (3/10)a_{24}$	1/5	3/10	0	5
	0	0	5	$a_{14} + a_{34} - (1/2)a_{24}$	1	−1/2	1	15
$\boldsymbol{r}^T$	0	0	0	$(1/5)a_{14} + (4/5)a_{24} + c_4$	1/5	4/5	0	11

如 $c_4 + (1/5)a_{14} + (4/5)a_{24} \geq 0$, 则上表达到最优, 且最优解保持不变. 也可以由(a)中的最后一张表读出最优基 B 的逆，然后由 $\boldsymbol{y}_4 = \boldsymbol{B}^{-1}\boldsymbol{a}_4$和 $r_4 = c_4 - \boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{y}_4$ 算出单纯形表中与 x_4 对应的数据，然后得出结论.

2.16 在两阶段法的第 I 阶段，假定给系统 $\boldsymbol{Ax} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ 的辅助问题应用单纯形法后，所得表格(忽略费用行)形如

$$\begin{array}{ccc|ccc|ccc|ccc|c}
x_1 & & x_k & x_{k+1} & \cdots & x_n & y_1 & \cdots & y_k & y_{k+1} & \cdots & y_m & \\ \hline
1 & & & & & & & & & 0 & \cdots & 0 & b'_1 \\
& \ddots & & & \boldsymbol{R}_1 & & & \boldsymbol{S}_1 & & \vdots & & \vdots & \vdots \\
& & 1 & & & & & & & 0 & \cdots & 0 & b'_k \\ \hline
0 & \cdots & 0 & & & & & & & 1 & & & 0 \\
\vdots & & \vdots & & \boldsymbol{R}_2 & & & \boldsymbol{S}_2 & & & \ddots & & \vdots \\
0 & \cdots & 0 & & & & & & & & & 1 & 0 \\ \hline
\end{array}$$

即基变量中有 $m - k$ 个人工变量，它们取零值.

(a) 证明 $\boldsymbol{R}_2$ 中的任何非零元素都可作为转轴元以消去人工基变量，这样将产生一个类似的表格，但 k 会增加 1.

(b) 重复(a)中的过程，直到 $\boldsymbol{R}_2 = \boldsymbol{0}$. 证明原始系统是冗余的，并说明可以删除底端的这些行，然后继续第 II 阶段.

(c) 利用上面的方法(即两阶段法)求解线性规划

$$\begin{aligned}\text{minimize}\quad & 2x_1+6x_2+x_3+x_4\\ \text{subject to}\quad & x_1+2x_2+x_4=6\\ & x_1+2x_2+x_3+x_4=7\\ & x_1+3x_2-x_3+2x_4=7\\ & x_1+x_2+x_3=5\\ & x_i\ge 0, i=1,2,3,4.\end{aligned}$$

解： (c) 第 I 阶段：引入人工变量 y_1,y_2,y_3,y_4 ，构造辅助问题

$$\begin{aligned}\text{minimize}\quad & y_1+y_2+y_3+y_4\\ \text{subject to}\quad & x_1+2x_2+x_4+y_1=6,\\ & x_1+2x_2+x_3+x_4+y_2=7,\\ & x_1+3x_2-x_3+2x_4+y_3=7,\\ & x_1+x_2+x_3+y_4=5,\\ & x_i\ge 0, y_i\ge 0, i=1,2,3,4.\end{aligned}$$

辅助问题的初始表为

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	$\boldsymbol{x_B}$
y_1	1	2	0	1	1	0	0	0	6
y_2	1	2	1	1	0	1	0	0	7
y_3	1	3	-1	2	0	0	1	0	7
y_4	1	1	1	0	0	0	0	1	5
$\boldsymbol{c}^T$	0	0	0	0	1	1	1	1	0

将最后一行与基变量对应的系数化为零，得第一张单纯形表后，因此进行转轴运算，得

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	$\boldsymbol{x_B}$
y_1	1	2	0	1	1	0	0	0	6
y_2	1	2	1	1	0	1	0	0	7
y_3	1	$\boxed{3}$	-1	2	0	0	1	0	7
y_4	1	1	1	0	0	0	0	1	5
$\boldsymbol{r}^T$	-4	-8	-1	-4	0	0	0	0	-25

$\to$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	$\boldsymbol{x_B}$
y_1	$1/3$	0	$2/3$	$-1/3$	1	0	$-2/3$	0	$4/3$
y_2	$1/3$	0	$\boxed{5/3}$	$-1/3$	0	1	$-2/3$	0	$7/3$
x_2	$1/3$	1	$-1/3$	$2/3$	0	0	$1/3$	0	$7/3$
y_4	$2/3$	0	$4/3$	$-2/3$	0	0	$-1/3$	1	$8/3$
$\boldsymbol{r}^T$	$-4/3$	0	$-11/3$	$4/3$	0	0	$8/3$	0	$-19/3$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	$\boldsymbol{x_B}$
y_1	$\boxed{1/5}$	0	0	$-1/5$	1	$-2/5$	$-2/5$	0	$2/5$
x_3	$1/5$	0	1	$-1/5$	0	$3/5$	$-2/5$	0	$7/5$
x_2	$2/5$	1	0	$3/5$	0	$1/5$	$1/5$	0	$14/5$
y_4	$2/5$	0	0	$-2/5$	0	$-4/5$	$1/5$	1	$4/5$
$\boldsymbol{r}^T$	$-3/5$	0	0	$3/5$	0	$11/5$	$6/5$	0	$-6/5$

$\rightarrow$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	$\boldsymbol{x_B}$
x_1	1	0	0	-1	5	-2	-2	0	2
x_3	0	0	1	0	-1	1	0	0	1
x_2	0	1	0	1	-2	1	1	0	2
y_4	0	0	0	0	-2	0	1	1	0
$\boldsymbol{r}^T$	0	0	0	0	3	1	0	0	0

因为该单纯形表第 4 行与原问题数据对应的数全为零，从而不能再次转轴将人工变量 y_4 赶出基；此时，第 4 个约束为冗余的, 将其删除后，得基本可行解 $(2,2,1,0)^T$.

第 II 阶段：以上述基本可行解作为初始基本可行解，利用单纯形法求解原问题. 首先得到如下初始表

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_1	1	0	0	-1	2
x_3	0	0	1	0	1
x_2	0	1	0	1	2
$\boldsymbol{c}^T$	2	6	1	1	0

将最后一行与基变量对应的系数化为零后，得第一张单纯形表后，因此进行转轴运算，得

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_1	1	0	0	-1	2
x_3	0	0	1	0	1
x_2	0	1	0	$\boxed{1}$	2
$\boldsymbol{r}^T$	0	0	0	-3	-17

$\rightarrow$

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_1	1	1	0	0	4
x_3	0	0	1	0	1
x_4	0	1	0	1	2
$\boldsymbol{r}^T$	0	3	0	0	-11

因为 $r_j \geq 0$, 所以得最优解 $\boldsymbol{x}^* = (4,0,1,2)^T$ ， $f^* = 11$.

2.17 利用修正单纯形法找出下列系统的一个基本可行解

$$
\begin{aligned}
&x_1+2x_2-x_3+x_4=3\\
&2x_1+4x_2+x_3+2x_4=12\\
&x_1+4x_2+2x_3+x_4=9\\
&x_i\geq 0, i=1,2,3,4.
\end{aligned}
$$

解： 引入人工变量 x_5,x_6,x_7 ，构造辅助问题. 得辅助问题的如下表格

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$
x_5	1	0	0	3
x_6	0	1	0	12
x_7	0	0	1	9

由此表格得 $\boldsymbol{\lambda}^T=(1,1,1)$, $\boldsymbol{r}_{\boldsymbol{N}}^T=\boldsymbol{c_N}-\boldsymbol{\lambda}^T\boldsymbol{N}=(-4,-10,-2.4)$. 让 x_2 进基，计算 $\boldsymbol{y}_2=\boldsymbol{B}^{-1}\boldsymbol{a}_2=(2,4,4)^T$ ，得首张修正单纯形表，并转轴，得

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$	$\boldsymbol{y}_2$
x_5	1	0	0	3	$\boxed{2}$
x_6	0	1	0	12	4
x_7	0	0	1	9	4
$\boldsymbol{\lambda}^T$	1	1	1	24	10

$\rightarrow$

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$
x_2	1/2	0	0	3/2
x_6	−2	1	0	6
x_7	−2	0	1	3
$\boldsymbol{\lambda}^T$	−4	1	1	9

计算相对费用系数，得 $r_1=1,r_3=-7,r_4=1,r_5=5$. 选取 x_3 进基，计算 $\boldsymbol{y}_3=\boldsymbol{B}^{-1}\boldsymbol{a}_3=(-1/2,3,4)^T$. 故有

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$	$\boldsymbol{y}_3$
x_2	1/2	0	0	3/2	−1/2
x_6	−2	1	0	6	3
x_7	−2	0	1	3	$\boxed{4}$
$\boldsymbol{\lambda}^T$	−4	1	1	9	7

$\rightarrow$

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$
x_2	1/4	0	1/8	15/8
x_6	−1/2	1	−3/4	15/4
x_3	−1/2	0	1/4	3/4
$\boldsymbol{\lambda}^T$	−1/2	1	−3/4	15/4

计算相对费用系数，得 $r_1=-3/4,r_4=-3/4,r_5=3/2,r_7=7/4$. 选取 x_1 进基，计算 $\boldsymbol{y}_1=\boldsymbol{B}^{-1}\boldsymbol{a}_1=(3/8,3/4,-1/4)^T$. 故有

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$	$\boldsymbol{y}_1$
x_2	1/4	0	1/8	15/8	3/8
x_6	−1/2	1	−3/4	15/4	$\boxed{3/4}$
x_3	−1/2	0	1/4	3/4	−1/4
$\boldsymbol{\lambda}^T$	−1/2	1	−3/4	15/4	3/4

$\rightarrow$

	$\boldsymbol{B}^{-1}$			$\boldsymbol{x_B}$
x_2	1/2	−1/2	1/2	0
x_1	−2/3	4/3	−1	5
x_3	−2/3	1/3	0	2
$\boldsymbol{\lambda}^T$	0	0	0	0

计算相对费用系数，得 $r_4=0,r_5=r_6=r_7=1$. 至此得到辅助问题的最优基本可行解. 因为基变量不含人工变量，所以得原问题的基本可行解 $\boldsymbol{x}=(5,0,2,0)^T$.

2.20 利用单纯形法求解问题(2.2.14)，其中初始点 $\boldsymbol{x}^{(0)}=(0,0,0)^T$ ，要求每次迭代选既约费用系数最负的变量进基.

解: 将原始问题化为标准形,利用单纯形法求解，因此得如下单纯形表格

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_4	$\boxed{1}$	0	0	1	0	0	1
x_5	4	1	0	0	1	0	100
x_6	8	4	1	0	0	1	10000
$\boldsymbol{r}^T$	-4	-2	-1	0	0	0	0

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_1	1	0	0	1	0	0	1
x_5	0	$\boxed{1}$	0	-4	1	0	96
x_6	0	4	1	-8	0	1	9992
$\boldsymbol{r}^T$	0	-2	-1	4	0	0	4

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_1	1	0	0	$\boxed{1}$	0	0	1
x_2	0	1	0	-4	1	0	96
x_6	0	0	1	8	-4	1	9608
$\boldsymbol{r}^T$	0	0	-1	-4	2	0	196

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_4	1	0	0	1	0	0	1
x_2	4	1	0	0	1	0	100
x_6	-8	0	$\boxed{1}$	0	-4	1	9600
$\boldsymbol{r}^T$	4	0	-1	0	2	0	200

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_4	$\boxed{1}$	0	0	1	0	0	1
x_5	4	1	0	0	1	0	100
x_6	-8	0	1	0	-4	1	9600
$\boldsymbol{r}^T$	-4	0	0	0	-2	1	9800

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_1	1	0	0	1	0	0	1
x_2	0	1	0	-4	$\boxed{1}$	0	96
x_3	0	0	1	8	-4	1	9608
$\boldsymbol{r}^T$	0	0	0	4	-2	1	9804

		x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
	x_1	1	0	0	$\boxed{1}$	0	0	1
$\longrightarrow$	x_5	0	1	0	-4	1	0	96
	x_3	0	4	1	-8	0	1	9992
	$\boldsymbol{r}^T$	0	2	0	-4	0	1	9996

		x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
	x_4	1	0	0	1	0	0	1
$\longrightarrow$	x_5	4	1	0	0	1	0	100
	x_3	8	4	1	0	0	1	10000
	$\boldsymbol{r}^T$	4	2	0	0	0	1	10000

整个迭代遍历了原问题在三维空间中的超立方体可行域的全部 8 个顶点. 得最优解 $\boldsymbol{x}^* = (0, 0, 10000)^T$; 对应的目标函数值是 10000.

2.24 构造一个原始问题和对偶问题都没有可行解的例子.

解: 构造原问题如下

$$\begin{aligned} \text{minimize} \quad & -x_1 - x_2 \\ \text{subject to} \quad & x_1 \geq 1, \\ & -x_2 \geq 1, \\ & x_1 \geq 0, \quad x_2 \geq 0. \end{aligned}$$

其对偶问题为

$$\begin{aligned} \text{maximize} \quad & \lambda_1 + \lambda_2 \\ \text{subject to} \quad & \lambda_1 \leq -1, \\ & -\lambda_2 \leq -1, \\ & \lambda_1 \geq 0, \quad \lambda_2 \geq 0. \end{aligned}$$

易见二者均无可行解.

2.25 设 $\boldsymbol{A}$ 是 $m \times n$ 矩阵, $\boldsymbol{b}$ 是 n 维向量. 证明 $\boldsymbol{Ax} \leq \boldsymbol{0}$ 蕴含着 $\boldsymbol{c}^T\boldsymbol{x} \geq 0$ 当且仅当存在 $\boldsymbol{\lambda} \geq 0$ 使得 $\boldsymbol{c} + \boldsymbol{A}^T\boldsymbol{\lambda} = \boldsymbol{0}$. 给出该结论的一个几何解释.

证: 充分性. 假设存在 $\boldsymbol{\lambda} \geq \boldsymbol{0}$ ，使得 $\boldsymbol{c}^T + \boldsymbol{\lambda}^T\boldsymbol{A} = \boldsymbol{0}$ 成立，则对任意的 $\boldsymbol{x} \in \mathbb{R}^n$ ，有 $\boldsymbol{c}^T\boldsymbol{x} = -\boldsymbol{\lambda}^T(\boldsymbol{Ax})$ 成立. 如果 $\boldsymbol{x}$ 使得 $\boldsymbol{Ax} \leq \boldsymbol{0}$ ，则由 $\boldsymbol{\lambda} \geq \boldsymbol{0}$ 必有 $\boldsymbol{c}^T\boldsymbol{x} \geq 0$ 成立.

必要性. 构造如下线性规划问题，即

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}^T\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{Ax} \leq \boldsymbol{0}. \end{aligned}$$

它的对偶问题为

$$\begin{aligned} \text{maximize} \quad & 0 \\ \text{subject to} \quad & \boldsymbol{\lambda}^T\boldsymbol{A} = \boldsymbol{c}^T \\ & \boldsymbol{\lambda} \leq \boldsymbol{0}. \end{aligned}$$

易见 $\boldsymbol{x}=\boldsymbol{0}$ 是原始问题的一个可行解，且由 $\boldsymbol{Ax}\leq 0$ 蕴含着 $\boldsymbol{c}^T\boldsymbol{x}\geq 0$ 可知，原始问题的目标函数有下界. 根据弱对偶性，对偶问题也有可行解，即存在 $\boldsymbol{\lambda}'\leq\boldsymbol{0}$ 且满足 $(\boldsymbol{\lambda}')^T\boldsymbol{A}=\boldsymbol{c}^T$. 令 $\boldsymbol{\lambda}=-\boldsymbol{\lambda}'$ ，易见 $\boldsymbol{\lambda}$ 即为满足条件的向量.

注：这个结论即著名的Farkas引理，详见课本163页的引理7.3.4. 关于该结论的证明常见的有两种，一种即该作业的基于线性规划对偶理论的证明，还有一种就是课本上给出的基于点与闭凸锥的分离定理(引理7.3.5)的证明. 课本上给出的证明是构造性的，即给出了具体的满足要求的向量 $\boldsymbol{\lambda}$. 将该结论用在非线性规划的最优性条件的推导中，这个结论中的向量 $\boldsymbol{\lambda}$ 就是著名的Lagrange乘子. 此外，如果要用对偶性构造出所需的向量 $\boldsymbol{\lambda}$ ，需要用线性规划强对偶定理的证明，即37 页的定理2.3.2，那里也用了凸集分离定理.

几何解释：定义集合 $C=\{\boldsymbol{c}|\boldsymbol{c}=\boldsymbol{A}^T\boldsymbol{\lambda},\ \boldsymbol{\lambda}\geq\boldsymbol{0}\}$ ，此即 $\boldsymbol{A}$ 的行向量生成的锥集合(行向量的所有非负线性组合形成的集合)，以及 $C^\circ=\{\boldsymbol{x}|\boldsymbol{Ax}\leq 0\}$ ，此集合由与 C 中每个向量所夹角均不是锐角的向量组成，也称为 C 的**极锥**(polar cone). Farkas引理表明：如果向量 $\boldsymbol{c}$ 与 C° 中的每个向量的夹角为锐角或者直角，即满足

$$\boldsymbol{c}^T\boldsymbol{x}\geq 0,\ \ \forall\boldsymbol{x}\in C^\circ,$$

那么 $-\boldsymbol{c}$ 必属于 C ，反之亦成立. 令 $\boldsymbol{A}=\begin{bmatrix}1 & 1\\ 1 & 3\end{bmatrix}$ ，Farkas引理的几何解释见图2.1.

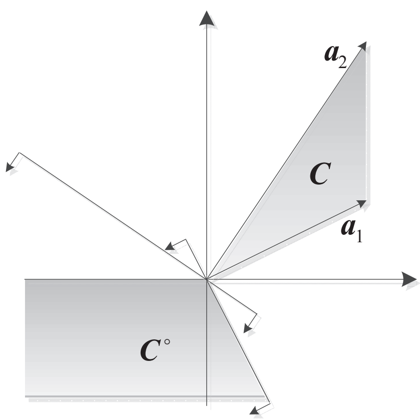


图 2.1: 习题2.25 Farkas引理的几何解释

Farkas引理的另一种描述：

系统I：$\boldsymbol{Ax}\leq\boldsymbol{0},\boldsymbol{c}^T\boldsymbol{x}<0$ 与

系统II：$\boldsymbol{\lambda}\geq\boldsymbol{0},\boldsymbol{c}^T+\boldsymbol{\lambda}^T\boldsymbol{A}=\boldsymbol{0}$

有且只能有一个有解.

系统I有解等价于存在向量 $\boldsymbol{x}$ ，它与 $\boldsymbol{A}$ 的每个行向量的夹角至少为 90° ，且它与向量 $-\boldsymbol{c}$ 的夹角小于 90° ；系统II有解等价于向量 $-\boldsymbol{c}$ 在由 $\boldsymbol{A}$ 的行向量生成的凸锥中.

基于此描述，Farkas引理的几何解释如下：给定一个由 $\boldsymbol{A}$ 的行向量生成的凸锥 C 和向量 $-\boldsymbol{c}$ ，要么向量 $-\boldsymbol{c}$ 属于锥，要么存在一个超平面

$$H:=\{y\in\mathbb{R}^n:\boldsymbol{y}^T\boldsymbol{x}^*=0\},\text{其中}\boldsymbol{x}^*\text{为系统I的解}$$

分离该凸锥和这个向量(即 C 在这个超平面的非正半空间，向量 $-\boldsymbol{c}$ 在这个超平面确定的正半空间)，且二者不会同时成立.

2.26 通常在优化理论和自由竞争之间有很强的联系，这可通过经营实体选址的理想模型进行说明. 假设存在 n 种经营实体(各种工厂、公司、商场等)，准备将它们单独地设在 n 块不同的土地上. 如果将实体 i 设在第 j 块上，能够产生 s_{ij} 单位(元)的价值.

如果给实体指派土地的工作由专家来作，一种作法是使生产价值最大来作出决定. 换句话说，确定的指派方式要极大化 $\sum_i \sum_j s_{ij}x_{ij}$ ，其中

$$x_{ij} = \begin{cases} 1, & \text{如果将实体 } i \text{ 指派到土地 } j \\ 0, & \text{否则.} \end{cases}$$

更明确地说，该方法需要求解优化问题

$$\begin{array}{ll} \text{maximize} & \sum_i \sum_j s_{ij}x_{ij} \\ \text{subject to} & \sum_j x_{ij} = 1, i = 1, 2, \cdots, n \\ & \sum_i x_{ij} = 1, j = 1, 2, \cdots, n \\ & x_{ij} \geq 0, x_{ij} \in \{0, 1\}, i, j = 1, 2, \cdots, n. \end{array}$$

实际上，能够证明：通过约束定义的集合的任一极点是自动满足最终的要求($x_{ij} = 0$ 或 1) 的，所以通过线性规划的单纯形法能够找到最优指派.

如果一个人从自由竞争的观点来考虑问题，假定不是由专家来确定指派，而是建立价格机制，由各个实体对土地进行投标.

(a) 证明：存在实体的价格 $p_i, i = 1, \cdots, n$ 和土地价格 $q_j, j = 1, \cdots, n$ 使得 $p_i + q_j \geq s_{ij}, i = 1, \cdots, n, \quad j = 1, \cdots, n$. 进一步，如果一个最优解指派实体 i 到土地 j ，在这组价格使上述不等式中的等号成立.

(b) 证明(a)蕴含着：如果最优解指派实体 i 到土地 j ，且如果 j' 是任一其他的土地，有 $s_{ij} - q_j \geq s_{ij'} - q_{j'}$. 给出该结论的一个经济解释，并以此为背景解释自由竞争和最优性之间的关系.

(c) 假定每一个 s_{ij} 是正的，请证明(a)中存在的价格是非负的.

解： (a) 令 $\boldsymbol{c} = (s_{11} \cdots s_{1n}\ s_{21} \cdots s_{2n} \cdots s_{n1} \cdots s_{nn})^T, \boldsymbol{y} = (x_{11} \cdots x_{1n}\ x_{21} \cdots x_{2n} \cdots x_{n1} \cdots x_{nn})^T$.

$$\boldsymbol{A} = \left(\begin{array}{cccc} 1\,1 \cdots 1 & & & \\ & 1\,1 \cdots 1 & & \\ & & \ddots & \\ & & & 1\,1 \cdots 1 \\ \hline 1\,0 \cdots 0 & 1\,0 \cdots 0 & \cdots & 1\,0 \cdots 0 \\ 0\,1 \cdots 0 & 0\,1 \cdots 0 & \cdots & 0\,1 \cdots 0 \\ \ddots & \ddots & \cdots & \ddots \\ 0\,0 \cdots 1 & 0\,0 \cdots 1 & \cdots & 0\,0 \cdots 1 \end{array}\right) = \left(\begin{array}{c} \boldsymbol{A}_1 \\ \hline \boldsymbol{A}_2 \end{array}\right) \quad \boldsymbol{b} = \left(\begin{array}{c} 1 \\ 1 \\ \vdots \\ 1 \\ \hline 1 \\ 1 \\ \vdots \\ 1 \end{array}\right) = \left(\begin{array}{c} \boldsymbol{b}_1 \\ \hline \boldsymbol{b}_2 \end{array}\right)$$

则所给的优化问题可写为

$$\begin{array}{ll} \text{maximize} & \boldsymbol{c}^T\boldsymbol{y} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{y} = \boldsymbol{b} \\ & \boldsymbol{y} \geq \boldsymbol{0}, \boldsymbol{y} \in \{0, 1\}^{n^2} \end{array} \tag{2.1}$$

这个问题的松弛问题(去掉整性约束)为

$$\begin{array}{ll}\text{maximize} & \boldsymbol{c}^T\boldsymbol{y} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{y}=\boldsymbol{b} \\ & \boldsymbol{y}\geq \boldsymbol{0}\end{array} \tag{2.2}$$

用单纯形法求解这个具有标准形的线性规划问题. 因为该问题有可行解，且目标函数在可行域上有上界，所以能得到基本可行(极点)解$\boldsymbol{x}^*$，$\boldsymbol{x}^*$会自动满足整性要求. 所以这个松弛问题和原始问题(2.1) 的最优值相等.

考虑松弛问题(2.2)的对偶问题，即

$$\begin{array}{ll}\text{minimize} & \boldsymbol{\lambda}^T\boldsymbol{b} \\ \text{subject to} & \boldsymbol{A}^T\boldsymbol{\lambda}\geq \boldsymbol{c}\end{array}$$

具体地，即

$$\begin{array}{ll}\text{minimize} & \sum_i u_i+\sum_j v_j \\ \text{subject to} & u_i+v_j\geq s_{ij}, i,j=1,2,\cdots,n.\end{array} \tag{2.3}$$

因为原始问题(2.2)有最优解，由强对偶定理，对偶问题(2.3)有最优解. 设

$$\boldsymbol{\lambda}^*=(p_1\ p_2\ \cdots\ p_n\ q_1\ q_2\ \cdots\ q_n)^T$$

是一个最优解，这里q_j是为第j块土地定出的价格. 则由可行性有$p_i+q_j\geq s_{ij}$. 如果最优解指派实体i到土地j，即$x_{ij}^*=1>0$，则由互补性有$p_i+q_j=s_{ij}$.

(b) $s_{ij}-q_j$表示当实体i投标土地j时所能获得的净利润. 如果最优解指派实体i到土地j，(a)中最后的结论表明$p_i+q_j=s_{ij}$(互补性). 若实体i未投标土地j'，则由可行性有$p_i+q_{j'}\geq s_{ij'}$. 联立这两个式子，得到

$$s_{ij}-q_j\geq s_{ij'}-q_{j'},$$

这表明按照题目中所给的问题的最优解配置时，既能使整个收益最大化，也能使每个实体i 获得最大利润. 即专家按照总体收益最大化进行配置与自由竞争下各实体追求各自的利润最大化得到的结果相一致. 从而，专家可以通过对土地进行合理定价(求解问题(2.3))，所得价格可以诱导各实体在自由竞争的情况下追求各自利润最大的同时使得总体收益最大.

(c) 考虑问题

$$\begin{array}{ll}\text{maximize} & \boldsymbol{c}^T\boldsymbol{y} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{y}\leq\boldsymbol{b} \\ & \boldsymbol{y}\geq \boldsymbol{0}, \boldsymbol{y}\in\{0,1\}^{n^2}\end{array} \tag{2.4}$$

$$\begin{array}{ll}\text{maximize} & \boldsymbol{c}^T\boldsymbol{y} \\ \text{subject to} & \boldsymbol{A}\boldsymbol{y}\leq\boldsymbol{b} \\ & \boldsymbol{y}\geq \boldsymbol{0}\end{array} \tag{2.5}$$

如果问题(2.4)的最优解使得不等式约束$\boldsymbol{A}\boldsymbol{y}\leq\boldsymbol{b}$取等号，那么其松弛问题(2.5)的对偶问题与问题 2.3相比，会增加非负约束. 利用与(a)和(b)中完全类似的分析，可以得到(a)和(b)中的结果.

下面说明问题(2.4) 的最优解使得不等式约束中的等号成立. 假设$\boldsymbol{x}^*$是最优解，且对应的向量是$\boldsymbol{y}^*$. 则$\boldsymbol{A}\boldsymbol{y}^* \le \boldsymbol{b}$. 若这$2n$ 个不等式中至少有一个严格成立，不妨设是前n个不等式中的第1个不等式严格成立. 再由整性约束，必有$\sum_{j=1}^n x_{ij}^* = 0$，从而$x_{1j}^* = 0, j = 1, 2, \cdots, n$. 进一步，由$\boldsymbol{A}$的特殊性，后$n$个不等式中至少有一个是严格成立的，设为$n$. 同前，得$x_{in}^* = 0, i = 1, 2, \cdots, n$. 令

$$\hat{x}_{1n} = 1,\ \hat{x}_{ij} = x_{ij}^*, i \neq 1, j \neq n,$$

易见$\hat{\boldsymbol{x}}$是(2.4)的可行解，且其对应的目标值比最优值大s_{1n}，又因为s_{1n}是正的，这与$\boldsymbol{x}^*$的最优性矛盾.

综上，总可以假定价格是非负的.

2.32 (a) 利用单纯形法求解

$$\begin{aligned}
\text{minimize}\quad & 2x_1 + 3x_2 + 2x_3 + 2x_4 \\
\text{subject to}\quad & x_1 + 2x_2 + x_3 + 2x_4 = 3 \\
& x_1 + 1x_2 + 2x_3 + 4x_4 = 5 \\
& x_i \ge 0, i = 1, 2, 3, 4.
\end{aligned}$$

(b) 利用 (a) 中所作工作和对偶单纯形法求解相同的问题,除了方程组的右端向量变成 $(1\ 8)^T$.

解:

(a) 因为该问题没有显然的基本可行解，利用两阶段法求解该问题.

第 I 阶段：首先引入人工变量 x_5, x_6 ，构造辅助问题

$$\begin{aligned}
\text{minimize}\quad & x_5 + x_6 \\
\text{subject to}\quad & x_1 + 2x_2 + x_3 + 2x_4 + x_5 = 3, \\
& x_1 + x_2 + 2x_3 + 4x_4 + x_6 = 5, \\
& x_i \ge 0, i = 1, 2, 3, 4, 5, 6.
\end{aligned}$$

利用单纯形法求解辅助问题. 首先得初始表

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_5	1	2	1	2	1	0	3
x_6	1	1	2	4	0	1	5
$\boldsymbol{c}^T$	0	0	0	0	1	1	0

将基变量 x_5, x_6 对应的最后一行的系数化为零，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_5	1	2	1	2	1	0	3
x_6	1	1	2	$\boxed{4}$	0	1	5
$\boldsymbol{r}^T$	-2	-3	-3	-6	0	0	-8

以 4 为转轴元，即 x_4 进基，x_6 出基，得

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_5	1/2	$\boxed{3/2}$	0	0	1	−1/2	1/2
x_4	1/4	1/4	1/2	1	0	1/4	5/4
$\boldsymbol{r}^T$	−1/2	−3/2	0	0	0	3/2	−1/2

再次转轴之后，得

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_2	1/3	1	0	0	2/3	−1/3	1/3
x_4	1/6	0	1/2	1	−1/6	1/3	7/6
$\boldsymbol{r}^T$	0	0	0	0	1	1	0

因为辅助问题的最优值为 0，所以原始问题有可行解. 因为基变量不含人工变量，故 $\boldsymbol{x} = (0, 1/3, 0, 7/6)^T$ 是一个基本可行解.

第 II 阶段：在第 I 阶段的最后一个单纯形表中删去人工变量所对应列和最后一行，可得原始问题的一张初始表

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_2	1/3	1	0	0	1/3
x_4	1/6	0	1/2	1	7/6
$\boldsymbol{c}^T$	2	3	2	2	0

将基变量 x_2 和 x_4 下面的系数化为零，得第一张单纯形表

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_2	1/3	1	0	0	1/3
x_4	1/6	0	1/2	1	7/6
$\boldsymbol{r}^T$	2/3	0	1	0	−10/3

因为相对费用系数向量非负，故得最优解 $\boldsymbol{x}^* = (0, 1/3, 0, 7/6)^T$ ，最优值 $z^* = 10/3$.

(b) 由 (a) 可知 $\boldsymbol{B} = (\boldsymbol{a}_2, \boldsymbol{a}_4)$ ，所以 $\boldsymbol{B}^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ -1/6 & 1/3 \end{bmatrix}$. 右端项改为 $(1\ 8)^T$ 后，只需将上表的最后一列改为 $\boldsymbol{B}^{-1}\begin{bmatrix} 1 \\ 8 \end{bmatrix} = \begin{bmatrix} -2 \\ \frac{15}{6} \end{bmatrix}$. 这时，得新问题的单纯形表

	x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
x_2	1/3	1	0	0	−2
x_4	1/6	0	1/2	1	$\frac{15}{6}$
$\boldsymbol{r}^T$	2/3	0	1	0	1

此时，由对偶单纯形法，x_2 出基，但是第一行没有负元素. 由单纯形表第一行的数据知，问题的第一个约束等价于 $\frac{1}{3}x_1 + x_2 = -2$. 而变量非负，故修改右端项后所得新问题没有可行解.

2.33 对于问题

$$\begin{array}{ll} \text{minimize} & 5x_1 - 3x_2 \\ \text{subject to} & 2x_1 - x_2 + 4x_3 \leq 4 \\ & x_1 + x_2 + 2x_3 \leq 5 \\ & 2x_1 - x_2 + x_3 \geq 1 \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \end{array}$$

(a) 以 1 为转轴元, 仅**转轴一次**找到一个可行解;

(b) 利用单纯形法求解问题;

(c) 对偶问题是什么?

(d) 对偶问题的解怎么样?

解:

(a) 引入松驰变量 x_4, x_5 和盈余变量 x_6 后，得标准形问题

$$\begin{array}{ll} \text{minimize} & 5x_1 - 3x_2 \\ \text{subject to} & 2x_1 - x_2 + 4x_3 + x_4 = 4, \\ & x_1 + x_2 + 2x_3 + x_5 = 5, \\ & 2x_1 - x_2 + x_3 - x_6 = 1, \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{array}$$

得与等式约束对应的表格如下

x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{b}$
2	−1	4	1	0	0	4
1	1	2	0	1	0	5
2	−1	$\boxed{1}$	0	0	−1	1

以 1 为转轴元转轴(即从第三个方程解出 x_3 ，代入第一个和第二个方程消去 x_3)后，得

x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
−6	3	0	1	0	4	0
−3	3	0	0	1	2	3
2	−1	1	0	0	−1	1

该表对应着原始问题的一个基本可行解.

(b) 下面利用单纯形法求解原始问题. 首先得初始表为

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_4	−6	$\boxed{3}$	0	1	0	4	0
x_5	−3	3	0	0	1	2	3
x_3	2	−1	1	0	0	−1	1
$\boldsymbol{c}^T(\boldsymbol{r}^T)$	5	−3	0	0	0	0	0

因为最后一行与基变量对应的系数已经为零，故上面的初始表即为第一张单纯形表. 以 3 为转轴元进行转轴后，得

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_2	-2	1	0	1/3	0	4/3	0
x_5	$\boxed{3}$	0	0	-1	1	-2	3
x_3	0	0	1	1/3	0	1/3	1
$\boldsymbol{r}^T$	-1	0	0	1	0	4	0

再以 3 为转轴元转轴后，得

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_2	0	1	0	$-1/3$	2/3	0	2
x_1	1	0	0	$-1/3$	1/3	$-2/3$	1
x_3	0	0	1	1/3	0	1/3	1
$\boldsymbol{r}^T$	0	0	0	2/3	1/3	10/3	1

因为相对费用系数向量非负，故得原始问题最优解 $\boldsymbol{x}^* = (1,2,1)^T$，最优值 $z^* = -1$.

(c) 原始问题的对偶问题为

$$
\begin{aligned}
\text{maximize} \quad & 4w_1 + 5w_2 + w_3 \\
\text{subject to} \quad & 2w_1 + w_2 + 2w_3 \le 5, \\
& -w_1 + w_2 - w_3 \le -3, \\
& 4w_1 + 2w_2 + w_3 \le 0, \\
& w_1 \le 0, w_2 \le 0, w_3 \ge 0.
\end{aligned}
$$

(d) 易于看到原始问题引入松驰变量 x_4, x_5 和盈余变量 x_6 后，所得标准形问题的对偶问题也是该问题. 从而标准形问题的最优乘子 $(\boldsymbol{\lambda}^*)^T = \boldsymbol{c}_{\boldsymbol{B}}^T \boldsymbol{B}^{-1}$ 即为对偶问题的最优解，其中 $\boldsymbol{B}$ 为最优解对应的基. 而

$$
\begin{aligned}
\begin{bmatrix} r_4 & r_5 & r_6 \end{bmatrix} &= \begin{bmatrix} c_4 & c_5 & c_6 \end{bmatrix} - \boldsymbol{c}_{\boldsymbol{B}}^T \boldsymbol{B}^{-1} \begin{bmatrix} \boldsymbol{a}_4 & \boldsymbol{a}_5 & \boldsymbol{a}_6 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} \lambda_1^* & \lambda_2^* & \lambda_3^* \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\
&= \begin{bmatrix} -\lambda_1^* & -\lambda_2^* & \lambda_3^* \end{bmatrix}
\end{aligned}
$$

所以由(b)中的最优单纯形表，得 $\boldsymbol{\lambda}^* = (-2/3, -1/3, 10/3)^T$. 计算易得最优值为 -1 .

第三章　线性规划：应用及扩展

习题设计说明：

1. 习题3.1 用以熟悉网络单纯形法中各个量的计算方式；习题3.4用以熟悉网络单纯形法求解运输问题时，其中各个量的存储和计算方式.

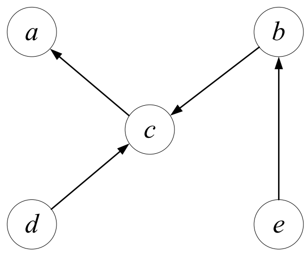
2. 习题3.2和习题3.3是关于网络流问题的理论性质的练习.

3. 习题3.5，习题3.6和习题3.7介绍典型的可以建模为整数规划的运筹问题，其中背包问题和车辆路由问题都可建模为0-1线性规划、二次指派是0-1二次规划，利用线性化技术也可表示成0-1线性规划问题. 并以背包问题为例，说明求解该问题的分枝定界法的计算复杂度是指数级的.

4. 习题3.8是用以体会对偶单纯形法在求解整数线性规划的分支限界法中的特殊应用，体会根据问题性质和应用场景选取算法的重要性!

5. 习题3.9用以熟悉分枝定界法，如剪枝、定界、分支等.

3.1. 考虑图3.1.1 所示的网络流问题，令 $\mathcal{T} = \{(b,c),(c,a),(d,c),(e,b)\}$ ，即如下生成树.



请以 e 为根节点, 完成以下工作:
(a) 求每个树弧上的原始流量;
(b) 求与每个节点对应的单纯形乘子;
(c) 求与每个非树弧对应的既约费用系数.

解: (a) 原始流量为 $x_{ca} = 2$ ，$x_{dc} = 5$ ，$x_{bc} = 1$ ，$x_{eb} = 1$.

(b) 以 e 为根节点，则单纯形乘子依次为 $y_b = -11$ ，$y_c = -13$ ，$y_d = 12$ ，$y_a = -27$.

(c) 非树弧的既约费用系数依次为 $r_{ab} = 28, r_{ce} = 23, r_{da} = -29, r_{de} = -2$.

3.4 考虑表3.2.1 所给的运输问题.
(a) 表3.2.2 给出的树解原始可行吗？对偶可行吗？
(b) 求解该运输问题.

解: (a)树解非负，故为原始可行，相对费用系数含负数，故不是最优解.

(b) 首先假设 $\mathcal{S}=\{1,2,3,4\},\mathcal{D}=\{5,6,7\}$.

表3.2.1 费用信息

$\mathcal{S}\setminus\mathcal{D}$	10	27	15
7	5	6	*
11	8	4	3
18	*	9	*
16	*	3	6

表3.2.2 运输问题的树解

$y_i\setminus y_j$	-5	-1	-4
0	$\boxed{7}$	5	*
3	$\boxed{3}$	$\boxed{8}$	-4
8	*	$\boxed{18}$	*
2	*	$\boxed{1}$	$\boxed{15}$

因为只有 $r_{27}=-4$ 为负，所以选 x_{27} 为入弧，这样弧 $x_{26},x_{27},x_{46},x_{47}$ 形成一个圈. 与 x_{27} 异向的弧有 x_{26},x_{47} ，选取流量最小的弧 x_{26} 为出弧；更新圈上弧的流量为 $x_{26}=0,x_{27}=8,x_{46}=9,x_{47}=7$. 然后分别计算子树3,6的单纯形乘子和非树弧 x_{16},x_{26} 的相对费用系数 $y_3=12,y_6=3$, (入弧指向不含根节点的子树，单纯形乘子均减少 -4)；$r_{16}=9,r_{26}=4$. (非树弧均与入弧同向，相对费用系数都减少 -4) 最后得到更新后的树解：

$y_i\setminus y_j$	-5	3	0
0	$\boxed{7}$	9	*
3	$\boxed{3}$	4	$\boxed{8}$
12	*	$\boxed{18}$	*
6	*	$\boxed{9}$	$\boxed{7}$

因为相对费用系数都非负，所以当前树解是最优的；最小运输费 $f^*=314$.

3.7 **二次指派问题**(quadratic assignment problem). 令 $\mathcal{F}$ 是 n 个工厂的集合, $\mathcal{C}$ 是 n 个城市的集合. 要求在每一城市中设置且只设置一个工厂, 并要使工厂两两之间总的通讯费用最小. 每对工厂 (i,k) 之间一定时间内通讯的次数为 t_{ik} , 每对城市 (j,l) 之间的距离为 d_{jl} . 通讯费用 $c_{ijkl}=t_{ik}d_{jl}$.

(a) 试为该问题建立目标函数为二次函数的整数规划模型.

(b) 将上述模型中的非线性目标函数线性化.

解: (a) 建模：设 $x_{ij}=1$ 表示在城市 j 建工厂 i ， $x_{ij}=0$ 表示不在城市 j 建工厂 i ，则 $i\in\mathcal{F}$ ， $j\in\mathcal{C}$. 这样， c_{ijkl} （即 $t_{ik}d_{jl}$ ）表示在城市 j 的工厂 i 到在城市 l 的工厂 k 之间的通讯费. 这样得到二次的目标函数，即

$$\sum_{i,k\in\mathcal{F}}\sum_{j,l\in\mathcal{C}}c_{ijkl}x_{ij}x_{kl}.$$

显然有 0, 1 约束： $x_{ij}\in\{0,1\}$. 由于每一个城市中有且只有一个工厂，故得到约束 $\sum_{\mathcal{F}}x_{ij}=1,\forall j\in\mathcal{C}$. 而每一个工厂也只能在一个城市中，故得到约束 $\sum_{\mathcal{C}}x_{ij}=1,\forall i\in\mathcal{F}$.

综上，二次指派问题的优化模型为

$$\begin{array}{ll} \text{minimize} & \sum_{i,k\in\mathcal{F}}\sum_{j,l\in\mathcal{C}} c_{ijkl}x_{ij}x_{kl} \\ \text{subject to} & \sum_{i\in\mathcal{F}} x_{ij}=1, \forall j\in\mathcal{C}, \\ & \sum_{j\in\mathcal{C}} x_{ij}=1, \forall i\in\mathcal{F}, \\ & x_{ij}\in\{0,1\}, \forall i\in\mathcal{F}, j\in\mathcal{C}. \end{array}$$

(b) 线性化目标函数：由于诸 $c_{ijkl}\geq 0$ ，故可以利用新变量 y_{ijkl} 来替换 $x_{ij}x_{kl}$ ，其中 y_{ijkl} 满足

$$y_{ijkl}\geq x_{ij}+x_{kl}-1,\ y_{ijkl}\geq 0.$$

这样，二次指派问题就转化为具有线性目标函数的 0-1 线性规划问题，即

$$\begin{array}{ll} \text{minimize} & \sum_{i,k\in\mathcal{F};j,l\in\mathcal{C}} c_{ijkl}y_{ijkl} \\ \text{subject to} & \sum_{i\in\mathcal{F}} x_{ij}=1, \forall j\in\mathcal{C}, \\ & \sum_{j\in\mathcal{C}} x_{ij}=1, \forall i\in\mathcal{F}, \\ & y_{ijkl}\geq x_{ij}+x_{kl}-1,\ y_{ijkl}\geq 0, \forall i,k\in\mathcal{F}, j,l\in\mathcal{C} \\ & x_{ij}\in\{0,1\}, \forall i\in\mathcal{F}, j\in\mathcal{C}. \end{array}$$

3.8 考虑例3.4.4 的枚举树上节点 2 的线性规划松弛问题 P_2 . 请从 P_0 的最优单纯形表开始, 利用对偶单纯形法求解 P_2 .

解:首先写出 P_0 的最优单纯形表

	x_1	x_2	x_3	$\boldsymbol{B}^{-1}\boldsymbol{b}$
	1	0.5	-0.25	1.25
$\boldsymbol{r}^T$	0	1.5	0.25	-1.25

给 P_0 添加约束 $x_1\geq 2$ 就得到 P_2 ，再对该约束引进一个松弛变量 x_4 ，继而写出表格

	x_1	x_2	x_3	x_4	$\boldsymbol{B}^{-1}\boldsymbol{b}$
	1	0.5	-0.25	0	1.25
	-1	0	0	1	-2
$\boldsymbol{r}^T$	0	1.5	0.25	0	-1.25

将第一行加到第二行后，得到单纯形表

	x_1	x_2	x_3	x_4	$\boldsymbol{B}^{-1}\boldsymbol{b}$
	1	0.5	-0.25	0	1.25
	0	0.5	$\boxed{-0.25}$	1	-0.75
$\boldsymbol{r}^T$	0	1.5	0.25	0	-1.25

此时得到的 P_2 的基本解是对偶可行的，所以利用对偶单纯形法求解. 以 -0.25 为转轴元转轴后，得到

	x_1	x_2	x_3	x_4	$\boldsymbol{B}^{-1}\boldsymbol{b}$
	1	0	0	-1	2
	0	-2	1	-4	3
$\boldsymbol{r}^T$	0	2	0	1	-2

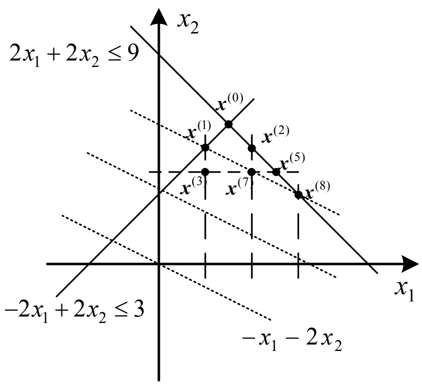
此为最优表，得最优解 $\boldsymbol{x}^* = (2,0)^T$.

3.9 对线性规划

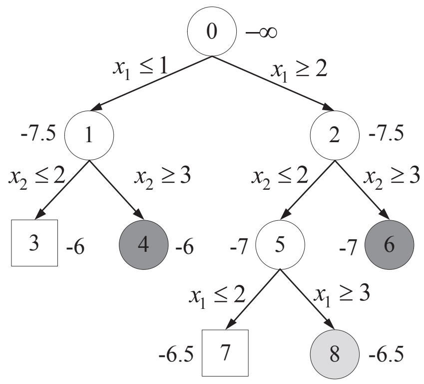
$$\begin{array}{lrcrl} \text{minimize} & -x_1 & - & 2x_2 & \\ \text{subject to} & -2x_1 & + & 2x_2 & \le 3 \\ & 2x_1 & + & 2x_2 & \le 9, \end{array}$$

分(i)无整数限制, (ii) x_1 为整数, (iii) x_1, x_2 均为整数三种情况，用图解法求解相应的问题. 并给出用分枝定界法求解(iii)的过程, 画出枚举树.

解: 根据题意，由图3.1(a)可知，三种情况的最优解分别是 $\boldsymbol{x}^* = (1.5,3)^T$; (ii) $\boldsymbol{x}' = (2,2.5)^T$; (iii) $\boldsymbol{x}'' = (2,2)^T$.



(a) 整数规划的图解.



(b) 分枝定界法的枚举树(采用的是深度优先的枚举策略).

图 3.1: 习题3.9图解

图3.1(b)表示用分枝定界法求解(iii)的枚举树，其中枚举策略是深度优先. 方法依次求解问题的详细信息见下表，其中 P_i 表示子问题编号，L 表示为子问题确定的下界，$\boldsymbol{x}^*$ 和 f^* 分别表示松弛子问题后所得线性规划问题的最优解和最优值，$\hat{\boldsymbol{x}}$ 和 $\hat{f}$ 分别表示当前最好解和最好值.

(P_i, L)	$\boldsymbol{x}^{*T}$	f^*	采取措施	$\hat{\boldsymbol{x}}^T$	$\hat{f}$
$(P_0, -\infty)$	$(1.5, 3)$	-7.5	分枝，得$(P_1, -7.5)$&$(P_2, -7.5)$	[ ]	$+\infty$
$(P_1, -7.5)$	$(1, 2.5)$	-6	分枝，得$(P_3, -6)$&$(P_4, -6)$		
$(P_3, -6)$	$(1, 2)$	-5	得可行解，更新$\hat{\boldsymbol{x}}$，剪枝	$(1, 2)$	-5
$(P_4, -6)$	[ ]	$+\infty$	子问题不可行，剪枝		
$(P_2, -7.5)$	$(2, 2.5)$	-7	分枝，得$(P_5, -7)$&$(P_6, -7)$		
$(P_5, -7)$	$(2.5, 2)$	-6.5	分枝，得$(P_7, -6.5)$&$(P_8, -6.5)$		
$(P_7, -6.5)$	$(2, 2)$	-6	得可行解，更新$\hat{\boldsymbol{x}}$，剪枝	$(2, 2)$	-6
$(P_8, -6.5)$	$(3, 1.5)$	-6	$f^* \ge L$, 剪枝		
$(P_6, -7)$	[ ]	$+\infty$	子问题不可行，剪枝		

第四章　无约束优化：基础

习题设计说明：

1. 最优性条件：习题4.1-习题4.5.
2. 凸函数：习题4.6. 收敛速度：习题4.7.
3. 一维搜索：4.8-4.13.

4.2 考虑问题 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{Ax}-\boldsymbol{b}\|^2$, 其中 $\boldsymbol{A}$ 是 $m\times n$ 矩阵, $\boldsymbol{b}$ 是 m 维向量，$m>n$.

(a) 给出问题的几何解释.

(b) 写出最优性的必要条件. 它也是一个充分条件吗？

(c) 最优解唯一吗？理由是什么？

(d) 你能给出最优解的一种闭合形式(解析式)吗？在满足题目所给信息下，允许规定任何你所需的假设.

(e) 求解该问题, 其中

$$\boldsymbol{A}=\begin{bmatrix}1&-1&0\\0&2&1\\0&1&0\\1&0&1\end{bmatrix},\quad \boldsymbol{b}=\begin{pmatrix}2\\1\\1\\0\end{pmatrix}.$$

解: (a) 几何解释：问题在求向量 $\boldsymbol{b}$ 到 $\boldsymbol{A}$ 的值空间R($\boldsymbol{A}$)(即由 $\boldsymbol{A}$ 的列生成的子空间)的最小距离. 设 $\boldsymbol{b}$ 在R($\boldsymbol{A}$) 上的投影为 $\boldsymbol{c}$，则满足 $\boldsymbol{Ax}=\boldsymbol{c}$ 的解 $\boldsymbol{x}^*$ 即为最优解. 具体例子见(e)，这时有三个未知数，四个方程，所给方程没有普通意义的解，称这里求出的 $\boldsymbol{x}^*$ 是方程组的最小二乘解.

(b) 令 $f(\boldsymbol{x})=\|\ \boldsymbol{Ax}-\boldsymbol{b}\ \|^2=\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{Ax}-2\boldsymbol{b}^T\boldsymbol{Ax}+\boldsymbol{b}^T\boldsymbol{b}$, $\nabla f(\boldsymbol{x})=2\boldsymbol{A}^T\boldsymbol{Ax}-2\boldsymbol{A}^T\boldsymbol{b}$. 一阶必要条件为：

$$\boldsymbol{A}^T\boldsymbol{Ax}=\boldsymbol{A}^T\boldsymbol{b}. \tag{4.1}$$

又因为 $\nabla^2 f(\boldsymbol{x})=2\boldsymbol{A}^T\boldsymbol{A}$ 为半正定矩阵，所以 $f(\boldsymbol{x})$ 为凸函数，从而方程(4.1)也是充分条件.

(c) 最优解不一定唯一. 当 $\boldsymbol{b}\in\mathrm{R}(\boldsymbol{A})$ 且 $\mathrm{rank}(\boldsymbol{A})<\min(m,n)$ 时，$\boldsymbol{Ax}=\boldsymbol{b}$ 有无穷多解，而此时满足 $\boldsymbol{Ax}=\boldsymbol{b}$ 的解均为最优解.

(d) 如果设 $\boldsymbol{A}^T\boldsymbol{A}$ 是正定的(即 $\boldsymbol{A}$ 是列满秩的)，则可给出最优解的一种闭合形式 $\boldsymbol{x}^*=(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{b}$.

(e) 将所给数据代入方程(4.1)，求解得 $\boldsymbol{x}^*=(3,\frac{3}{2},-\frac{5}{2})^T$. ∎

4.3 对于二次函数 $q(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{Gx}-\boldsymbol{b}^T\boldsymbol{x}$, 证明:

(a) 当且仅当 $\boldsymbol{G}$ 半正定且 $\boldsymbol{Gx}=\boldsymbol{b}$ 有解时，极小点存在.

(b) 当且仅当 $\boldsymbol{G}$ 正定时，极小点唯一.

(c) 如果 $\boldsymbol{G}$ 半正定，则每一个稳定点都是全局极小点.

证明： (a) 极小点的必要条件为 $\nabla q(\boldsymbol{x}) = \boldsymbol{G}\boldsymbol{x} - \boldsymbol{b} = \boldsymbol{0}, \nabla^2 q(\boldsymbol{x}) = \boldsymbol{G} \succeq \boldsymbol{0}$ ，即仅当 $\boldsymbol{b}$ 在 $\boldsymbol{G}$ 的值域里，且 $\boldsymbol{G}$ 半正定. 此外，当 $\boldsymbol{G}$ 半正定时，$q(\boldsymbol{x})$ 是凸函数，从而稳定点是全局极小点.

(b) 当 $\boldsymbol{G}$ 正定时，方程组 $\boldsymbol{G}\boldsymbol{x} = \boldsymbol{b}$ 有惟一解 $\boldsymbol{x}^* = \boldsymbol{G}^{-1}\boldsymbol{b}$ ，由(a)知，这是函数的惟一极小点.

(c) 若 $\nabla^2 q(\boldsymbol{x}) = \boldsymbol{G}$ 半正定，则 $q(\boldsymbol{x})$ 是凸函数，所以其任意稳定点，即方程组 $\boldsymbol{G}\boldsymbol{x} = \boldsymbol{b}$ 的任意解，都是全局极小点. ∎

4.6 设 C 是 $\mathbb{R}^n$ 中的凸集, $f(\cdot): C \to \mathbb{R}$. 证明 f 是 C 上的凸函数当且仅当对任一整数 $k \geq 2$, $\boldsymbol{x}_i \in C, \theta_i \geq 0, i = 1, \cdots, k, \sum_{i=1}^k \theta_i = 1$ 蕴含着 $f\left(\sum_{i=1}^k \theta_i \boldsymbol{x}_i\right) \leq \sum_{i=1}^k \theta_i f(\boldsymbol{x}_i)$.

证明: **“必要性”：**用数学归纳法证. 若 f 是 C 上的凸函数, 结论对 $k = 2$ 显然成立. 假设结论对 k 成立，以下我们证明结论对 $k+1$ 仍然成立.

对任意的 $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k+1} \in C$, $\theta_1, \ldots, \theta_{k+1} \geq 0$ 且 $\sum_{i=1}^{k+1} \theta_i = 1$. 若 $\theta_{k+1} = 1$, 此时结论显然成立. 若 $0 \leq \theta_{k+1} < 1$ ， 则

$$
\begin{aligned}
f\left(\sum_{i=1}^{k+1} \theta_i \boldsymbol{x}_i\right) &= f\left((1-\theta_{k+1}) \sum_{i=1}^{k} \frac{\theta_i}{1-\theta_{k+1}} \boldsymbol{x}_i + \theta_{k+1} \boldsymbol{x}_{k+1}\right) \\
&\leq (1-\theta_{k+1}) f\left(\sum_{i=1}^{k} \frac{\theta_i}{1-\theta_{k+1}} \boldsymbol{x}_i\right) + \theta_{k+1} f(\boldsymbol{x}_{k+1}) (\text{ 凸函数定义}) \\
&\leq (1-\theta_{k+1}) \sum_{i=1}^{k} \frac{\theta_i}{1-\theta_{k+1}} f(\boldsymbol{x}_i) + \theta_{k+1} f(\boldsymbol{x}_{k+1}) \\
&= \sum_{i=1}^{k+1} \theta_i f(\boldsymbol{x}_i)
\end{aligned}
$$

其中第二个不等式是因为 $\frac{\theta_i}{1-\theta_{k+1}} \geq 0$ ，且 $\sum_{i=1}^k \frac{\theta_i}{1-\theta_{k+1}} = 1$ ，从而由归纳假设知结论对 k 成立得到的. 由归纳法, 必要性得证.

“充分性”: 取 $k = 2$, 即知 f 为凸函数. ∎

4.7 考虑以下序列

(a) $x^{(k)} = 1/k$,

(b) $x^{(k)} = (0.5)^{2^k}$,

(c) $x^{(k)} = 1/(k!)$,

给出它们的收敛阶及速率常数; 并说明为了 $x^{(k)} < 10^{-4}$, 每个序列的 k 至少取多大?

解: 三个序列均收敛到 $x^* = 0$. 定义误差序列 $e_k = x^{(k)} - x^*$ ，则

(a) $\lim_{k\to\infty} \frac{e_{k+1}}{e_k} = \frac{k}{k+1} = 1$. 这表明序列是次线性(sub-linear)收敛的. 为了 $\boldsymbol{x}^{(k)} < 10^{-4}$ ， k 至少取 $10^4 + 1$;

(b) $\lim_{k\to\infty} \frac{e_{k+1}}{e_k^2} = \frac{(0.5)^{2^{(k+1)}}}{(0.5^{2^k})^2} = 1$. 这表明序列是二次收敛的. 有 $(0.5)^{16} < 10^{-4}$ ，这里 $k = 4$;

(c) $\lim_{k\to\infty}\frac{e_{k+1}}{e_k}=\frac{k!}{(k+1)!}=0$. 这表明序列是超线性收敛的. k 取 8 即有 $x^{(k)}<10^{-4}$.

∎

4.8 考查函数 $f(x_1,x_2)=(x_1+x_2^2)^2$在点 $\boldsymbol{x}^{(k)}=(1,0)^T$ 处的搜索方向 $\boldsymbol{p}^{(k)}=(-1,1)^T$ ，证明 $\boldsymbol{p}^{(k)}$ 是所给函数在 $\boldsymbol{x}^{(k)}$ 处的下降方向, 并求出一维极小化问题 $\min\limits_{\alpha>0} f(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{p}^{(k)})$ 的所有极小点.

解: $\boldsymbol{g}(\boldsymbol{x})=\begin{bmatrix}2(x_1+x_2^2)\\4x_2(x_1+x_2^2)\end{bmatrix}$, $\boldsymbol{g}^{(k)T}\boldsymbol{p}^{(k)}=\begin{bmatrix}2\\0\end{bmatrix}^T\begin{bmatrix}-1\\1\end{bmatrix}=-2<0$, 故 $\boldsymbol{p}^{(k)}$ 是一个下降方向.

$$\boldsymbol{x}^{(k)}+\alpha\boldsymbol{p}^{(k)}=\begin{bmatrix}1\\0\end{bmatrix}+\alpha\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1-\alpha\\\alpha\end{bmatrix},$$

所以 $f(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{p}^{(k)})=(1-\alpha+\alpha^2)^2$, 而 $1-\alpha+\alpha^2=(\alpha-\frac{1}{2})^2+\frac{3}{4}\geq\frac{3}{4}$, 当且仅当 $\alpha=\frac{1}{2}$ 时取等号，所以当 $\alpha=\frac{1}{2}$ 时， $f(\boldsymbol{x}^{(k)}+\alpha\boldsymbol{p}^{(k)})$ 的值最小. ∎

4.11 对于函数 $\phi(\alpha)=1-\alpha e^{-\alpha^2}$, 对 $\sigma=\rho=\frac{1}{10}$ 及 $\sigma=\rho=\frac{1}{4}$两种情况分别确定满足Goldstein条件, Wolfe 条件以及强 Wolfe 条件的 α 值的可接受区间. 对于后一种情况, 以 $\alpha=0$ 和1为初值, 应用算法4.4.1 和算法4.4.2 得到一个可接受点.

解： 首先对函数 $\phi(\alpha)=1-\alpha e^{-\alpha^2}$ 求导，得 $\phi'(\alpha)=(2\alpha^2-1)e^{-\alpha^2}$，从而得到 $\phi(0)=1,\phi'(0)=-1$ ，再由 Armijo 条件 $\phi(\alpha)\leq\phi(0)+\alpha\rho\phi'(0)$ 可得

$$e^{-\alpha^2}\geq\rho,\ \alpha>0. \tag{4.2}$$

进一步，将所给数据代入 Goldstein 条件的第二个不等式 $\phi(\alpha)\geq\phi(0)+\alpha(1-\rho)\phi'(0)$ 可得

$$e^{-\alpha^2}\leq1-\rho. \tag{4.3}$$

将所给数据代入 Wolfe 条件中的斜率条件$\phi'(\alpha)\geq\sigma\phi'(0)$，得到

$$(2\alpha^2-1)e^{-\alpha^2}\geq-\sigma. \tag{4.4}$$

将所给数据代入强 Wolfe 条件中的双边斜率条件 $|\phi'(\alpha)|\leq-\sigma\phi'(0)$得

$$-\sigma\leq(2\alpha^2-1)e^{-\alpha^2}\leq\sigma. \tag{4.5}$$

从而，当 $\rho=0.1$ 时，解(4.2)得 $0<\alpha\leq1.5174$ ；解(4.3)得 $\alpha\geq0.3246$ ；解(4.4)得 $\alpha\geq0.6509$ ；解(4.5)得 $0.6509\leq\alpha\leq0.7682$.

这样，Goldstein 条件的可接受区间为 $[0.3246,1.5174]$ ；Wolfe 条件的可接受区间为 $[0.6509,1.5174]$ ；强 Wolfe 条件的可接受区间为 $[0.6509,0.7682]$.

当 $\rho=0.25$ 时，解(4.2)得 $0<\alpha\leq1.1774$ ；解(4.3)得 $\alpha>0.5364$ ；解(4.4)得 $\alpha\geq0.5716$ ；解(4.5)得 $0.5716\leq\alpha\leq0.8775$.

这样，Goldstein 条件的可接受区间为 $[0.5364,1.1774]$ ；Wolfe 条件的可接受区间为 $[0.5716,1.1774]$ ；强 Wolfe 条件的可接受区间为 $[0.5716,0.8775]$.

在划界与分割阶段，程序均用极小化二次插值多项式的方法选取 α . 对于初始条件 $\bar{\phi}=0,\alpha_0=0,\alpha_1=1,\sigma=\rho=0.25$ ，划界与分割阶段各迭代 1 次，就可得到 $\alpha_k=0.75$ ，由上可知它满足强 Wolfe 条件. ∎

第六章　无约束优化：线搜索法

习题设计说明：

1. 最速下降法：习题5.1-习题5.6，重点理解最速下降法有限步收敛的条件和收敛因子的上界等结论.

2. 牛顿法：习题5.7-习题5.14，重点理解牛顿法是局部二阶收敛的，当初始点不合适时，牛顿法的不适定性，从而需要修正牛顿法.

3. 共轭梯度法：习题5.15-习题5.17(向量共轭的定义)，习题5.18(共轭梯度法的性质)，习题5.19(共轭梯度法的有限步收敛性).

4. 拟牛顿法：习题5.22(关于曲率条件的理解，进而理解拟牛顿法通常与Wolfe或者强Wolfe线索搭配使用的原因)，习题2.23(DFP法的练习，帮助理解使用精确线搜索时，DPF法产生的方向是共轭的性质)，习题5.24(证明DFP校正能保证正定性的充要条件是曲率条件).

5. 说明各种方法之间关系的综合性练习：习题5.20(最速下降法、共轭梯度法和BFGS法的关系)，习题2.23(DFP法使用精确线搜索时产生的方向是共轭的)，习题5.25(非常有名的逆的校正公式，在数值计算中有重要的应用，利用他可以由BFGS的关于Hessian阵的校正公式推出关于Hessian阵逆的校正公式).

6. 最小二乘：习题5.26-习题5.29，练习GN法，也可以编写LM方法求解.

7. 理解方法性质的数值训练：习题5.6(理解最速下降法的收敛速度与收敛因子上界等结果，体会方法的收敛速度对初值的依赖性)，习题5.8(这是求解线性规划的对数障碍法中得到的子问题，体会函数有定义域时如何使用线搜索法，并体会障碍因子对方法收敛速度的影响)，习题5.9(体会最速下降法与牛顿法的异同，并体会线搜索牛顿法中线搜索策略的意义！)，习题5.19(理解共轭梯度法的 n 步之内收敛性质，体会理论结果与实际计算的差异；这里的 Hilbert 矩阵是有名的病态矩阵，它的坏条件数很大).

5.3 考虑将最速下降法应用于Hessian阵为 $\boldsymbol{G}$ 的正定二次函数. 设初始点 $\boldsymbol{x}^{(0)} \neq \boldsymbol{x}^*$ 可以表示为

$$\boldsymbol{x}^{(0)} = \boldsymbol{x}^* + \mu \boldsymbol{s},$$

其中 s 是 $\boldsymbol{G}$ 的相应于特征值 λ 的特征向量.

(a) 证明 $\boldsymbol{g}^{(0)} = \mu\lambda\boldsymbol{s}$, 且若沿最速下降方向进行精确一维搜索, 则方法经一次迭代后终止.

(b) 如果 $\boldsymbol{G}$ 为单位阵的倍数, 证明对任意的初始点 $\boldsymbol{x}^{(0)}$, 方法经一次迭代后终止.

(c) 若 $\boldsymbol{x}^{(0)}$ 不能表示成如上形式, 则 $\boldsymbol{x}^{(0)}$ 可表示为

$$\boldsymbol{x}^{(0)} = \boldsymbol{x}^* + \sum_{i=1}^{m} \mu_i \boldsymbol{s}_i,$$

其中 $m> 1$, 且对所有的 i 有 $\mu_i \neq 0, \boldsymbol{s}_i$ 是 $\boldsymbol{G}$ 的对应于不同特征值 λ_i 的特征向量. 证明此时方法经一次迭代后不能终止.

解： 根据题意，可假设目标函数$q(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}+\boldsymbol{b}^T\boldsymbol{x}+c$，则梯度$\boldsymbol{g}(\boldsymbol{x})=\boldsymbol{G}\boldsymbol{x}+\boldsymbol{b}$. 所以在最优点$\boldsymbol{x}^*$ 处，满足一阶条件

$$\boldsymbol{G}\boldsymbol{x}^*+\boldsymbol{b}=\boldsymbol{0}.$$

(a) 因为$\boldsymbol{x}^{(0)}=\boldsymbol{x}^*+\mu\boldsymbol{s}$，而$\boldsymbol{G}\boldsymbol{s}=\lambda\boldsymbol{s}$，所以

$$\begin{aligned}\boldsymbol{g}^{(0)}&=\boldsymbol{G}(\boldsymbol{x}^*+\mu\boldsymbol{s})+\boldsymbol{b}\\&=\boldsymbol{G}\boldsymbol{x}^*+\boldsymbol{b}+\mu\boldsymbol{G}\boldsymbol{s}\\&=\mu\lambda\boldsymbol{s}.\end{aligned}$$

从而 $\boldsymbol{g}^{(0)^T}\boldsymbol{g}^{(0)}=\mu^2\lambda^2\boldsymbol{s}^T\boldsymbol{s},\ \boldsymbol{g}^{(0)^T}\boldsymbol{G}\boldsymbol{g}^{(0)}=\mu^2\lambda^3\boldsymbol{s}^T\boldsymbol{s}$，则

$$\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(0)}-\frac{1}{\lambda}\boldsymbol{g}^{(0)}=\boldsymbol{x}^*.$$

(b)假设$\boldsymbol{G}=a\boldsymbol{I}$，$a$是一个常数. 对于任意初始点$\boldsymbol{x}^{(0)}$，可表示成

$$\boldsymbol{x}^{(0)}=\boldsymbol{x}^*+\boldsymbol{p},$$

在(a)中，令 $\boldsymbol{s}=\boldsymbol{p},\ \mu=1$，又因为 $\boldsymbol{G}\boldsymbol{s}=a\boldsymbol{s}$，由(a)的结论知，只需一次迭代，方法就终止.

(c)首先假设迭代一次后算法终止. 然后由题意知，$\boldsymbol{g}^{(0)}=\boldsymbol{G}\boldsymbol{x}^{(0)}+\boldsymbol{b}=\sum_{i=1}^m\mu_i\lambda_i\boldsymbol{s}_i$，再让$\alpha$表示第一次迭代的步长，从而$\boldsymbol{x}^{(1)}$为

$$\begin{aligned}\boldsymbol{x}^{(1)}&=\boldsymbol{x}^{(0)}-\alpha\boldsymbol{g}\\&=\boldsymbol{x}^*+\sum_{i=1}^m\mu_i\boldsymbol{s}_i-\alpha\sum_{i=1}^m\mu_i\lambda_i\boldsymbol{s}_i\\&=\boldsymbol{x}^*+\sum_{i=1}^m(\mu_i-\alpha\mu_i\lambda_i)\boldsymbol{s}_i\end{aligned}$$

因此，要让$\boldsymbol{x}^{(1)}=\boldsymbol{x}^*$，需满足

$$\sum_{i=1}^m(\mu_i-\alpha\mu_i\lambda_i)\boldsymbol{s}_i=\boldsymbol{0}.$$

又因为$\boldsymbol{s}_i$是矩阵$\boldsymbol{G}$对应于不同特征值λ_i的特征向量，所以$\boldsymbol{s}_1,\cdots,\boldsymbol{s}_m$线性无关. 从而可推出

$$\mu_i-\alpha\mu_i\lambda_i=0,\ i=1,2,\ldots,m.$$

最后得到$\lambda_i=1/\alpha,\ i=1,2,\ldots,m$，与$\lambda_i$相异矛盾，所以假设不成立. 迭代无法在一步终止.

5.5 假设我们需要极小化 $q(\boldsymbol{x})=5x_1^2+5x_2^2-x_1x_2-11x_1+11x_2+11$.

(a) 找到一个满足一阶必要条件的解.

(b) 说明该点是全局极小点.

(c) 针对该问题, 最速下降法的收敛因子最大不会超过多少?

(d) 从 $\boldsymbol{x}^{(0)}=(1,1)^T$ 出发, 最多需要多少步可将目标函数值减少到 10^{-11} ?

解： (a) 令 $\boldsymbol{g}(\boldsymbol{x}) = \begin{bmatrix} 10x_1 - x_2 - 11 \\ 10x_2 - x_1 + 11 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ ，可得满足一阶必要条件的解为 $\boldsymbol{x}^* = (1, -1)^T$.

(b) 因为 $\boldsymbol{G} = \begin{bmatrix} 10 & -1 \\ -1 & 10 \end{bmatrix}$ 正定，所以 $q(\boldsymbol{x})$ 是凸函数，从而 $\boldsymbol{x}^* = (1, -1)^T$ 也是全局极小点。

(c) 因为 $\boldsymbol{G}$的特征值为 $9, 11$ ，所以得线性收敛的收敛因子的上界 $a = (\frac{11-9}{11+9})^2 = 0.01$.

(d) 因为 $q(\boldsymbol{x}^*) = 0, q(\boldsymbol{x}^{(0)}) = 20,$

$$q(\boldsymbol{x}^{(k)}) - q(\boldsymbol{x}^*) \le a^k (q(\boldsymbol{x}^{(0)}) - q(\boldsymbol{x}^*)) \tag{6.1}$$

令 $10^{-11} = 10^{-2k} \cdot 20$ ，解得 $k = \lceil \frac{11 + \log 20}{2} \rceil = \lceil 6.150 \rceil = 7$. 这样， $k = 7$ 是使不等式(6.1)成立的最小正整数. 所以应取 $k = 7$ ，否则有可能达不到要求. 这样，最多需要 7 步即可达到要求. 请注意，这里的“最多需要”是因为利用条件数估计的是收敛因子的上界，也即最坏情况估计. ∎

5.12 应用牛顿法于函数

$$f(x) = \tfrac{11}{546} x^6 - \tfrac{38}{364} x^4 + \tfrac{1}{2} x^2,$$

其中 $x^{(0)} = 1.01$ 为初始点. 验证 $G^{(k)}$ 总是正定的, 且 $f^{(k)}$ 是单调递减的. 说明方法产生的序列 $\{x^k\}$ 的聚点 $x^\infty = \pm 1$, 且 $g^\infty \neq \mathbf{0}$. 验证对任意固定的正数 ρ , 不管其多么小, Armijo条件(4.3.1)对充分大的 k 总不成立(舍入误差起支配作用的情况除外).

证：对函数 $f(x) = \frac{11}{546} x^6 - \frac{38}{364} x^4 + \frac{1}{2} x^2$ ，易得：

$$\boldsymbol{g}(x) = f'(\boldsymbol{x}) = \frac{11}{91} x^5 - \frac{38}{91} x^3 + x$$

$$\boldsymbol{G}(x) = f''(\boldsymbol{x}) = \frac{55}{91} x^4 - \frac{114}{91} x^2 + 1$$

因而，基于步长为1的基本牛顿法，第k次迭代过程可以写为：

$$\boldsymbol{s}^{(k)} = -\frac{\boldsymbol{g}^{(k)}}{\boldsymbol{G}^{(k)}} = -\frac{11 {\boldsymbol{x}^{(k)}}^5 - 38 {\boldsymbol{x}^{(k)}}^3 + 91 \boldsymbol{x}^{(k)}}{55 {\boldsymbol{x}^{(k)}}^4 - 114 {\boldsymbol{x}^{(k)}}^2 + 91}$$

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \boldsymbol{s}^{(k)} = \frac{44 {\boldsymbol{x}^{(k)}}^5 - 76 {\boldsymbol{x}^{(k)}}^3}{55 {\boldsymbol{x}^{(k)}}^4 - 114 {\boldsymbol{x}^{(k)}}^2 + 91}$$

由上面的计算结果，易见$\boldsymbol{G}(x)$的最小值$1 - \frac{57^2}{91 \times 55} > 0$，所以$\boldsymbol{G}(x)$恒大于零. 将给定的初始点 $x^{(0)} = 1.01$ 代入上式进行迭代，前10 个迭代点的信息见表6.1. 从数值结果可以看出 $f^{(k)}$ 是单调递减的，并最终收敛到0.4158. 此外，对于$\{\boldsymbol{x}^{(k)}\}$，指标为偶数的子列收敛到 1 ，指标为奇数的子列收敛到 -1. 因而用牛顿法产生的迭代点列 $\{\boldsymbol{x}^{(k)}\}$ 有两个聚点：1 和 -1, 并且 $\boldsymbol{g}^\infty = \pm \frac{64}{91} \neq 0$.

对任意给定的正数 ρ ，将 $\rho \boldsymbol{g}^{(k)} \boldsymbol{s}^{(k)} = -\rho {\boldsymbol{g}^{(k)}}^2 / \boldsymbol{G}^{(k)} < 0$ 代入 Armijo条件，即 $f^{(k)} + \rho \boldsymbol{g}^{(k)} \boldsymbol{s}^{(k)} - f^{(k+1)} \ge 0$，得

$$\rho \le \frac{\boldsymbol{G}^{(k)}}{{\boldsymbol{g}^{(k)}}^2} (f^{(k)} - f^{(k+1)}). \tag{6.2}$$

表 6.1: 基本牛顿法求解习题 5.12 的部分迭代数据

k	$\boldsymbol{x}^{(k)}$	$f^{(k)}$	$\boldsymbol{g}^{(k)}$	k	$\boldsymbol{x}^{(k)}$	$f^{(k)}$	$\boldsymbol{g}^{(k)}$
0	1.0100	0.4228	0.7068	5	-1.0002	0.4159	-0.7034
1	-1.0037	0.4183	-0.7046	6	1.0001	0.4158	0.7033
2	1.0016	0.4169	0.7039	7	-1	0.4158	-0.7033
3	-1.0008	0.4163	-0.7036	8	1	0.4158	0.7033
4	1.0004	0.4160	0.7034	9	-1	0.4158	-0.7033

由前面的分析知：当 k 趋于无穷大时，不等式的右边大于零且趋于零(因为 $\frac{G(x^\infty)}{(g(x^\infty))^2}(f(x^\infty)-f(x^\infty))=\frac{91}{128}\times 0=0$). 从而对于给定的正数$\rho$，无论其多小，由极限的定义，存在指标 K，当 $k>K$ 时，不等式(6.2)的右边严格小于 ρ，此时Armijo条件不再成立(舍入误差除外).

说明：该问题说明，当我们使用基本牛顿法时，即使初始点使得迭代过程中每个迭代点的Hessian阵正定，且基本牛顿法产生的序列有聚点，也不一定是稳定点. 一种可能的原因是步长为 1 的基本牛顿法得到的迭代点不满足Armijo条件. 如果每个迭代点满足Armijo条件，可以得到基本牛顿法是收敛的，即迭代序列的聚点是稳定点.

5.21 考虑四元二次函数 $\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}+\boldsymbol{b}^T\boldsymbol{x}$, 其中 $\boldsymbol{G}$ 为习题5.17 中的三对角矩阵, $\boldsymbol{b}=(-1,0,2,\sqrt{5})^T$. 取初始点 $\boldsymbol{x}^{(0)}=\boldsymbol{0}$ ，应用共轭梯度法极小化该函数, 并验证经两次迭代后终止, 且向量组 $\boldsymbol{g}^{(0)},\boldsymbol{G}\boldsymbol{g}^{(0)},\boldsymbol{G}^2\boldsymbol{g}^{(0)},\boldsymbol{G}^3\boldsymbol{g}^{(0)}$ 中只有两个独立向量.

解： $\boldsymbol{G}=\begin{pmatrix}2&-1&0&0\\-1&2&-1&0\\0&-1&2&-1\\0&0&-1&2\end{pmatrix}$, $\boldsymbol{b}=(-1,0,2,\sqrt{5})^T$，由共轭梯度法求得$\boldsymbol{x}^*=(0.4472,1.8944,3.3416,2.7889)^T$, $f^*=18.7082$. 以下验证向量组$\boldsymbol{g}^{(0)},\boldsymbol{G}\boldsymbol{g}^{(0)},\boldsymbol{G}^2\boldsymbol{g}^{(0)},\boldsymbol{G}^3\boldsymbol{g}^{(0)}$中只有两个独立向量.

首先，$\boldsymbol{g}^{(0)}=\boldsymbol{G}\boldsymbol{x}^{(0)}+\boldsymbol{b}=(-1,0,2,\sqrt{5})^T$, $\boldsymbol{G}\boldsymbol{g}^{(0)}=(-2,-1,4-\sqrt{5},-2+2\sqrt{5})^T$ 线性无关. 此外，易验证 $\boldsymbol{G}^2\boldsymbol{g}^{(0)}=(-3,-4+\sqrt{5},11-4\sqrt{5},-8+5\sqrt{5})^T$ 和 $\boldsymbol{G}^2\boldsymbol{g}^{(0)}=\lambda_0\boldsymbol{g}^{(0)}+\lambda_1\boldsymbol{G}\boldsymbol{g}^{(0)}$, 其中$\lambda_0=-5+2\sqrt{5},\lambda_1=4-\sqrt{5}$. 进一步，可验证

$$\boldsymbol{G}^3\boldsymbol{g}^{(0)}=\boldsymbol{G}\cdot\boldsymbol{G}^2\boldsymbol{g}^{(0)}=\lambda_0\boldsymbol{G}\boldsymbol{g}^{(0)}+\lambda_1\boldsymbol{G}^2\boldsymbol{g}^{(0)}=\lambda_0\lambda_1\boldsymbol{g}^{(0)}+(\lambda_0+\lambda_1^2)\boldsymbol{G}\boldsymbol{g}^{(0)}.$$

由上分析可见所要验证的结论成立.

由上述验证过程知$\boldsymbol{g}^{(0)},\boldsymbol{G}\boldsymbol{g}^{(0)},\boldsymbol{G}^2\boldsymbol{g}^{(0)},\boldsymbol{G}^3\boldsymbol{g}^{(0)}$中只有两个独立向量，又因为

$$\mathrm{span}\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\boldsymbol{p}^{(2)},\boldsymbol{p}^{(3)}\}=\mathrm{span}\{\boldsymbol{g}^{(0)},\boldsymbol{G}\boldsymbol{g}^{(0)},\boldsymbol{G}^2\boldsymbol{g}^{(0)},\boldsymbol{G}^3\boldsymbol{g}^{(0)}\},$$

故$\{\boldsymbol{p}^{(0)},\boldsymbol{p}^{(1)},\boldsymbol{p}^{(2)},\boldsymbol{p}^{(3)}\}$中也只有两个独立变量，所以共轭梯度法经**两次迭代后终止**.

5.22 (a) 如果 f 是连续可微的, 则 f 严格凸当且仅当

$$f(\boldsymbol{y})>f(\boldsymbol{x})+\nabla f(\boldsymbol{x})^T(\boldsymbol{y}-\boldsymbol{x}),\quad \forall\ \boldsymbol{x}\neq\boldsymbol{y}.$$

利用该事实证明：对于连续可微的严格凸函数, **曲率条件**

$$\boldsymbol{s}^{(k)T}\boldsymbol{y}^{(k)} > 0$$

对任一向量 $\boldsymbol{x}^{(k)} \neq \boldsymbol{x}^{(k+1)}$ 成立, 其中 $\boldsymbol{s}^{(k)} = \boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}, \boldsymbol{y}^{(k)} = \boldsymbol{g}^{(k+1)} - \boldsymbol{g}^{(k)}$.

(b) 给出一个单变量函数 f 满足 $f(0) = -1$ 和 $f(1) = -1/4$, 并说明在这种情况下曲率条件不成立.

解: (a) 对于严格凸函数，由所给性质有

$$\begin{aligned} f^{(k)} &> f^{(k+1)} + \boldsymbol{g}^{(k+1)T}(\boldsymbol{x}^{(k)} - \boldsymbol{x}^{(k+1)}) \\ f^{(k+1)} &> f^{(k)} + \boldsymbol{g}^{(k)T}(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) \end{aligned}$$

将这两个不等式相加，得

$$(\boldsymbol{g}^{(k)} - \boldsymbol{g}^{(k+1)})^T(\boldsymbol{x}^{(k+1)} - \boldsymbol{x}^{(k)}) < 0,$$

给该不等式两边同时乘以 -1 得 $\boldsymbol{s}^{(k)T}\boldsymbol{y}^{(k)} > 0$.

(b) 取 $f(x) = -x^2 + \frac{7}{4}x - 1$ ，则 $f'(x) = -2x + \frac{7}{4}$. 且 $f(0) = -1, f(1) = -\frac{1}{4}, f'(0) = \frac{7}{4}, f'(1) = -\frac{1}{4}$. 如果取 $\boldsymbol{x}^{(k)} = 0, \boldsymbol{x}^{(k+1)} = 1$ ，则有 $\boldsymbol{s}^{(k)T}\boldsymbol{y}^{(k)} = (1-0)(-\frac{1}{4} - \frac{7}{4}) = -2 < 0$ ，所以曲率条件不成立. ∎

5.23 把精确步长的DFP法应用于习题5.4，其中 $\boldsymbol{H}^{(0)} = \boldsymbol{I}$.

(a) 验证：经过 n 次一维搜索后的二次终止性, 且 $\boldsymbol{H}^{(n)} = \boldsymbol{G}^{-1}$.

(b) 验证该方法与共轭梯度法的等价性.

解：(a) 原问题为最小化二次函数 $q(\boldsymbol{x}) = 10x_1^2 + x_2^2$ ，且

$$\boldsymbol{G} = \begin{bmatrix} 20 & 0 \\ 0 & 2 \end{bmatrix},\ x^{(0)} = (0.1, 1)^T.$$

应用DFP法，其中 $\boldsymbol{H}^{(0)} = \boldsymbol{I}$. 第一步迭代的相关数据如下

$$\begin{aligned}
&\boldsymbol{g}^{(0)} = \boldsymbol{G}\boldsymbol{x}^{(0)} = (2, 2)^T, \\
&\boldsymbol{p}^{(0)} = -\boldsymbol{H}^{(0)}\boldsymbol{g}^{(0)} = -(2, 2)^T, \\
&\alpha^{(0)} = -\frac{\boldsymbol{g}^{(0)\boldsymbol{T}}\boldsymbol{p}^{(0)}}{\boldsymbol{p}^{(0)\boldsymbol{T}}\boldsymbol{G}\boldsymbol{p}^{(0)}} = 1/11, \\
&\boldsymbol{x}^{(1)} = \boldsymbol{x}^{(0)} + \alpha^{(0)}\boldsymbol{p}^{(0)} = (-9/110, 9/11)^T, \\
&\boldsymbol{g}^{(1)} = \boldsymbol{G}\boldsymbol{x}^{(1)} = (-18/11, 18/11)^T, \\
&\boldsymbol{s}^{(0)} = \boldsymbol{x}^{(1)} - \boldsymbol{x}^{(0)} = -(2/11, 2/11)^T, \\
&\boldsymbol{y}^{(0)} = \boldsymbol{g}^{(1)} - \boldsymbol{g}^{(0)} = -(40/11, 4/11)^T, \\
&\boldsymbol{H}^{(1)} = \boldsymbol{H}^{(0)} - \frac{\boldsymbol{H}^{(0)}\boldsymbol{y}^{(0)}\boldsymbol{y}^{(0)\boldsymbol{T}}\boldsymbol{H}^{(0)}}{\boldsymbol{y}^{(0)\boldsymbol{T}}\boldsymbol{H}^{(0)}\boldsymbol{y}^{(0)}} + \frac{\boldsymbol{s}^{(0)}\boldsymbol{s}^{(0)\boldsymbol{T}}}{\boldsymbol{y}^{(0)\boldsymbol{T}}\boldsymbol{s}^{(0)}} = \frac{1}{2222}\begin{bmatrix} 123 & -119 \\ -119 & 2301 \end{bmatrix}.
\end{aligned}$$

第二步迭代的相关数据为

$$
\begin{aligned}
\boldsymbol{p}^{(1)} &= \frac{1}{101}(18, -180)^T, \\
\alpha^{(1)} &= 101/220, \\
\boldsymbol{x}^{(2)} &= (0,0)^T, \\
\boldsymbol{g}^{(2)} &= (0,0)^T, \\
\boldsymbol{s}^{(1)} &= (9/110, -9/11)^T, \\
\boldsymbol{y}^{(1)} &= (18/11, -18/11)^T, \\
\boldsymbol{H}^{(2)} &= \begin{bmatrix} 1/20 & 0 \\ 0 & 1/2 \end{bmatrix}.
\end{aligned}
$$

由于 $\boldsymbol{g}^{(2)} = \mathbf{0}$,因此迭代停止，且 $\boldsymbol{H}^{(2)} = \boldsymbol{G}^{-1}$ ，得证.

(b) 使用共轭梯度法求解该问题，第一步迭代的相关数据如下

$$
\begin{aligned}
\boldsymbol{p}^{(0)} &= -\boldsymbol{g}^{(0)} = -(2,2)^T, \\
\alpha^{(0)} &= -\frac{\boldsymbol{g}^{(0)\boldsymbol{T}}\boldsymbol{g}^{(0)}}{\boldsymbol{p}^{(0)\boldsymbol{T}}\boldsymbol{G}\boldsymbol{p}^{(0)}} = 1/11, \\
\boldsymbol{x}^{(1)} &= \boldsymbol{x}^{(0)} + \alpha_0 \boldsymbol{p}^{(0)} = (-9/110, 9/11)^T, \\
\boldsymbol{g}^{(1)} &= \boldsymbol{G}\boldsymbol{x}^{(1)} = (-18/11, 18/11)^T, \\
\beta^{(0)} &= \frac{\boldsymbol{g}^{(1)\boldsymbol{T}}\boldsymbol{g}^{(1)}}{\boldsymbol{g}^{(0)\boldsymbol{T}}\boldsymbol{g}^{(0)}} = 81/121, \\
\boldsymbol{p}^{(1)} &= -\boldsymbol{g}^{(1)} + \beta^{(0)}\boldsymbol{p}^{(0)} = (36/121, -360/121)^T.
\end{aligned}
$$

进行第二次迭代，得

$$
\begin{aligned}
\alpha^{(1)} &= 11/40, \\
\boldsymbol{x}^{(2)} &= \boldsymbol{x}^{(1)} + \alpha^{(1)}\boldsymbol{p}^{(1)} = (0,0)^T, \\
\boldsymbol{g}^{(2)} &= (0,0)^T.
\end{aligned}
$$

因为 $\boldsymbol{g}^{(2)} = \mathbf{0}$ ，迭代终止，所以两种方法具有同样的二次终止性. ∎

第六章　无约束优化：信赖域法

习题设计说明：

1. 习题6.1：熟悉信赖域子问题的定义，及两种情况(模型函数为开口向上的抛物面的和马鞍面型的)下信赖域子问题解的特点，理解柯西点的定义与确定方法.

2. 习题6.2：将信赖域半径当成参数，计算特殊的信赖域子问的解与柯西点，并判断二者的关系. 帮助理解刻画信赖域子问题解的结论(定理6.2.1).

3. 数值计算：习题6.3和习题6.4均是编写信赖域算法，其中习题6.3利用 dog-leg 折线法求解信赖域子问题，习题6.4利用Steihaug共轭梯度法求解信赖域子问题. 请根据数值结果体会信赖域策略的意义，并体会算法的主要计算量.

4. 信赖域子问题中的信赖域由无穷范数定义：习题6.6-习题6.8.

5. 理解信赖域子问题解的刻画和求解子问题的精确数值方法：习题6.9，习题6.14.

6. 理解LM轨道，即2范数信赖域子问题中半径变化时，子问题的解的轨迹：习题6.10-习题6.13.

6.1 设 $f(\boldsymbol{x}) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$. 并假定信赖子问题的二次模型函数中取 $\boldsymbol{B}^{(k)} = \boldsymbol{G}^{(k)}$.

(a) 完整写出信赖域法在点 $\boldsymbol{x}^{(k)} = (0, -1)^T$ 的二次子问题, 并画出子问题目标函数的等值线.

(b) 针对该子问题, 画出信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时, 信赖域子问题解族的示意图.

(c) 对 $\Delta = 1$, 求出该信赖域子问题的柯西点.

对于点 $\boldsymbol{x}^{(k)} = (0, \frac{1}{2})^T$ 重复上面的工作.

解: (a) 首先 $f(\boldsymbol{x}^{(k)}) = 11$, $\boldsymbol{g}^{(k)} = \begin{bmatrix} -2 \\ -20 \end{bmatrix}$, $\boldsymbol{G}^{(k)} = \begin{bmatrix} 42 & 0 \\ 0 & 20 \end{bmatrix}$. 所以原问题在 $\boldsymbol{x}^{(k)}$ 的信赖域子问题为

$$\begin{array}{ll} \text{minimize} & m_k(\boldsymbol{s}) := \frac{1}{2}(42s_1^2 + 20s_2^2) - 2s_1 - 20s_2 + 11 \\ \text{subject to} & s_1^2 + s_2^2 \le \Delta^2. \end{array}$$

因为 $\boldsymbol{G}^{(k)}$ 正定，所以信赖域子问题的目标函数的等值线是一族同心椭圆，求 $m_k(\boldsymbol{s})$ 的稳定点可得椭圆的中心 $\boldsymbol{s}^* = (\frac{1}{21}, 1)^T$. 具体的等值线族如图6.1，这里以 s_1 为横坐标，以 s_2 为纵坐标，以 $(0, -1)$ 为原点.

(b) 当信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时，可用图解法求解信赖域子问题，有三种情况. (i) 当 $\Delta = 0$ 时，解 $\boldsymbol{s}^{(k)} = (0, 0)^T$ ；(ii) 当 $0 < \Delta < \|\boldsymbol{s}^*\|_2 = \sqrt{1 + \frac{1}{21^2}}$ 时，以 $(0, 0)^T$ 为中心以 Δ 为半径的圆与 $m_k(\boldsymbol{s})$ 等值线的切点是信赖域子问题的解 $\boldsymbol{s}^{(k)}$ ；(iii) 当 $\Delta \ge \|\boldsymbol{s}^*\|_2$ 时，信赖域子问题的解即为 $\boldsymbol{s}^*$.

从而，当信赖域半径从 $\Delta = 0$ 变到 $\Delta = 2$ 时，信赖域子问题解族的示意图是一条连接 $(0, 0)^T$ 与 $\boldsymbol{s}^*$ 的曲线，为了示意图更准确，取 $\Delta = 1$ ，用图解法得到

子问题的解，三点就可以较精确地确定出该曲线. 当信赖域半径变化时，解族的轨迹如图6.1中绿线所示.

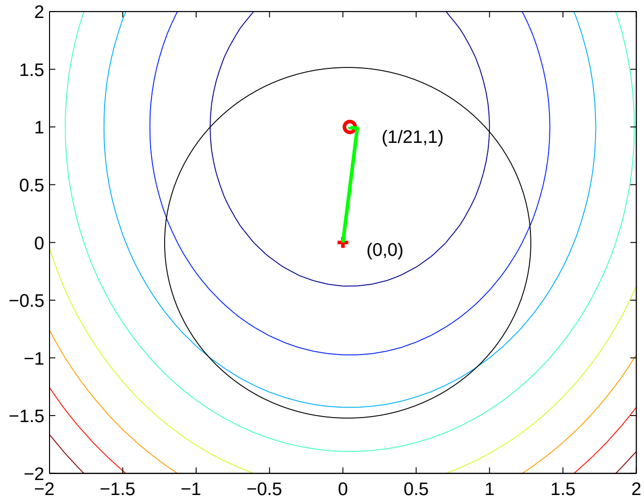


图 6.1: $\boldsymbol{x}^{(k)} = (0, -1)^T$ 处信赖域子问题解的轨迹

(c) 信赖域子问题的柯西点是 $m_k(\boldsymbol{s})$ 沿最速下降方向，且限制在信赖域内的极小点，即需要求解问题

$$\min_{\|-\alpha \boldsymbol{g}^{(k)}\| \le \Delta_k} m_k(-\alpha \boldsymbol{g}^{(k)}). \tag{6.1}$$

将该点处的梯度 $\boldsymbol{g}^{(k)} = (-2, -20)^T, \Delta_k = 1$ 代入(6.1)，得

$$\min_{0 \le \alpha \le \frac{1}{\sqrt{2^2+20^2}}} 4084\alpha^2 - 404\alpha + 11$$

求得目标函数的无约束极小点 $\alpha^* = \frac{404}{2\times 4084} = 0.0495$，其在约束区间 $[0, 1/\sqrt{2^2+20^2}] = [0, 0.0499]$ 内. 故该信赖域子问题的柯西点

$$\boldsymbol{s}_{\mathrm{C}}^{(k)} = -\alpha^* \boldsymbol{g}^{(k)} = \frac{404}{2\times 4084}(2, 20)^T = \frac{101}{1021}(1, 10)^T.$$

在点 $\boldsymbol{x}^{(k)} = (0, \frac{1}{2})^T$ 重复上面的工作:

(a) 对于点 $\boldsymbol{x}^{(k)} = [0, \frac{1}{2}]$，$f(\boldsymbol{x}^{(k)}) = 3.5$，$\boldsymbol{g}^{(k)} = \begin{bmatrix} -2 \\ 10 \end{bmatrix}$, $\boldsymbol{G}^{(k)} = \begin{bmatrix} -18 & 0 \\ 0 & 20 \end{bmatrix}$.

所以原问题在 $\boldsymbol{x}^{(k)}$ 的信赖域子问题为

$$\begin{aligned} \text{minimize} \quad & m_k(\boldsymbol{s}) = \tfrac{1}{2}(-18s_1^2 + 20s_2^2) - 2s_1 + 10s_2 + 3.5 \\ \text{subject to} \quad & s_1^2 + s_2^2 \le \Delta^2. \end{aligned}$$

因为 $\boldsymbol{G}^{(k)}$ 是不定的，从而信赖域子问题的目标函数的等值线是一族双曲线，渐近线为 $3(s_1 + \frac{1}{9}) \pm \sqrt{10}(s_2 + \frac{1}{2}) = 0$. 详见图6.2.

(b) 当 $\Delta = 0$ 时，解为 $\boldsymbol{s}^{(k)} = (0, 0)^T$；当 $0 < \Delta \le 2$ 时，以 $(0, 0)^T$ 为中心以 Δ 为半径的圆与 $m_k(\boldsymbol{s})$ 等值线(水平双曲线的右端分支)的切点是信赖域子问题的解.

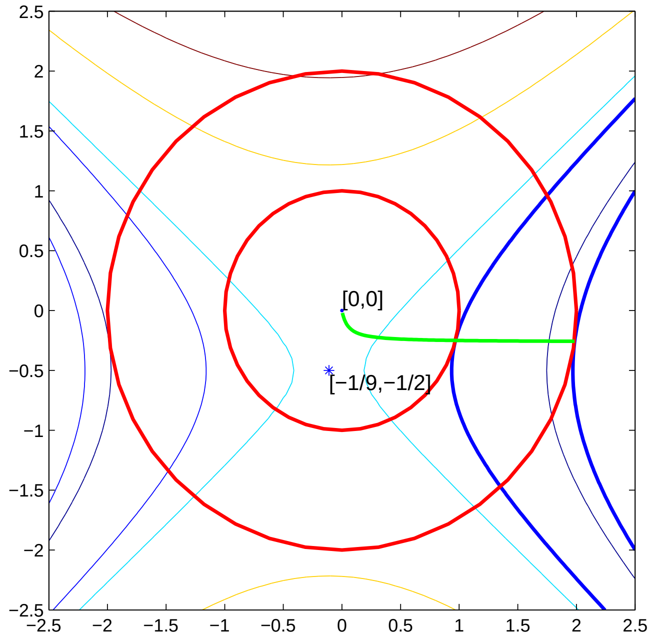


图 6.2: $\boldsymbol{x}^{(k)}=(0,\frac{1}{2})^T$ 处信赖域子问题解的轨迹

(c) 将 $\boldsymbol{g}^{(k)}=(-2,10)^T$ 代入(6.1)，得

$$\min_{0\le\alpha\le 1/\sqrt{2^2+10^2}} 964\alpha^2-104\alpha+3.5$$

求得目标函数的无约束极小点是 $\alpha^*=\frac{52}{964}$ ，其在区间 $[0,1/\sqrt{2^2+10^2}]$ 内. 故该信赖域子问题的柯西点

$$\boldsymbol{s}_{\mathrm{C}}^{(k)}=-\alpha^*\boldsymbol{g}^{(k)}=-\frac{52}{964}(-2,10)^T.$$

∎

6.2 假设在信赖域子问题(6.2.1)中取 $\boldsymbol{B}=\nu\boldsymbol{I}$，其中 ν 为参数. 记子问题(6.2.1)的解为 $\boldsymbol{s}^*$. 请完成以下问题

(a) 写出 $\nu=0$ 时 $\boldsymbol{s}^*$ 的表达式.

(b) 写出 $\nu<0$ 时信赖域子问题的柯西点 $\boldsymbol{s}_{\mathrm{C}}$ 的表达式.

(c) 画出 $\nu>0$ 时 $\frac{q(\boldsymbol{s}_{\mathrm{C}})}{q(\boldsymbol{s}^*)}$ 随信赖域半径 Δ 变化的图像.

解: (a) 将 $\nu=0$ 代入 $m(\boldsymbol{s})$ ，得 $m(\boldsymbol{s})=\boldsymbol{g}^T\boldsymbol{s}$ ，所以对其最小化只需取负梯度方向，并在信赖域边界上取得最优值，即 $\boldsymbol{s}^*=-\frac{\Delta}{\|\boldsymbol{g}\|_2}\boldsymbol{g}$.

(b) 记柯西点 $\boldsymbol{s}_{\mathrm{C}}=-\alpha_{\mathrm{C}}\boldsymbol{g}$. 为此，令

$$\phi(\alpha)=m(-\alpha\boldsymbol{g})=-\alpha\boldsymbol{g}^T\boldsymbol{g}+\frac{1}{2}\nu\alpha^2\boldsymbol{g}^T\boldsymbol{g}=(\frac{1}{2}\nu\alpha^2-\alpha)\boldsymbol{g}^T\boldsymbol{g}, \tag{6.2}$$

为了得到 α_{C} ，需要

$$\min_{0<\alpha<\frac{\Delta}{\|\boldsymbol{g}\|_2}} \phi(\alpha). \tag{6.3}$$

当 $\nu < 0$ ，由(6.2)知 $\phi(\alpha)$ 在 $(0, +\infty)$ 上严格单调递减. 所以柯西点在信赖域边界上，即 $\boldsymbol{s}_{\mathrm{C}} = -\frac{\Delta}{\|\boldsymbol{g}\|_2}\boldsymbol{g}$.

(c) 先求 $m(\boldsymbol{s}_{\mathrm{C}})$ ：当 $\nu > 0$ ，由(6.3)知 $\phi(\alpha)$ 在 $(0, +\infty)$ 上存在唯一极小点 $\alpha^* = \frac{1}{\nu}$. 另外，当信赖域内不包含极小点时，柯西点在其边界上，即 $\alpha_{\mathrm{C}} = \frac{\Delta}{\|\boldsymbol{g}\|_2}$ ；否则 $\alpha_{\mathrm{C}} = \alpha^*$ ，于是

$$m(\boldsymbol{s}_{\mathrm{C}}) = \phi(\alpha_{\mathrm{C}}) = \begin{cases} \frac{1}{2}\nu\Delta^2 - \sqrt{\boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{g}}\Delta, & \Delta < \frac{\|\boldsymbol{g}\|_2}{\nu} \\ -\frac{1}{2\nu}\boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{g}, & \Delta \geq \frac{\|\boldsymbol{g}\|_2}{\nu}. \end{cases}$$

再求 $m(\boldsymbol{s}^*)$ ：根据题意， $\boldsymbol{B} = \nu\boldsymbol{I}$ 正定，所以子问题是一个凸规划，且目标函数的稳定点(即无约束极小点) $\hat{\boldsymbol{s}} = -\frac{1}{\nu}\boldsymbol{g}$. 如果 $\hat{\boldsymbol{s}}$ 在信赖域内，则 $\boldsymbol{s}^* = \hat{\boldsymbol{s}}$. 否则，即 $\Delta^2 < \frac{1}{\nu^2}\boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{g}$ 时，极小点只能在信赖域边界上取得($(\boldsymbol{s}^{*\boldsymbol{T}}\boldsymbol{s}^* = \Delta^2)$)，且此时需考虑

$$m(\boldsymbol{s}) = \boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{s} + \frac{1}{2}\nu\Delta^2$$

由 (a) 可知，它的极小点是 $\boldsymbol{s}^* = -\frac{\Delta}{\|\boldsymbol{g}\|_2}\boldsymbol{g}$, 所以

$$m(\boldsymbol{s}^*) = \begin{cases} \frac{1}{2}\nu\Delta^2 - \sqrt{\boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{g}}\Delta, & \Delta < \frac{\|\boldsymbol{g}\|_2}{\nu} \\ -\frac{1}{2\nu}\boldsymbol{g}^{\boldsymbol{T}}\boldsymbol{g}, & \Delta \geq \frac{\|\boldsymbol{g}\|_2}{\nu}. \end{cases}$$

最后可得 $\frac{m(\boldsymbol{s}_c)}{m(\boldsymbol{s}^*)} = 1$. 这样，当 $\nu > 0$ 时 $\frac{q(\boldsymbol{s}_{\mathrm{C}})}{q(\boldsymbol{s}^*)}$ 随信赖域半径 Δ 变化如图6.3所示. ∎

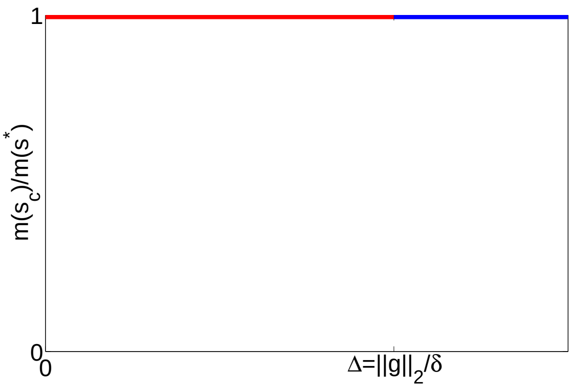


图 6.3: $\frac{m(\boldsymbol{s}_c)}{m(\boldsymbol{s}^*)}$ 与 Δ 的关系