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

2016 年 4 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、杨茜和杨欣，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

2016.4 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

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

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

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

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

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

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

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

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

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

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

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

该问题的最优值与

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

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：

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

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

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

将此问题化成线性规划.

解: 引入变量 t，所给问题等价于

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

考虑问题

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

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

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

2.5 考虑问题

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

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) 引入松弛变量 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	0

$\rightarrow$

	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	1

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
→	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
	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
→	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) 引入松弛变量 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

$\rightarrow$

	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

$\rightarrow$

	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 $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \boldsymbol{x} \geq \boldsymbol{0}$ 的辅助问题应用单纯形法后，所得表格(忽略费用行)形如

x_1		x_k	x_{k+1}	$\cdots$	x_n	y_1	$\cdots$	y_k	y_{k+1}	$\cdots$	y_m	
1									0	$\cdots$	0	b'_1
	$\ddots$			$\boldsymbol{R}_1$			$\boldsymbol{S}_1$		$\vdots$		$\vdots$	$\vdots$
		1							0	$\cdots$	0	b'_k
0	$\cdots$	0							1			0
$\vdots$		$\vdots$		$\boldsymbol{R}_2$			$\boldsymbol{S}_2$			$\ddots$		$\vdots$
0	$\cdots$	0									1	0

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

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

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

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

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

$\rightarrow$

	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	$7/5$
$\to$	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	$-6/5$

		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
$\to$	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

		x_1	x_2	x_3	x_4	$\boldsymbol{x_B}$
	x_1	1	1	0	0	4
$\to$	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 $\boldsymbol{x}^* = (4,0,1,2)^T$ ， $f^* = 11$.

2.19 下面的表格是用单纯形法求解极小化问题的一个中间表格

	$\boldsymbol{y}_1$	$\boldsymbol{y}_2$	$\boldsymbol{y}_3$	$\boldsymbol{y}_4$	$\boldsymbol{y}_5$	$\boldsymbol{y}_6$	$\boldsymbol{y}_0$
	1	2/3	0	0	4/3	0	4
	0	−7/3	3	1	−2/3	0	2
	0	−2/3	−2	0	2/3	1	2
$\boldsymbol{r}^T$	0	8/3	−11	0	4/3	0	−8

(a) 确定下一个转轴元.

(b) 给定当前基的逆

$$\boldsymbol{B}^{-1} = [\boldsymbol{a}_1, \boldsymbol{a}_4, \boldsymbol{a}_6]^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

和对应的费用系数 $\boldsymbol{c}_{\boldsymbol{B}} = (c_1, c_4, c_6)^T = (-1, -3, 1)^T$. 确定原问题的数据 $(\boldsymbol{c}, \boldsymbol{A}, \boldsymbol{b})$.

解:

(a) 下一个转轴元为3(第2行第3列)

(b) $\boldsymbol{B} = (\boldsymbol{a}_1, \boldsymbol{a}_4, \boldsymbol{a}_6) = (\boldsymbol{B}^{-1})^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$.

由 $\boldsymbol{y}_0 = \boldsymbol{B}^{-1}\boldsymbol{b}$ 可得 $\boldsymbol{b} = \boldsymbol{B}\boldsymbol{y}_0 = \begin{bmatrix} 10 \\ 6 \\ 4 \end{bmatrix}$.

由 $\boldsymbol{y}_q = \boldsymbol{B}^{-1}\boldsymbol{a}_q$ 可得 $[\boldsymbol{a}_2, \boldsymbol{a}_3, \boldsymbol{a}_5] = \boldsymbol{B}[\boldsymbol{y}_2, \boldsymbol{y}_3, \boldsymbol{y}_5] = \begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 2 \\ -3 & 1 & 0 \end{bmatrix}$.

所以

$$\boldsymbol{A} = [\boldsymbol{a}_1 \quad \boldsymbol{a}_2 \quad \boldsymbol{a}_3 \quad \boldsymbol{a}_4 \quad \boldsymbol{a}_5 \quad \boldsymbol{a}_6] = \begin{bmatrix} 2 & -1 & 3 & 1 & 2 & 0 \\ 1 & 0 & -2 & 0 & 2 & 1 \\ 0 & -3 & 1 & 1 & 0 & 1 \end{bmatrix}$$

由 $r_q = c_q - \boldsymbol{c}_{\boldsymbol{B}}^T\boldsymbol{y}_q$ 可得 $(c_2, c_3, c_5)^T = (25/3, -22, 8/3)^T$. 所以原问题为

$$\begin{aligned} \text{minimize} \quad & -x_1 + (25/3)x_2 - 22x_3 - 3x_4 + (8/3)x_5 + x_6 \\ \text{subject to} \quad & 2x_1 - 1x_2 + 3x_3 + x_4 + 2x_5 = 10, \\ & x_1 - 2x_3 + 2x_5 + x_6 = 6, \\ & -3x_2 + x_3 + x_4 + x_6 = 4, \\ & x_i \geq 0, i = 1, 2, 3, 4, 5, 6. \end{aligned}$$

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 $\boldsymbol{x}=\mathbf{0}$ 是原始问题的一个可行解，且由 $\boldsymbol{A}\boldsymbol{x}\leq 0$ 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 $\boldsymbol{\lambda}'\leq\mathbf{0}$ 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 $C=\{\boldsymbol{c}\in\mathbb{R}^n|\boldsymbol{c}=\boldsymbol{A}^T\boldsymbol{\lambda},\ \boldsymbol{\lambda}\geq\mathbf{0}\}$ ，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 $C^\circ=\{\boldsymbol{x}|\boldsymbol{A}\boldsymbol{x}\leq 0\}$ ，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**极锥或者法锥**(polar cone or normal cone)[1]. 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

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

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.

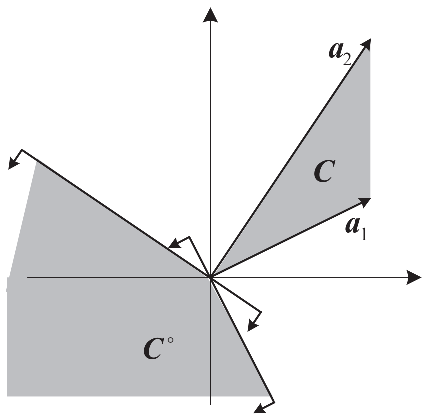


图 2.1: 习题2.25 Farkas引理的几何解释

Farkas引理的另一种描述：

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

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

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

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

[1] $C^\circ=\{\boldsymbol{y}\in\mathbb{R}^n:\boldsymbol{y}^T\boldsymbol{c}\leq 0,\forall\boldsymbol{c}\in C\}$

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.

2.27 **博弈论**(game theory)部分地涉及线性规划理论. 考虑如下博弈，其中参与人甲可以在 m 种策略(strategy)中任选一种，参与人乙可以在 n 种策略中任选一种；如果甲选择策略 i ，乙选择策略 j ，则甲从乙处赢得的数量为 a_{ij} . 博弈通常会重复多次. 假设参与人甲设计了一种**混合**(mixed)策略，记为向量 $\boldsymbol{x}=(x_1,\cdots,x_m)^T$ ，其中分量 $x_i\geq 0$ 表示参与人选择策略 i 的概率，要求 $\sum_{i=1}^m x_i=1$. 同样的，乙的混合策略记为 $\boldsymbol{y}=(y_1,\cdots,y_n)^T$ ，其中 $y_i\geq 0,\sum_{i=1}^n y_i=1$ ，这时甲的平均收益 $P(\boldsymbol{x},\boldsymbol{y})=\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}$.

(a) 考虑线性规划问题

$$
\begin{array}{ll}
\text{maximize} & \alpha \\
\text{subject to} & \sum_{i=1}^m x_i=1 \\
& \sum_{i=1}^m x_i a_{ij}\geq\alpha,\quad j=1,2,\cdots,n \\
& x_i\geq 0,\qquad\qquad\ \ i=1,2,\cdots,m.
\end{array}
$$

设 $(\alpha,\boldsymbol{x})$ 是该问题的可行解. 证明当参与人甲选择可行解中的 $\boldsymbol{x}$ 作为自己的策略时，不管乙选取何种策略 $\boldsymbol{y}$ ，甲的支付至少为 α .

(b) 证明上面问题的对偶是

$$
\begin{array}{ll}
\text{minimize} & \beta \\
\text{subject to} & \sum_{j=1}^n y_j=1 \\
& \sum_{j=1}^n a_{ij}y_j\leq\beta,\ i=1,2,\cdots,m \\
& y_j\geq 0, j=1,2,\cdots,n.
\end{array}
$$

(c) 证明 $\max\alpha=\min\beta$. (称这个共同的值是博弈的**值**(value)，对应的解 $\boldsymbol{x}^*$ 和 $\boldsymbol{y}^*$ 是这个博弈的平衡点).

(d) 考虑“匹配博弈”. 每个参与人选择正面或反面. 如果选择是匹配的话，即一个选正面的同时另外一个选反面，则甲从乙那儿赢得 1 元；如果它们不匹配，乙从甲那儿赢得 1 元. 求出该博弈的值和平衡点.

(e) 考虑博弈，其中每个参与人可以选择 1，2，3 中的某个，相应的支付规则是当参与人的数字相等时，不发生支付；否则，当参与人的数字恰好比另一个参与人的数字大 1 时，该参与人输 3 元；当参与人的数字恰好比另一个参与人的数字大 2 时，参与人赢 1 元. 确定该博弈的值和平衡点.

(a) **证明：**

$$
\begin{aligned}
P(\boldsymbol{x},\boldsymbol{y}) &= \boldsymbol{x}^T\boldsymbol{A}\boldsymbol{y}=(x_1,x_2,...,x_m)\begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ ... & ... & ... & ... \\ a_{m1} & a_{m2} & ... & a_{mn}\end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_n\end{pmatrix} \\
&= y_1\sum_{i=1}^m x_i a_{i1}+y_2\sum_{i=1}^m x_i a_{i2}+...+y_n\sum_{i=1}^m x_i a_{in}
\end{aligned}
$$

由已知，$\sum_{i=1}^m x_i a_{ij}\geq\alpha, j=1,2,...,n$ 可得：

$$
P(\boldsymbol{x},\boldsymbol{y})\geq y_1\alpha+y_2\alpha+...+y_n\alpha=(y_1+y_2+...+y_n)\alpha=\alpha
$$

所以得证 X 的支付至少是 α .

(b) 问题 (a) 可以用向量形式表示为

$$\begin{aligned}
&\text{maximize} \quad (1,0,0,...,0)\begin{pmatrix}\alpha\\ x_1\\ ...\\ x_m\end{pmatrix}\\
&\text{subject to} \quad \begin{pmatrix}\alpha & x_1 & \cdots & x_m\end{pmatrix}\begin{pmatrix}0 & 1 & ... & 1\\ 1 & -a_{11} & ... & -a_{n1}\\ ... & ... & ... & ...\\ 1 & -a_{m1} & ... & -a_{mn}\end{pmatrix}\begin{pmatrix} = & \leq & \cdots & \leq\end{pmatrix}\begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}\\
&\qquad\qquad x_i \geq 0, \quad i=1,2,...,m
\end{aligned}$$

根据一般性对偶原则可写出上述问题的对偶问题是

$$\begin{aligned}
&\text{minimize} \quad \begin{pmatrix}1 & 0 & \cdots & 0\end{pmatrix}\begin{pmatrix}\beta\\ y_1\\ ...\\ y_n\end{pmatrix}\\
&\text{subject to} \quad \begin{pmatrix}0 & 1 & ... & 1\\ 1 & -a_{11} & ... & -a_{n1}\\ ... & ... & ... & ...\\ 1 & -a_{m1} & ... & -a_{mn}\end{pmatrix}\begin{pmatrix}\beta\\ y_1\\ ...\\ y_n\end{pmatrix}\begin{matrix}=\\ \geq\\ \cdots\\ \geq\end{matrix}\begin{pmatrix}1\\ 0\\ ...\\ 0\end{pmatrix}\\
&\qquad\qquad y_j \geq 0, \quad j=1,2,...,n
\end{aligned}$$

化简得

$$\begin{aligned}
\text{minimize} \quad & \beta\\
\text{subject to} \quad & \textstyle\sum_{j=1}^n y_j = 1\\
& \beta - \textstyle\sum_{j=1}^n a_{ij}y_j \geq 0, \quad i=1,2,...,m\\
& y_j \geq 0, \quad j=1,2,...,n,
\end{aligned}$$

此即 (b) 中所给问题.

(c) 该问题为线性规划，并且再由已知条件可知 (a) 问题必有有限最优解. 再由线性规划的强对偶定理可得 $\max\alpha = \min\beta$.

(d) 用 1 表示正面，用 2 表示反面. 根据所给规则， $a_{11}=a_{22}=-1, a_{12}=a_{21}=1$. 从而参与人甲对应的问题是

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

该问题的最优解 $\boldsymbol{x}^* = (\frac{1}{2},\frac{1}{2})^T$ ，最优值为 0；同理可得 $\boldsymbol{y}^* = (\frac{1}{2},\frac{1}{2})^T$ ，最优值也为 0.

(e) 根据问题的描述，该博弈的支付矩阵为

$$\boldsymbol{A}=\begin{bmatrix} 0 & 3 & -1 \\ -3 & 0 & 3 \\ 1 & -3 & 0 \end{bmatrix}$$

对应的(a)中的线性规划问题是

$$\begin{aligned} \text{maximize} \quad & \alpha \\ \text{subject to} \quad & x_1+x_2+x_3=1 \\ & -3x_2+x_3\geq\alpha \\ & 3x_1-3x_3\geq\alpha \\ & -x_1+3x_2\geq\alpha \\ & x_1\geq 0, x_2\geq 0, x_3\geq 0 \end{aligned}$$

该问题的最优解 $\boldsymbol{x}^*=(\frac{3}{7},\frac{1}{7},\frac{3}{7})^T$ ，最优值为 0；同理可得 $\boldsymbol{y}^*=(\frac{3}{7},\frac{1}{7},\frac{3}{7})^T$ ，最优值也为 0.

2.32 (a) 利用单纯形法求解

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

(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

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

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.

2.34 考虑

$$\begin{aligned} \text{minimize} \quad & 12x_1 + 8x_2 + 16x_3 + 12x_4 \\ \text{subject to} \quad & -2x_1 - x_2 - 4x_3 + x_5 = -2 \\ & -2x_1 - 2x_2 - 4x_4 + x_6 = -3 \\ & x_i \geq 0, \quad i = 1, \cdots, 6. \end{aligned}$$

(a) 利用对偶单纯形法求解问题；

(b) 写出对偶问题，并用图解法求解；

(c) 写出(a)中与每张单纯形表对应的单纯形乘子，并将这些点标在(b)中的图上．写出你发现的事实.

2.34 **解：** (a) 利用对偶单纯形法求解原始问题的步骤如下，

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_5	-2	-1	$\boxed{-4}$	0	1	0	-2
x_6	-2	-2	0	-4	0	1	-3
$\boldsymbol{r}^T$	12	8	16	12	0	0	0

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_3	1/2	1/4	1	0	-1/4	0	1/2
x_6	$\boxed{-2}$	-2	0	-4	0	1	-3
$\boldsymbol{r}^T$	4	4	0	12	4	0	-8

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_3	0	$\boxed{-1/4}$	1	-1	-1/4	1/4	-1/4
x_1	1	1	0	2	0	-1/2	3/2
$\boldsymbol{r}^T$	0	0	0	4	4	2	-14

$\longrightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	$\boldsymbol{x_B}$
x_2	0	1	-4	4	1	-1	1
x_1	1	0	4	-2	-1	1/2	1/2
$\boldsymbol{r}^T$	0	0	0	4	4	2	-14

所以最优解 $\boldsymbol{x}^* = (1/2, 1, 0, 0, 0, 0)^T$ ，最优目标值 $z^* = 14$.

(b) 其对偶问题如下. 根据图2.2，可知它的最优解 $\boldsymbol{\lambda}^* = (-4, -2)^T$ ，最优目标值是 14 .

$$\begin{aligned} \text{maximize} \quad & -2\lambda_1 - 3\lambda_2 \\ \text{subject to} \quad & -2\lambda_1 - 2\lambda_2 \leq 12 \\ & -\lambda_1 - 2\lambda_2 \leq 8 \\ & -4\lambda_1 \leq 16 \\ & -4\lambda_2 \leq 12 \\ & \lambda_1 \leq 0 \\ & \lambda_2 \leq 0. \end{aligned}$$

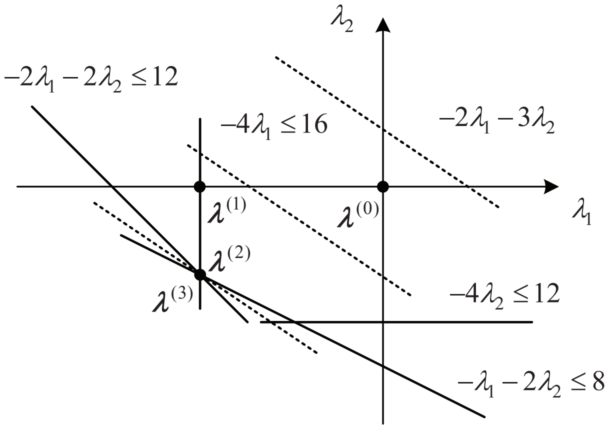


图 2.2: 习题2.34中所给问题的对偶问题的可行域

(c) 由 $\boldsymbol{r}^T = \boldsymbol{c} - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{A}$ 可得

$$\begin{aligned}(r_5, r_6) &= (0, 0) - \boldsymbol{c}_B^T \boldsymbol{B}^{-1} \boldsymbol{I} \\ &= -\boldsymbol{c}_B^T \boldsymbol{B}^{-1}.\end{aligned}$$

而单纯形乘子 $\boldsymbol{\lambda}^T = \boldsymbol{c}_B^T \boldsymbol{B}^{-1}$ ，所以与每张单纯形表对应的乘子分别是

$$\boldsymbol{\lambda}^{(0)} = (0, 0)^T,\ \boldsymbol{\lambda}^{(1)} = (-4, 0)^T,\ \boldsymbol{\lambda}^{(2)} = (-4, -2)^T,\ \boldsymbol{\lambda}^{(3)} = (-4, -2)^T.$$

它们的位置如图2.2，可发现单纯形乘子沿着对偶问题可行域的边界移动，最后到达对偶问题的最优点.

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.

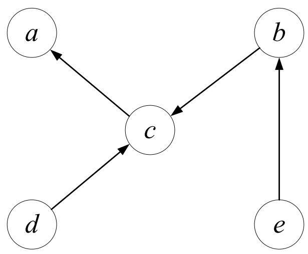
2. 习题3.2和习题3.3是关于网络流问题的理论性质的练习.

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

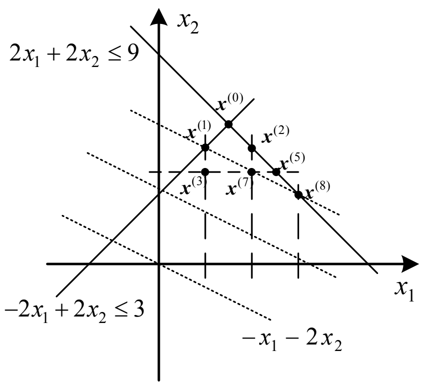
此为最优表，得最优解 $\boldsymbol{x}^* = (2, 0)^T$.

3.9 对线性规划

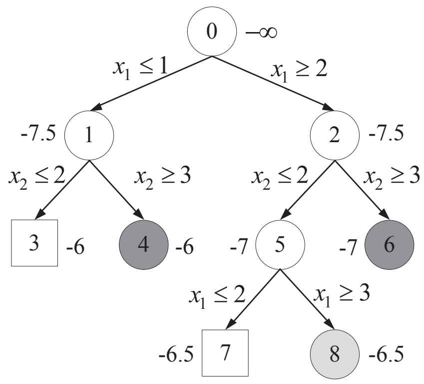
$$\begin{array}{llcll} \text{minimize} & -x_1 & - & 2x_2 & \\ \text{subject to} & -2x_1 & + & 2x_2 & \leq 3 \\ & 2x_1 & + & 2x_2 & \leq 9, \end{array}$$

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	$f^* \geq L$, 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 $\min\limits_{\boldsymbol{x}\in\mathbb{R}^n} \|\boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\|^2$, 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 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{c}$ 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 $f(\boldsymbol{x})=\|\ \boldsymbol{A}\boldsymbol{x}-\boldsymbol{b}\ \|^2=\boldsymbol{x}^T\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}-2\boldsymbol{b}^T\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}^T\boldsymbol{b}$, $\nabla f(\boldsymbol{x})=2\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{x}-2\boldsymbol{A}^T\boldsymbol{b}$. 一阶必要条件为：

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

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，$\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 有无穷多解，而此时满足 $\boldsymbol{A}\boldsymbol{x}=\boldsymbol{b}$ 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 $q(\boldsymbol{x})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{G}\boldsymbol{x}-\boldsymbol{b}^T\boldsymbol{x}$, 证明:

(a) 当且仅当 $\boldsymbol{G}$ 半正定且 $\boldsymbol{G}\boldsymbol{x}=\boldsymbol{b}$ 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

$$e^{-\alpha^2}\leq 1-\rho. \tag{4.3}$$

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 $0<\alpha\leq 1.5174$ ；解(4.3)得 $\alpha\geq 0.3246$ ；解(4.4)得 $\alpha\geq 0.6509$ ；解(4.5)得 $0.6509\leq\alpha\leq 0.7682$.

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 $0<\alpha\leq 1.1774$ ；解(4.3)得 $\alpha>0.5364$ ；解(4.4)得 $\alpha\geq 0.5716$ ；解(4.5)得 $0.5716\leq\alpha\leq 0.8775$.

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).

5.4 考虑使用精确线搜索的最速下降法, 证明对所有的 k , 搜索方向 $\boldsymbol{p}^{(k+1)}$ 正交于 $\boldsymbol{p}^{(k)}$. 将该方法应用于函数 $q(\boldsymbol{x}) = 10x_1^2 + x_2^2$, 选取初始点 $\boldsymbol{x}^{(0)} = (1/10, 1)^T$. 从数值上验证本例可以达到最速下降法的最坏收敛速度, 即方法产生的序列 $\{q^{(k)} - q^*\}$ 是线性收敛，且收敛因子是 $(\lambda_1 - \lambda_n)^2/(\lambda_1 + \lambda_n)^2$, 其中 λ_1 , λ_n 分别为 $\boldsymbol{G}$ 的最大和最小特征值. 注意, 如果 $\boldsymbol{G}$ 相当病态, 则这里的收敛因子能任意接近于 1.

解： 记当前点为 $\boldsymbol{x}^{(k)}$. 则由精确线搜索可知

$$\nabla q(\boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)})^T \boldsymbol{p}^{(k)} = 0$$

由于有 $\nabla q(\boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{p}^{(k)}) = -\boldsymbol{p}^{(k+1)}$ ，故 ${\boldsymbol{p}^{(k+1)}}^T \boldsymbol{p}^{(k)} = 0$. 易见最优解 $\boldsymbol{x}^* = \boldsymbol{0}$. 又因为 $\boldsymbol{G} = \begin{bmatrix} 20 & 0 \\ 0 & 2 \end{bmatrix}$，所以 $\lambda_1 = 20, \lambda_2 = 2$; 从而得收敛因子的上界为

$$\frac{(\lambda_1 - \lambda_2)^2}{(\lambda_1 + \lambda_2)^2} = 0.6694.$$

选取初始点 $\boldsymbol{x}^{(0)} = (1/10, 1)^T$ ，数值结果(后4次迭代)见表6.1. ∎

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

表 6.1: 说明最速下降法的收敛因子可以达到最坏情况的例子

k	$10^{-5}\cdot\boldsymbol{x}^{(k)}$	$\frac{q^{(k+1)}-q^*}{q^{(k)}-q^*}$
49	$(0.656, 6.558)$	0.6693
50	$(-0.537, 5.365)$	0.6694
51	$(0.439, 4.390)$	0.6695
52	$(-0.3592, 3.5920)$	

(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

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

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 $x^\infty=\pm1$, 且 $g^\infty\neq\boldsymbol{0}$. 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$1-\frac{57^2}{91\times55} > 0$，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6.2. 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 $\boldsymbol{g}^{\infty} = \pm\frac{64}{91} \neq 0$.

表 6.2: 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

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 $f^{(k)} + \rho\boldsymbol{g}^{(k)}\boldsymbol{s}^{(k)} - f^{(k+1)} \geq 0$，得

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

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

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

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 $\boldsymbol{g}^{(2)}=\boldsymbol{0}$,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 $\boldsymbol{g}^{(2)}=\boldsymbol{0}$ ，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.

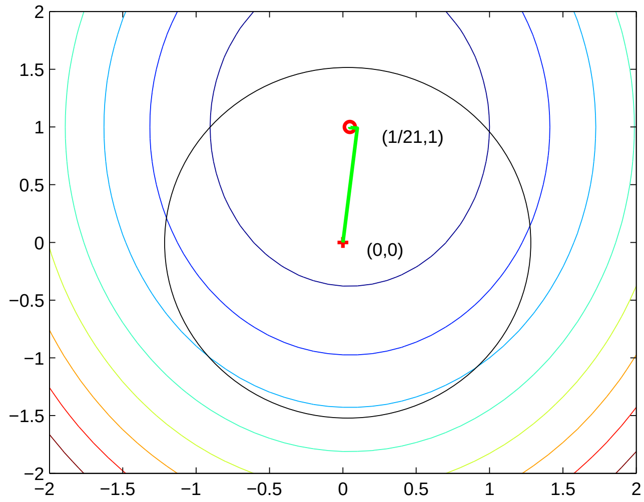


图 6.1: $\boldsymbol{x}^{(k)}=(0,-1)^T$ 处信赖域子问题解的轨迹

(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.

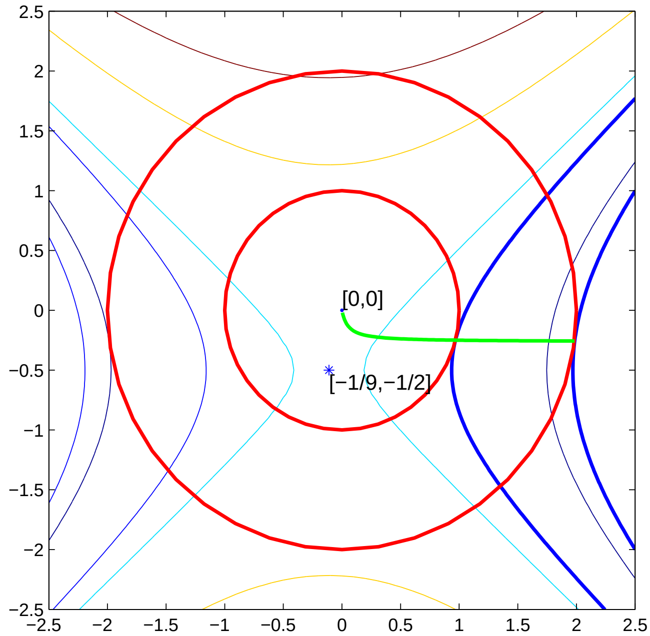


图 6.2: $\boldsymbol{x}^{(k)}=(0,\frac{1}{2})^T$ 处信赖域子问题解的轨迹

(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 $\boldsymbol{s}_{\rm C} = -\frac{\Delta}{||\boldsymbol{g}||_2}\boldsymbol{g}$.

(c) 先求 $m(\boldsymbol{s}_{\rm C})$ ：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 $\alpha_{\rm C} = \frac{\Delta}{||\boldsymbol{g}||_2}$ ；否则 $\alpha_{\rm C} = \alpha^*$ ，于是

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

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 $\frac{q(\boldsymbol{s}_{\rm C})}{q(\boldsymbol{s}^*)}$ 随信赖域半径 Δ 变化如图6.3所示. ∎

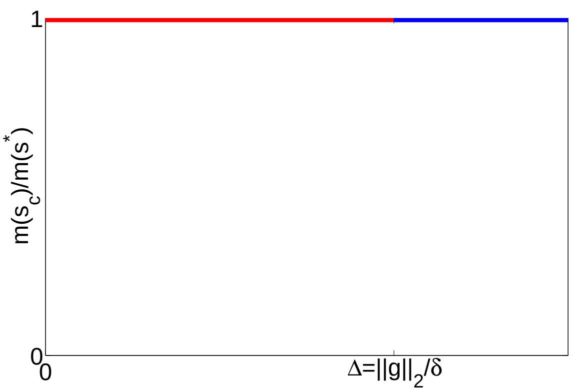


图 6.3: $\frac{m(\boldsymbol{s}_c)}{m(\boldsymbol{s}^*)}$ 与 Δ 的关系

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

2016 年 4 月

第七章　约束优化：理论

习题设计说明：

1. KKT条件的理解和应用：习题7.1-习题7.3，习题7.5，习题7.7，习题7.21(凸规划).

2. 深刻理解KKT条件：KKT条件中的约束规范的作用(习题7.6，习题7.7(a))； Lagrange乘子的意义(习题7.4)；如果积极约束是线性无关的，即LICQ成立，如何设计数值稳定的算法得到Lagrange乘子(习题7.8)；非积极约束对问题局部解和全局解的影响(习题7.20).

3. 二阶最优性条件：习题7.9.

4. 凸函数的性质与次梯度：习题7.10-习题7.11.

5. Lagrange对偶：习题7.12-习题7.14(如何写对偶)，其中习题7.13体会写对偶时的灵活性(即确定那些约束松弛到目标中)，习题7.15(结合具体例子理解对偶定理).

6. SDP的相关练习：习题7.16(SDP中常用的事实)，习题7.17(熟悉SDP的对偶理论)，习题7.18(该结论在工程中很常用).

7. 表述练习：习题7.19给出了工程中常用的三种最优化问题，所给问题是非光滑函数的无约束最优化，常用的软件不能求解这些问题. 利用解决习题2.2和习题2.3时的技巧，可以把这些问题重新表述成光滑的约束最优化，从而利用常用的优化软件进行求解.

7.3 考虑问题

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^2}{\text{minimize}} \quad & (x_1-\tfrac{9}{4})^2+(x_2-2)^2 \\
\text{subject to} \quad & -x_1^2+x_2 \ge 0 \\
& x_1+x_2 \le 6 \\
& x_1 \ge 0,\ x_2 \ge 0.
\end{aligned}$$

(a) 写出 KKT 条件, 并验证点 $\boldsymbol{x}^* = (\frac{3}{2}, \frac{9}{4})^T$ 满足这些条件.

(b) 给出 $\boldsymbol{x}^*$ 处KKT条件的几何解释.

(c) 说明 $\boldsymbol{x}^*$ 是该问题的最优解.

解: (a) $\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda}) = (x_1-\frac{9}{4})^2+(x_2-2)^2+\lambda_1(x_1^2-x_2)+\lambda_2(x_1+x_2-6)+\lambda_3(-x_1)+\lambda_4(-x_2)$.

$$\nabla_x\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda}) = \begin{bmatrix} 2(x_1-\frac{9}{4})+2\lambda_1x_1+\lambda_2-\lambda_3 \\ 2(x_2-2)-\lambda_1+\lambda_2-\lambda_4 \end{bmatrix}.$$

KKT条件:

$$
\begin{aligned}
2(1+\lambda_1)x_1+\lambda_2-\lambda_3-\tfrac{9}{2} &= 0\\
2x_2-\lambda_1+\lambda_2-\lambda_4-4 &= 0\\
x_1^2-x_2 &\le 0\\
x_1+x_2-6 &\le 0\\
x_1\ge 0, x_2 &\ge 0\\
\lambda_1(-x_1^2+x_2) &= 0\\
\lambda_2(6-x_1-x_2) &= 0\\
\lambda_3x_1=0, \lambda_4x_2 &= 0\\
\lambda_i\ge 0, i=1,2, & \;\; 3,4.
\end{aligned}
$$

首先，因为 $\boldsymbol{x}^*=(\frac{3}{2},\frac{9}{4})^T$ 满足约束条件，且有 $\mathcal{I}(\boldsymbol{x}^*)=\{1\}$, 所以由互补条件有 $\lambda_2^*=\lambda_3^*=\lambda_4^*=0$. 再将

$$\boldsymbol{g}^*=(-\frac{3}{2},\frac{1}{2})^T,\quad \boldsymbol{a}_1^*=(3,-1)^T.$$

代入上边的梯度条件，即KKT条件的前两个方程，解得 $\lambda_1^*=\frac{1}{2}$. 从而 $\boldsymbol{x}^*$ 满足KKT条件.

(b) 在 $\boldsymbol{x}^*$ 处， $c_1(\boldsymbol{x})=x_1^2-x_2$ 为积极约束，且梯度 $\boldsymbol{a}_1^*=(3,-1)^T$ ，而 $\boldsymbol{g}^*=(-3/2,1/2)^T$ ，易见 $\boldsymbol{g}^*=-\frac{1}{2}\boldsymbol{a}_1^*$. 所以 $\boldsymbol{x}^*$ 处目标函数与积极约束的梯度共线，且方向相反.

(c) 目标函数显然为凸函数. $c_1(\boldsymbol{x})=x_1^2-x_2$ 为凸函数，其余约束是线性函数，所以此问题为凸规划，从而 KKT 点即是全局解. 所以 $\boldsymbol{x}^*$ 是该问题的全局解. ∎

7.4 计划修建一个长 x_1, 高 x_2, 宽 x_3 (单位: m), 容积 1500 m^3 的仓库. 每平方米的修建费用是: 墙 4 元, 屋顶 6 元, 地板加地面处理共 12 元. 由于美学原因, 长应该是高的两倍. 为了寻找花费最小的设计方案,

(a) 将该问题表述成优化问题, 写出 KKT 条件, 并由 KKT 条件确定解 $\boldsymbol{x}^*$ 和Lagrange乘子 $\boldsymbol{\lambda}^*$;

(b) 从所得优化问题中消去 x_1 和 x_3, 说明距所得解最近的整数 $x_2=10$ 使得费用最小, 然后回代算出 x_1 和 x_3 ;

(c) 设容积约束为 $c_1(\boldsymbol{x})=0$. 在问题中将容积约束变成 $c_1(\boldsymbol{x})=\epsilon$ 时, 用(a) 中的方法求出$f(\boldsymbol{x})$ 在所得解处的改变量 $h(\epsilon)$ 并验证 $h'(0)=-\lambda_1^*$，计算目标值的改变量 $h(-150)$ ，即将所需容积缩减10%时成本的改变量; 比较 $h(-150)$ 与由Lagrange乘子得到的估计值 $-\lambda_1^*\epsilon$.

解: (a) 设费用为 $f(\boldsymbol{x})$ ，依据题意，得

$$
\begin{aligned}
\text{minimize}\quad & f(\boldsymbol{x})=8x_1x_2+8x_2x_3+18x_1x_3\\
\text{subject to}\quad & x_1x_2x_3-1500=0\\
& x_1-2x_2=0.
\end{aligned}
$$

Lagrange 函数 $\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda})=8x_1x_2+8x_2x_3+18x_1x_3+\lambda_1(x_1x_2x_3-1500)+\lambda_2(x_1-2x_2)$.

KKT条件为

$$
\begin{aligned}
8x_2+18x_3+\lambda_1 x_2x_3+\lambda_2&=0,\\
8x_1+8x_3+\lambda_1 x_1x_3-2\lambda_2&=0,\\
8x_2+18x_1+\lambda_1 x_1x_2&=0,\\
x_1x_2x_3-1500&=0,\\
x_1-2x_2&=0.
\end{aligned}
$$

由两个等式约束得 $x_1=2x_2, x_3=\frac{750}{x_2^2}$, 从而原问题可既约为

$$
\begin{array}{ll}
\text{minimize} & 16x_2^2+\frac{33000}{x_2}\\
\text{subject to} & x_2>0.
\end{array}
$$

求稳定点，得 $x_2^*=10\sqrt[3]{\frac{33}{32}}$.

将 $\boldsymbol{x}^*$ 代入 KKT 条件中的梯度条件(第一个和第三个方程)，解得 Lagrange 乘子 $\boldsymbol{\lambda}^*=(\lambda_1^*,\lambda_2^*)^T=(-\frac{22}{x_2^*},-8x_2^*+\frac{3000}{{x_2^*}^2})^T$.

(b) $x_2^*=10\sqrt[3]{33/32}\in(10,11)$. 由于

$$
f(\boldsymbol{x})=16x_2^2+\frac{33000}{x_2}
$$

在 $[0,10\sqrt[3]{33/32}]$ 上递减，且 $f(10)=4900$, 在 $[10\sqrt[3]{33/32},\infty]$ 上递增且 $f(11)=4936$ ，从而 $x_2=10$ 使得费用最小. 此时 $x_1=20, x_3=7.5$ ，这样，我们得到一个实用的设计方案 $(20,10,7.5)$.

(c) 若将 $c_1(\boldsymbol{x})=0$ 改变为 $c_1(\boldsymbol{x})=\varepsilon$ ，得扰动问题

$$
\begin{array}{ll}
\text{minimize} & f(\boldsymbol{x})=8x_1x_2+8x_2x_3+18x_1x_3\\
\text{subject to} & x_1x_2x_3-1500=\varepsilon\\
 & x_1-2x_2=0.
\end{array}
$$

类似于 (a) 中的解法，可得

$$
x_2(\varepsilon)=\sqrt[3]{\frac{33000+22\varepsilon}{32}},\ x_1(\varepsilon)=2x_2,\ x_3(\varepsilon)=\frac{1500+\varepsilon}{2x_2^2}.
$$

代入目标函数，得

$$
h(\varepsilon)=f(\boldsymbol{x}(\varepsilon))-f(\boldsymbol{x}^*)=16(x_2(\varepsilon))^2+22\frac{1500+\varepsilon}{x_2(\varepsilon)}-16(x_2^*)^2-\frac{33000}{x_2^*}
$$

关于 ε 求导数，得

$$
h'(0)=\left(32x_2^*-\frac{33000}{(x_2^*)^2}\right)x_2'(0)+\frac{22}{x_2^*},
$$

其中 $x_2'(0)$ 表示 $x_2(\varepsilon)$ 关于 ε 的导数在 0 处的值. 由 x_2^* 的求法知 $32x_2^*-\frac{33000}{(x_2^*)^2}=0$ ，所以 $h'(0)=-\lambda_1^*$. 这验证了教材上的式(7.2.9)所给出的结论.

当容积减到 $1350\mathrm{m}^3$ 时，$\varepsilon=-150$，此时 $x_2=9.7544, h(-150)=-332.3333$. 根据灵敏度分析，我们也可以得到

$$
h(-150)\approx-\lambda_1^*(-150)=(-150)\frac{22}{x_2^*}=-326.6324. \tag{7.1}
$$

比较这个近似值与精确值，发现近似程度是很高的. 这表明：利用灵敏度分析，我们不用求容积改变后的新问题，而直接利用近似式(7.1) 知道容积减少 10% 后，成本将近似减少 326.6324 元. ∎

7.6 考虑找曲线 $(x-1)^3=y^2$ 上哪个点离原点最近(在Euclidean范数意义下)的问题. 可以将该问题表述为

$$\begin{aligned}&\underset{x,y}{\text{minimize}} && f(x,y)=x^2+y^2\\ &\text{subject to} && (x-1)^3=y^2.\end{aligned}$$

(a) 用图解法求解该问题. 消去 y 求解问题, 所得到的函数有极小点吗? 给出结果不一致的可能原因. 消去 x 求解问题, 得到怎样的结果?

(b) 对于该问题, 找到所有的KKT点. LICQ 成立吗?请给出结果可能的解释.

解: (a) 问题的可行集和目标函数的等值线如图7.1，其中目标函数的等值线是以圆点为中心的同心圆，易见问题的最优解 $x^*=1, y^*=0$.

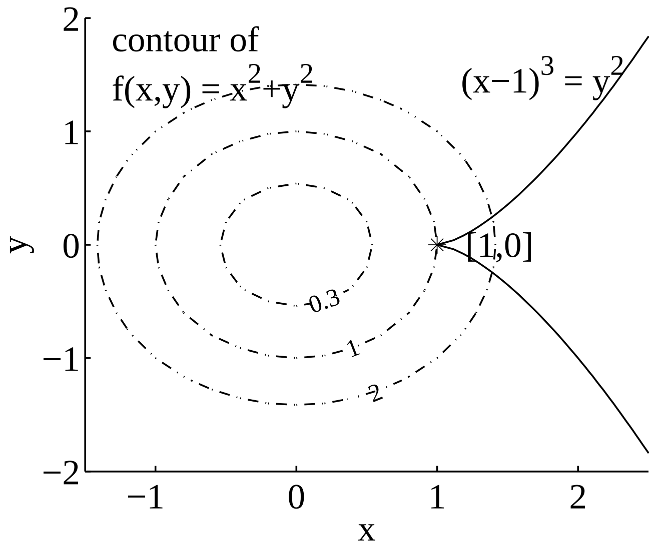


图 7.1: 使用消元法需要注意的问题

由等式约束得到 $y^2=(x-1)^3$ ，将其代入目标函数，得到 $\min\limits_x x^2+(x-1)^3$ ，此时问题无界；这种消元法在将 $y^2=(x-1)^3$ 代入目标函数时忽略了 $x\geq 1$ 这个等式约束所蕴含的条件.

由等式约束得到 $x=y^{2/3}+1$ ，将其代入目标函数，得到 $\min\limits_y (y^{2/3}+1)^2+y^2$. 易见最优解 $y^*=0$ ，于是得原问题的解 $x^*=1, y^*=0$.

(b) 该问题的 Lagrange 函数 $\mathcal{L}(x,y,\lambda)=x^2+y^2+\lambda((x-1)^3-y^2)$. KKT条件为：

$$\begin{aligned}2x+3\lambda(x-1)^2&=0\\ 2y-2\lambda y&=0\\ (x-1)^3-y^2&=0.\end{aligned}$$

第二个方程可写成 $2y(1-\lambda)=0$. 对其分情况讨论可求解该方程组. 具体地，若

$y=0$, 则 $x=1$, 此时第一个方程无解；若 $\lambda=1$, 此时第一个方程无解. 所以该问题没有KKT点!

在该问题的最优解 $x^*=1, y^*=0$ 处，约束梯度 $\boldsymbol{a}^*=(0,0)^T$ ，它不满足LICQ，故不能保证问题的解 $x^*=1, y^*=0$ 一定满足 KKT 条件.

注记:该问题的(a) 说明消元时需要谨慎，否则可能会忽略掉一些隐含条件；从而得出错误的结论；(b)说明引入约束规范的意义，即对于一个优化问题，当最优解处约束规范不成立时，即使是最优解，也可能不满足 KKT 条件，即不存在 Lagrange乘子 . ∎

7.7 考虑问题

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^2}{\text{minimize}} \quad & -x_1 \\ \text{subject to} \quad & x_1^2+x_2^2 \le 1, \\ & (x_1-1)^3 - x_2 \le 0. \end{aligned}$$

(a) 说明线性无关约束规范(LICQ)条件在点 $\boldsymbol{x}^*=(1,0)^T$ 处成立.

(b) 说明点 $\boldsymbol{x}^*=(1,0)^T$ 是一个 KKT 点. 该点是全局极小点吗？请给出理由.

(c) 考虑盒子(界)约束优化问题

$$\begin{aligned} \underset{\boldsymbol{x}\in\mathbb{R}^n}{\text{minimize}} \quad & f(\boldsymbol{x}) \\ \text{subject to} \quad & l_i \le x_i \le u_i, i=1,2,\cdots,n, \end{aligned} \tag{7.2}$$

其中 $l_i, u_i (i=1,2,\cdots,n)$ 是给定的常数，$f(\boldsymbol{x})$ 是连续可微函数. 请问该问题一定有KKT点吗？为什么？如果有KKT点，它一定是问题的最优解吗？请给出理由.

解: (a) 由 $\boldsymbol{a}_1(\boldsymbol{x})=(2x_1,2x_2)^T$, $\boldsymbol{a}_2(\boldsymbol{x})=(3(x_1-1)^2,-1)^T$ 得，$\boldsymbol{a}_1^*=(2,0)^T$, $\boldsymbol{a}_2^*=(0,-1)^T$. 明显地，a_1^*, a_2^* 线性无关，所以LICQ条件成立.

(b) 原问题的Lagrange函数是 $L(\boldsymbol{x},\boldsymbol{\lambda})=-x_1+\lambda_1(x_1^2+x_2^2-1)+\lambda_2[(x_1-1)^3-x_2]$. 因为 $\boldsymbol{x}^*$ 是可行点，积极约束指标集 $\mathcal{I}^*=\{1,2\}$ ，$\boldsymbol{g}^*=(-1,0)^T$，从而判断 KKT 条件是否成立简化为判断关于 $\boldsymbol{\lambda}$ 的方程组

$$\nabla_x L(\boldsymbol{x}^*,\boldsymbol{\lambda})=\boldsymbol{g}^*+\lambda_1\boldsymbol{a}_1^*+\lambda_2\boldsymbol{a}_2^*=\mathbf{0}$$

是否有非负解. 解之得 $\boldsymbol{\lambda}^*=(\frac{1}{2},0)$ ，得到非负的Lagrange乘子，且显然满足互补条件，从而 $\boldsymbol{x}^*$ 是 KKT 点.

此外，由约束 $x_1^2+x_2^2-1\le 0$ 可知 $-1\le x_1\le 1$ ，所以目标函数 $-x_1$ 在边界1上取最小值，因而 $\boldsymbol{x}^*$ 是全局极小点.

(c) 问题(7.2)的约束集是有界闭集，从而若目标函数 $f(\boldsymbol{x})$ 连续，则连续函数必可在有界闭集上取到极小值. 从而(7.2)存在极小点. 因为这个问题的所有约束是线性的，故满足线性约束规范(LCQ)条件，从而极小点必是KKT点. 从而当 $f(\boldsymbol{x})$ 连续时，必存在 KKT 点. 假设 $\boldsymbol{x}^*$ 是 KKT 点，它不一定是问题的最优解. 仅当这个问题是凸规划时，$\boldsymbol{x}^*$ 才是问题的全局最优解. 此外，当 KKT 点惟一时，它也是问题的全局极小点.

7.9 利用一阶和二阶最优性条件找到函数 $f(\boldsymbol{x}) = x_1x_2$ 在单位圆 $x_1^2 + x_2^2 = 1$ 上的极小点，并用图解法求解该问题.

解: 问题的可行集和目标函数的等值线如图7.2，其有两个局部极小点，分别是$(\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})^T$ 和$(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})^T$.

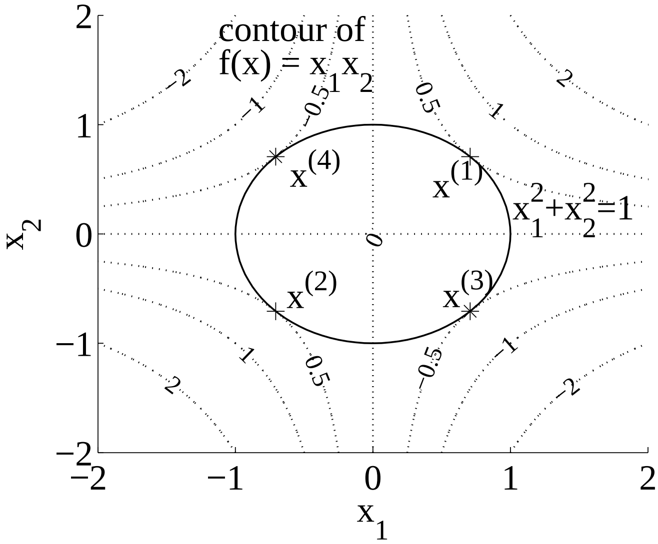


图 7.2: 习题7.9的图示

该问题的 Lagrange 函数 $\mathcal{L}(\boldsymbol{x}, \lambda) = x_1x_2 + \lambda(x_1^2 + x_2^2 - 1)$, KKT条件为：

$$\begin{aligned} x_2 + 2\lambda x_1 &= 0 \\ x_1 + 2\lambda x_2 &= 0 \\ x_1^2 + x_2^2 &= 1. \end{aligned}$$

解上述系统，得KKT点及对应的Lagrange乘子依次为：

$$\begin{aligned} &\boldsymbol{x}^{(1)} = (\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2})^T, && \lambda^{(1)} = -\tfrac{1}{2}, \\ &\boldsymbol{x}^{(2)} = (-\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2})^T, && \lambda^{(2)} = -\tfrac{1}{2}, \\ &\boldsymbol{x}^{(3)} = (\tfrac{\sqrt{2}}{2}, -\tfrac{\sqrt{2}}{2})^T, && \lambda^{(3)} = \tfrac{1}{2}, \\ &\boldsymbol{x}^{(4)} = (-\tfrac{\sqrt{2}}{2}, \tfrac{\sqrt{2}}{2})^T, && \lambda^{(4)} = \tfrac{1}{2}, \end{aligned}$$

对 $i = 1$ ，计算易得：

$$\boldsymbol{a}^{(1)} = \nabla c(\boldsymbol{x}^{(1)}) = \left(\sqrt{2}, \sqrt{2}\right), \quad \boldsymbol{W}^{(1)} = \nabla^2_{\boldsymbol{xx}}\mathcal{L}(\boldsymbol{x}^{(1)}, \lambda^{(1)}) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

$$G(\boldsymbol{x}^{(1)}, \lambda^{(1)}) = \{\boldsymbol{p} \in \mathbb{R}^2 \mid \boldsymbol{p}^T\boldsymbol{a}^{(1)} = 0\} = \{k(1, -1)^T \mid k \in \mathbb{R}\}.$$

因为对 $k \neq 0$ ，有 $k^2(1, -1)\boldsymbol{W}^{(1)}(1, -1)^T = -4k^2 < 0$ ，从而 $\boldsymbol{x}^{(1)}$ 不满足二阶必要条件，所以 $\boldsymbol{x}^{(1)}$ 不是局部解. 同理可验证 $\boldsymbol{x}^{(2)}$ 不是局部解. 对 $i = 3$ ，计算易得：

$$\boldsymbol{a}^{(3)} = \left(\sqrt{2}, -\sqrt{2}\right), \quad \boldsymbol{W}^{(3)} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$G(\boldsymbol{x}^{(3)},\lambda^{(3)}) = \{\boldsymbol{p}\in\mathbb{R}^2 \mid \boldsymbol{p}^T\boldsymbol{a}^{(3)}=0\} = \{k(1,1)^T \mid k\in\mathbb{R}\}.$$

因为对 $k\neq 0$ ，有 $k^2(1,1)\boldsymbol{W}^{(3)}(1,1)^T = 4k^2 > 0$ ，从而 $\boldsymbol{x}^{(3)}$ 满足二阶充分条件，所以 $\boldsymbol{x}^{(3)}$ 是局部解. 同理可验证 $\boldsymbol{x}^{(4)}$ 是局部解.

目标函数的等值线是双曲线，以横轴和纵轴为渐近线. 且在第 I 和第 III 卦限目标函数的值趋于正无穷；在第 II 和第 IV 卦限目标函数的值趋于负无穷.

∎

7.11 设 f 是定义在 $\mathbb{R}^n$ 上的凸函数. 如果向量 $\boldsymbol{p}\in\mathbb{R}^n$ 满足

$$f(\boldsymbol{y}) \geq f(\boldsymbol{x}) + \boldsymbol{p}^T(\boldsymbol{y}-\boldsymbol{x}), \quad \forall \boldsymbol{y}\in\mathbb{R}^n,$$

称 $\boldsymbol{p}$ 是 f 在 $\boldsymbol{x}$ 处的**次梯度**(subgradient)；f 在 $\boldsymbol{x}$ 处所有次梯度的集合记为 $\partial f(\boldsymbol{x})$, 称为 f 在 $\boldsymbol{x}$ 处的**次微分**(subdifferential).

(a) 证明：如果 f 在 $\boldsymbol{x}$ 处可微，则 $\partial f(\boldsymbol{x}) = \{\nabla f(\boldsymbol{x})\}$.

(b) 设 $f(x)=|x|$ ，求 $\partial f(x)$.

解: (a) 证明：一方面，因为 f 在 $\boldsymbol{x}$ 处可微，从而 $\forall\, \boldsymbol{0}\neq\boldsymbol{h}\in\mathbb{R}^n$ 有

$$\lim_{\alpha\to 0}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{h})-f(\boldsymbol{x})}{\alpha} = \boldsymbol{h}^T\nabla f(\boldsymbol{x}).$$

另一方面，任取 $\boldsymbol{p}\in\partial f(\boldsymbol{x})$ ，由次梯度的定义有 $f(\boldsymbol{x}+\alpha\boldsymbol{h})-f(\boldsymbol{x}) \geq \alpha\boldsymbol{p}^T\boldsymbol{h}$. 这样，我们有

$$\boldsymbol{h}^T\nabla f(\boldsymbol{x}) = \lim_{\alpha\searrow 0}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{h})-f(\boldsymbol{x})}{\alpha} \geq \boldsymbol{p}^T\boldsymbol{h}.$$

同时也有

$$\boldsymbol{h}^T\nabla f(\boldsymbol{x}) = \lim_{\alpha\nearrow 0}\frac{f(\boldsymbol{x}+\alpha\boldsymbol{h})-f(\boldsymbol{x})}{\alpha} \leq \boldsymbol{p}^T\boldsymbol{h}.$$

从而得 $\boldsymbol{p}^T\boldsymbol{h} = \nabla f(\boldsymbol{x})^T\boldsymbol{h}$. 再由 $\boldsymbol{h}$ 的任意性，有 $\boldsymbol{p}=\nabla f(\boldsymbol{x})$.

(b) 对 $x\neq 0$，由 (a) 易得 $\partial f(x) = \{\mathrm{sign}(x)\}$. 对$x=0$，由次梯度的定义，可演算得到 $\partial f(0)=[-1,1]$. 从而

$$\partial f(x) = \begin{cases} \{1\}, & x>0 \\ [-1,1], & x=0 \\ \{-1\}, & x<0 \end{cases}$$

∎

7.13 考虑问题

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

分别基于集合约束 $x\in X=\{\boldsymbol{x}\in\mathbb{R}^n \mid \boldsymbol{x}\geq\boldsymbol{0}\}$ 和 $x\in X=\mathbb{R}^n$ 写出该问题的对偶问题.

解： (a) Lagrange 函数 $\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda})=\boldsymbol{c}^T\boldsymbol{x}+\boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x})=(\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A})\boldsymbol{x}+\boldsymbol{\lambda}^T\boldsymbol{b}$. 记对偶函数 $\varphi(\boldsymbol{\lambda})=\min\limits_{\boldsymbol{x}\geq 0}\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda})$. 如果 $\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A}$ 的某分量为负，则 $\varphi(\boldsymbol{\lambda})=-\infty$. 如果 $\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A}\geq\boldsymbol{0}$ ，则 $\varphi(\boldsymbol{\lambda})=\boldsymbol{b}^T\boldsymbol{\lambda}$. 所以对偶问题为

$$\begin{aligned}\underset{\boldsymbol{\lambda}\in\mathbb{R}^m}{\text{maximize}}\quad & \boldsymbol{b}^T\boldsymbol{\lambda}\\ \text{subject to}\quad & \boldsymbol{\lambda}^T\boldsymbol{A}\leq\boldsymbol{c}^T\\ & \boldsymbol{\lambda}\geq\boldsymbol{0}.\end{aligned}$$

(b) Lagrange 函数 $\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})=\boldsymbol{c}^T\boldsymbol{x}+\boldsymbol{\lambda}^T(\boldsymbol{b}-\boldsymbol{A}\boldsymbol{x})+\boldsymbol{\mu}^T(-\boldsymbol{x})=(\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A}-\boldsymbol{\mu}^T)\boldsymbol{x}+\boldsymbol{\lambda}^T\boldsymbol{b}$. 记对偶函数 $\varphi(\boldsymbol{\lambda},\boldsymbol{\mu})=\min\limits_{\boldsymbol{x}\in\mathbb{R}^n}\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\mu})$. 如果 $\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A}-\boldsymbol{\mu}^T$ 的某分量非零，则 $\varphi(\boldsymbol{\lambda},\boldsymbol{\mu})=-\infty$. 如果 $\boldsymbol{c}^T-\boldsymbol{\lambda}^T\boldsymbol{A}-\boldsymbol{\mu}^T=\boldsymbol{0}$ ，则 $\varphi(\boldsymbol{\lambda})=\boldsymbol{b}^T\boldsymbol{\lambda}$. 故对偶问题为

$$\begin{aligned}\underset{\boldsymbol{\lambda}\in\mathbb{R}^m,\boldsymbol{\mu}\in\mathbb{R}^n}{\text{maximize}}\quad & \boldsymbol{b}^T\boldsymbol{\lambda}\\ \text{subject to}\quad & \boldsymbol{\lambda}^T\boldsymbol{A}+\boldsymbol{\mu}^T=\boldsymbol{c}^T,\\ & \boldsymbol{\lambda}\geq\boldsymbol{0},\ \boldsymbol{\mu}\geq\boldsymbol{0}.\end{aligned}$$

注记： 由此题体会写对偶问题的灵活性，即通常仅把难处理的约束松弛到目标函数中形成Lagrange函数，这样对偶问题的变量要少一些. 但是求对偶函数时可能要稍微复杂些，只要能得到显式解，说明选取是成功的. 对于线性规划问题，如本题目，所得到的问题显然是等价的，只不过第二个是标准形问题而已，但是变量的个数比第一个多 n 个. ∎

7.15 写出问题

$$\begin{aligned}\text{minimize}\quad & \tfrac{1}{2}\sigma x_1^2+\tfrac{1}{2}x_2^2+x_1\\ \text{subject to}\quad & x_1\geq 0\end{aligned}$$

的 Lagrange 对偶问题. 分 $\sigma=1$ 和 $\sigma=-1$ 两种情况讨论对偶问题的解是否是原问题最优解处的Lagrange乘子, 并解释两种情况下的结果.

解: 该问题的 KKT 条件为

$$\begin{pmatrix}\sigma x_1+1\\ x_2\end{pmatrix}+\lambda\begin{pmatrix}-1\\ 0\end{pmatrix}=\boldsymbol{0},$$

$$x_1\geq 0,\ \lambda\geq 0,$$

$$x_1\lambda=0.$$

当 $\sigma=1$ 时，解之，得到问题的KKT点 $\boldsymbol{x}^*=(0,0)^T$ ，Lagrange 乘子 $\lambda^*=1$. 当 $\sigma=-1$ 时，解之，得到问题的两个KKT点. 第一个为 $\boldsymbol{x}^*=(0,0)^T$ ，Lagrange 乘子 $\lambda^*=1$；第二个为$\boldsymbol{x}^*=(1,0)^T$ ，Lagrange 乘子 $\lambda^*=0$.

问题的对偶函数

$$\varphi(\lambda)=\min_{\boldsymbol{x}\in\mathbb{R}^2}\frac{1}{2}\sigma x_1^2+\frac{1}{2}x_2^2+x_1-\lambda x_1.$$

若 $\sigma=1$，原始问题是凸规划，从而 KKT 点也是全局解. 进一步，$\min\limits_{\boldsymbol{x}\in\mathbb{R}^2}(\frac{1}{2}\sigma x_1^2+\frac{1}{2}x_2^2+x_1-\lambda x_1)$ 的最优解 $\boldsymbol{x}^*(\lambda)=(\lambda-1,0)^T$ ，得对偶函数 $\varphi(\lambda)=-\frac{1}{2}(\lambda-1)^2$.

所以对偶问题是 $\max\limits_{\lambda\ge 0} -\frac{1}{2}(\lambda-1)^2$ ，其解为 $\lambda^*=1$. 此时对偶问题的解和原始问题 KKT 点处的 Lagrange 乘子相等.

当 $\sigma=-1$ 时，问题无下界. 这时得到的KKT点 $\boldsymbol{x}^*=(0,0)^T$ 是问题的鞍点. 这种情况下，对于任何 λ 值，Lagrange函数中固定$x_2=0$，得到的关于x_1的二次函数的图形开口向下，从而得对偶函数 $\varphi(\boldsymbol{\lambda})=-\infty$. 这可以理解成对偶问题不可行(因为要极大化，但每个可行点处的目标值均为 $-\infty$).

两种情况下结果的解释：当 $\sigma=1$ 时，该问题是凸规划，且满足 Slater 约束规范，故强对偶性成立. 这时，对偶问题的最优解即原始问题的Lagrange乘子，且在 Lagrange 函数中固定对偶最优解 λ^* ，所得函数的无约束极小点是原始问题的解，即 $\boldsymbol{x}^*=\boldsymbol{x}^*(\lambda^*)=(0,0)^T$.

当 $\sigma=-1$ 时，对偶问题是极大化，且对任意的 λ 有 $g(\lambda)=-\infty$ ，这种情况相当于对偶问题没有可行解，从而原始问题无下界. 该事实符合弱对偶性. 然而这种情况下，该问题不是凸规划，所以不能保证强对偶性成立. 该问题的强对偶性的确不成立. ∎

对两个矩阵 $\boldsymbol{A},\boldsymbol{B}\in\mathbb{R}^{k\times l}$ 定义

7.17 设 λ_n 表示 $n\times n$ 对称矩阵 $\boldsymbol{Q}$ 的最小特征值. 说明下面的三个优化问题的最优值均是 λ_n ：

$$\begin{array}{lll} \mathrm{P}_1 & \underset{\boldsymbol{d}\in\mathbb{R}^n}{\text{minimize}} & \boldsymbol{d}^T\boldsymbol{Q}\boldsymbol{d} \\ & \text{subject to} & \boldsymbol{d}^T\boldsymbol{d}=1. \end{array}$$

$$\begin{array}{lll} \mathrm{P}_2 & \underset{\lambda\in\mathbb{R}}{\text{maximize}} & \lambda \\ & \text{subject to} & \boldsymbol{Q}\succeq\lambda\boldsymbol{I}. \end{array}$$

$$\begin{array}{lll} \mathrm{P}_3 & \underset{\boldsymbol{X}\in S^n}{\text{minimize}} & \boldsymbol{Q}\cdot\boldsymbol{X} \\ & \text{subject to} & \boldsymbol{I}\cdot\boldsymbol{X}=1 \\ & & \boldsymbol{X}\succeq\boldsymbol{0}. \end{array}$$

证明: 因为 $\boldsymbol{Q}$ 是对称矩阵，由特征值分解，可以设

$$\boldsymbol{Q}=\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T, \tag{7.3}$$

其中 $\boldsymbol{\Lambda}$ 是对角矩阵，且对角线元素 $\lambda_1\ge\lambda_2\ge\ldots\ge\lambda_n$ 为 $\boldsymbol{Q}$ 的 n 个特征值，$\boldsymbol{P}$ 为正交矩阵.

P_1 : 令 $\boldsymbol{x}=\boldsymbol{P}\boldsymbol{d}$ ，则 $\boldsymbol{d}=\boldsymbol{P}^T\boldsymbol{x}$. 于是约束转化为 $\boldsymbol{x}^T\boldsymbol{x}=1$ ，从而

$$\boldsymbol{d}^T\boldsymbol{Q}\boldsymbol{d}=\sum_{i=1}^n\lambda_i x_i^2\ge\sum_{i=1}^n\lambda_n x_i^2=\lambda_n.$$

且只要取 $\boldsymbol{x}=\boldsymbol{e}_n$ ，即取 $\boldsymbol{d}$ 为 $\boldsymbol{P}$ 的第 n 列，即可以使得上式中的等式成立，从而得到问题 P_1 的最优值为 λ_n .

P_2 : 约束即 $\boldsymbol{Q}-\lambda\boldsymbol{I}$ 为半正定矩阵. 又因为 $\boldsymbol{Q}-\lambda\boldsymbol{I}$ 的特征值为 $\lambda_i-\lambda(i=1,\cdots,n)$ ，且一个矩阵半正定当且仅当它的特征值均非负，所以问题 P_2 的约束条件等价于

$$\lambda_i\ge\lambda,\ i=1,2,\cdots,n.$$

显然最小的使得该条件成立的 λ 即是 λ_n .

P_3 :将 $\boldsymbol{Q}$ 的分解(7.3)代入问题 P_3 ，由习题 7.16 的(a)，目标函数可表示成 $\sum_{i=1}^n \lambda_i(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})_{ii}$. 再由 $\boldsymbol{P}$ 是正交矩阵，约束 $\boldsymbol{I}\bullet\boldsymbol{X}=1$ 等价于 $(\boldsymbol{P}^T\boldsymbol{P})\bullet\boldsymbol{X}=1$. 再由习题 7.16 的(a)和(b)，约束等价于 $\mathrm{trace}(\boldsymbol{P}\boldsymbol{P}^T\boldsymbol{X})=\mathrm{trace}(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})=1$. 此外，由约束 $\boldsymbol{X}\succeq \underline{0}$ 可以推出 $\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P}\succeq\boldsymbol{0}$ ，这意味着 $(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})_{ii}\geq 0$. 从而，目标函数

$$\sum_{i=1}^n \lambda_i(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})_{ii}\geq\sum_{i=1}^n \lambda_n(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})_{ii}=\lambda_n\mathrm{trace}(\boldsymbol{P}^T\boldsymbol{X}\boldsymbol{P})=\lambda_n.$$

且目标函数可以在 $\boldsymbol{X}=\boldsymbol{P}\mathrm{diag}(0,0,0,\ldots,1)\boldsymbol{P}^T$ 时取到 λ_n ，这说明问题 P_3 的最优值是 λ_n .

另外，易见问题 P_2 是问题 P_3 的对偶问题，而 Slater 约束规范显然成立，从而由半定规划的强对偶定理知这两个问题的最优值相等. ∎

7.18 考虑矩阵 $\boldsymbol{M}=\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B}^T & \boldsymbol{C}\end{bmatrix}$, 其中 $\boldsymbol{A},\boldsymbol{C}$ 是对称矩阵, 且 $\boldsymbol{A}$ 是正定的. 证明 $\boldsymbol{M}\succeq\boldsymbol{0}$ 当且仅当**Schur**补 $\boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\succeq\boldsymbol{0}$.

作为该结论的一个简单应用，请将关于 (x,y) 的无约束优化问题，即

$$\underset{x,y\in\mathbb{R}}{\mathrm{minimize}}\sum_{i=1}^n\sqrt{(x-x_i)^2+(y-y_2)^2},$$

重新表述成等价的半定规划问题，这里 $(x_i,y_i)\in\mathbb{R}^2, i=1,2,\cdots,n,$ 是给定的参数.

解: 设 $\boldsymbol{M}\in\mathbb{R}^{n\times n},\boldsymbol{C}\in\mathbb{R}^{m\times m}$ ，考虑非奇异矩阵 $\boldsymbol{T}=\begin{bmatrix}\boldsymbol{I} & -\boldsymbol{A}^{-1}\boldsymbol{B}\\ \boldsymbol{0} & \boldsymbol{I}\end{bmatrix}$, 则 $\boldsymbol{T}^T\boldsymbol{M}\boldsymbol{T}=\begin{bmatrix}\boldsymbol{A} & 0\\ 0 & \boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\end{bmatrix}$ ，即对 $\boldsymbol{M}$ 作初等行变换和相应的初等列变换，将其化成对角块矩阵.

“ $\Rightarrow$ ”设 $\boldsymbol{M}\succeq\boldsymbol{0}$. 则 $\forall\,\boldsymbol{x}\in\mathbb{R}^m$, 令 $\boldsymbol{y}=\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{x}\end{pmatrix}$ ，则

$$\begin{aligned}
&\boldsymbol{x}^T(\boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B})\boldsymbol{x}\\
=\;&(\boldsymbol{0}\ \boldsymbol{x}^T)\begin{bmatrix}\boldsymbol{A} & \boldsymbol{0}\\ \boldsymbol{0} & \boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\end{bmatrix}\begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{x}\end{pmatrix}\\
=\;&\boldsymbol{y}^T\boldsymbol{T}^T\boldsymbol{M}\boldsymbol{T}\boldsymbol{y}\\
=\;&(\boldsymbol{T}\boldsymbol{y})^T\boldsymbol{M}(\boldsymbol{T}\boldsymbol{y})\geq 0,
\end{aligned}$$

最后一个不等式是因为 $\boldsymbol{M}\succeq\boldsymbol{0}$. 故 $\boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\succeq\boldsymbol{0}$.

“ $\Leftarrow$ ”设 $\boldsymbol{C}-\boldsymbol{B}^T\boldsymbol{A}^{-1}\boldsymbol{B}\succeq\boldsymbol{0}$, $\boldsymbol{A}\succeq\boldsymbol{0}$. 则 $\forall\boldsymbol{y}\in\mathbb{R}^{n\times n}$, 记 $\boldsymbol{T}^{-1}\boldsymbol{y}=\begin{pmatrix}\boldsymbol{x}_1\\ \boldsymbol{x}_2\end{pmatrix}$, 其

中 $\boldsymbol{x}_1 \in \mathbb{R}^{n-m}, \boldsymbol{x}_2 \in \mathbb{R}^m$. 于是

$$
\begin{aligned}
\boldsymbol{y}^T \boldsymbol{M} \boldsymbol{y} = &\ \boldsymbol{y}^T \boldsymbol{T}^{-T} \begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B} \end{bmatrix} \boldsymbol{T}^{-1} \boldsymbol{y} \\
= &\ (\boldsymbol{x}_1^T \quad \boldsymbol{x}_2^T) \begin{bmatrix} \boldsymbol{A} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{C} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B} \end{bmatrix} \begin{pmatrix} \boldsymbol{x}_1 \\ \boldsymbol{x}_2 \end{pmatrix} \\
= &\ \boldsymbol{x}_1^T \boldsymbol{A} \boldsymbol{x}_1 + \boldsymbol{x}_2^T (\boldsymbol{C} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B}) \boldsymbol{x}_2
\end{aligned}
$$

最后一个不等式是因为 $\boldsymbol{C} - \boldsymbol{B}^T \boldsymbol{A}^{-1} \boldsymbol{B} \succeq \mathbf{0}$ ， $\boldsymbol{A} \succeq \mathbf{0}$ 正定. 所以 $\boldsymbol{M} \succeq \mathbf{0}$. ∎

所给的二维极小化问题可等价表述为

$$
\begin{array}{ll}
\underset{x,y,d_i \in \mathbb{R}, i=1,\dots,n}{\text{minimize}} & \sum_{i=1}^n d_i \\
\text{subject to} & \sqrt{(x-x_i)^2 + (y-y_2)^2} \leq d_i, i = 1, 2, \cdots, n,
\end{array}
$$

进一步等价于

$$
\begin{array}{ll}
\underset{x,y,d_i \in \mathbb{R}, i=1,\dots,n}{\text{minimize}} & \sum_{i=1}^n d_i \\
\text{subject to} & (x-x_i)^2 + (y-y_2)^2 \leq d_i^2, i = 1, 2, \cdots, n,
\end{array}
$$

进一步等价于

$$
\begin{array}{ll}
\underset{x,y,d_i \in \mathbb{R}, i=1,\dots,n}{\text{minimize}} & \sum_{i=1}^n d_i \\
\text{subject to} & d_i - \begin{bmatrix} x - x_i \\ y - y_i \end{bmatrix}^T \begin{bmatrix} 1/d_i & 0 \\ 0 & 1/d_i \end{bmatrix} \begin{bmatrix} x - x_i \\ y - y_i \end{bmatrix} \geq 0, i = 1, 2, \cdots, n,
\end{array}
$$

由**Schur**补可进一步将问题等价表述为

$$
\begin{array}{ll}
\underset{x,y,d_i \in \mathbb{R}, i=1,\dots,n}{\text{minimize}} & \sum_{i=1}^n d_i \\
\text{subject to} & \begin{bmatrix} d_i & 0 & x - x_i \\ 0 & d_i & y - y_i \\ x - x_i & y - y_i & d_i \end{bmatrix} \succeq \mathbf{0}, i = 1, 2, \cdots, n.
\end{array}
$$

7.19 设 $\boldsymbol{r} : \mathbb{R}^n \to \mathbb{R}^m$ 是光滑的向量值函数. 考虑 $\min\limits_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x})$, 其中

(a) $f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_\infty$;

(b) $f(\boldsymbol{x}) = \max\{r_i(\boldsymbol{x}), i = 1, \cdots, m\}$;

(c) $f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_1$.

请将这些(通常是非光滑的)问题重新表述成光滑的优化问题.

解: 常用技巧：

$$
\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x}) \longleftrightarrow \begin{array}{ll} \underset{\boldsymbol{x} \in \mathbb{R}^n, t \in \mathbb{R}}{\text{minimize}} & t \\ \text{subject to} & f(\boldsymbol{x}) \leq t. \end{array}
$$

(a) 由于 $f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_\infty = \max\limits_{1 \leq i \leq m} |r_i(\boldsymbol{x})|$，而

$$
\max_{1 \leq i \leq m} |r_i(\boldsymbol{x})| \leq t \Longleftrightarrow |r_i(\boldsymbol{x})| \leq t, \ i = 1, \cdots, m,
$$

从而所给问题转化为如下的光滑问题

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}}{\text{minimize}} \quad & t \\
\text{subject to} \quad & -t \le r_i(\boldsymbol{x}) \le t,\ i=1,\cdots,m.
\end{aligned}$$

(b) 由于 $f(\boldsymbol{x}) = \max\limits_{1\le i\le m} r_i(\boldsymbol{x})$，而

$$\max_{1\le i\le m} r_i(\boldsymbol{x}) \le t \Longleftrightarrow r_i(\boldsymbol{x}) \le t,\ i=1,\cdots,m,$$

从而所给问题转化为如下的光滑问题

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}}{\text{minimize}} \quad & t \\
\text{subject to} \quad & r_i(\boldsymbol{x}) \le t,\ i=1,\cdots,m.
\end{aligned}$$

(c) 常用技巧：

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} \sum_{i=1}^m |r_i(\boldsymbol{x})| \longleftrightarrow \begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n, t\in\mathbb{R}^m}{\text{minimize}} \quad & \sum_{i=1}^m t_i \\
\text{subject to} \quad & |r_i(\boldsymbol{x})| \le t_i,\ i=1,\cdots,m.
\end{aligned}$$

由于 $f(\boldsymbol{x}) = \|\boldsymbol{r}(\boldsymbol{x})\|_1 = \sum\limits_{i=1}^m |r_i(\boldsymbol{x})|$，从而所给问题可以转化为如下的光滑问题

$$\begin{aligned}
\underset{\boldsymbol{x}\in\mathbb{R}^n, \boldsymbol{t}\in\mathbb{R}^m}{\text{minimize}} \quad & \sum_{i=1}^m t_i \\
\text{subject to} \quad & -t_i \le r_i(\boldsymbol{x}) \le t_i,\ i=1,\cdots,m.
\end{aligned}$$

7.20 求出 $\min\limits_{x\ge 0} \frac{1}{2}(\frac{1}{2}x^2-14)^2 - x^2 + 3x$ 的解 x^* 和 $\mathcal{A}^*$. 去掉 x^* 处的非积极约束后再次求解问题. 给出你发现的结论.

解：设目标函数为 $f(x)$ ，则 $f'(x) = \frac{1}{2}x^3 - 16x + 3, f''(x) = \frac{3}{2}x^2 - 16$. 令 $f'(x) = 0$ 得 $x_1 = 5.56, x_2 = 0.19, x_3 = -5.7$ (要求 $x \ge 0$ ，故舍去)，又 $f''(x_1) > 0, f''(x_2) < 0$ ，故得局部极小点 $x^* = 5.56, \mathcal{A}^* = \varnothing$.

去掉 x^* 处的非积极约束 $x \ge 0$ 后再次求解，得到两个局部极小点 x_1 和 x_3 ，比较后知 $x_3 = -5.7$ 是这个无约束问题的全局极小点.

结论：若 x^* 是一个问题的局部解，则去掉 x^* 处非积极约束后所得问题而言， x^* 仍是这个问题的局部解，但不一定是全局解. 若 x^* 是凸规划的解，则它仍然是去掉 x^* 处非积极约束问题的全局解. ∎

第八章　约束优化：线性约束规划

习题设计说明：

1. 熟悉消元法：习题8.1，习题8.15，习题8.16，习题8.20.
2. 消元法的深入理解：习题8.3(二次规划)，习题8.11(一般的线性等式约束规划).
3. Lagrange乘子法(值空间法)：习题8.2，习题8.17.
4. 二次规划的积极集法：习题8.7-8.11，习题8.18-8.19.
5. 线性约束规划的积极集法：习题8.13，习题8.14.

8.1 考虑等式二次规划问题

$$\begin{aligned} &\text{minimize} \quad q(\boldsymbol{x}) = \tfrac{1}{2}\boldsymbol{x}^T \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}^T \boldsymbol{x} \\ &\text{subject to} \quad x_1 + 2x_2 + x_3 = 4. \end{aligned}$$

消去变量x_1 后，得到目标函数

$$\frac{1}{2}\boldsymbol{y}^T \boldsymbol{U}\boldsymbol{y} + \boldsymbol{v}^T\boldsymbol{y},$$

其中 $\boldsymbol{U}$ 是一个对称的矩阵, $\boldsymbol{v}$ 是一个固定的向量. 由此给出二次规划问题的解 $\boldsymbol{x}^*$ ，并给出等式约束的拉格朗日乘子 λ^* . 请问 $\boldsymbol{x}^*$ 是否是

$$\begin{aligned} &\text{minimize} \quad q(\boldsymbol{x}) \\ &\text{subject to} \quad x_1 + 2x_2 + x_3 \ge 4, \\ &\qquad\qquad\quad \boldsymbol{x} \ge \boldsymbol{0} \end{aligned}$$

的解.

解. 为了消去x_1，令$\boldsymbol{y} = [x_2\ x_3]^T$，并将约束条件写作$x_1 = 4 - 2x_2 - x_3$，代入目标函数，得

$$q(\boldsymbol{y}) = \frac{1}{2}\boldsymbol{y}^T \begin{bmatrix} 18 & 6 \\ 6 & 4 \end{bmatrix} \boldsymbol{y} - \begin{bmatrix} 29 \\ 12 \end{bmatrix}^T \boldsymbol{y} + 28, \tag{8.1}$$

即将原问题等价表述成一个无约束极小化问题. 由局部极小点的一阶必要条件得方程组

$$\begin{aligned} 18x_2 + 6x_3 - 29 &= 0 \\ 6x_2 + 4x_3 - 12 &= 0 \end{aligned}$$

解得$\boldsymbol{y}^* = [11/9\ \ 7/6]^T$. 易见$q(\boldsymbol{y})$的Hessian阵$\nabla^2 q(\boldsymbol{y})$正定，因此$q(\boldsymbol{y})$是严格凸函数，所以$\boldsymbol{y}^*$是$q(\boldsymbol{y})$的全局极小点.

将$\boldsymbol{y}^*$代入$x_1 = 4 - 2x_2 - x_3$，得$x_1^* = 7/18$. 因此，原问题的解$\boldsymbol{x}^* = [7/18\ 11/9\ 7/6]^T$. 并由KKT条件求得等式约束的拉格朗日乘子$\lambda^* = -17/18$.

易见$\boldsymbol{x}^*$是这个新的不等式约束问题的KKT点，对应的Lagrange乘子是17/18. 又因为$\nabla^2 q(\boldsymbol{x})$正定，因此这个不等式约束问题是凸规划，所以$\boldsymbol{x}^* = [7/18\ \ 11/9\ \ 7/6]^T$是这个问题的全局极小点.

8.2 考虑点 $\boldsymbol{x}^{(0)}$ 到多面集 $\{\boldsymbol{x}\in\mathbb{R}^n:\ \boldsymbol{A}^T\boldsymbol{x}=\boldsymbol{b}\}$ (其中 $\boldsymbol{A}$ 是列满秩的)的最短距离问题. 它可以表述为二次规划

$$\begin{array}{ll}\text{minimize} & \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}^{(0)})^T(\boldsymbol{x}-\boldsymbol{x}^{(0)})\\ \text{subject to} & \boldsymbol{A}^T\boldsymbol{x}=\boldsymbol{b}.\end{array}$$

证明解 $\boldsymbol{x}^*=\boldsymbol{x}^{(0)}-\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{x}^{(0)}-\boldsymbol{b})$. 说明当 $\boldsymbol{A}=\boldsymbol{a}$ 是一个列向量时, 从 $\boldsymbol{x}^{(0)}$ 到超平面 $\boldsymbol{a}^T\boldsymbol{x}=b$ 的最短距离是 $\frac{|\boldsymbol{a}^T\boldsymbol{x}^{(0)}-b|}{\|\boldsymbol{a}\|_2}$.

解: 所给问题的 Lagrange 函数

$$\mathcal{L}(\boldsymbol{x},\boldsymbol{\lambda})=\frac{1}{2}\boldsymbol{x}^T\boldsymbol{x}-\boldsymbol{x}^{(0)T}\boldsymbol{x}+\frac{1}{2}\boldsymbol{x}^{(0)T}\boldsymbol{x}^{(0)}+\boldsymbol{\lambda}^T(\boldsymbol{A}^T\boldsymbol{x}-\boldsymbol{b}).$$

KKT条件为

$$\begin{aligned}\boldsymbol{x}-\boldsymbol{x}^{(0)}+\boldsymbol{A}\boldsymbol{\lambda}&=\boldsymbol{0},\\ \boldsymbol{A}^T\boldsymbol{x}-\boldsymbol{b}&=\boldsymbol{0}.\end{aligned}$$

因为 $\boldsymbol{A}^T$ 行满秩，故该条件的第一个方程等价于 $\boldsymbol{A}^T\boldsymbol{x}-\boldsymbol{A}^T\boldsymbol{x}^{(0)}+\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\lambda}=\boldsymbol{0}$. 将 $\boldsymbol{A}^T\boldsymbol{x}=\boldsymbol{b}$ 代入，得 $\boldsymbol{A}^T\boldsymbol{A}\boldsymbol{\lambda}=\boldsymbol{A}^T\boldsymbol{x}^{(0)}-\boldsymbol{b}$. 再由 $\boldsymbol{A}^T$ 行满秩知 $\boldsymbol{A}^T\boldsymbol{A}$ 可逆. 从而 Lagrange 乘子 $\boldsymbol{\lambda}^*=(\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{x}^{(0)}-\boldsymbol{b})$ ，最优解

$$\boldsymbol{x}^*=\boldsymbol{x}^{(0)}-\boldsymbol{A}\boldsymbol{\lambda}^*=\boldsymbol{x}^{(0)}-\boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}(\boldsymbol{A}^T\boldsymbol{x}^{(0)}-\boldsymbol{b}).$$

特别地，当 $\boldsymbol{A}=\boldsymbol{a}$ 是列向量时，$\{\boldsymbol{x}\in\mathbb{R}^n\ :\ \boldsymbol{a}^T\boldsymbol{x}=b\}$ 表示超平面. 此时 $\boldsymbol{a}^T\boldsymbol{a}=\|\boldsymbol{a}\|_2^2$. 所以

$$\boldsymbol{x}^*=\boldsymbol{x}^{(0)}-\frac{\boldsymbol{a}^T\boldsymbol{x}^{(0)}-b}{\|\boldsymbol{a}\|_2^2}\boldsymbol{a}.$$

$\boldsymbol{x}^{(0)}$ 到超平面 $\{\boldsymbol{x}\in\mathbb{R}^n\ :\ \boldsymbol{a}^T\boldsymbol{x}=b\}$ 的最短距离

$$\sqrt{(\boldsymbol{x}^*-\boldsymbol{x}^{(0)})^T(\boldsymbol{x}^*-\boldsymbol{x}^{(0)})}=\sqrt{\frac{(\boldsymbol{a}^T\boldsymbol{x}^{(0)}-b)^2}{\|\boldsymbol{a}\|_2^4}\boldsymbol{a}^T\boldsymbol{a}}=\frac{|\boldsymbol{a}^T\boldsymbol{x}^{(0)}-b|}{\|\boldsymbol{a}\|_2}.$$

注记：这是点到直线和点到平面的距离公式的推广. 特别地，这里还可以求出垂足 $\boldsymbol{x}^*$. ∎

8.12 在问题(8.3.1)中, 假设 $\operatorname{rank}(\boldsymbol{A})=m$. 考虑由(8.3.4)和(8.3.5)得到的条件 $\boldsymbol{Z}^T\boldsymbol{g}^*=\boldsymbol{0}$ 和 $\boldsymbol{Z}^T\boldsymbol{G}^*\boldsymbol{Z}$ 是正定的与约束优化问题(8.3.1)的一阶和二阶条件之间的联系，即

(a) 证明 $\boldsymbol{Z}^T\boldsymbol{g}^*=\boldsymbol{0}$ 等价于 $\boldsymbol{g}^*+\boldsymbol{A}\boldsymbol{\lambda}^*=\boldsymbol{0}$，

(b) 说明 $\boldsymbol{Z}$ 的列为线性子空间 $N=\{\boldsymbol{p}\in\mathbb{R}^n:\boldsymbol{A}^T\boldsymbol{p}=\boldsymbol{0}\}$ 的一个基, 因此 $\boldsymbol{W}^*$ 在线性子空间 N 上正定等价于 $\boldsymbol{Z}^T\boldsymbol{G}^*\boldsymbol{Z}$ 是正定的，这里$\boldsymbol{W}^*$是(8.3.1)的Lagrange函数$L(\boldsymbol{x},\boldsymbol{\lambda}^*)$在$\boldsymbol{x}^*$的的Hessian阵.

解：(a) 因为 $\boldsymbol{A}^T\boldsymbol{z}_i=0(i=1,2,\cdots,n-m)$ 且 $\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_{n-m}$ 线性无关，所以 $\mathrm{N}(\boldsymbol{A}^T):=\{\boldsymbol{p}\in\mathbb{R}^n:\boldsymbol{A}^T\boldsymbol{p}=\boldsymbol{0}\}=\operatorname{span}\{\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_{n-m}\}$. 又因为 $\boldsymbol{Z}^T\boldsymbol{g}^*=\boldsymbol{0}$ 等价于 $\boldsymbol{g}^*\in(\mathrm{N}(\boldsymbol{A}^T))^\perp$. 再由线性代数基本定理 $\mathrm{N}(\boldsymbol{A}^T)\bigoplus\mathrm{R}(\boldsymbol{A})=\mathbb{R}^n$ 知 $\boldsymbol{Z}^T\boldsymbol{g}^*=\boldsymbol{0}$ 当且仅当 $\boldsymbol{g}^*\in\mathrm{R}(\boldsymbol{A})$ ，即存在 $\boldsymbol{\lambda}^*$ 满足 $\boldsymbol{g}^*+\boldsymbol{A}\boldsymbol{\lambda}^*=\boldsymbol{0}$.

(b) 因为 $\boldsymbol{z}_i$ 使得 $\boldsymbol{A}^T\boldsymbol{z}_i=0(i=1,2,\cdots,n-m)$ 且 $\boldsymbol{z}_1,\boldsymbol{z}_2,\cdots,\boldsymbol{z}_{n-m}$ 线性无关，所以 $\boldsymbol{Z}$ 的列是 $\mathbb{R}^n$ 的子空间 N (即 $\mathrm{N}(\boldsymbol{A}^T)$) 的基. 从而 $\boldsymbol{0}\neq\boldsymbol{p}\in N$ 当且仅当存在 $\boldsymbol{0}\neq\boldsymbol{y}\in\mathbb{R}^{n-m}$ s.t. $\boldsymbol{p}=\boldsymbol{Z}\boldsymbol{y}$. 而

$$\boldsymbol{p}^T\boldsymbol{W}^*\boldsymbol{p}=\boldsymbol{y}^T\boldsymbol{Z}^T\boldsymbol{W}^*\boldsymbol{Z}\boldsymbol{y};$$

又因为这里的约束是线性的，从而$\boldsymbol{W}^* = \boldsymbol{G}^*$，所以 $\boldsymbol{W}^*$ 在线性子空间 N 上正定等价于 $\boldsymbol{Z}^T\boldsymbol{G}^*\boldsymbol{Z}$ 是正定的. ∎

第九章　约束优化：非线性约束规划

习题设计说明：

1. Courant罚函数：习题9.1(等式约束问题)，习题9.16和习题9.17(不等式约束问题)，习题9.18(等式约束问题)

2. 乘子罚函数：习题9.4，习题9.7.

3. SQP法：习题9.10，习题9.11.

4. 对数障碍函数：习题9.3，习题9.8

9.1 考虑用罚函数(9.1.3)求解问题

$$\begin{aligned} &\text{minimize} \quad -x_1x_2x_3 \\ &\text{subject to} \quad 72 - x_1 - 2x_2 - 2x_3 = 0. \end{aligned}$$

验证由 $x_2 = x_3 = 24/(1+\sqrt{(1-8/\sigma)}), x_1 = 2x_2$ 给出的 $\boldsymbol{x}(\sigma)$ 满足表达式 $\nabla\phi(\boldsymbol{x}(\sigma),\sigma) = \mathbf{0}$. 还验证当 $\sigma \to \infty$ 时有 $\boldsymbol{x}(\sigma) \to \boldsymbol{x}^*$. 确定 $\sigma = 9$ 时的 $\boldsymbol{x}(\sigma)$ ，并验证 $\nabla^2\phi(\boldsymbol{x}(9),9)$ 是正定的.

解： 原问题的 Lagrange 函数 $\mathcal{L}(\boldsymbol{x},\lambda) = -x_1x_2x_3 + \lambda(72 - x_1 - 2x_2 - 2x_3)$ ，进而得 KKT 条件为

$$\begin{aligned} -x_2x_3 - \lambda &= 0 \\ -x_1x_3 - 2\lambda &= 0 \\ -x_1x_2 - 2\lambda &= 0 \\ 72 - x_1 - 2x_2 - 2x_3 &= 0 \end{aligned}$$

解得 $\boldsymbol{x}^* = (24,12,12)^T, \lambda^* = -144.$

该问题的 Courant 罚函数问题为

$$\min_{\boldsymbol{x}\in\mathbb{R}^3} \phi(\boldsymbol{x},\sigma) = -x_1x_2x_3 + \frac{1}{2}\sigma(72 - x_1 - 2x_2 - 2x_3)^2,$$

它的梯度条件为

$$\begin{aligned} &-x_2x_3 - \sigma(72 - x_1 - 2x_2 - 2x_3) = 0, \\ &-x_1x_3 - 2\sigma(72 - x_1 - 2x_2 - 2x_3) = 0, \\ &-x_1x_2 - 2\sigma(72 - x_1 - 2x_2 - 2x_3) = 0. \end{aligned} \tag{9.1}$$

将要验证的解 $x_1(\sigma)=\frac{48}{1+\sqrt{1-8/\sigma}}, x_2(\sigma)=x_3(\sigma)=\frac{24}{1+\sqrt{1-8/\sigma}}$ 代入式(9.1)，可验证方程组成立. 从而 $\nabla\phi(\boldsymbol{x}(\sigma),\sigma)=\boldsymbol{0}$. 当 $\sigma\to\infty$ 时，易见 $\boldsymbol{x}(\sigma)\to\boldsymbol{x}^*$.

当 $\sigma=9$ 时，$\boldsymbol{x}(9)=(36,18,18)^T$ ，$\phi(\boldsymbol{x},9)=-x_1x_2x_3+\dfrac{9}{2}(72-x_1-2x_2-2x_3)^2$,

$$\nabla^2\phi(\boldsymbol{x}(9),9)=\begin{bmatrix}9&0&0\\0&36&0\\0&0&36\end{bmatrix} \tag{9.2}$$

易见 $\nabla^2\phi(\boldsymbol{x}(9),9)$ 是正定的. ∎

9.4 考虑用乘子罚函数(9.2.3)求解习题9.1中的问题. 设所取的初始控制参数 $\lambda^{(0)}=0,\sigma=9$. 这时极小点 $\boldsymbol{x}(0,9)$ 与习题 9.1 中已确定的极小点 $\boldsymbol{x}(9)$ 相同, 用公式(9.2.13)与(9.2.15)计算出 $\lambda^{(1)}$, 哪一个公式给出的结果作为 λ^* 的估计较好?

解： 习题9.1中问题的乘子罚函数

$$\phi(\boldsymbol{x},\lambda,\sigma)=-x_1x_2x_3+\lambda(72-x_1-2x_2-2x_3)+\frac{1}{2}\sigma(72-x_1-2x_2-2x_3)^2$$

此时极小点 $\boldsymbol{x}(0,9)=(36,18,18)^T$ ，又由于 $\lambda^{(0)}=0,\sigma=9$ 时，乘子罚函数与习题 9.1 中的 Courant 罚函数相同，所以由式(9.2)，得

$$\boldsymbol{W}_9=\nabla^2\phi(\boldsymbol{x}(9),0,9)=\begin{bmatrix}9&0&0\\0&36&0\\0&0&36\end{bmatrix}.$$

此外 $c(\boldsymbol{x}(9),0)=-36, \boldsymbol{A}^{(0)^T}=(-1,-2,-2), \boldsymbol{A}^{(0)^T}\boldsymbol{W}_9^{-1}\boldsymbol{A}^{(0)}=1/3$. 从而，按牛顿法迭代(9.2.13)得

$$\lambda_{\mathrm{N}}^{(1)}=\lambda^{(0)}+\Big(\frac{1}{3}\Big)^{-1}*(-36)=-108.$$

按 Powell-Hestenes 迭代(9.2.15)得

$$\lambda_{\mathrm{PH}}^{(1)}=\lambda^{(0)}+9*(-36)=-324.$$

因为 $\lambda^*=-144$，所以第一种估计的近似程度优于第二种的近似程度.

9.7 考虑

$$\begin{aligned}&\text{minimize} && f(\boldsymbol{x})=\tfrac{1}{2}(x_1^2-x_2^2)-3x_2\\&\text{subject to} && x_2=0.\end{aligned}$$

(a) 计算最优解和 Lagrange 乘子.

(b) 对 $k=0,1,2$ 和罚参数 $\sigma_k=10^{k+1}$计算Courant罚函数和乘子罚函数($\lambda^{(0)}=0$)所得到的点.

(c) 针对该问题, 画出目标函数的等高线, 在其上标出这两种方法产生的迭代点, 从几何上理解这两种方法.

(d) 假设在乘子罚函数中将罚参数取为常数 σ. 对于 σ 的哪些值, 增广Lagrange函数将有一个极小点? 对于 σ 的哪些值, 方法将是收敛的?

解: (a) 原问题的解显然是 $\boldsymbol{x}^* = (0,0)^T$. 而 $\boldsymbol{g}^* = (0,-3)^T, \boldsymbol{a}^* = (0,1)^T$ ，所以与 $\boldsymbol{x}^*$ 对应的 Lagrange 乘子 $\lambda^* = 3$.

(b) 该问题的 Courant 罚函数 $\phi(\boldsymbol{x},\sigma) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \frac{\sigma}{2}x_2^2$. 令

$$\nabla\phi(\boldsymbol{x},\sigma) = \begin{pmatrix} x_1 \\ -x_2 - 3 + \sigma x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

解得 $x_1(\sigma) = 0, x_2(\sigma) = \frac{3}{\sigma-1}$. 又因为 $\sigma_k = 10^{k+1} > 1$ ，从而

$$\nabla^2\Phi(\boldsymbol{x},\sigma_k) = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_k - 1 \end{bmatrix}$$

正定，所以 $(0, \frac{3}{\sigma_k-1})^T$ 是 $\phi(\boldsymbol{x},\sigma_k)$ 的极小点.

该问题的乘子罚函数 $\phi(\boldsymbol{x},\lambda,\sigma) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \frac{\sigma}{2}x_2^2 + \lambda x_2$. 令

$$\nabla_x\phi(\boldsymbol{x},\lambda,\sigma) = \begin{pmatrix} x_1 \\ -x_2 - 3 + \sigma x_2 + \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},$$

解得 $x_1(\sigma) = 0, x_2(\sigma) = \frac{3-\lambda}{\sigma-1}$. 又因为 $\sigma_k = 10^{k+1} > 1$ ，从而

$$\nabla_{xx}^2\phi(\boldsymbol{x},\lambda^{(k)},\sigma_k) = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_k - 1 \end{bmatrix}$$

正定，所以 $(0, \frac{3-\lambda^{(k)}}{\sigma_k-1})^T$ 是 $\phi(\boldsymbol{x},\lambda^{(k)},\sigma_k)$ 的极小点.

表9.1给出了 $k = 0,1,2$ 时两种方法得到的结果，其中对于 Courant 罚函数，$\lambda^{(k)} = \sigma_k c^{(k)}$；对于乘子罚函数，$\lambda^{(k+1)} = \lambda^{(k)} + \sigma_k c^{(k)}$. 这些数据表明乘子法收敛到解的速度要比 Courant 罚函数法快.

表 9.1: Courant 罚函数法与乘子罚函数法结果的比较

		Courant 罚函数法			乘子罚函数法		
k	σ_k	$(\boldsymbol{x}^{(k)})^T$	$c^{(k)}$	$\lambda^{(k)}$	$\lambda^{(k)}$	$(\boldsymbol{x}^{(k)})^T$	$c^{(k)}$
0	10	$(0, \frac{1}{3})$	$\frac{1}{3}$	$\frac{10}{3}$	0	$(0, \frac{1}{3})$	$\frac{1}{3}$
1	100	$(0, \frac{1}{33})$	$\frac{1}{33}$	$\frac{100}{33}$	$\frac{10}{3}$	$(0, -\frac{1}{297})$	$-\frac{1}{297}$
2	1000	$(0, \frac{1}{333})$	$\frac{1}{333}$	$\frac{1000}{333}$	$\frac{890}{297}$	$(0, \frac{1}{297\times999})$	$\frac{1}{297\times999}$

(c) 图示略.

(d) 对于固定的 $\sigma > 0$ ，因为

$$\nabla_{xx}^2\phi(\boldsymbol{x},\lambda^{(k)},\sigma) = \begin{bmatrix} 1 & 0 \\ 0 & \sigma - 1 \end{bmatrix},$$

所以当 $\sigma > 1$ 时，$\nabla_{\boldsymbol{xx}}^2\mathcal{L}_A(\boldsymbol{x},\lambda^{(k)},\sigma)$ 正定，从而 $(0, \frac{3-\lambda^{(k)}}{\sigma-1})^T$ 是 $\phi(\boldsymbol{x},\lambda^{(k)},\sigma)$ 的极小点. 为了 $k \to \infty$ 迭代的点列趋于 $(0,0)^T$ ，需要 $\lambda^{(k)} \to 3$.

此时，由乘子罚函数法的迭代

$$\lambda^{(k+1)} = \lambda^{(k)} + \sigma x_2^{(k)} = \lambda^{(k)} + \frac{\sigma}{\sigma-1}(3-\lambda^{(k)}) = \frac{1}{1-\sigma}\lambda^{(k)} + \frac{3\sigma}{\sigma-1}.$$

易见对于固定的 σ ，只要序列 $\{\lambda^{(k)}\}$ 收敛，它的极限即为 3 . 为了收敛，需要

$$\left|\frac{1}{1-\sigma}\right| < 1.$$

又因为已经要求 $\sigma > 1$ ，故解得 $\sigma > 2$ 时，$\{\lambda^{(k)}\}$ 是收敛的，且 $\lambda^{(k)} \to 3(k\to\infty)$. 此时 $(0, \frac{3-\lambda^{(k)}}{\sigma-1})^T \to \boldsymbol{x}^*(k\to\infty)$.

综上，对于此问题而言，乘子法不用 $\sigma\to\infty$ ，只需 $\sigma > 2$ 即是收敛的. 从而可以避免病态问题. ∎

9.8 考虑极小化问题

$$\begin{aligned} \underset{x\in\mathbb{R}}{\text{minimize}} \quad & x \\ \text{subject to} \quad & x^2 \ge 0, \\ & x+1 \ge 0. \end{aligned}$$

该问题的解 $x^* = -1$. 写出该问题的对数障碍函数 $\phi(x,\mu)$，并找到它的局部极小点. 说明对数障碍函数的全局极小点收敛到 $x^* = -1$；但是非全局极小点的局部极小点收敛到零, 它不是所给问题的解.

解: 该问题的对数障碍函数 $\phi(x,\mu) = x - \mu\ln x^2 - \mu\ln(x+1)$.令 $\phi'(x,\mu) = 1 - \frac{2\mu}{x} - \frac{\mu}{x+1} = 0$ ，解得

$$x_1(\mu) = \tfrac{3\mu-1+\sqrt{9\mu^2+2\mu+1}}{2},\quad x_2(\mu) = \tfrac{3\mu-1-\sqrt{9\mu^2+2\mu+1}}{2}.$$

因为在区间 $(-1,0)$ 和 $(0,\infty)$ 上分别有 $\phi''(x,\mu) = \frac{2\mu}{x^2} + \frac{\mu}{(x+1)^2} > 0$ ，所以 $\phi(x,\mu)$ 的稳定点 $x_1(\mu)$ 和 $x_2(\mu)$ 分别是函数在这两个区间上的全局极小点，由函数 $\phi(x,\mu)$ 的图形知 $x_1(\mu)$ 是局部极小点(见图9.1). 此外，易见 $\mu\downarrow 0$ 时， $x_1(\mu)\to 0$, $x_2(\mu)\to -1$. ∎

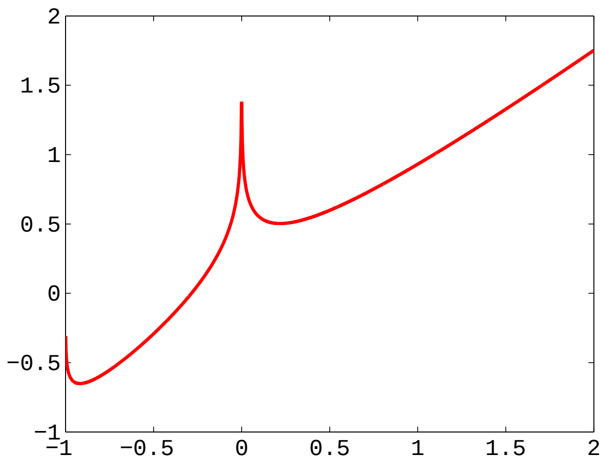


图 9.1: 对数障碍函数法

9.10 考虑约束条件 $c(\boldsymbol{x}) = x_1^2 + x_2^2 - 1 = 0$. 分别写出由 $c(\boldsymbol{x})$ 在点 $(0,0)^T, (0,1)^T, (0.1, 0.02)^T, (-0.1, -0.02)^T$ 处的线性近似所确定的约束条件.

解: 令 $c(\boldsymbol{x}) = x_1^2 + x_2^2 - 1$, $\boldsymbol{x}^1 = (0,0)^T$, $\boldsymbol{x}^2 = (0,1)^T$ ，则

$$\nabla c(\boldsymbol{x}) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix},\ c(\boldsymbol{x}^1) = -1,\ c(\boldsymbol{x}^2) = 0,\ \nabla c(\boldsymbol{x}^1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \nabla c(\boldsymbol{x}^2) = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

所以约束 $x_1^2 + x_2^2 = 1$ 在 $\boldsymbol{x}^1$ 处的线性化约束为 $c(\boldsymbol{x}^1) + \nabla c(\boldsymbol{x}_1)^T(\boldsymbol{x} - \boldsymbol{x}^1) = 0$ ，即 $-1 = 0$. 这是一个矛盾约束. 该例说明在 SQP 法中，每次迭代形成的二次规划子问题有可能出现不可行的情况，从而需要采取相应的措施进行处理，才能使得 SQP 法成为实用算法.

约束 $x_1^2 + x_2^2 = 1$ 在 $\boldsymbol{x}^2$ 处的线性化约束为 $c(\boldsymbol{x}^2) + \nabla c(\boldsymbol{x}^2)^T(\boldsymbol{x} - \boldsymbol{x}^2) = 0$ ，即 $x_2 = 1$; 类似可得约束 $x_1^2 + x_2^2 = 1$ 在其它点的线性化约束. ∎

9.11 用基本SQP法求解问题

$$\begin{array}{ll} \underset{\boldsymbol{x}\in\mathbb{R}^2}{\text{minimize}} & x_1 + x_2 \\ \text{subject to} & x_1^2 - x_2 \le 0. \end{array}$$

(a) 取 $\boldsymbol{x}^{(0)} = \boldsymbol{0}, \lambda^{(0)} = 0$. 给出所得结果的一个可能的原因;

(b) 取 $\boldsymbol{x}^{(0)} = \boldsymbol{0}, \lambda^{(0)} = 1$.

解: 该问题的 Lagrange 函数 $\mathcal{L}(\boldsymbol{x}, \lambda) = x_1 + x_2 + \lambda(x_1^2 - x_2)$. 易得 KKT 点 $\boldsymbol{x}^* = (-\frac{1}{2}, \frac{1}{4})^T, \lambda^* = 1$. 对于 $\boldsymbol{x}^{(k)}$ ，有 $c^{(k)} = (x_1^{(k)})^2 - x_2^{(k)}$ ，

$$\boldsymbol{a}^{(k)} = \begin{bmatrix} 2x_1^{(k)} \\ -1 \end{bmatrix},\ \boldsymbol{g}^{(k)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ \boldsymbol{W}^{(k)} = \nabla_{xx}^2 \mathcal{L}(\boldsymbol{x}^{(k)}, \lambda^{(k)}) = \begin{bmatrix} 2\lambda^{(k)} & 0 \\ 0 & 0 \end{bmatrix}.$$

此时，SQP 法中的二次规划子问题

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\boldsymbol{s}^T\boldsymbol{W}^{(k)}\boldsymbol{s} + \boldsymbol{g}^{(k)^T}\boldsymbol{s} \\ \text{subject to} & c^{(k)} + \boldsymbol{a}^{(k)^T}\boldsymbol{s} \le 0, \end{array}$$

具体为

$$\begin{array}{ll} \text{minimize} & \lambda^{(k)} s_1^2 + s_1 + s_2 \\ \text{subject to} & c^{(k)} + \boldsymbol{a}^{(k)^T}\boldsymbol{s} \le 0, \end{array} \tag{9.3}$$

当 $k = 0, \boldsymbol{x}^{(k)} = (0,0)^T$ 时，

$$c^{(0)} = 0, \quad \boldsymbol{a}^{(0)} = (0, -1)^T, \quad \boldsymbol{g}^{(0)} = (1,1)^T, \quad \boldsymbol{W}^{(0)} = \begin{bmatrix} 2\lambda^{(0)} & 0 \\ 0 & 0 \end{bmatrix}.$$

将上述数据代入子问题(9.3)，得

$$\begin{array}{ll} \underset{\boldsymbol{s}\in\mathbb{R}^2}{\text{minimize}} & \lambda^{(0)} s_1^2 + s_1 + s_2 \\ \text{subject to} & s_2 \ge 0. \end{array}$$

(a) 若取 $\lambda^{(0)}=0$, 则二次规划子问题(9.3)显然无界；从而 SQP 法失败. 这是由初值选取不当造成的.

(b) 若取 $\lambda^{(0)}=1$, 解二次规划子问题(9.3)得解 $\boldsymbol{s}^{(0)}=(-0.5,0)^T$, 对应的 Lagrange 乘子 $\lambda^{(1)}=1$. 于是 $\boldsymbol{x}^{(1)}=\boldsymbol{x}^{(0)}+\boldsymbol{s}^{(0)}=(-0.5,0)^T$.

第二次迭代有

$$c^{(1)}=0.25,\quad \boldsymbol{a}^{(1)}=(-1,-1)^T,\quad \boldsymbol{g}^{(1)}=(1,1)^T,\quad \boldsymbol{W}^{(1)}=\begin{bmatrix}2&0\\0&0\end{bmatrix}.$$

将上述向量代入二次规划子问题(9.3)后，得

$$\begin{aligned}&\text{minimize} && s_1^2+s_1+s_2\\ &\text{subject to} && 0.25-s_1-s_2\le 0.\end{aligned}$$

解这个二次规划子问题得解 $\boldsymbol{s}^{(1)}=(0,0.25)^T$, 对应的 Lagrange 乘子 $\lambda^{(2)}=1$. 于是 $\boldsymbol{x}^{(2)}=\boldsymbol{x}^{(1)}+\boldsymbol{s}^{(1)}=(-0.5,0.25)^T$. 即得到问题的解 $\boldsymbol{x}^*$ 和 Lagrange 乘子 λ^* . 故从该初始点出发，由基本 SQP 法迭代两次即可得到问题的解.

最优化理论与算法 (数值优化) – – 第 1 章作业参考解答

1. (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 (barrel) 原油 A 和 5 百万桶原油 B 进行下个月的生产. 这些资源可以用来生产汽油，其售价为 38 美元 / 桶；或者生产家用加热油，其售价为 33 美元 / 桶. 这里有三种生产过程，其特征如下：

	process 1	process 2	process 3
input crude A	3	1	5
input crude B	5	1	3
output gasoline	4	1	3
onput heating oil	3	1	4
cost	\$51	\$11	\$40

所有的量均以桶为单位. 例如, 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. 表格中的成本指总的成本 (即原油成本和生产过程的成本). 将此问题表述成线性规划, 其可以帮助管理者来极大化下个月的净利润.

请利用 Excel 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，均以百万桶为单位. 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

$$\begin{array}{ll} \text{minimize} & 200x + 60y + 206z \\ \text{subject to} & 3x + y + 5z \le 8 \\ & 5x + y + 3z \le 5 \\ & x, y, z \ge 0. \end{array}$$

最优解是 $x = 0, y = 0.5, z = 1.5$ 百万.

2. 利用你所能想到的方法求解下列问题

$$\begin{array}{ll} \texttt{minimization} & (x_1-2)^2 + (x_2-1)^2 \\ \texttt{Subject to} & x_1^2 - x_2 \le 0, \\ & x_1 + x_2 \le 2. \end{array}$$

提示：可用图解法，即在平面上画出约束集及目标函数的等值线，并观察求解.

3. 设 a 为给定的 $n-$ 维向量， A 为给定的 $n\times n$ 对称矩阵. 计算函数 $f_1(x) = a^T x$ 和 $f_2(x) = x^T A x$ 的梯度和 Hessian 阵.

解答：设 $a=(a_1,\cdots,a_n)^T$, $A=(a_{ij})_{n\times n}$, $i,j=1,2,\cdots,n$, 则 $f_1(x)=a^Tx=a_1x+a_2x+\cdots+a_nx$

所以 $\nabla f_1(x)=(a_1,\cdots,a_n)^T=a$, $\nabla^2 f_1(x)=\theta_{n\times n}$($n$ 阶零方阵)

$f_x(x)=x^TAx=a_{11}x_1^2+a_{22}x_2^2+\cdots+a_{nn}x_n^2+2a_{12}x_1x_2+\cdots+2a_{1n}x_1x_n+2a_{23}x_2x_3+\cdots+2a_{2n}x_2x_n+\cdots+2a_{n-1n}x_{n-1}x_n$

所以 $\frac{\partial f_2}{\partial x_1}=2a_{11}x_1+2a_{12}x_2+\cdots+2a_{1n}x_n=2(a_{11},\cdots,a_{1n})x.$

$\frac{\partial f_2}{\partial x_2}=2a_{21}x_1+2a_{22}x_2+\cdots+2a_{2n}x_n=2(a_{21},\cdots,a_{2n})x.$

$\frac{\partial f_2}{\partial x_2}=2a_{21}x_1+2a_{22}x_2+\cdots+2a_{2n}x_n=2(a_{21},\cdots,a_{2n})x.$

$\cdots$

$\frac{\partial f_2}{\partial x_n}=2a_{n1}x_1+2a_{n2}x_2+\cdots+2a_{nn}x_n=2(a_{n1},\cdots,a_{nn})x.$

所以 $\nabla f_2(x)=2Ax$, 又 $\frac{\partial^2 f_2}{\partial x_i \partial x_j}=2a_{ij}$ 所以 $\nabla^2 f_2(x)=2A$

4. 写出函数 $\cos(1/x)$ 在非零点 x 附近的二阶 Taylor 展式；写出函数 $\cos x$ 在任一点 x 附近的三阶 Taylor 展式；并针对具体值 $x=1$ 计算二阶展开？

解答： $(cos\frac{1}{x})'=-sin\frac{1}{x}(-\frac{1}{x^2})=\frac{1}{x^2}sin\frac{1}{x}$

$(cos\frac{1}{x})''=-\frac{2}{x^3}sin\frac{1}{x}+\frac{1}{x^2}[cos\frac{1}{x}(-\frac{1}{x^2})]=-\frac{2}{x^3}sin\frac{1}{x}-\frac{1}{x^4}cos\frac{1}{x}$

所以 $cos\frac{1}{x}$ 在非零点 x_0 附近的二阶 Taylor 展开式为：

$(cos\frac{1}{x})=cos\frac{1}{x_0}+\frac{1}{x^2}sin\frac{1}{x_0}(x-x_0)-\frac{1}{2}(\frac{2}{x_0^3}sin\frac{1}{x_0}+\frac{1}{x_0^4}cos\frac{1}{x_0})(x-x_0)^2$

所以 $cosx$ 在任意一点 x_0 附近的三阶 Taylor 展开式为：

$cosx=cosx_0-sinx_0(x-x_0)-\frac{1}{2}cosx_0(x-x_0)^2+\frac{1}{6}sinx_0(x-x_0)^3.$

最优化理论与算法 (数值优化) – – 第 2 章作业参考解答

1. 第 2 章 2.2　将下面的问题转化成标准形

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

解：令 $x^+ = \max\{x, 0\}$, $x^- = \max\{-x, 0\}$, 则 $x^+ \ge 0$, $x^- \ge 0$, 且 $x^+ \cdot x^- = 0$, $x = x^+ - x^-$, $|x| = x^+ + x^-$. 由此可将原问题化为

$$\begin{aligned} \text{minimize} \quad & x^+ + x^- + y^+ + y^- + z^+ + z^- \\ \text{subject to} \quad & x^+ - x^- + y^+ - y^- + s = 1, \\ & 2x^+ - 2x^- + z^+ - z^- = 3, \\ & x^+, x^-, y^+, y^-, z^+, z^-, s \ge 0, \\ & x^+x^- = 0, y^+y^- = 0, y^+y^- = 0. \end{aligned}$$

如用单纯形法解此问题，则互补条件自动满足，从而此时便可化为线性规划的标准形式

$$\begin{aligned} \text{minimize} \quad & x^+ + x^- + y^+ + y^- + z^+ + z^- \\ \text{subject to} \quad & x^+ - x^- + y^+ - y^- + s = 1, \\ & 2x^+ - 2x^- + z^+ - z^- = 3, \\ & x^+, x^-, y^+, y^-, z^+, z^-, s \ge 0. \end{aligned}$$

2. 第 2 章 2.3　一类逐段线性函数能够表示为 $f(x) = \texttt{Maximum}(c_1^T x + d_1, c_2^T x + d_2, \cdots, c_p^T x + d_p)$. 针对这样的函数，考虑问题

$$\begin{aligned} \text{minimize} \quad & f(x) \\ \text{subject to} \quad & Ax = b, \\ & x \ge 0. \end{aligned}$$

说明如何将此问题转化成线性规划问题.

3. 第 2 章 2.4　给出

$$\begin{aligned} & x_1 + \tfrac{8}{3}x_2 \le 4, \\ & x_1 + x_2 \le 2, \\ & 2x_1 \le 3, \\ & x_1 \ge 0, x_2 \ge 0. \end{aligned}$$

所对应标准形系统的所有基本解.

解：引入松弛变量 x_3, x_4, x_5, 将原问题化为标准形

$$\begin{cases} x_1 + \frac{8}{3}x_2 + x_3 = 4 \\ x_1 + x_2 + x_4 = 2 \\ 2x_1 + x_5 = 3 \\ x_i \geq 0, \quad i = 1,2,3,4,5 \end{cases}$$

新问题的系数矩阵为 $A = [\, a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \,] = \begin{bmatrix} 1 & \frac{8}{3} & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 & 1 \end{bmatrix}$.

序号	基B	基本解	可行性
1	$[\, a_1 \quad a_2 \quad a_3 \,]$	$(\frac{3}{2}, \frac{1}{2}, \frac{7}{6}, 0, 0)$	可行
2	$[\, a_1 \quad a_2 \quad a_4 \,]$	$(\frac{3}{2}, \frac{15}{16}, 0, -\frac{7}{16}, 0)$	不可行
3	$[\, a_1 \quad a_2 \quad a_5 \,]$	$(\frac{3}{2}, \frac{6}{5}, 0, 0, \frac{7}{5})$	可行
4	$[\, a_1 \quad a_3 \quad a_4 \,]$	$(\frac{3}{2}, 0, \frac{5}{2}, \frac{1}{2}, 0)$	可行
5	$[\, a_1 \quad a_3 \quad a_5 \,]$	$(2, 0, 2, 0, -1)$	不可行
6	$[\, a_1 \quad a_4 \quad a_5 \,]$	$(4, 0, 0, -2, -5)$	不可行
7	$[\, a_2 \quad a_3 \quad a_5 \,]$	$(0, 2, -\frac{4}{3}, 0, 3)$	不可行
8	$[\, a_2 \quad a_4 \quad a_5 \,]$	$(0, \frac{3}{2}, 0, \frac{1}{2}, 3)$	可行
9	$[\, a_3 \quad a_4 \quad a_5 \,]$	$(0, 0, 4, 2, 3)$	可行

4. 第 2 章 2.6　考虑问题

$$\begin{aligned} \text{minimize} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_1nx_n \leq b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_2nx_n \leq b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \\ & x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0, \end{aligned}$$

这种情况下, 约束集完全由线性不等式确定. 为了将其化为标准形, 引入松弛变量 $y_1, y_2, \cdots, y_m$，该问题可以另外表示为

$$\begin{aligned} \text{minimize} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_1nx_n + y_1 = b_1 \\ & a_{21}x_1 + a_{22}x_2 + \cdots + a_2nx_n + y_2 = b_2 \\ & \vdots \\ & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + y_m = b_m \\ & x_1 \geq 0, x_2 \geq 0, \cdots, x_n \geq 0, \\ & y_1 \geq 0, y_2 \geq 0, \cdots, y_m \geq 0. \end{aligned}$$

通过考虑有 $n+m$ 个未知数 $x_1, x_2, \cdots, x_n, y_1, y_2, \cdots, y_m$ 的问题，原问题转化成了标准形．描述线性等式约束的 $m \times (n+m)$ 矩阵具有特殊形式 $[A, I]$(即可将列剖分成几个集合；前 n 列由原

来的矩阵 A 组成，后 m 列由 $m \times m$ 的单位矩阵组成).

上述的两个线性规划问题，一个在 E^n ，另一个在 E^{n+m}. 证明这两个问题的极点之间存在一一对应.

证明. 设原问题的可行集为 S, 化为标准形后其可行集为 $\bar{S}$, 即

$$S = \{x \in \mathbb{R}^n \mid Ax \le b, x \ge 0\}, \quad \bar{S} = \{(x, y) \in \mathbb{R}^{n+m} \mid Ax + y = b, x \ge 0, y \ge 0\}.$$

首先设 x 为 S 的一个极点，令 $y = b - Ax$ ，下面证明 (x, y) 是 $\bar{S}$ 的极点. 如不然，则存在 (x^1, y^1), $(x^2, y^2) \in \bar{S}$ 且 $(x^1, y^1) \neq (x^2, y^2)$ 及 $\alpha \in (0, 1)$ 使得

$$(x, y) = \alpha(x^1, y^1) + (1 - \alpha)(x^2, y^2). \tag{1}$$

一方面，因为 (x^1, y^1), $(x^2, y^2) \in \bar{S}$ 且 $(x^1, y^1) \neq (x^2, y^2)$ ，所以 $y^1 = b - Ax^1 \ge 0$, $y^2 = b - Ax^2 \ge 0$, 从而 $x^1, x^2 \in S$ ，且 $x^1 \neq x^2$(否则， $y^1 = y^2$ ，这与 $(x^1, y^1) \neq (x^2, y^2)$ 矛盾); 另一方面，由 (1) 有 $x = \alpha x^1 + (1 - \alpha)x^2$ ，这与 x 为 S 的极点矛盾！

反之，设 (x, y) 是 S 的极点，则 x 也应该为 S 的极点. 否则，存在 $x^1, x^2 \in S$, 且 $x^1 \neq x^2$, 及 $\alpha \in (0, 1)$, 使得

$$x = \alpha x^1 + (1 - \alpha)x^2. \tag{2}$$

令

$$y^1 = b - Ax^1, \quad y^2 = b - Ax^2, \tag{3}$$

则一方面易见 (x^1, y^1), $(x^2, y^2) \in \bar{S}$ 且 $(x^1, y^1) \neq (x^2, y^2)$ ，另外，由 (2) 和 (3) 有 $(x, y) = \alpha(x^1, y^1) + (1 - \alpha)(x^2, y^2)$ ，这与 (x, y) 是 $\bar{S}$ 的极点矛盾！

综上可知两个问题的极点之间存在一一对应关系.

5. 第 2 章 2.7　利用转轴求解联立方程组：

$$\begin{aligned} 3x_1 + 2x_2 &= 5 \\ 5x_1 + x_2 &= 9. \end{aligned}$$

解：构造初始基得增广表格为

$$\begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b \\ \hline 1 & 0 & \boxed{3} & 2 & 5 \\ 0 & 1 & 5 & 1 & 9 \end{array} \rightarrow \begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b \\ \hline 1/3 & 0 & 1 & 2/3 & 5/3 \\ -5/3 & 1 & 0 & \boxed{-7/3} & 2/3 \end{array} \rightarrow \begin{array}{cccc|c} e_1 & e_2 & a_1 & a_2 & b \\ \hline -1/7 & 2/7 & 1 & 0 & 13/7 \\ 5/7 & -3/7 & 0 & 1 & -2/7 \end{array}$$

解为 $x_1 = 13/7$, $x_2 = -2/7$.

利用转轴求解联立方程组：

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 7 \\ 2x_1 - x_2 + 2x_3 &= 9 \\ x_1 + x_2 + 3x_3 &= 12. \end{aligned}$$

解：构造初始基得增广表格为

e_1	e_2	e_3	a_1	a_2	a_3	b
1	0	0	$\boxed{1}$	2	1	7
0	1	0	2	-1	2	9
0	0	1	1	1	3	12

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
1	0	0	1	2	1	7
-2	1	0	0	$\boxed{-5}$	0	-5
-1	0	1	0	-1	2	5

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
$1/5$	$2/5$	0	1	0	1	5
$2/5$	$-1/5$	0	0	1	0	1
$-3/5$	$-1/5$	1	0	0	$\boxed{2}$	6

$\rightarrow$

e_1	e_2	e_3	a_1	a_2	a_3	b
$1/2$	$1/2$	$-1/2$	1	0	0	2
$2/5$	$-1/5$	0	0	1	0	1
$-3/10$	$-1/10$	$1/2$	0	0	1	3

解为 $x_1 = 2, x_2 = 1, x_3 = 3$.

6. 第 2 章 2.12 将下面的问题转化成标准形并且求解

   (a) 利用单纯形过程，求解

$$\begin{aligned} \text{maximize} \quad & -x_1 + x_2 \\ \text{subject to} \quad & x_1 - x_2 \leq 2, \\ & x_1 + x_2 \leq 6, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

   (b) 画出问题在 x_1, x_2 空间的图形解释，并标明单纯形步骤的迭代路径.

   (c) 针对下面问题重复

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

解： a) 引入松驰变量 x_3, x_4 ，化为标准形

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

写出初始表，其已是第一张单纯形表

	x_1	x_2	x_3	x_4	b
x_3	1	-1	1	0	2
x_4	1	$\boxed{1}$	0	1	6
$c^T(r^T)$	1	-1	0	0	0

$\rightarrow$

	x_1	x_2	x_3	x_4	b
x_3	2	0	1	1	8
x_2	1	1	0	1	6
r^T	2	0	0	1	6

因为 $r_j \geq 0$, 所以得最优解 $x^* = (0, 6)$, 最优值为 $f^* = -6$. 从而原问题最优值为 6.

b)(此处为一曲线图) 单纯形步骤的迭代路径为由 0 到 A, 其中 A 即为最优解.

c) 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	b
x_3	-2	1	1	0	1
x_4	$\boxed{1}$	-1	0	1	1
$c^T(r^T)$	-1	-1	0	0	1

$\rightarrow$

	x_1	x_2	x_3	x_4	b
x_3	0	-1	1	2	3
x_1	1	-1	0	1	1
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.

7. 第 2 章 2.13 解：将原问题化为标准形得初始表为：

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	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
$c^T(r^T)$	-2	-4	-1	-1	0	0	0	0

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	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
r^T	0	-3	-1	-1	0	1	0	3

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_2	0	1	0	$2/5$	$2/5$	$-1/5$	0	1
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
r^T	0	0	-1	$1/5$	$6/5$	$2/5$	0	6

$\rightarrow$

	a_1	a_2	a_3	a_4	a_5	a_6	a_7	b
x_2	0	1	0	$2/5$	$2/5$	$-1/5$	0	1
x_1	1	0	0	$-1/5$	$-1/5$	$3/5$	0	1
x_3	0	0	1	$3/20$	$-1/10$	$1/20$	$1/4$	$1/2$
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 $x^* = (1, 1, 1/2, 0)$.

(a) 设 b 的改变量为 $\Delta b = (\Delta b_1, \Delta b_2, \Delta b_3)$, 要使最优基不变，则需要 $B^{-1}(b + \Delta b) \geq 0$, 即

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

当 b 只有一个分量改变时，由 (1) 可得其各分量的变化范围为

$$\Delta b_1 \in [-5/2, 5],\ \Delta b_2 \in [-5/3, 5],\ \Delta b_3 \in [-2/3, +\infty).$$

当 b 不只一个分量改变时， Δb 应满足 (1). 当 b 的三个分量变化相同时，其变化范围应为 $\epsilon_b = \Delta b_1 \bigcap \Delta b_2 \bigcap \Delta b_3 = [-2/3, 5]$.

(b) 设 c 的改变量为 $\Delta c = (\Delta c_1, \Delta c_2, \Delta c_3, \Delta c_4)$, 要使最优基不变，则需要 $c_D - c_B B^{-1} D \geq 0$, 即

$$(-1 + \Delta c_4, 0, 0, 0) - (-4 + \Delta c_2, -2 + \Delta c_1, -1 + \Delta c_3) \begin{bmatrix} 2/5 & 2/5 & -1/5 & 0 \\ -1/5 & -1/5 & 3/5 & 0 \\ 3/20 & -1/10 & 1/20 & 1/4 \end{bmatrix} \geq 0$$

当 c 只有一个分量改变时，由上式可得其各分量的变化范围为

$$\Delta c_1 \in [-7/4, 3/4],\ \Delta c_2 \in [-9/4, 7/8],\ \Delta c_3 \in [-11, 1],\ \Delta c_4 \in [-7/20, +\infty).$$

当 c 不只一个分量改变时， Δc 应满足上式. 当 c 的三个分量变化相同时，其变化范围应为 $\epsilon_c = \Delta c_1 \bigcap \Delta c_2 \bigcap \Delta c_3 \bigcap \Delta c_4 = [7/20, 3/4]$.

(c) 对于 b 小的改变，由 (a) 可知其最优解将不会改变.

(d) 对于 c 小的改变，由 (b) 可知其最优解将不会改变，但其最优费用会变，即当 c 增加时，最优费用增加； c 减少时，最优费用减少.

8. 第 2 章 2.16 利用两阶段单纯形过程求解

(a)

$$\begin{aligned} \text{minimize} \quad & -3x_1 + x_2 + 3x_3 - x_4 \\ \text{subject to} \quad & x_1 + 2x_2 - x_3 + x_4 = 0, \\ & 2x_1 - 2x_2 + 3x_3 + 3x_4 = 9, \\ & x_1 - x_2 + 2x_3 - x_4 = 6, \\ & x_i \geq 0, i = 1, 2, 3, 4. \end{aligned}$$

(b)

$$\begin{aligned} \text{minimize} \quad & x_1 + 6x_2 - 7x_3 + x_4 + 5x_5 \\ \text{subject to} \quad & 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 = 20, \\ & x_1 - x_2 + 5x_3 - x_4 + x_5 = 8, \\ & x_i \geq 0, i = 1, 2, 3, 4, 5. \end{aligned}$$

解： a) 第 I 阶段：引入人工变量 x_5, x_6, x_7 ，构造辅助问题，得

$$\begin{aligned} \text{minimize} \quad & x_5 + x_6 + x_7 \\ \text{subject to} \quad & x_1 + 2x_2 - x_3 + x_4 + x_5 = 0, \\ & 2x_1 - 2x_2 + 3x_3 + 3x_4 + x_6 = 9, \\ & x_1 - x_2 + 2x_3 - x_4 + x_7 = 6, \\ & x_i \geq 0, i = 1, 2, 3, \cdots, 7. \end{aligned}$$

则辅助问题的初始单纯形表为

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_5	1	2	-1	1	1	0	0	0
x_6	2	-2	3	3	0	1	0	9
x_7	1	-1	2	-1	0	0	1	6
c^T	0	0	0	0	1	1	1	0

将最后一行与基变量 x_5, x_6, x_7 对应的系数化为 0，得第一张单纯形表

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_5	$\boxed{1}$	2	-1	1	1	0	0	0
x_6	2	-2	3	3	0	1	0	9
x_7	1	-1	2	-1	0	0	1	6
r^T	-4	1	-4	-3	0	0	0	-15

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	2	-1	1	1	0	0	0
x_6	0	-6	$\boxed{5}$	1	-2	1	0	9
x_7	0	-3	3	-2	-1	0	1	6
r^T	0	9	-8	1	4	0	0	-15

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	$4/5$	0	$6/5$	$3/5$	$1/5$	0	$9/5$
x_3	0	$-6/5$	1	$1/5$	$-2/5$	$1/5$	0	$9/5$
x_7	0	$\boxed{3/5}$	0	$-13/5$	$1/5$	$-3/5$	1	$3/5$
r^T	0	$-3/5$	0	$13/5$	$4/5$	$8/5$	0	$-3/5$

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_1	1	0	0	$14/3$	$1/3$	1	$-4/3$	1
x_3	0	0	1	-5	0	-1	2	3
x_2	0	1	0	$-13/3$	$1/3$	-1	$5/3$	1
r^T	0	0	0	0	1	1	1	0

由此得原问题的基本可行解 $(1, 1, 3, 0)$.

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	b
x_1	1	0	0	$14/3$	1
x_3	0	0	1	-5	3
x_2	0	1	0	$-13/3$	1
c^T	-3	1	3	-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

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	$14/3$	1
x_3	0	0	1	-5	3
x_2	0	1	0	$-13/3$	1
r^T	0	0	0	$97/3$	-7

已达最优，所以 $x^* = (1, 1, 3, 0)$, 最优值 $f^* = 7$.

b) 第 I 阶段：引入人工变量 x_6, x_7 ，构造辅助问题

$$\begin{aligned}
\text{minimize} \quad & x_6 + x_7 \\
\text{subject to} \quad & 5x_1 - 4x_2 + 13x_3 - 2x_4 + x_5 + x_6 = 20, \\
& x_1 - x_2 + 5x_3 - x_4 + x_5 + x_7 = 8, \\
& x_i \geq 0, i = 1, 2, \cdots, 7.
\end{aligned}$$

辅助问题的初始单纯形表为

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	5	-4	13	-2	1	1	0	20
x_7	1	-1	5	-1	1	0	1	8
c^T	0	0	0	0	0	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

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	5	-4	13	-2	1	1	0	20
x_7	1	-1	5	-1	$\boxed{1}$	0	1	8
r^T	-6	5	-18	3	-2	0	0	-28

(注：如选 13 为转轴元，则计算比较繁琐.)

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_6	4	-3	$\boxed{8}$	-1	0	1	-1	12
x_5	1	-1	5	-1	1	0	1	8
r^T	-4	3	-8	1	0	0	2	-12

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	x_6	x_7	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$1/8$	$-1/8$	$3/2$
x_5	$-3/2$	$7/8$	0	$-3/8$	1	$-5/8$	$13/8$	$1/2$
r^T	0	0	0	0	0	1	1	0

至此得基本可行解 $(0, 0, 3/2, 0, 1/2)$.

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	x_5	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$3/2$
x_5	$-3/2$	$7/8$	0	$-3/8$	1	$1/2$
c^T	1	6	-7	1	5	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

	x_1	x_2	x_3	x_4	x_5	b
x_3	$1/2$	$-3/8$	1	$-1/8$	0	$3/2$
x_5	$-3/2$	$\boxed{7/8}$	0	$-3/8$	1	$1/2$
r^T	12	-1	0	2	0	8

$\rightarrow$

	x_1	x_2	x_3	x_4	x_5	b
x_3	$-1/7$	0	1	$-2/7$	$3/7$	$12/7$
x_2	$-12/7$	1	0	$-3/7$	$8/7$	$4/7$
r^T	$72/7$	0	0	$11/7$	$8/7$	$60/7$

因为 $r_j \geq 0$, 所以已达最优解 $x^* = (0, 4/7, 12/7, 0, 0)$, 且 $f^* = -60/7$.

9. 第 2 章 2.18 假定给系统 $Ax = b, x \geq 0$ 应用单纯形过程，且所得表格 (忽略费用行) 形如

$$
\begin{array}{|ccc|ccc|ccc|ccc|c|}
\hline
x_1 & & x_k & x_{k+1} & \cdots & x_n & y_1 & \cdots & y_k & y_{k+1} & \cdots & y_m & \\
\hline
1 & & & & & & & & & 0 & \cdots & 0 & \bar{b}_1 \\
& \ddots & & & R_1 & & & S_1 & & \vdots & & \vdots & \vdots \\
& & 1 & & & & & & & 0 & \cdots & 0 & \bar{b}_k \\
\hline
0 & \cdots & 0 & & & & & & & 1 & & & 0 \\
\vdots & & \vdots & & R_2 & & & S_2 & & & \ddots & & \vdots \\
0 & \cdots & 0 & & & & & & & & & 1 & 0 \\
\hline
\end{array}
$$

对应地有 $m-k$ 个基变量在零水平.

(a) 证明 R_2 中的任何非零元素都可作为转轴元以消去一个人工基变量，这样将产生一个类似的表格，但 k 会增加 1.

(b) 假设 (a) 中的过程被重复，直到 $R_2 = \mathbf{0}$. 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

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

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

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

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	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
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

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	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
r^T	-4	-8	-1	-4	0	0	0	0	-25

$\rightarrow$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	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	$8/3$
r^T	$-4/3$	0	$-11/3$	$4/3$	0	0	$8/3$	0	$-19/3$

$\rightarrow$

	x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	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	$4/5$
r^T	$-3/5$	0	0	$3/5$	0	$11/5$	$6/5$	0	$-6/5$

		x_1	x_2	x_3	x_4	y_1	y_2	y_3	y_4	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
	r^T	0	0	0	0	3	1	0	0	0

由 b) 可知第 4 个约束为冗余的，可删除．至此得基本可行解 $(2,2,1,0)$.

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	b
x_1	1	0	0	−1	2
x_3	0	0	1	0	1
x_2	0	1	0	1	2
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

	x_1	x_2	x_3	x_4	b
x_1	1	0	0	−1	2
x_3	0	0	1	0	1
x_2	0	1	0	$\boxed{1}$	2
r^T	0	0	0	−3	−17

		x_1	x_2	x_3	x_4	b
	x_1	1	1	0	0	4
→	x_3	0	0	1	0	1
	x_4	0	1	0	1	2
	r^T	0	3	0	0	−11

因为 $r_j \geq 0$, 所以已达最优解 $x^* = (4,0,1,2)$, $f^* = 11$.

10. 第 2 章 2.19 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

	B^{-1}			x_5
x_5	1	0	0	3
x_6	0	1	0	12
x_7	0	0	1	9

$\lambda^T = (1,1,1), r_D^T = c_D - \lambda^T D = (-4,-10,-2.4)$.
让 x_2 进基，计算 $y_2 = B^{-1}a_2 = (2,4,4)$，得

	B^{-1}			x_B	y_2
x_5	1	0	0	3	$\boxed{2}$
x_6	0	1	0	12	4
x_7	0	0	1	9	4
λ^T	1	1	1	24	10

$\rightarrow$

	B^{-1}			x_B
x_2	1/2	0	0	3/2
x_6	-2	1	0	6
x_7	-2	0	1	3
λ^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 $y_3 = B^{-1}a_3 = (-1/2,3,4)$. 故有

	B^{-1}			x_B	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	$\boxed{4}$
λ^T	-4	1	1	9	7

$\rightarrow$

	B^{-1}			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
λ^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 $y_1 = B^{-1}a_1 = (3/8,3/4,-1/4)$. 故有

	B^{-1}			x_B	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	$-1/4$
λ^T	$-1/2$	1	$-3/4$	15/4	3/4

$\rightarrow$

	B^{-1}			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
λ^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 $x = (5,0,2,0)$.

11. 第 2 章 2.21 下面的表格是求解极小化问题的一个中间阶段：

	y_1	y_2	y_3	y_4	y_5	y_6	y_0
	1	2/3	0	0	4/3	0	4
	0	$-7/3$	3	1	$-2/3$	0	2
	0	$-2/3$	-2	0	2/3	1	2
r^T	0	8/3	-11	0	4/3	0	-8

(a) 确定下一个转轴元.

(b) 给定当前基的逆

$$B^{-1} = [a_1, a_4, a_6]^{-1} = \frac{1}{3}\begin{bmatrix} 1 & 1 & -1 \\ 1 & -2 & 2 \\ -1 & 2 & 1 \end{bmatrix}$$

和对应的费用系数

$$c_B^T = (c_1, c_4, c_6) = (-1, -3, 1),$$

求原始问题.

解：

(a) 下一个转轴元为 3(第 2 行第 3 列)

(b) $B = (a_1, a_4, a_6) = (B^{-1})^{-1} = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}$. 由 $y_0 = B^{-1}b$ 可得 $b = By_0 = \begin{bmatrix} 10 \\ 6 \\ 4 \end{bmatrix}$.

由 $y_q = B^{-1}a_q$ 可得 $(a_2, a_3, a_5) = B(y_2, y_3, y_5) = \begin{bmatrix} -1 & 3 & 2 \\ 0 & -2 & 2 \\ -3 & 1 & 0 \end{bmatrix}$ 所以

$$A = (a_1 \quad a_2 \quad a_3 \quad a_4 \quad a_5 \quad a_6) = \begin{bmatrix} 2 & -1 & 3 & 1 & 2 & 0 \\ 1 & 0 & -2 & 0 & 2 & 1 \\ 0 & -3 & 1 & 1 & 0 & 1 \end{bmatrix}.$$

由 $r_q = c_q - c_B^T y_q$ 可得 $(c_2, c_3, c_5) = (25/3, -22, 8/3)$. 所以原问题为

$$\begin{aligned} \text{minimize} \quad & -x_1 + (25/3)x_2 - 22x_3 - 3x_4 + (8/3)x_5 + x_6 \\ \text{subject to} \quad & 2x_1 - 1x_2 + 3x_3 + x_4 + 2x_5 = 10, \\ & x_1 - 2x_3 + 2x_5 + x_6 = 6, \\ & -3x_2 + x_3 + x_4 + x_6 = 4, \\ & x_i \geq 0, i = 1, 2, 3, 4, 5, 6. \end{aligned}$$

最优化理论与算法 (数值优化) ––第 3 章作业参考解答

第 4 章作业： 4.1 详细验证线性规划问题对偶的对偶是原始问题.

解：原问题和对偶问题分别如下：

$$
\begin{aligned}
\text{(P)}\quad & \text{minimize} && c^T x \\
& \text{subject to} && Ax = b, \\
& && x \ge 0.
\end{aligned}
\qquad\qquad
\begin{aligned}
\text{(D)}\quad & \text{maximize} && \lambda^T b \\
& \text{subject to} && A^T\lambda \le c.
\end{aligned}
$$

方法 I：将对偶问题 (D) 化成标准形，写出对偶，再转化成与原问题等价的线性规划．具体操作见讲义.

第 4 章作业： 4.5 构造一个原问题和对偶问题都没有可行解的例子.

解：构造原问题如下

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

其对偶问题为

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

易见二者均无可行解.

第 4 章作业： 4.9 考虑原线性规划问题

$$
\begin{aligned}
\text{minimize} \quad & c^T x \\
\text{subject to} \quad & Ax = b, \\
& x \ge 0.
\end{aligned}
$$

假设该问题和它的对偶是可行的．设 $\bar{\lambda}$ 是对偶的一个最优解.

1. 如果给原问题的第 k 个方程乘以 $\mu \neq 0$，确定这个新问题对偶的一个最优解.
2. 假设，在原问题中，我们给第 r 个方程加上第 k 个方程的 μ 倍．对应对偶问题的最优解 λ 怎么样？
3. 假设，在原问题中，我们给 c 加上 A 的第 k 行的 μ 倍．对应对偶问题的最优解 λ 怎么样？

解：该问题的对偶问题为

$$\begin{aligned} &\text{maximize} \quad b^T\lambda \\ &\text{subject to} \quad A^T\lambda \le c. \end{aligned}$$

设原问题的最优解为 $\bar{x}$ ，对偶问题的最优解为 $\bar{\lambda}$ ，则其满足如下的互补松弛条件

$$\begin{aligned} A\bar{x} - b &= 0 \\ \bar{x} &\ge 0 \\ c - A^T\bar{\lambda} &\ge 0 \\ \bar{x}^T(c - A^T\bar{\lambda}) &= 0 \end{aligned} \tag{1}$$

以下设 $A = \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}$. 则原问题的等式约束 $Ax = b$ 即为 $a^1x = b_1, \cdots, a^mx = b_m$. 对偶问题的约束 $c - A^T\lambda \ge 0$ 即 $c^T - \sum_{i=1}^m \lambda_i a^i \ge 0$.

1. 给原问题的第 k 个方程乘以 $\mu \ne 0$ 后，第 k 个约束变为 $\mu a^k x = \mu b_k$ ，易见原问题的最优解是新问题的最优解，但是对偶最优解变了．所得新问题的互补松弛条件为

$$\begin{aligned} c^T - \mu\lambda_k a^k - \textstyle\sum_{i=1,i\ne k}^m \lambda_i a^i &\ge 0 \\ a^ix = b_i, i \ne k, \quad \mu a^k x &= \mu b_k \\ x &\ge 0 \\ (c^T - \mu\lambda_k a^k - \textstyle\sum_{i=1,i\ne k}^m \lambda_i a^i)x &= 0. \end{aligned} \tag{2}$$

   因为 (1) ，故可验证原问题最优解 $\bar{x}$ 和 $\lambda = (\bar{\lambda}_1, \cdots, \frac{1}{\mu}\bar{\lambda}_k, \cdots, \bar{\lambda}_m)^T$ 满足 (2). 由互补松弛定理，λ 即为新问题的对偶问题的最优解.

2. 给第 r 个方程加上第 k 个方程的 μ 倍后，第 r 个约束变为 $(a^r + \mu a^k)x = b_r + \mu b_k$ ，易见原问题的最优解是新问题的最优解，但是对偶最优解变了．所得新问题的 KKT 条件为

$$\begin{aligned} c^T - (a^r + \mu a^k)\lambda_r - \textstyle\sum_{i=1,i\ne r}^m \lambda_i a^i &\ge 0 \\ a^ix = b_i, i \ne r, \quad (a^r + \mu a^k)x &= b_r + \mu b_k \\ x &\ge 0 \\ (c^T - (a^r + \mu a^k)\lambda_r - \textstyle\sum_{i=1,i\ne r}^m \lambda_i a^i)x &= 0 \end{aligned} \tag{3}$$

   令 $\lambda_i = \bar{\lambda}_i, i \ne k, \lambda_k = \bar{\lambda}_k - \mu\bar{\lambda}_r$ ，同上可以验证 $\bar{x}$ 和 λ 满足 (3). 所以 λ 即为新问题对偶问题的最优解.

3. 给 c 加上 A 的第 k 行的 μ 倍后，目标函数变为 $c^Tx + \mu a^k x$. 所得新问题的 KKT 条件为

$$\begin{aligned} c^T + \mu a^k - \textstyle\sum_{i=1}^m \lambda_i a^i &\ge 0 \\ a^ix = b_i, i = 1, \cdots, m& \\ x &\ge 0 \\ (c^T + \mu a^k - \textstyle\sum_{i=1}^m \lambda_i a^i)x &= 0 \end{aligned} \tag{4}$$

   令 $\lambda_i = \bar{\lambda}_i, i \ne k, \lambda_k = \bar{\lambda}_k + \mu$ ，同上可以验证 $\bar{x}$ 和 λ 满足 (4). 所以 λ 即为新问题对偶问题的最优解.

第 4 章作业： 4.12 解： (a) 先将原问题化为标准形．由提示 $x=(2,0,0,1,0)$ 为基本可行解．可得与其对应的规范形数据如下：

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	−1	3	1	−2	1
x_1	1	−1	1	0	−1	2

进而得到与此基本可行解对应的初始表格

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	−1	3	1	−2	1
x_1	1	−1	1	0	−1	2
c^T	2	−1	0	0	0	0

通过初等行变换将基变量对应的相对费用系数变为零，得第一张单纯形表，并依次进行单纯形转轴

→

	a_1	a_2	a_3	a_4	a_5	b
x_4	0	−1	$\boxed{3}$	1	−2	1
x_1	1	−1	1	0	−1	2
r^T	0	1	−2	0	2	−4

→

	a_1	a_2	a_3	a_4	a_5	b
x_3	0	−1/3	1	1/3	−2/3	1/3
x_1	1	−2/3	0	−1/3	−1/3	5/3
r^T	0	1/3	0	2/3	2/3	−10/3

至此所有 $r_j \geq 0$, 得到最优解 $x^*=(1,1,1/2,0)$.
(b) 原问题的对偶问题为

$$\begin{aligned}
\text{maximize} \quad & 3\lambda_1+2\lambda_2 \\
\text{subject to} \quad & 2\lambda_1+\lambda_2 \leq 2, \\
& -\lambda_1-\lambda_2 \leq -1, \\
& -\lambda_1+\lambda_2 \leq 0, \\
& \lambda_1 \geq 0,\ \lambda_2 \geq 0.
\end{aligned}$$

由 (a) 的最后一张表可知对偶问题的最优解为 $\lambda^*=(2/3,2/3)$.

第 4 章作业： 4.13

1. 利用单纯形法求解

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

2. 利用 (a) 中所作工作和对偶单纯形法求解相同的问题，除过方程组的右端项分别变成 8 和 7.

解. a) 第一阶段：首先引入人工变量 x_5, x_6 ，构造辅助问题

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

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	x_B
x_5	1	2	1	2	1	0	3
x_6	1	1	2	4	0	1	5
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	x_B
x_5	1	2	1	2	1	0	3
x_6	1	1	2	$\boxed{4}$	0	1	5
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	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	$1/4$	$5/4$
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	x_B
x_2	$1/3$	1	0	0	$2/3$	$-1/3$	$1/3$
x_4	$1/6$	0	$1/2$	1	$-1/6$	$1/3$	$7/6$
r^T	0	0	0	0	1	1	0

因为辅助问题的最优值为 0 ，所以原问题有可行解. 因为基变量不含人工变量，故 $x = (0, 1/3, 0, 7/6)$ 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，可得原问题的一张初始表

	x_1	x_2	x_3	x_4	x_B
x_2	$1/3$	1	0	0	$1/3$
x_4	$1/6$	0	$1/2$	1	$7/6$
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	x_B
x_2	$1/3$	1	0	0	$1/3$
x_4	$1/6$	0	$1/2$	1	$7/6$
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 $x^* = (0, 1/3, 0, 7/6)$，最优值 $f^* = 10/3$.

b) 由 a) 可知 $B = (a_2, a_4)$，所以 $B^{-1} = \begin{bmatrix} 2/3 & -1/3 \\ -1/6 & 1/3 \end{bmatrix}$. 右端项改为 8 和 7 后，只需将上表的最后一列改为 $B^{-1}\begin{bmatrix} 8 \\ 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, 即为

	x_1	x_2	x_3	x_4	x_B
x_2	$1/3$	1	0	0	3
x_4	$1/6$	0	$1/2$	1	1
r^T	$2/3$	0	1	0	-11

因为相对费用系数向量非负，故得最优解 $x^* = (0, 3, 0, 1)$，最优值 $f^* = 11$.

第 4 章作业： 4.14.

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

1. 以 1 为转轴元，利用单个转轴运算找到一个可行解.
2. 利用单纯形法求解问题.
3. 对偶问题是什么？
4. 对偶的解怎么样？

解.

1. 引入松驰变量 x_4, x_5 和盈余变量 x_6 后，得标准形问题

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

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	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	b
−6	3	0	1	0	4	0
−3	3	0	0	1	2	3
2	−1	1	0	0	−1	1

该表对应着原问题的一个基本可行解.

2. 下面利用单纯形法求解原问题. 首先得初始表为

	x_1	x_2	x_3	x_4	x_5	x_6	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
c^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	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
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	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
r^T	0	0	0	2/3	1/3	10/3	1

因为相对费用系数向量非负，故得原问题最优解 $x^* = (1, 2, 1)$, 最优值 $f^* = -1$.

3. 原问题的对偶问题为

$$
\begin{array}{lll}
 & \text{maximize} & 4w_1 + 5w_2 + w_3 \\
\text{(D)} & \text{subject to} & 2w_1 + w_2 + 2w_3 \le 5, \\
 & & -w_1 + w_2 - w_3 \le -3, \\
 & & 4w_1 + 2w_2 + w_3 \le 0, \\
 & & w_1 \le 0, w_2 \le 0, w_3 \ge 0.
\end{array}
$$

4. 易于看到原问题引入松弛变量 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 $(\lambda^*)^T = c_B^T B^{-1}$ 即为对偶问题的最优解，其中 B 为最优解对应的基. 而

$$\begin{aligned}
[\, r_4 \quad r_5 \quad r_6 \,] &= \; [\, c_4 \quad c_5 \quad c_6 \,] - c_B^T B^{-1} [\, a_4 \quad a_5 \quad a_6 \,] \\
&= \; [\, 0 \quad 0 \quad 0 \,] - [\, \lambda_1^* \quad \lambda_2^* \quad \lambda_3^* \,] \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \\
&= \; [\, -\lambda_1^* \quad -\lambda_2^* \quad \lambda_3^* \,]
\end{aligned}$$

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

最优化理论与算法 B(数值优化) －－第 4 章及第 5 章部分作业参考解答

1. 说明函数 $f(x) = 8x_1 + 12x_2 + x_1^2 - 2x_2^2$ 仅有一个驻点，且其既不是最小点也不是最大点，而是一个鞍点．画出 f 的等值线.

   解答： $\nabla f = \begin{bmatrix} 8+2x_1 \\ 12-4x_2 \end{bmatrix}$， 令 $\triangle f = 0$ 则 $x^* = \begin{bmatrix} -4 \\ 3 \end{bmatrix}$ 为唯一解，从而仅有一个驻点。又 $\nabla^2 f = \begin{bmatrix} 2 & 0 \\ 0 & -4 \end{bmatrix}$ 是不定的，所以 x^* 为鞍点，它既不是最大值点，也不是最小值点。

   图形：双曲线的中心在 $(-4, 3)$ ，渐近线不平行于坐标轴。

2. 考虑问题 $\min \|Ax - b\|^2$ ，其中 A 是 $m \times n$ 矩阵， b 是 m 维向量.

   (a) 给出问题的几何解释.

   (b) 写出最优性的必要条件．这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吗？可用规定任何你所需的假设.

   (e) 针对如下给出的 A 和 b:

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

   求解问题.

   解答： (a) 几何解释： b 到 A 的值空间 $R(A)$(即由 A 的列生成的子空间) 的最小距离。设 b 在 $R(A)$ 上的投影为 c ，则满足 $Ax = c$ 的解 x^* 即为最优解。请同学们自己举一个例子，如三个方程，两个未知数的矛盾方程。

   (b) 令 $f(x) = \| Ax - b \|^2 = x^T A^T Ax - 2b^T Ax + b^T b, \quad \nabla f(x) = 2A^T Ax - 2A^T b$

   一阶必要条件为： $\nabla f = 0, \quad$ 即 $A^T Ax = A^T b$

   又 $\nabla^2 f(x) = 2A^T A$ 为半正定矩阵，所以 $f(x)$ 为凸函数，从而也为充分条件。

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

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

   (e) 解 $A^T Ax = A^T b$ 即可。 $x^* = (3, \frac{3}{2}, \frac{-5}{2})^T$.

3. 对标量 β 的每一个值，找到下列函数的所有驻点

   $$f(x, y) = x^2 + y^2 + \beta xy + x + 2y.$$

   这些驻点中的哪些是全局极小点.

解答： $\nabla f(x,y)=\begin{bmatrix}2x+\beta y+1\\ 2y+\beta x+2\end{bmatrix}$， $\nabla^2 f(x,y)=\begin{bmatrix}2 & \beta\\ \beta & 2\end{bmatrix}$.

令 $\nabla f(x,y)=0$, 即 $\begin{cases}2x+\beta y+1=0\cdots\cdots(1)\\ 2y+\beta x+2=0\cdots\cdots(2)\end{cases}$

由 $(1), x=-\frac{1}{2}-\frac{\beta}{2}y$ 代入 (2) 得 $(2-\frac{\beta^2}{2})y=\frac{\beta}{2}-2\cdots\cdots(3)$

当 $\beta=\pm2$ 时， (3) 无解，从而无驻点；

当 $\beta\neq\pm2$ 时， $y=\frac{\beta-4}{4-\beta^2}$， $x=-\frac{1}{2}-\frac{\beta}{2}y=\frac{2\beta-2}{4-\beta^2}$， $|\nabla^2 f(x,y)|=4-\beta^2$;

当 $\beta\in(-2,2)$ 时， (因为 $2>0$, 且 $|\nabla^2 f(x,y)|>0$), 此时 f 为凸函数，所有此时 $(\frac{2\beta-2}{4-\beta^2},\frac{\beta-4}{4-\beta^2})^T$ 为全局极小点；

当 $\beta<-2$ 或者 $\beta>2$ 时， $\nabla^2 f(x,y)$ 不定且有负特征值，所有 f 可以任意小，即没有极小点。

综上， $f(x,y)$ 的驻点中只有当 $\beta\in(-2,2)$ 时， $(\frac{2\beta-2}{4-\beta^2},\frac{\beta-4}{4-\beta^2})^T$ 为全局极小点。

4. 第 4 章作业： 4.8

说明如果非零向量 $p_0,p_1,\cdots,p_l$ 关于对称且正定的矩阵 A 是共轭的, 则这些向量是线性无关的. (该结论蕴含着 A 最多有 n 个共轭方向.)

解答：设存在数 $k_0,k_1,\cdots,k_l$, 使得 $k_0p_0+k_1p_1+\cdots+k_lp_l=0$, 则

$$p_i^TA(k_0p_0+k_1p_1+\cdots+k_lp_l)=0,\quad i=0,1,\cdots,l\quad\cdots\cdots(1)$$

又 $p_0,p_1,\cdots,p_l$ 关于 A 共轭，所以 $p_i^TAp_j=0,\quad(i\neq j)$. 由 (1) 式可以得到 $k_ip_i^TAp_i=0$, 又 A 正定，所以 $p_i^TAp_i>0$, 由此得到 $k_i=0,\quad i=0,1,\cdots,l$, 从而 $p_0,p_1,\cdots,p_l$ 线性无关。

5. 第 4 章作业： 4.9

考查函数 $f(x_1,x_2)=(x_1+x_2^2)^2$. 在点 $x_k=(1,0)$ 处考虑搜索方向 $p_k=(-1,1)$ ，证明 p_k 是一个下降方向，并且求出一维极小化问题

$$\min_{\alpha>0} f(x_k+\alpha p_k)$$

的所有极小点.

解答： $\nabla f=\begin{bmatrix}2(x_1+x_2^2)\\ 4x_2(x_1+x_2^2)\end{bmatrix}$， $\nabla f(x_k)^Tp_k=\begin{bmatrix}2\\0\end{bmatrix}^T\begin{bmatrix}-1\\1\end{bmatrix}=-2<0$, 故 p_k 是一个下降方向。

$x_k+\alpha p_k=\begin{bmatrix}1\\0\end{bmatrix}+\alpha\begin{bmatrix}-1\\1\end{bmatrix}=\begin{bmatrix}1-\alpha\\ \alpha\end{bmatrix}$, 所以 $f(x_k+\alpha p_k)=(1-\alpha+\alpha^2)^2$, 又 $1-\alpha+\alpha^2=(\alpha-\frac{1}{2})^2+\frac{3}{4}\geq\frac{3}{4}$, 故当且仅当 $\alpha=\frac{1}{2}$ 时取等号。所以当 $\alpha=\frac{1}{2}$ 时， $f(x_k+\alpha p_k)$ 的值最小。

6. 第 5 章作业： 5.1 考虑将带精确线搜索的最速下降法应用到凸二次函数，即

$$f(x)=\frac{1}{2}x^TQx+b^Tx,$$

其中 Q 对称半正定，最优解为 x^* ，则利用本章给出的性质说明：如果初始点使得 x_0-x^* 与 Q 的一个特征向量平行，则最速下降法将在一步内找到解.

解答：因为 x_0-x^* 与 Q 的一个特征向量平行，所以 $Q(x_0-x^*)=\lambda(x_0-x^*)$， λ 为与 (x_0-x^*) 平行的特征向量所属的特征值。

因为 x^* 为最优解，从而 $Qf(x^*)=Qx^*+b=0$, 所以 $b=-Qx^*$, 从而 $\nabla f(x_0)=Qx_0+b=Q(x_0-x^*)=\lambda(x_0-x^*)$.

case1： $\lambda=0$, 则 $\nabla f(x_0)=0$, 而 f 为凸函数，从而 x_0 即为最优解；

case2： $\lambda\neq 0$, 则 $x_0=\frac{\nabla f_0^T\nabla f_0}{\nabla f_0^T Q\nabla f_0}=\frac{\lambda^2(x-x^*)^T(x-x^*)}{\lambda^3(x-x^*)^T(x-x^*)}=\frac{1}{\lambda}$;

所以 $x_1=x_0+x_0p_0=x_0-\frac{1}{\lambda}\nabla f_0=x_0-\frac{1}{\lambda}\lambda(x_0-x^*)=x^*$.

所以，最速下降法将在一步内找到解。

7. 第 5 章作业： 5.2

假设我们需求极小化

$$f(x_1,x_2)=5x_1^2+5x_2^2-x_1x_2-11x_1+11x_2+11.$$

(a) 找到一个满足一阶必要条件的解.

(b) 说明该点是一个全局极小点.

(c) 针对该问题最速下降法的最坏速率将会是多少？

(d) 从 $(x_1,x_2)=(0,0)$ 出发，最多需要多少步可用将目标函数值减少到 10^{-11}?

解答：

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

b) 因为 $\nabla^2 f=\begin{bmatrix}10&-1\\-1&10\end{bmatrix}$ 正定，所以 f 为凸函数，从而 $x^*=(1,-1)^T$ 也是全局极小点。

c)$\nabla^2 f$ 的特征值为 $9,11$ ，所以最坏速率 $v=(\frac{11-9}{11+9})^2=0.01$.

d)$f(x^*)=0$, $\quad f(0,0)=11$, $\quad f(x_k)-f(x^*)\leq r^k[f(x_0)-f(x^*)]\cdots\cdots(*)$, 即 $10^{-11}=10^{-2k}\cdot 11$, $\quad 2k=lg11+11$, $\quad k=\frac{lg11+11}{2}$, 令 $(*)$ 式中等号成立，则应取 $k=7$，否则可能达不到要求。所以最多需要 7 步即可达到要求。

最优化理论与算法 B(数值优化) ––第 5 章部分作业参考解答

5.1 考虑将精确步长的最速下降法应用于二次函数，即

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^{\mathrm{T}}\mathbf{G}\mathbf{x} - \mathbf{b}^{\mathrm{T}}\mathbf{x},$$

其中 $\mathbf{G}$ 对称半正定，最优解 $\mathbf{x}^* = \mathbf{G}^{-1}\mathbf{b}$ ，利用最速下降法的迭代说明：如果初始点 $\mathbf{x}^{(0)}$ 满足 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 平行于 $\mathbf{G}$ 的某个特征向量，则使用精确步长的最速下降法仅迭代一次即可找到解.

解答：因为 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 与 $\mathbf{G}$ 的一个特征向量平行，所以 $\mathbf{G}(\mathbf{x}^{(0)} - \mathbf{x}^*) = \lambda(\mathbf{x}^{(0)} - \mathbf{x}^*)$， λ 为与 $\mathbf{x}^{(0)} - \mathbf{x}^*$ 平行的特征向量所属的特征值。

因为 $\mathbf{x}^*$ 为最优解，从而 $\mathbf{g}^* = \mathbf{G}\mathbf{x}^* - \mathbf{b} = \mathbf{0}$, 所以 $\mathbf{b} = \mathbf{G}\mathbf{x}^*$, 从而 $\mathbf{g}^{(0)} = \mathbf{G}\mathbf{x}^{(0)} - \mathbf{b} = \mathbf{G}(\mathbf{x}^{(0)} - \mathbf{x}^*) = \lambda(\mathbf{x}^{(0)} - \mathbf{x}^*)$.

case1： $\lambda = 0$, 则 $\mathbf{g}^{(0)} = \mathbf{0}$, 而 f 为凸函数，从而 $\mathbf{x}^{(0)}$ 即为最优解;

case2： $\lambda \neq 0$, 则 $\alpha_0 = \frac{\mathbf{g}^{(0)\mathrm{T}}\mathbf{g}^{(0)}}{\mathbf{g}^{(0)\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}} = \frac{\lambda^2(\mathbf{x}^{(0)}-\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}^{(0)}-\mathbf{x}^*)}{\lambda^3(\mathbf{x}^{(0)}-\mathbf{x}^*)^{\mathrm{T}}(\mathbf{x}^{(0)}-\mathbf{x}^*)} = \frac{1}{\lambda}$;

所以 $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \alpha_0\mathbf{p}^{(0)} = \mathbf{x}^{(0)} - \frac{1}{\lambda}\mathbf{g}^{(0)} = \mathbf{x}^{(0)} - \frac{1}{\lambda}\lambda(\mathbf{x}^{(0)} - \mathbf{x}^*) = \mathbf{x}^*$.

所以，最速下降法将在一步内找到解。

5.2 假设我们需求极小化

$$f(x_1, x_2) = 5x_1^2 + 5x_2^2 - x_1x_2 - 11x_1 + 11x_2 + 11.$$

a) 找到一个满足一阶必要条件的解.

b) 说明该点是一个全局极小点.

c) 针对该问题最速下降法的最坏速率将会是多少？

d) 从 $(x_1, x_2) = (0, 0)$ 出发，最多需要多少步可用将目标函数值减少到 10^{-11}?

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

b) 因为 $\nabla^2 f=\begin{bmatrix}10 & -1\\-1 & 10\end{bmatrix}$ 正定，所以 f 为凸函数，从而 $x^*=(1,-1)$ 也是全局极小点。

c) $\nabla^2 f$ 的特征值为 $9,11$ ，所以最坏速率 $v=(\frac{11-9}{11+9})^2=0.01$.

d) $f(x^*)=0,\quad f(0,0)=11,$

$$f(\mathbf{x}^{(k)})-f(\mathbf{x}^*)\le r^k[f(x_0)-f(x^*)]\cdots\cdots(*)$$

即 $10^{-11}=10^{-2k}\cdot 11,\quad 2k=\ln 11+11,\quad k=\frac{\ln 11+11}{2}$, 令 $(*)$ 式中等号成立, 则应取 $k=7$, 否则可能达不到要求。所以最多需要 7 步即可达到要求。

5.7

(a) 如果 f 是连续可微的且严格凸，则

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

并利用该事实证明：对于连续可微的严格凸函数， **曲率条件**(curvature condition)

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

对任一向量 $\mathbf{x}^{(k)}\neq\mathbf{x}^{(k+1)}$ 成立.

(b) 给出一个单变量函数 f ，其满足 $f(0)=-1$ 和 $f(1)=-\frac{1}{4}$ ，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

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

将这两个不等式相加，得

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

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

(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 $x_k=0,x_{k+1}=1$ ，则有 $s_k^Ty_k=(1-0)(-\frac{1}{4}-\frac{7}{4})=-2<0$ ，所以曲率条件不成立.

5. 8 说明如果 $f(\mathbf{x})$ 是严格凸二次函数, 则函数 $h(\sigma)\overset{\text{def}}{=}f(\mathbf{x}^{(0)}+\sigma_1\mathbf{p}^{(0)}+\cdots+\sigma_{k-1}\mathbf{p}^{(k-1)})$ 也是关于 $\sigma=(\sigma_0,\cdots,\sigma_{k-1})$ 的严格凸函数.

解答：因为 $f(\mathbf{x})$ 是严格凸二次函数，所以海森矩阵为常矩阵且正定，设为 $\mathbf{G}$. 令 $\mathbf{P}=\begin{bmatrix}\mathbf{p}^{(0)} & \mathbf{p}^{(1)} & \cdots & \mathbf{p}^{(k-1)}\end{bmatrix}$, 则 $h(\sigma)=f(\mathbf{P}\sigma+\mathbf{x}^{(0)})$ ，且

$$\nabla^2 h=\mathbf{P}^{\mathrm{T}}\mathbf{G}\mathbf{P}=\begin{bmatrix}{\mathbf{p}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(0)} & \cdots & 0\\ 0 & {\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(1)} & 0\\ \cdots & \cdots & \cdots\\ 0 & \cdots & {\mathbf{p}^{(k-1)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(k-1)}\end{bmatrix},$$

又因为 $\mathbf{G}$ 正定，所以 ${\mathbf{p}^{(i)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(i)}>0,\quad i=0,1,\cdots,k-1$. 从而 $\nabla^2 h$ 也正定，故 $h(\sigma)$ 也是关于 σ 的严格凸函数。

5.9 利用 CG 的迭代形式直接证明共轭梯度法的性质对 $k=1$ 成立.

证明：由线性共轭梯度法，有 $\mathbf{g}^{(0)}=\mathbf{G}\mathbf{x}^{(0)}-\mathbf{b},\quad \mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$,

$$\alpha_0=\frac{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}}{{\mathbf{p}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{p}^{(0)}},\ \mathbf{x}^{(1)}=\mathbf{x}^{(0)}+\alpha_0\mathbf{p}^{(0)},\ \mathbf{g}^{(1)}=\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)},$$

$$\beta_0=\frac{{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}}{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}},\quad \mathbf{p}^{(1)}=-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)}.$$

1) 将 β_0 代入，得

的选取即可知 $p_1^T A p_0=0$.

2) ${\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(0)}=(\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)})^{\mathrm{T}}\mathbf{g}^{(0)}$. 将 $\mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$ 和 α_0 代入，得

${\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(0)}=(\mathbf{g}^{(0)}-\alpha_0\mathbf{G}\mathbf{g}^{(0)})^{\mathrm{T}}\mathbf{g}^{(0)}={\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}-\frac{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(0)}}{{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}}{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{G}\mathbf{g}^{(0)}=0$

3) ${\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}=(-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)})^{\mathrm{T}}\mathbf{g}^{(1)}=-{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}-\beta_0{\mathbf{g}^{(0)}}^{\mathrm{T}}\mathbf{g}^{(1)}$. 由 2) 得 ${\mathbf{p}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}=-{\mathbf{g}^{(1)}}^{\mathrm{T}}\mathbf{g}^{(1)}$.

4) 因为 $\mathbf{g}^{(1)}=\mathbf{g}^{(0)}+\alpha_0\mathbf{G}\mathbf{p}^{(0)}=\mathbf{g}^{(0)}-\alpha_0\mathbf{G}\mathbf{g}^{(0)}$, 所以 $\mathbf{g}^{(1)}\in[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$, 从而 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]\subset[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

又因为 $\mathbf{G}\mathbf{g}^{(0)}=\frac{\mathbf{g}^{(1)}-\mathbf{g}^{(0)}}{-\alpha_0}$, 所以 $\mathbf{G}\mathbf{g}^{(0)}\in[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$. 从而 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]\supset[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

综上有 $[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

5) 因为 $\mathbf{p}^{(0)}=-\mathbf{g}^{(0)}$, $\mathbf{p}^{(1)}=-\mathbf{g}^{(1)}+\beta_0\mathbf{p}^{(0)}$ ，所以 $\mathbf{p}^{(1)}=-\mathbf{g}^{(1)}-\beta_0\mathbf{g}^{(0)}$. 从而 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]\supset[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$.

此外， $\mathbf{g}^{(0)}=-\mathbf{p}^{(0)},\mathbf{g}^{(1)}=\beta_0\mathbf{p}^{(0)}-\mathbf{p}^{(1)}$ ，所以 $[\mathbf{g}^{(1)},\mathbf{g}^{(0)}]\supset[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]$.

综上，有 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{g}^{(1)}]$. 再由 4) ，有 $[\mathbf{p}^{(0)},\mathbf{p}^{(1)}]=[\mathbf{g}^{(0)},\mathbf{G}\mathbf{g}^{(0)}]$.

最优化理论与算法 B(数值优化) －－第 6 章部分作业参考解答

6.1 设 $f(x) = 10(x_2 - x_1^2)^2 + (1 - x_1)^2$. 考虑信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 的信赖域子问题，并假定信赖二次模型中 $\mathbf{B}^{(k)} = \nabla^2 f(0, -1)$.

(a) 完整写出信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 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 Cauchy 点.

在点 $\mathbf{x}^{(k)} = (0, 0.5)$ 重复上面的工作.

解：$\mathbf{g}(\mathbf{x}) = \begin{bmatrix} 40x_1^3 - 40x_1x_2 + 2x_1 - 2 \\ -20x_1^2 + 20x_2 \end{bmatrix}$, $\mathbf{G}(\mathbf{x}) = \begin{bmatrix} 120x_1^2 - 40x_2 + 2 & -40x_1 \\ -40x_1 & 20 \end{bmatrix}$,
考虑信赖域法在点 $\mathbf{x}^{(k)} = (0, -1)$ 的信赖域子问题：

(a) 首先

$$f(\mathbf{x}^{(k)}) = 11, \quad \mathbf{g}^{(k)} = \begin{bmatrix} -2 \\ -20 \end{bmatrix}, \quad \mathbf{G}^{(k)} = \begin{bmatrix} 42 & 0 \\ 0 & 20 \end{bmatrix}.$$

所以原问题在 $\mathbf{x}^{(k)}$ 的局部模型，也即信赖域子问题为

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

(b) 因为点 $\mathbf{x}^{(k)} = (0, -1)$ 处的海森矩阵正定，所以信赖域子问题的目标函数的等值线是一族同心椭圆, 先求 $m_k(\mathbf{s})$ 的驻点, 即得中心 $\mathbf{s}^* = (\frac{1}{21}, 1)$(以 s_1 为横坐标，以 s_2 为纵坐标). 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，解为 $\mathbf{s}^{(k)} = (0, 0)$;

II) 当 $0 < \Delta < \|\mathbf{s}^*\|_2 = \sqrt{1 + \frac{1}{21^2}}$ 时，解 $\mathbf{s}^{(k)}$ 是以 $(0, 0)$ 为中心，以 Δ 为半径的圆与 $m_k(\mathbf{s})$ 等值线的切点；

III) 当 $\Delta \geq \|\mathbf{s}^*\|_2$ 时，信赖域子问题的解即为 $\mathbf{s}^*$.

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 $(0,0)$ 与 $\mathbf{s}^*$ 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.

(c) 该点处的最速下降方向 $-\mathbf{g}^{(k)} = (2,20)$，信赖域子问题的 Cauchy 点即 $m_k(\mathbf{s})$ 沿最速下降方向，且限制在信赖域内的极小点，可用图解法求，也可用解析法求。解析法即求解问题 $\text{minimize}_{\|-\alpha\mathbf{g}^{(k)}\|\leq 1} m_k(-\alpha\mathbf{g}^{(k)})$. 将 $\mathbf{g}^{(k)}$ 代入，即

$$\text{minimize}_{0\leq\alpha\leq 1/\sqrt{2+20^2}} 4084\alpha^2 - 404\alpha + 11$$

求得目标函数的无约束极小点是 $\alpha^* = \frac{404}{2\times 4084} = 0.0495$，其显然在约束区间 $[0, 1/\sqrt{2+20^2}] = [0, 0.0499]$. 故该信赖域子问题的 Cauchy 点 $\mathbf{s}_{\mathrm{C}}^{(k)} = -\alpha^*\mathbf{g}^{(k)} = \frac{404}{2\times 4084}(2,20) = \frac{101}{1021}(1,10)$.

最优化理论与算法 B： 第 8 章部分作业参考解答

8.3 考虑下面的问题

$$\begin{array}{ll} \text{minimum}_{\mathbf{x}\in\mathbb{R}^2} & (x_1-\frac{9}{4})^2+(x_2-2)^2 \\ \text{subject to} & -x_1^2+x_2\geq 0 \\ & x_1+x_2\leq 6 \\ & x_1\geq 0, x_2\geq 0. \end{array}$$

(a) 写出 KKT 最优性条件，并验证点 $\mathbf{x}^*=(\frac{3}{2},\frac{9}{4})$ 满足这些条件.

(b) 给出 $\mathbf{x}^*$ 处 KKT 条件的一个几何解释.

(c) 证明 $\mathbf{x}^*$ 是该问题的最优解.

解：

(a) $\mathcal{L}(\mathbf{x},\lambda)=(x_1-\frac{9}{4})^2+(x_2-2)^2-\lambda_1(-x_1^2+x_2)-\lambda_2(6-x_1-x_2)-\lambda_3x_1-\lambda_4x_2.$

$$\nabla_x\mathcal{L}(\mathbf{x},\lambda)=\begin{bmatrix} 2(x_1-\frac{9}{4})+2\lambda_1x_1+\lambda_2-\lambda_3 \\ 2(x_2-2)-\lambda_1+\lambda_2-\lambda_4\end{bmatrix}.$$

KKT 条件：

$$\begin{array}{rl} 2(1+\lambda_1)x_1+\lambda_2-\lambda_3-\frac{9}{2} & =0 \\ 2x_2-\lambda_1+\lambda_2-4 & =0 \\ -x_1^2+x_2 & \geq 0 \\ 6-x_1-x_2 & \geq 0 \\ x_1\geq 0, x_2 & \geq 0 \\ \lambda_1(-x_1^2+x_2) & =0 \\ \lambda_2(6-x_1-x_2) & =0 \\ \lambda_3x_1=0, \lambda_4x_2 & =0 \\ \lambda_i\geq 0, i=1,2, & 3,4. \end{array}$$

对 $\mathbf{x}^*=(\frac{3}{2},\frac{9}{4})$，有 $\mathcal{I}(\mathbf{x}^*)=\{1\}$. 代入上边的条件，得 $\lambda_1=\frac{1}{2},\lambda_2=\lambda_3=\lambda_4=0$ 符合条件. 从而 $\mathbf{x}^*$ 满足 KKT 条件.

(b) 在 $\mathbf{x}^*$ 处，积极约束为 $c_1(\mathbf{x})=-x_1^2+x_2$，在 $\mathbf{x}^*$ 处的梯度为 $\nabla c_1(\mathbf{x}^*)=(-3,1)$，而 $\nabla f(\mathbf{x}^*)=(-3/2,1/2)$，易见 $\nabla f(\mathbf{x}^*)=\frac{1}{2}\nabla c_1(\mathbf{x}^*)$. 所以 $\mathbf{x}^*$ 处目标函数与积极约束的梯度共线，且同方向.

(c) 目标函数显然为凸函数. $c_1(\mathbf{x})=-x_1^2+x_2$ 为凹函数，易见其余为线性约束，从而也是凸函数. 所以此问题为凸规划. 从而 KKT 点即是全局最优解. 所以 $\mathbf{x}^*$ 是该问题的最优解.

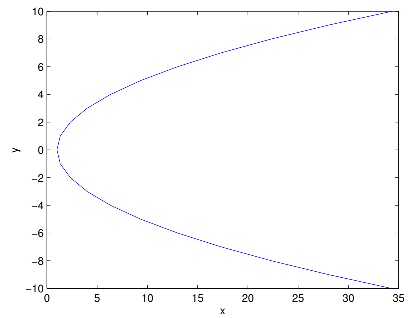
8.6 考虑找抛物线上哪个点离原点最近 (在 Euclidean 范数意义下) 的问题. 我们可以将该问题表述为

$$\text{minimize } f(x,y)=x^2+y^2 \quad \text{subject to} \quad (x-1)^3=y^2.$$

(a) 图解法求解该问题. 消去 y 求解问题，所得到的函数有极小点吗？给出结果不一致的可能原因。消去 x 求解问题会得到怎样的结果？

(b) 对于该问题，找到所有的 KKT 点. LICQ 成立吗？这些点中的那些是解？

解: (a) 图解法：该问题的可行集如下图 所示， 目标函数的等值线是以圆



点为中心的同心圆，故问题的最优解 $x^*=1, y^*=0$.

用 $y^2=(x-1)^3$ 代入，得到 $\min x^2+(x-1)^3$ ，此时问题无界；这种消元法在将 $y^2=(x-1)^3$ 代入时忽略了 $x\geq 1$ 这个暗含的条件.

用 $x=y^{2/3}+1$ 代入，得到 $\min(y^{2/3}+1)^2+y^2$, 得到最优解 $y^*=0$ ，于是原问题的最优解 $x^*=1, y^*=0$.

(b) 该问题的 Lagrange 函数为 $\mathcal{L}(x,y;\lambda)=x^2+y^2-\lambda((x-1)^3-y^2)$.

KKT 条件为：

$$\begin{aligned} 2x-3\lambda(x-1)^2&=0\\ 2y+2\lambda y&=0\\ (x-1)^3-y^2&=0. \end{aligned}$$

由第二个方程入手，分情况讨论该方程组的解。

$y=0$, 则 $x=1$, 此时第一个方程 $2x-3\lambda(x-1)^2=0$ 不满足;

$\lambda=-1$, 此时 $2x+3(x-1)^2=0$ 无解。

所以该问题没有 KKT 点！

在该问题的最优解 $x^*=1, y^*=0$ 处, 约束梯度 $\mathbf{a}=(0,0)$, 不满足 LICQ , 故不能保证 $x^*=1, y^*=0$ 一定满足 KKT 条件。

注记：该问题的 (a) 说明消元时需要谨慎，否则可能会忽略掉一些隐含条件；从而得出错误的结论； (b) 说明约束规范的重要性，即对于一个优化问

题，当最优解处约束规范不成立时，即使是最优解，也可能不满足 KKT 条件。

8.7 考虑问题

$$\min_{\mathbf{x}\in\mathbb{R}^2} f(\mathbf{x}) = -2x_1 + x_2 \quad \text{subject to} \quad \begin{cases} (1-x_1)^3 - x_2 & \geq & 0. \\ x_2 + 0.25x_1^2 - 1 & \geq & 0. \end{cases}$$

最优解是 $\mathbf{x}^* = (0,1)$, 其中两个约束都是积极的. LICQ 在该点成立吗？KKT 条件满足吗？

解. $\nabla f(\mathbf{x}) = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$, $\nabla c_1(\mathbf{x}) = \begin{bmatrix} -3(1-x_1)^2 \\ -1 \end{bmatrix}$, $\nabla c_2(\mathbf{x}) = \begin{bmatrix} 0.5x_1 \\ 1 \end{bmatrix}$.
在 $x^* = (0,1)$ 处，有

$$\nabla c_1(\mathbf{x}^*) = \begin{bmatrix} -3 \\ -1 \end{bmatrix}, \quad \nabla c_2(\mathbf{x}^*) = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

易见 $\nabla c_1(\mathbf{x}^*)$ 和 $\nabla c_2(\mathbf{x}^*)$ 线性无关，所以 LICQ 成立.

此外，对于 $\mathbf{x}^* = (0,1)$ ，有 $\mathcal{I}(\mathbf{x}^*) = \{1,2\}$ ，易见 $\nabla f(\mathbf{x}^*) = \frac{2}{3}\nabla c_1(\mathbf{x}^*) + \frac{5}{3}\nabla c_2(\mathbf{x}^*)$. 所以 $\mathbf{x}^*$ 满足 KKT 条件.

8.9 利用一阶和二阶最优性条件找到函数 $f(\mathbf{x}) = x_1x_2$ 在单位圆 $x_1^2 + x_2^2 = 1$ 上的极小值点. 从几何上说明该问题.

解. 该问题的 Lagrange 函数为

$$\mathcal{L}(\mathbf{x},\lambda) = x_1x_2 - \lambda(x_1^2 + x_2^2 - 1),$$

KKT 条件为：

$$\begin{aligned} x_2 - 2\lambda x_1 &= 0 \\ x_1 - 2\lambda x_2 &= 0 \\ x_1^2 + x_2^2 &= 1. \end{aligned}$$

解上述系统，得 KKT 点：

$$\begin{array}{ll} \mathbf{x}^{(1)} = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), & \lambda^{(1)} = \frac{1}{2}, \\ \mathbf{x}^{(2)} = (-\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}), & \lambda^{(2)} = \frac{1}{2}, \\ \mathbf{x}^{(3)} = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}), & \lambda^{(3)} = -\frac{1}{2}, \\ \mathbf{x}^{(4)} = (-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}), & \lambda^{(4)} = -\frac{1}{2}, \end{array}$$

对 $i = 1$ ，计算易得：

$$\nabla_{xx}^2\mathcal{L}(\mathbf{x}^{(1)}, \lambda^{(1)}) = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix},$$

$$G(\mathbf{x}^{(1)}, \lambda^{(1)}) = \{\mathbf{p} \in \mathbb{R}^2 \mid \mathbf{p}^{\mathrm{T}}\nabla c(\mathbf{x}^{(1)}) = 0\} = \{k(1,-1) \mid k \in \mathbb{R}\}.$$

因为对 $k \neq 0$ ，有 $k^2(1,-1)^{\mathrm{T}}\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(1)},\lambda^{(1)})(1,-1) = -4k^2 < 0$ ，所以，在 $\mathbf{x}^{(1)}$ 处不满足二阶必要条件．所以 $\mathbf{x}^{(1)}$ 不是局部解．同理可验证 $\mathbf{x}^{(2)}$ 不是局部解.

对 $i = 3$ ，计算易得：

$$\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(3)},\lambda^{(3)}) = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

$$G(\mathbf{x}^{(3)},\lambda^{(3)}) = \{\mathbf{p} \in \mathbb{R}^2 \mid \mathbf{p}^{\mathrm{T}}\nabla c(\mathbf{x}^{(3)}) = 0\} = \{k(1,1) \mid k \in \mathbb{R}\}.$$

因为对 $k \neq 0$ ，有 $k^2(1,1)^T\nabla^2_{xx}\mathcal{L}(\mathbf{x}^{(3)},\lambda^{(3)})(1,1) = 4k^2 > 0$ ，所以，在 $\mathbf{x}^{(3)}$ 处满足二阶充分条件．所以 $\mathbf{x}^{(3)}$ 是局部解．同理可验证 $\mathbf{x}^{(4)}$ 是局部解.

目标函数的等值线是双曲线，以横轴和纵轴为渐近线．且在第 I 和第 III 卦限目标函数的值趋于正无穷；在第 II 和第 IV 卦限目标函数的值趋于负无穷.

8.10 设 K 是 $\mathbb{R}^n$ 中的凸集，设 $f(\cdot): K \to \mathbb{R}$. 证明 f 是 K 上的凸函数当且仅当对任一整数 $k \geq 2$ ，下面结论成立：

$$\lambda_1, \cdots, \lambda_k \geq 0, \sum_{j=1}^{k}\lambda_j = 1, \mathbf{x}_1, \cdots, \mathbf{x}_k \in K \Rightarrow f\left(\sum_{j=1}^{k}\lambda_j\mathbf{x}_j\right) \leq \sum_{j=1}^{k}\lambda_j f(\mathbf{x}_j)$$

证明．“必要性”：用数学归纳法证．若 f 是 S 上的凸函数，结论对 $k = 2$ 显然成立．假设 $k = n$ 时结论成立．以下我们证明对于 $k = n+1$ 时结论仍然成立.

对任意的 $\mathbf{x}_1, \ldots, \mathbf{x}_{n+1} \in S$, $\lambda_1, \ldots, \lambda_{n+1} \geq 0$ 且 $\sum_{j=1}^{n+1}\lambda_j = 1$. 若 $\lambda_{n+1} = 1$, 此时结论显然成立．若 $0 \leq \lambda_{n+1} < 1$, 　则

$$\begin{aligned}
f(\sum_{j=1}^{n+1}\lambda_j\mathbf{x}_j) &= f\left((1-\lambda_{n+1})\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}\mathbf{x}_j + \lambda_{n+1}\mathbf{x}_{n+1}\right) \\
&\leq (1-\lambda_{n+1})f\left(\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}\mathbf{x}_j\right) + \lambda_{n+1}f(\mathbf{x}_{n+1})(\text{ 凸函数定义}) \\
&\leq (1-\lambda_{n+1})\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}}f(\mathbf{x}_j) + +\lambda_{n+1}f(\mathbf{x}_{n+1}) \\
&= \sum_{j=1}^{n+1}\lambda_j f(\mathbf{x}_j)
\end{aligned}$$

其中第二个不等式是因为 $\frac{\lambda_j}{1-\lambda_{n+1}} \geq 0$ ，且 $\sum_{j=1}^{n}\frac{\lambda_j}{1-\lambda_{n+1}} = 1$ ，从而由归纳假设 $k = n$ 时结论成立得到的．由归纳法，必要性得证.

"充分性"：取 $k=2$, 即为凸函数的定义.

8.14 写出问题

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\sigma x_1^2 + \frac{1}{2}x_2^2 + x_1 \\ \text{subject to} & x_1 \geq 0 \end{array}$$

的 Lagrange 对偶问题。分 $\sigma = 1$ 和 $\sigma = -1$ 两种情况讨论对偶问题的解是否是原问题最优解处的 Lagrange 乘子，并解释两种情况下的结果。

解：问题的 Lagrange 对偶为

$$\max_{\lambda \geq 0} \min_{\mathbf{x} \in \mathbb{R}^2} [\frac{1}{2}\sigma x_1^2 + \frac{1}{2}x_2^2 + x_1 - \lambda x_1].$$

若 $\sigma = 1$，则 $\min_{\mathbf{x} \in \mathbb{R}^2} [\frac{1}{2}\sigma x_1^2 + \frac{1}{2}x_2^2 + x_1 - \lambda x_1]$ 的最优解为 $\mathbf{x}^*(\lambda) = (\lambda - 1, 0)$。所以对偶问题即 $\max_{\lambda \geq 0} -\frac{1}{2}(\lambda - 1)^2$，其解为 $\lambda^* = 1$.

该问题的 KKT 条件即

$$\begin{bmatrix} x_1 + 1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \lambda \\ 0 \end{bmatrix} = 0,$$

$$x_1 \geq 0,$$

$$x_1 \lambda = 0.$$

解之，得到得到问题的最优解 $\mathbf{x}^* = (0, 0)$， Lagrange 乘子为 1. 可见对偶问题的解是原问题最优解处得 Lagrange 乘子。

同样的步骤可知 $\sigma = -1$ 时对于任何 λ 值，对偶问题目标值都是 $-\infty$. 没有最优解。原问题无界。

两种情况下结果的解释：当 $\sigma = 1$ 时，该问题是凸规划，且满足 Slater 约束规范，故强对偶定理成立。这时，对偶问题的最优解即原问题的 Lagrange 乘子，且在 Lagrange 函数中固定对偶最优解 λ^*，所得函数的无约束极小点是原问题的最优解，即 $\mathbf{x}^* = \mathbf{x}^*(\lambda^*) = (0, 0)$.

8.15 针对 (方) 阵 $\mathbf{M} \in \mathbb{R}^{n \times n}$，定义迹 $\text{trace}(\mathbf{M}) = \sum_{j=1}^n m_{jj}$，并对两个矩阵 $\mathbf{A}, \mathbf{B} \in \mathbb{R}^{k \times l}$ 定义

$$\mathbf{A} \bullet \mathbf{B} := \sum_{i=1}^k \sum_{j=1}^l a_{ij} b_{ij},$$

其中 a_{ij}, b_{ij} 分别为 $\mathbf{A}$ 和 $\mathbf{B}$ 的第 (i, j) 个元素．证明

(a) $\mathbf{A} \bullet \mathbf{B} = \text{trace}(\mathbf{A}^{\mathrm{T}}\mathbf{B})$.

(b) $\text{trace}(\mathbf{MN}) = \text{trace}(\mathbf{NM})$.

证明： (a)

$$
\begin{aligned}
\mathbf{A}^{\mathrm{T}}\mathbf{B} &= \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{k1} \\ a_{12} & a_{22} & \cdots & a_{k2} \\ \cdots & \cdots & \cdots & \\ a_{1l} & a_{2l} & \cdots & a_{kl} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1l} \\ b_{21} & b_{22} & \cdots & a_{2l} \\ \cdots & \cdots & \cdots & \\ a_{k1} & a_{k2} & \cdots & a_{kl} \end{bmatrix} \\
&= \begin{bmatrix} a_{11}b_{11} + a_{21}b_{21} + \cdots + a_{k1}b_{k1} & \cdots & \cdots & \cdots \\ \cdots & a_{12}b_{12} + \cdots + a_{k2}b_{k2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & a_{1l}b_{1l} + \cdots + a_{kl}b_{kl} \end{bmatrix}.
\end{aligned}
$$

所以

$$
\begin{aligned}
\operatorname{trace}(\mathbf{A}^{\mathrm{T}}\mathbf{B}) &= a_{11}b_{11} + \cdots + a_{k1}b_{k1} + a_{12}b_{12} + \cdots + a_{k2}b_{k2} + \cdots + a_{1l}b_{1l} + \cdots + a_{kl}b_{kl} \\
&= \Sigma_{i=1}^{k}\Sigma_{j=1}^{l} a_{ij}b_{ij} \\
&= \mathbf{A} \bullet \mathbf{B}.
\end{aligned}
$$

(b) $\operatorname{trace}(\mathbf{MN}) = \Sigma_{i=1}^{n}\Sigma_{j=1}^{n} m_{ij}n_{ji} = \Sigma_{i=1}^{n}\Sigma_{j=1}^{n} n_{ij}m_{ji} = \operatorname{trace}(\mathbf{NM})$.

8.17 考虑矩阵

$$
\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^{\mathrm{T}} & \mathbf{C} \end{bmatrix},
$$

其中 $\mathbf{A}, \mathbf{C}$ 是对称矩阵, 且 $\mathbf{A}$ 是正定的. 证明 $\mathbf{M} \succeq \mathbf{0}$ 当且仅当 $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \succeq \mathbf{0}$.

解：设 $\mathbf{M} \in \mathbb{R}^{n\times n}, \mathbf{C} \in \mathbb{R}^{m\times m}$ ，考虑非奇异矩阵 $\mathbf{T} = \begin{bmatrix} \mathbf{I} & -\mathbf{A}^{-1}\mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$, 则 $\mathbf{T}^{\mathrm{T}}\mathbf{M}\mathbf{T} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix}$. 即对 $\mathbf{M}$ 作初等行变换和相应的初等列变换，即其化成对角块矩阵.

“$\Rightarrow$” 设 $\mathbf{M} \succeq \mathbf{0}$. 则 $\forall \mathbf{x} \in \mathbb{R}^{m}$, 令 $\mathbf{y} = \begin{bmatrix} \mathbf{0} \\ \mathbf{x} \end{bmatrix}$ ，则

$$
\begin{aligned}
& \mathbf{x}^{\mathrm{T}}(\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B})\mathbf{x} \\
= \;& [\mathbf{0}\ \mathbf{x}^{\mathrm{T}}] \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{x} \end{bmatrix} \\
= \;& \mathbf{y}^{\mathrm{T}}\mathbf{T}^{\mathrm{T}}\mathbf{M}\mathbf{T}\mathbf{y} \\
= \;& (\mathbf{T}\mathbf{y})^{\mathrm{T}}\mathbf{M}(\mathbf{T}\mathbf{y}) \geq 0,
\end{aligned}
$$

最后一个不等式是因为 $\mathbf{M} \succeq \mathbf{0}$ 。故 $\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \succeq \mathbf{0}$.

“$\Leftarrow$” 设 $\mathbf{C} - \mathbf{B}^{T}\mathbf{A}^{-1}\mathbf{B} \succeq \mathbf{0}$, $\mathbf{A} \succeq \mathbf{0}$. 则 $\forall \mathbf{y} \in \mathbb{R}^{n\times n}$, 记 $\mathbf{T}^{-1}\mathbf{y} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}$, 其中 $\mathbf{x}_1 \in \mathbb{R}^{n-m}, \mathbf{x}_2 \in \mathbb{R}^{m}$. 于是

$$
\begin{aligned}
\mathbf{y}^{\mathrm{T}}\mathbf{M}\mathbf{y} = \;& \mathbf{y}^{\mathrm{T}}\mathbf{T}^{-\mathrm{T}} \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \mathbf{T}^{-1}\mathbf{y} \\
& [\mathbf{x}_1^{\mathrm{T}}\ \ \mathbf{x}_2] \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \\
& = \mathbf{x}_1^{\mathrm{T}}\mathbf{A}\mathbf{x}_1 + \mathbf{x}_2^{\mathrm{T}}\mathbf{C} - \mathbf{B}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{B}\mathbf{x}_2
\end{aligned}
$$

最后一个不等式是因为 $\mathbf{C} - \mathbf{B}^{T}\mathbf{A}^{-1}\mathbf{B} \succeq \mathbf{0}$, $\mathbf{A} \succeq \mathbf{0}$ 正定。故 $\mathbf{M} \succeq \mathbf{0}$.

最优化理论与算法 B： 第 9 章部分作业参考解答

9.1 下面的问题有解吗？给出解释.

$$
\begin{array}{llll}
(a) & \text{minimize } x_1+x_2 & \text{subject to} & x_1^2+x_2^2=2, 0\le x_1\le 1, 0\le x_2\le 1;\\
(b) & \text{minimize } x_1+x_2 & \text{subject to} & x_1^2+x_2^2\le 1, x_1+x_2=3;\\
(c) & \text{minimize } x_1x_2 & \text{subject to} & x_1^2+x_2^2=2.
\end{array}
$$

解： (a) 只有一个可行点 $(1,1)$, 从而 (a) 有解，且为 $(1,1)$.

(b) 的可行集为 ϕ, 从而 (b) 无解.

在 (c) 中，令 $x_1=\sqrt{2}\cos\theta,\quad x_2=\sqrt{2}\sin\theta,\quad \theta\in[0,2\pi]$, 则

$$x_1x_2=2\cos\theta\sin\theta=\sin(2\theta)$$

显然当 $\theta=\frac{1}{4}\pi$ 或 $\frac{7}{4}\pi$ 时， x_1x_2 最小，且为 -1. 所以 (c) 有解 $x^*=(1,-1)$ 或 $(-1,1)$.

9.2 考虑点 $\mathbf{x}^{(0)}$ 到多面集 $\{\mathbf{x}\mid\mathbf{A}^{\mathrm{T}}\mathbf{x}=\mathbf{b}\}$(其中 $\mathbf{A}$ 列满秩) 的最短距离问题，该问题可以表述为二次规划

$$
\begin{array}{ll}
\text{minimize} & \frac{1}{2}(\mathbf{x}-\mathbf{x}^{(0)})^{\mathrm{T}}(\mathbf{x}-\mathbf{x}^{(0)}),\\
\text{subject to} & \mathbf{A}^{\mathrm{T}}\mathbf{x}=\mathbf{b}.
\end{array}
$$

证明最优乘子

$$\boldsymbol{\lambda}^*=(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{b}-\mathbf{A}^{\mathrm{T}}\mathbf{x}^{(0)}),$$

最优解

$$\mathbf{x}^*=\mathbf{x}^{(0)}+\mathbf{A}(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}(\mathbf{b}-\mathbf{A}^{\mathrm{T}}\mathbf{x}^{(0)}).$$

说明当 $\mathbf{A}=\mathbf{a}$ 是一个列向量的特殊情况下，从 $\mathbf{x}^{(0)}$ 到超平面 $\mathbf{a}^{\mathrm{T}}\mathbf{x}=b$ 的最短距离是

$$\frac{|b-\mathbf{a}^{\mathrm{T}}\mathbf{x}^{(0)}|}{\|\mathbf{a}\|}.$$

证明：

$$\mathcal{L}(x,\lambda)=\frac{1}{2}x^Tx-x_0^Tx+\frac{1}{2}x_0^Tx_0-\lambda^T(Ax-b)$$

$$\nabla_x\mathcal{L}(x,\lambda)=x-x_0-A^T\lambda$$

KKT 条件是：

$$\begin{aligned} x - x_0 - A^T\lambda &= 0, \\ Ax - b &= 0. \end{aligned}$$

因为 A 行满秩，故该条件的第一个方程等价于 $Ax - Ax_0 - AA^T\lambda = 0$. 将 $Ax = b$ 代入, 得 $AA^T\lambda = b - Ax_0$. 再由 A 行满秩, 有 AA^T 可逆. 从而最优乘子 $\lambda^* = (AA^T)^{-1}(b - Ax_0)$. 最优解 $x^* = x_0 + A^T\lambda^* = x_0 + A^T(AA^T)^{-1}(b - Ax_0)$.

特别地，当 A 为行向量时，$\{x \mid Ax = b\}$ 表示超平面. 此时 $AA^T = \|A\|_2^2$. 所以 $x^* = x_0 + \frac{b - Ax_0}{\|A\|^2}A^T$. x_0 到超平面 $\{x \mid Ax = b\}$ 的最短距离为：

$$\|x^* - x_0\|_2 = \sqrt{(x^* - x_0)^T(x^* - x_0)} = \sqrt{\frac{(b - Ax_0)^2}{\|A\|_2^4}AA^T} = \frac{|b - Ax_0|}{\|A\|}.$$

9.3 考虑例 9.2

$$\begin{array}{ll} \text{minimize}_{\mathbf{x}\in\mathbb{R}^2} & q(\mathbf{x}) = (x_1 - 1)^2 + (x_2 - 2.5)^2, \\ \text{subject to} & x_1 - 2x_2 + 2 \geq 0, \\ & -x_1 - 2x_2 + 6 \geq 0, \\ & -x_1 + 2x_2 + 2 \geq 0, \\ & x_1 \geq 0, \\ & x_2 \geq 0. \end{array}$$

以 $\mathbf{x}^{(0)} = (2, 0), \mathcal{W}_0\{3, 5\}$ 的迭代结果如下：

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{3, 5\}$	$(0, 0)$			$(-2, -1)$	3
1	$(2, 0)$	$\{5\}$	$(-1, 0)$	1			
2	$(1, 0)$	$\{5\}$	$(0, 0)$			-5	5
3	$(1, 0)$	$\emptyset$	$(0, 2.5)$	0.6	1		
4	$(1, 1.5)$	$\{1\}$	$(0.4, 0.2)$	1			
5	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

保持初始点不变，工作集依次为 $\mathcal{W}_0 = \{3\}$， $\mathcal{W}_0 = \{5\}$ 和 $\mathcal{W}_0 = \emptyset$，完成积极集法的前两次迭代.

$\mathcal{W}_0 = \{5\}$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{5\}$	$(-1, 0)$	1			
1	$(1, 0)$	$\{5\}$	$(0, 0)$			-5	5
2	$(1, 0)$	$\emptyset$	$(0, 2.5)$	0.6	1		
3	$(1, 1.5)$	$\{1\}$	$(0.4, 0.2)$	1			
4	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

$\mathcal{W}_0 = \{3\}$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\{3\}$	$(0.2, 0.1)$	1			
1	$(2.2, 0.1)$	$\{3\}$	$(0, 0)$			-2.4	3
2	$(2.2, 0.1)$	$\emptyset$	$(-1.2, 2.4)$	2/3	1		
3	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

$\mathcal{W}_0 = \emptyset$

k	$\mathbf{x}^{(k)}$	$\mathcal{W}_k$	$\mathbf{s}^{(k)}$	α_k	p	$\boldsymbol{\lambda}^{(k)}$	q
0	$(2, 0)$	$\emptyset$	$(-1, 2.5)$	$2/3$	1		
1	$(4/3, 5/3)$	$\{1\}$	$(0.0667, 0.333)$	1			
3	$(1.4, 1.7)$	$\{1\}$	$(0, 0)$			0.8	

9.14 用基本逐步二次规划法求解问题

$$\text{minimize } x_1 + x_2 \text{ subject to } x_2 \geq x_1^2,$$

(a) 取 $\mathbf{x}^{(0)} = \mathbf{0}, \lambda^{(0)} = 0$; 得到怎样的结果？解释方法失败的原因；

(b) 取 $\mathbf{x}^{(0)} = \mathbf{0}, \lambda^{(0)} = 1$; 得到怎样的结果？收敛快吗？

解：该问题的 Lagrange 函数

$$\mathcal{L}(\mathbf{x}, \lambda) = x_1 + x_2 - \lambda(x_2 - x_1^2).$$

对于 $\mathbf{x}^{(k)}$ ，有 $c(\mathbf{x}^{(k)}) = x_2^{(k)} - (x_1^{(k)})^2$,

$$\mathbf{a}^{(k)} = \begin{bmatrix} -2x_1^{(k)} \\ 1 \end{bmatrix}, \quad \mathbf{g}^{(k)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \mathbf{W}^{(k)} = \nabla^2 \mathcal{L}(\mathbf{x}^{(k)}, \lambda^{(k)}) = \begin{bmatrix} 2\lambda^{(k)} & 0 \\ 0 & 0 \end{bmatrix}.$$

在 SQP 中，二次规划子问题为

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{s}^{\mathrm{T}}\mathbf{W}^{(k)}\mathbf{s} + \mathbf{g}^{(k)^{\mathrm{T}}}\mathbf{s} \\ \text{subject to} & c(\mathbf{x}^{(k)}) + \mathbf{a}^{(k)^{\mathrm{T}}}\mathbf{s} \geq 0, \end{array}$$

当 $k = 0, \mathbf{x}^{(k)} = (0, 0)$ 时，

$$c(\mathbf{x}^{(0)}) = 0, \quad \mathbf{a}^{(0)} = (0, 1), \quad \mathbf{g}^{(0)} = (1, 1), \quad \mathbf{W}^{(0)} = \begin{bmatrix} 2\lambda^{(0)} & 0 \\ 0 & 0 \end{bmatrix}.$$

将上述向量代入后，得

$$\begin{array}{ll} \text{minimize} & \lambda_0 s_1^2 + s_1 + s_2 \\ \text{subject to} & s_2 \geq 0. \end{array}$$

(a) 若取 $\lambda^{(0)} = 0$, 则二次规划子问题显然无界；故初值选取不当。

(b) 若取 $\lambda^{(0)} = 1$, 解上述二次规划子问题得解 $\mathbf{s}^{(0)} = (-0.5, 0)$, 对应的 Lagrange 乘子 $\lambda^{(1)} = 1$. 于是 $\mathbf{x}^{(1)} = \mathbf{x}^{(0)} + \mathbf{s}^{(0)} = (-0.5, 0)$.

第二次迭代：

$$c(\mathbf{x}^{(1)}) = -0.25, \quad \mathbf{a}^{(0)} = (-1, 1), \quad \mathbf{g}^{(0)} = (1, 1), \quad \mathbf{W}^{(1)} = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}.$$

将上述向量代入二次规划子问题后，得

$$\begin{array}{ll} \text{minimize} & s_1^2 + s_1 + s_2 \\ \text{subject to} & -0.25 - s_1 + s_2 \geq 0. \end{array}$$

解上述二次规划子问题得解 $\mathbf{s}^{(1)} = (0, 0.25)$, 对应的 Lagrange 乘子 $\lambda^{(2)} = 1$. 于是 $\mathbf{x}^{(2)} = \mathbf{x}^{(1)} + \mathbf{s}^{(1)} = (-0.5, 0.25)$.

易验证这是原问题的最优解。故从该初始点出发，由基本 SQP 法迭代两次即可得到问题的解。

9.14 考虑约束 $x_1^2 + x_2^2 = 1$. 写出在下列点 $(0,0), (0,1), (0.1, 0.02), -(0.1, 0.02)$ 处的线性化约束. 解. 令 $c(x) = x_1^2 + x_2^2 - 1, x^1 = (0,0), x^2 = (0,1)$ ，则 $\nabla c(x) = \begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix}$.

$$c(x^1) = -1, c(x^2) = 0, \nabla c(x^1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \nabla c(x^2) = \begin{bmatrix} 0 \\ 2 \end{bmatrix},$$

所以约束 $x_1^2 + x_2^2 = 1$ 在 x^1 处的线性化约束为 $c(x^1) + \nabla c(x_1)^T (x - x^1) = 0$ ，即 $-1 = 0$;
约束 $x_1^2 + x_2^2 = 1$ 在 x^2 处的线性化约束为 $c(x^2) + \nabla c(x_2)^T (x - x^2) = 0$ ，即 $x_2 = 1$; 类似可得约束 $x_1^2 + x_2^2 = 1$ 在其它点的线性化约束.

最优化理论与算法 B： 第 10 章部分作业参考解答

10.1 考虑问题

$$\begin{array}{ll} \text{minimize} & f(x) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 \\ \text{subject to} & x_2 = 0. \end{array}$$

(a) 计算最优解和 Lagrange 乘子.

(b) 对 $k = 0, 1, 2$ 和惩罚参数 $\frac{1}{\mu_k} = 10^{k+1}$，计算二次惩罚法迭代的点和增广 Lagrange 函数法迭代的点 (其中取 $\lambda^0 = 0$).

(c) 针对该问题，画出图形从几何上解释方法，并在该图形上画出这两种方法对 $k = 0, 1, 2$ 的迭代.

(d) 假设在乘子法中将惩罚参数取为常数. 对于 μ 的哪些值，增广 Lagrange 函数将有一个极小点？对于 μ 的哪些值，方法将是收敛的？

解.

(a) 显然原问题的最优解为 $x^* = (0, 0)$. 而 $\nabla f(x^*) = (0, -3), \nabla c(x^*) = (0, 1)$，所以最优解 x^* 对应的 Lagrange 乘子为 $\lambda^* = -3$.

(b) **二次惩罚法**：二次惩罚函数 $P(x, \mu) = \frac{1}{2}(x_1^2 - x_2^2) - 3x_2 + \frac{1}{2\mu}x_2^2$.

令

$$\nabla P(x, \mu) = \begin{bmatrix} x_1 \\ -x_2 - 3 + \frac{1}{\mu}x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

解得 $x_1(\mu) = 0, x_2(\mu) = \frac{3\mu}{1-\mu}$. 又因为 $\mu_k = 10^{-(k+1)} < 1$，所以

$$\nabla^2 P(x, \mu_k) = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_k} - 1 \end{bmatrix}$$

正定. 所以 $(0, \frac{3\mu_k}{1-\mu_k})$ 是 $P(x, \mu_k)$ 的极小值点.

当 $k = 0$ 时，　$\mu_0 = 0.1$,　$x^0 = (0, 1/3)$ ；
当 $k = 1$ 时，　$\mu_1 = 0.01$,　$x^1 = (0, 1/33)$ ；

当 $k=2$ 时， $\mu_2=0.001$， $x^2=(0,1/333)$ ；

乘子法:
增广 Lagrange 函数 $\mathcal{L}_A(x,\lambda,\mu)=\frac{1}{2}(x_1^2-x_2^2)-3x_2+\frac{1}{2\mu}x_2^2-\lambda x_2$. 令

$$\nabla\mathcal{L}_A(x,\lambda,\mu)=\begin{bmatrix} x_1 \\ -x_2-3+\frac{1}{\mu}x_2-\lambda \end{bmatrix}=\begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

解得 $x_1(\mu)=0, x_2(\mu)=\frac{\mu}{1-\mu}(3+\lambda)$. 又因为 $\mu_k=10^{-(k+1)}<1$ ，所以

$$\nabla^2\mathcal{L}_A(x,\lambda_k,\mu_k)=\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu_k}-1 \end{bmatrix}$$

正定. 所以 $(0,\frac{\mu_k(3+\lambda_k)}{1-\mu_k})$ 是 $\mathcal{L}_A(x,\lambda_k,\mu_k)$ 的极小值点.

当 $k=0$ 时， $\mu_0=0.1$， $\lambda_0=0$， $x^0=(0,1/3)$ ；
当 $k=1$ 时， $\mu_1=0.01$， $\lambda_1=\lambda_0-\frac{c(x^0)}{\mu_0}=-\frac{10}{3}$， $x^1=(0,-1/297)$ ；
当 $k=2$ 时， $\mu_2=0.001$， $\lambda_2=\lambda_1-\frac{c(x^1)}{\mu_1}=-\frac{890}{297}$， $x^2=(0,\frac{1}{999\times 297})$ ；
由此可见，乘子法迭代收敛到最优解的速率要比二次惩罚法快.

(c) 图示略。

(d) 因为

$$\nabla^2\mathcal{L}_A(x,\lambda_k,\mu)=\begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{\mu}-1 \end{bmatrix},$$

所以当 $0<\mu<1$ 时， $\nabla^2\mathcal{L}_A(x,\lambda_k,\mu)$ 正定. 从而 $(0,\frac{\mu(3+\lambda_k)}{1-\mu})$ 是 $\mathcal{L}_A(x,\lambda_k,\mu)$ 的极小值点.

此时，由乘子法的迭代

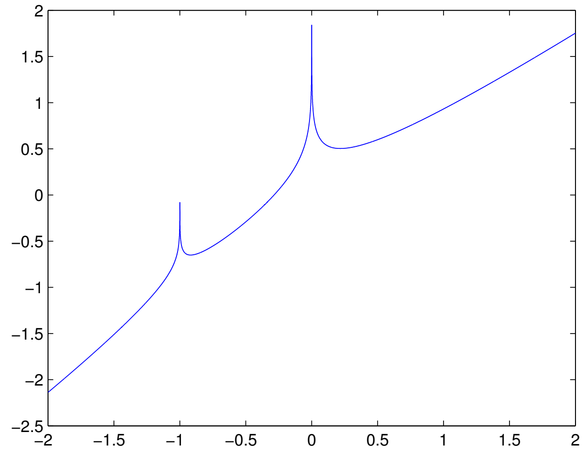
$$\lambda_{k+1}=\lambda_k-\frac{x_2^k}{\mu}=\lambda_k-\frac{1}{\mu}\frac{\mu}{1-\mu}(\lambda_k+3)=\frac{-\mu}{1-\mu}\lambda_k-\frac{3}{1-\mu}$$

易见 $|\frac{-\mu}{1-\mu}|<1$ ，即 $0<\mu<1/2$ 时， $\{\lambda_k\}$ 是收敛的，且 $\lambda_k\to-3(k\to\infty)$. 此时 $(0,\frac{\mu(3+\lambda_k)}{1-\mu})\to(0,0)(k\to\infty)$. 所以对于此问题而言，乘子法不用 $\mu_k\to 0$ ，只需 $0<\mu<1/2$ ，乘子法即是收敛的. 从而可以避免出现病态问题.

10.2 考虑标量极小化问题：

$$\min_x x, \quad \text{subject to } x^2\geq 0, x+1\geq 0,$$

该问题的解是 $x^*=-1$. 写出针对该问题的 $P(x;\mu)$ ，并找到它的局部极小点. 证明对数障碍函数的全局极小点收敛到约束全局极小点 $x^*=-1$ ，但是非全局极小点的局部极



小点收敛到零，其不是约束极小点.

解. 该问题的对数障碍函数为 $P(x,\mu)=x-\mu\ln x^2-\mu\ln(x+1)$. 令 $P'(x,\mu)=1-\frac{2\mu}{x}-\frac{\mu}{x+1}=0$ ，解得

$$\begin{aligned} x_1(\mu) &= \tfrac{3\mu-1+\sqrt{9\mu^2+2\mu+1}}{2}, \\ x_2(\mu) &= \tfrac{3\mu-1-\sqrt{9\mu^2+2\mu+1}}{2}. \end{aligned}$$

因为 $P''(x,\mu)=\frac{2\mu}{x^2}+\frac{\mu}{(x+1)^2}>0$ ，所以 $P(x,\mu)$ 的驻点 $x_1(\mu)$ 和 $x_2(\mu)$ 即为其局部极小点. 易见当 $\mu_k\downarrow 0$ 时， $x_1(\mu_k)\to 0, x_2(\mu_k)\to -1$. 上面是 $P(x,0.1)$ 的图形.