

数学 · 软件 · 应用系列教材

二次规划

——非线性规划与投资组合的算法

Quadratic Programming

Algorithms for

Nonlinear Programming and Portfolio Selection

张忠桢 著

武汉大学出版社

前言

二次规划是最简单的非线性规划,通常通过解其库恩-塔克条件(KT 条件)获取一个解. KT 条件的解称为 KT 对,其中与原问题的变量对应的部分称为 KT 点. 二次规划分为凸二次规划和非凸二次规划,前者的 KT 点便是其全局极小点,而后的 KT 点可能连局部极小点都不是.

有很多算法可用于解凸二次规划,如 Wolfe, Dantzig, Lemke, van de Panne and Whinston, Fletcher, Keller, Best and Ritter, Gill and Murray, von Hohenbalken, Beale, Goldfarb, Benveniste, Sacher, Jensen and King 等人的方法. 还有不少内点算法可用于解凸二次规划. 应用最广泛的是 Fletcher 的有效集法. 这种方法可用于非凸二次规划,但是也只能求出 KT 点,不能保证它是局部极小点. 为了判断 KT 点是不是局部极小点,通常的做法是随机地产生许多距离 KT 点很近的可行解,看是否有一个解对应的目标函数值更小,称为摄动(perturbation)法. 这种方法有时如同大海捞针,当 KT 点处可行下降方向的集合是一个十分狭窄的区域时,很难发现一个更好的解,而且可能不存在可行下降方向. 例如 $x_1 = 0$, $x_2 = 0$ 是二次规划

minimize $f(x) = x_1^2 - 20x_1x_2 + (100 - \delta)x_2^2$ subject to $x_1 + x_2 \leq 1$, $x_1, x_2 \geq 0$ 的 KT 点. 当 δ 是很小正数时,目标函数在 KT 点处下降方向的区域很狭窄;当 δ 小于或等于零时没有下降方向.

本书重点介绍非凸二次规划的旋转算法,由简单到复杂逐步讨论以下几个问题:

- (i) 求二次齐次函数的非负下降方向;
- (ii) 求二次齐次函数在锥内的下降方向;
- (iii) 求非凸二次规划的局部极小点;
- (iv) 求非凸二次规划的全局极小点.

问题(i)是在 n 维空间直角坐标系的第一象限找一个点使一个不定二次型小于零,可用数学规划模型表示为

$$\text{minimize } f(x) = x^T H x \text{ subject to } x \geq 0,$$

其中 H 是 n 阶不定对称矩阵, x 是 n 维未知列向量. 这是一个 NP hard 问题,见

Murty 和 Kabadi 的著述(1987). 如果 H 有一个负的对角元素, 或者有一个 2 阶主子矩阵, 其两个对角元素为 0, 非对角元素是负数, 很容易确定一个非负的 x 使得 $f(x) < 0$. 问题难在 H 的所有对角元素大于零且每行(列)都有负的非对角元素的情形. 本书利用旋转运算, 结合一种参数化技术, 证明了此问题等价于在一个有限集中寻找 I 型或 II 型非负下降方向. 问题(ii)与问题(i)类似, 不过形式要复杂一些. 问题(iii)的关键在于判断 KT 点处是否有可行下降方向, 有赖于问题(ii)的解决. 本书以较大篇幅介绍非凸二次规划局部极小点, 即问题(iii)的计算方法. 至于问题(iv), 只有当变量数目较少时才可能以较小的计算量获得全局极小点. 随着问题规模不断增加, 计算量成指数膨胀. Hansen 等人(1993)、Vandenbussche 和 Nemhauser(2005)提出用分支定界或分支切割算法解一个简单的非凸二次规划, 其约束条件仅仅是每个变量在 0 和 1 之间. 也就是说, 用类似于解混合整数规划的方法解一个简单的非凸二次规划, 其计算量之大可想而知. 本书简单介绍如何从一个局部极小点出发寻找其他局部极小点, 为进一步研究非凸二次规划的全局极小点提供一种思路.

序列二次规划(SQP)法是求解复杂非线性规划最有效的方法. SQP 法将一个复杂的非线性规划在某一点处用一个二次规划近似, 前者称为原问题, 后者称为子问题. 通过求解二次规划子问题获取一个新的解(点), 然后再将原问题在新的点用一个二次规划近似并求解, 如此等等, 以获取原问题的一个解. 用 SQP 法解非线性规划关键在于二次规划模型的形式及其计算方法. 序列二次规划法分为牛顿法和拟牛顿法. 在无约束极值问题的牛顿法中, 子问题目标函数的海色(Hesse)矩阵是原问题目标函数 $f(x)$ 在某点 $x^{(k)}$ 处的海色矩阵 $\nabla^2 f(x^{(k)})$. 对于约束极值问题, 子问题目标函数的海色矩阵通常是拉格朗日函数在 $(x^{(k)}, \lambda^{(k)})$ 处的海色矩阵 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)})$. 如果原问题的目标函数和约束条件较为复杂, 计算二阶偏导数是一项烦琐的工作, 而且得到的海色矩阵可能不是正定的, 导致一个非凸的二次规划子问题. 于是, 人们常用一个正定矩阵 B_k 代替目标函数或拉格朗日函数的海色矩阵, 并且在迭代过程中不断校正 B_k , 这种方法便是拟牛顿法. 近几十年来, 拟牛顿法的理论研究成果很多, 实际计算效果也较好. 校正 B_k 的公式有 DFP 公式、BFGS 公式、Broyden 族、Huang 族等, 其中最有名的是 BFGS 公式. 这些公式力图使 B_k 正定, 或者说使子问题成为凸二次规划. 可以证明, 当 B_k 正定且充分接近 $\nabla^2 f(x^{(k)})$ 或 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)})$ 时, 算法的收敛性很好. 对于无约束极值问题, 容易做到所有 B_k 正定, 例如使用精确线搜索, 但是这一点并不能保证算法对于所有问题收敛. Dai Yuhong(2002)和 Mascarenhas W. F.(2004)就给出了 BFGS 法用于非凸函数失效的一些例子. 至于某些约束极值问题, 想在迭代过程中让每个 B_k 正定都很困难. 早在 1987 年,

Murty 和 Kabadi 就指出了这些算法的局限性,他们说:“许多关于非线性规划的教科书留给读者一种印象:其算法收敛于凸非线性规划的全局极小点,收敛于非凸非线性规划的局部极小点,许多专业非线性规划的软件包文件也制造同样的印象,这是十分错误的。”

本书也用序列二次规划法解更为复杂的非线性规划。但与拟牛顿法不同,二次规划子问题目标函数的海色矩阵就是原问题目标函数的海色矩阵,或者是拉格朗日函数的海色矩阵。当约束条件较为复杂时,还可能使用增广拉格朗日函数的海色矩阵。这时的子问题可能是凸二次规划也可能是非凸的。非线性规划的形式多种多样,不同的算法有各自的适用范围。有的问题用序列线性规划法计算效果也很好。希望本书介绍的方法能弥补传统算法的不足。

本书还介绍若干种资产组合选择模型的计算方法。这些模型都以方差作为风险的度量,其中一种模型考虑凹交易成本,另有一种考虑 V 形交易成本,用序列二次规划法计算。据我们所知,现有文献都是用线性规划解这两种问题,用绝对偏差等指标作为风险的度量,而不是实际中通常使用的方差风险。

全书分为六章。第一章介绍线性代数和微分学的一些基础知识以及最优化的有关定理。第二章介绍线性不等式组的解法及参数化技术。第三章介绍二次规划旋转算法的基本概念和凸二次规划的解法。第四章由简单到复杂逐步介绍非凸二次规划的计算方法。第五章介绍约束非线性规划的序列二次规划法。第六章介绍若干均值方差资产组合选择问题的解法。

全书结构严谨,深入浅出,所有算法都有相应的定理为基础。为便于初学者掌握,仅用到一些较基本的数学知识,并且大多数问题有简单的例子说明算法的计算步骤。阅读本书只需具备线性代数、概率论和微分学的一些基础知识。本书可作为应用数学、经济学、管理科学以及许多工程技术学科研究生的教材或教学参考书。涉足运筹学优化的教师和研究人員更应该知道非线性规划特别是一些较为简单的非线性规划如何计算。

张忠桢

2006 年 3 月于武汉

目 录

第一章 基础知识.....	1
1.1 矩阵的旋转运算及分块矩阵公式	1
1.2 向量与矩阵的范数及其性质	8
1.3 凸锥及 Farkas 引理	10
1.4 约束最优化及 Kuhn-Tucker 条件	15
1.5 约束最优化的其他一些定理.....	20
第二章 线性不等式组	26
2.1 基本概念及有关定理	26
2.2 线性不等式组的旋转算法	29
2.3 循环与反循环	37
2.4 参数化技术.....	40
2.5 线性不等式组的极点解	44
第三章 凸二次规划	53
3.1 二次规划的旋转算法	53
3.2 凸二次规划的计算步骤	63
3.3 变量有上界的凸二次规划	66
第四章 非凸二次规划	78
4.1 二次齐次函数的下降方向	78
4.2 I 型和 II 型非负下降方向	79
4.3 二次齐次函数在锥内的下降方向	95
4.4 非凸二次规划的局部极小点	117
4.5 局部极小点的改进.....	140
第五章 约束非线性规划.....	152

5.1 序列二次规划法简介	152
5.2 序列线性规划法	161
5.3 牛顿法	167
5.4 拉格朗日法	176
5.5 增广拉格朗日法	182
5.6 拟牛顿法	189
 第六章 均值方差投资组合选择模型的解法	 195
6.1 标准均值方差投资组合选择模型	195
6.2 资产有上界限制的均值方差模型	208
6.3 具有交易成本的均值方差模型	212
6.4 自融资均值方差模型	216
6.5 V形交易成本的均值方差模型	222
 参考文献	 226

第一章 基础知识

1.1 矩阵的旋转运算及分块矩阵公式

1.1.1 矩阵的一些概念

设 $A = (a_{ij})$ 是 $m \times n$ 矩阵, 其中 a_{ij} 是第 i 行第 j 列的元素, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$. A 的 m 个行向量的秩称为 A 的行秩, A 的 n 个列向量的秩称为 A 的列秩. 若 A 的 m 行线性无关, 则称 A 行满秩. 同样, 若 A 的 n 列线性无关, 则称 A 列满秩. 可以证明, A 的行秩等于 A 的列秩. A 的行秩或列秩称为 A 的秩. 秩等于行数或列数的矩阵称为满秩矩阵.

若 $A = (a_{ij})$ 的行数等于列数等于 n , 称 A 是一个 n 阶方阵, 其中 $a_{11}, a_{22}, \dots, a_{nn}$ 称为对角元素, 其他元素称为非对角线元素. 一个满秩方阵称为非奇异矩阵. 若存在方阵 B , 使得 $AB = BA = I$, 则称 A 可逆, 并称 B 为 A 的逆矩阵, 记为 A^{-1} , 这里的 I 是单位矩阵. 可以证明, A 为非奇异矩阵的必要充分条件是 A 可逆.

n 阶方阵 $A = (a_{ij})$ 的元素如果满足 $a_{ij} = a_{ji} (i \neq j)$ 或者说 $A^T = A$, 则称 A 为对称矩阵, 这里的 T 是转置符. 设 A 为 n 阶对称矩阵. 如果对于任何 n 维向量 x (列向量) 均有 $x^T A x \geq 0$, 则称 A 正半定. 若等号仅当 $x = 0$ 时成立, 则称 A 正定.

可以证明以下结论成立.

结论 1.1.1 设 A 为对称矩阵, 则以下诸命题等价:

- (i) A 为正半定矩阵;
- (ii) 有矩阵 B 使 $A = B^T B$;
- (iii) A 的所有(顺序)主子式非负;
- (iv) A 的所有特征值非负.

结论 1.1.2 设 A 为对称矩阵, 则以下诸命题等价:

- (i) A 为正定矩阵；
- (ii) 有列满秩矩阵 B 使 $A = B^T B$ ；
- (iii) A 的所有(顺序)主子式为正；
- (iv) A 的所有特征值大于 0.

1.1.2 矩阵的旋转运算

设 $A = (a_{ij})$ 是 $m \times n$ 矩阵. 用 a_i 表示 A 的第 i 行, e_i 表示 n 阶单位矩阵的第 i 行, 则

$$a_i = a_{i1}e_1 + a_{i2}e_2 + \dots + a_{in}e_n, \quad i = 1, 2, \dots, m. \quad (1.1.1)$$

对于某个 $r \in \{1, 2, \dots, m\}$ $s \in \{1, 2, \dots, n\}$ 如果 $a_{rs} \neq 0$ 则可由式(1.1.1)的第 r 式解得

$$e_s = \frac{1}{a_{rs}}a_r + \sum_{j=1, j \neq s}^n \left(-\frac{a_{rj}}{a_{rs}} \right) e_j. \quad (1.1.2)$$

将其代入式(1.1.1)的其余各式整理后可得

$$a_i = \frac{a_{is}}{a_{rs}} a_r + \sum_{j=1, j \neq s}^n \left(a_{ij} - \frac{a_{is}}{a_{rs}} a_{rj} \right) e_j, \quad i = 1, \dots, r-1, r+1, \dots, m. \quad (1.1.3)$$

以上运算过程称为旋转运算, 运算前后的数据可用表 1.1、表 1.2 表示.

表 1.1	初 始 表	
	e_s	e_j
a_r	a_{rs}	a_{rj}
a_i	a_{is}	a_{ij}

表 1.2	旋转运算的结果	
	a_r	e_j
e_s	$\frac{1}{a_{rs}}$	$-\frac{a_{rj}}{a_{rs}}$
a_i	$\frac{a_{is}}{a_{rs}}$	a'_{ij}

其中 $a'_{ij} = a_{ij} - \frac{a_{is}}{a_{rs}} a_{rj}, i = 1, \dots, r-1, r+1, \dots, m, j = 1, \dots, s-1, s+1, \dots, n$.

如果将向量组 $a_1, \dots, a_m, e_1, \dots, e_n$ 的一个最大线性无关组称为基, 那么 $e_1, e_2, \dots, e_n$ 是其中一个基. 由于 $a_{rs} \neq 0$, 向量组 $e_1, \dots, e_{s-1}, a_r, e_{s+1}, \dots, e_n$ 也线性无关, 因而也是一个基. 在以上运算中, 我们说 a_r 入基 e_s 出基, 两者位置的交换记为 $a_r \leftrightarrow e_s, a_{rs}$ 称为枢轴(元素). 表 1.1 中(即矩阵 A 中)枢轴元素 a_{rs} 所在行称为枢轴行, 其所在列称为枢轴列. 去掉表 1.2 中 $\frac{1}{a_{rs}}$ 所在行和列后, 剩下的元素

构成的矩阵 (a'_{ij}) 称为 A 关于 a_{rs} 的残余矩阵.

给定一组向量 $a_i = (a_{i1} \ a_{i2} \ \dots \ a_{in})$, $i = 1 \ 2 \ \dots \ m$, 其中有多少向量线性无关, 或者一个向量能否表示为其余若干向量的线性组合, 这个问题在许多问题的讨论中具有十分重要的作用, 下面用旋转运算研究这个问题.

为叙述方便起见, 将向量组 $a_1 \ \dots \ a_m \ e_1 \ \dots \ e_n$ 的一个基中的向量称为基向量, 将其余向量称为非基向量, 并将 $a_1 \ a_2 \ \dots \ a_m$ 称为原向量, $e_1 \ e_2 \ \dots \ e_n$ 称为外来向量. 首先以 $e_1 \ e_2 \ \dots \ e_n$ 作为基向量, 并将所有原向量表示为这些向量的线性组合, 如表 1.1 所示. 然后尽可能地将非基向量中的原向量与基中的外来向量交换, 能够进行这种交换的必要充分条件是残余矩阵中至少有一个非零元素. 最后将出现以下两种情况:

(i) 所有原向量 $a_1 \ a_2 \ \dots \ a_m$ 进入基中, 因此这些向量线性无关;

(ii) $a_1 \ a_2 \ \dots \ a_m$ 中的部分向量进入基中, 而所有非基向量中的原向量已表示为这些向量的线性组合.

由于每进行一次旋转运算有一个原向量入基, 所以计算结束时, 旋转运算的次数等于向量组 $a_1 \ a_2 \ \dots \ a_m$ 的秩. 实际计算时, 某个外来向量一旦出基便可删掉. 例如可不要表 1.2 中的 e_s 行. 如果目的仅在于求向量组 $a_1 \ a_2 \ \dots \ a_m$ 的秩或其中的一个最大无关组, 还可不要表 1.2 中的 a_r 列, 因为原向量进入基中后, 其所在列各元素不能再作枢轴元素.

现设 $A = (a_{ij})$ 是个 n 阶方阵. 同样将其 n 行 $a_1 \ a_2 \ \dots \ a_n$ 表示为 n 个单位向量 $e_1 \ e_2 \ \dots \ e_n$ 的线性组合, 如表 1.1 所示. 如果 $a_{11} \neq 0$, 则可进行一次旋转运算. 在得到的表中, 如果主对角线上第二个元素也不为零, 则可进行第二次旋转运算, 依此类推. 假设共进行了 n 次旋转运算, 得到的结果用式子表示即为

$$\begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix} = \begin{pmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ u_{21} & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & & \vdots \\ u_{n1} & u_{n2} & \dots & u_{nn} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix},$$

可见得到的矩阵 (u_{ij}) 是 A 的逆矩阵.

注 如果用计算机求一个 n 阶方阵 A 的逆阵, 为减少计算误差, 最好在残余矩阵(开始时为 A)中选一个绝对值最大的元素. 如果大于一个充分小的正数 ε (一般为 $10^{-6} \sim 10^{-4}$) 便以其为枢轴作一次旋转运算. 只要能进行 n 次旋转运算, A 可逆. 这时, 初始的基向量 $e_1 \ e_2 \ \dots \ e_n$ 成为非基向量, 非基向量 $a_1 \ a_2 \ \dots \ a_m$ 成为基向量, 但它们的次序往往会发生变化. 为得到 A 的逆, 只需将

e_i 行和 a_j 列上的元素放在第 i 行和第 j 列, $i, j = 1, 2, \dots, n$. 挑选绝对值最大的元素作枢轴相当于线性方程组的全主元高斯消元法, 计算量大但解的精度高. 当 n 很大时, 可使用部分选主元的方法. 如果 A 正定, 沿其主对角线进行旋转运算便可, 数值稳定性非常好(见曹志浩、张玉得和李瑞遐(1979)). 这意味着, 当 A 是一个具有正对角元素的对称矩阵时, 仅以对角线上的正元素为枢轴作旋转运算, 数值稳定性将会非常好.

定理 1.1.1(旋转运算公式) 设一个 $m \times n$ 矩阵的分块矩阵为 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$, 其中 A 为可逆方阵. 那么, 用旋转运算在此分块矩阵中求 A 的逆, 其结果为 $\begin{pmatrix} A^{-1} & -A^{-1}B \\ CA^{-1} & D - CA^{-1}B \end{pmatrix}$.

证 用 a_i 表示此 $m \times n$ 矩阵的第 i 行, e_i 表示 n 阶单位矩阵的第 i 行, 设 A 的阶数为 k , 那么

$$\begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} = A \begin{pmatrix} e_1 \\ \vdots \\ e_k \end{pmatrix} + B \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix},$$

两边左乘 A^{-1} 并移项得

$$\begin{pmatrix} e_1 \\ \vdots \\ e_k \end{pmatrix} = A^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} - A^{-1}B \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix},$$

所以对 A 求逆后上方的两个子块变为 A^{-1} 和 $-A^{-1}B$. 另外

$$\begin{aligned} \begin{pmatrix} a_{k+1} \\ \vdots \\ a_m \end{pmatrix} &= C \begin{pmatrix} e_1 \\ \vdots \\ e_k \end{pmatrix} + D \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix} = C \left[A^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} - A^{-1}B \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix} \right] + D \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix} \\ &= CA^{-1} \begin{pmatrix} a_1 \\ \vdots \\ a_k \end{pmatrix} + (D - CA^{-1}B) \begin{pmatrix} e_{k+1} \\ \vdots \\ e_n \end{pmatrix}, \end{aligned}$$

所以下方的两个子块变为 CA^{-1} 和 $D - CA^{-1}B$.

1.1.3 若干分块矩阵公式

以下是一些较为常用的分块矩阵公式.

(i) 若 A_{11} 可逆, 则

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I_1 & O \\ A_{21}A_{11}^{-1} & I_2 \end{pmatrix} \begin{pmatrix} A_{11} & O \\ O & A_{22} - A_{21}A_{11}^{-1}A_{12} \end{pmatrix} \begin{pmatrix} I_1 & A_{11}^{-1}A_{12} \\ O & I_2 \end{pmatrix}, \quad (1.1.4a)$$

其中 I_1 和 I_2 为单位矩阵, 若 A_{22} 可逆, 则

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} I_1 & A_{12}A_{22}^{-1} \\ O & I_2 \end{pmatrix} \begin{pmatrix} A_{11} - A_{12}A_{22}^{-1}A_{21} & O \\ O & A_{22} \end{pmatrix} \begin{pmatrix} I_1 & O \\ A_{22}^{-1}A_{21} & I_2 \end{pmatrix}. \quad (1.1.4b)$$

(ii) 若 A_{11} 可逆, 则

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{可逆} \Leftrightarrow (A_{22} - A_{21}A_{11}^{-1}A_{12}) \text{可逆},$$

其中“ $\Leftrightarrow$ ”表示必要充分条件, 并且

$$\begin{aligned} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} &= \begin{pmatrix} I_1 & -A_{11}^{-1}A_{12} \\ O & I_2 \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & O \\ O & (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1} \end{pmatrix} \begin{pmatrix} I_1 & O \\ -A_{21}A_{11}^{-1} & I_2 \end{pmatrix} \\ &= \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \end{aligned} \quad (1.1.5a)$$

其中

$$\begin{aligned} B_{11} &= A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1}, \\ B_{12} &= -A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}, \\ B_{21} &= -(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1}, \\ B_{22} &= (A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}. \end{aligned}$$

若 A_{22} 可逆, 则

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \text{可逆} \Leftrightarrow (A_{11} - A_{12}A_{22}^{-1}A_{21}) \text{可逆},$$

并且

$$\begin{aligned} \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} &= \begin{pmatrix} I_1 & O \\ -A_{12}A_{22}^{-1} & I_2 \end{pmatrix} \begin{pmatrix} (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} & O \\ O & A_{22}^{-1} \end{pmatrix} \begin{pmatrix} I_1 & -A_{12}A_{22}^{-1} \\ O & I_2 \end{pmatrix} \\ &= \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}, \end{aligned} \quad (1.1.5b)$$

其中

$$B_{11} = (A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1},$$

$$B_{12} = - (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} A_{12} A_{22}^{-1},$$

$$B_{21} = - A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1},$$

$$B_{22} = A_{22}^{-1} + A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} A_{12} A_{22}^{-1}.$$

(iii) 若 $A_{11}, A_{22}, \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 均可逆, 则

$$(A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1},$$

$$(A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} A_{12} A_{22}^{-1} = A_{11}^{-1} A_{12} (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1},$$

$$A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1} A_{21} A_{11}^{-1},$$

$$A_{22}^{-1} + A_{22}^{-1} A_{21} (A_{11} - A_{12} A_{22}^{-1} A_{21})^{-1} A_{12} A_{22}^{-1} = (A_{22} - A_{21} A_{11}^{-1} A_{12})^{-1}. \quad (1.1.6)$$

(iv) 若 A_{12}, A_{21} 可逆, 则

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & O \end{pmatrix}^{-1} = \begin{pmatrix} O & A_{21}^{-1} \\ A_{12}^{-1} & -A_{12}^{-1} A_{11} A_{21}^{-1} \end{pmatrix}. \quad (1.1.7)$$

以上公式选自须田信英等的相关著述(1979).

1.1.4 其他旋转运算公式

定理 1.1.2 设 $\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$ 是秩为 r 的 n 阶正半定矩阵, 其中 A 为 k 阶可逆方

阵, $k < n$. 则 $D - B^T A^{-1} B$ 是秩为 $r - k$ 的正半定矩阵.

证 在下式两边右乘 $(0, x)$ 左乘 $(0, x)^T$, 其中 0 是 k 维零向量, x 是任意 $n - k$ 维向量(都是行向量):

$$\begin{pmatrix} I & O \\ -B^T A^{-1} & I \end{pmatrix} \begin{pmatrix} A & B \\ B^T & D \end{pmatrix} \begin{pmatrix} I & -A^{-1} B \\ O & I \end{pmatrix} = \begin{pmatrix} A & O \\ O & D - B^T A^{-1} B \end{pmatrix}.$$

由于 $\begin{pmatrix} A & B \\ B^T & D \end{pmatrix}$ 正半定, $x(D - B^T A^{-1} B)x^T$ 非负, 因而 $D - B^T A^{-1} B$ 正半定, 且秩为 $r - k$. ■

定理 1.1.1 和定理 1.1.2 说明, 如果矩阵 A 正半定, 则以主对角线上的正元素为枢轴进行旋转运算得到的残余矩阵也正半定.

定理 1.1.3 设 A 是一个 n 阶对称矩阵, 则 A 正定的必要充分条件是可沿其主对角线进行 n 次旋转运算并且所有枢轴元素都是正数.

证 考虑 A 的 $k+1$ 阶主子矩阵并将其表示为分块矩阵 $\begin{pmatrix} A_k & b \\ b^T & d \end{pmatrix}$, 其中 A_k 是 A 中左上角的 k 阶矩阵, $1 \leq k < n$, 并设 A_k 的行列式大于 0. 由于 $\begin{vmatrix} A_k & b \\ b^T & d \end{vmatrix} = |A_k| (d - b^T A_k^{-1} b)$ 并且 $d - b^T A_k^{-1} b$ 是在 A 中沿 A_k 的主对角线进行 k 次旋转运算所得矩阵主对角线上第 $k+1$ 个元素, 由归纳不难看出此定理成立. ■

定理 1.1.4 在分块矩阵 $\begin{pmatrix} B & -A^T \\ A & D \end{pmatrix}$ 中 B 正定, D 正半定, 那么沿 B 的对角线进行旋转运算得到的矩阵 $\begin{pmatrix} B^{-1} & B^{-1}A^T \\ AB^{-1} & D + AB^{-1}A^T \end{pmatrix}$ 正半定并且其秩等于 B 与 D 的秩之和.

证 因为

$$\begin{pmatrix} B^{-1} & B^{-1}A^T \\ AB^{-1} & D + AB^{-1}A^T \end{pmatrix} = \begin{pmatrix} I_1 & O \\ A & I_2 \end{pmatrix} \begin{pmatrix} B^{-1} & O \\ O & D \end{pmatrix} \begin{pmatrix} I_1 & A^T \\ O & I_2 \end{pmatrix}. \quad \blacksquare$$

设 A 和 B 是同阶方阵. 如果存在非奇异矩阵 P , 使得 $B = P^T A P$, 称 A 合同于 B , 或者说, 按这种方式由 A 得到 B 的运算是合同变换. 合同变换不改变实对称矩阵的正、负和零特征值的数目.

定理 1.1.5 设 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix}$, A_{11} 是非奇异方阵, 那么在分块矩阵

$$\begin{pmatrix} B_{11} & B_{12} & -A_{11}^T \\ B_{12}^T & B_{22} & -A_{12}^T \\ A_{11} & A_{12} & 0 \end{pmatrix}$$

中先后求 A_{11} 和 $-A_{11}^T$ 的逆矩阵是对 B 作合同变换.

证 按照定理 1.1.1, 在以上 3×3 分块矩阵中先后求 A_{11} 和 $-A_{11}^T$ 的逆矩阵, 结果为

$$\begin{pmatrix} A_{11}^{-T} B_{11} A_{11}^{-1} & A_{11}^{-T} B_{12} - A_{11}^{-T} B_{11} A_{11}^{-1} A_{12} & -A_{11}^{-T} \\ B_{12}^T A_{11}^{-1} - A_{12}^T A_{11}^{-T} B_{11} A_{11}^{-1} & B_{22} & A_{12}^T A_{11}^{-T} \\ A_{11}^{-1} & -A_{11}^{-1} A_{12} & O \end{pmatrix},$$

其中

$$B'_{22} = B_{22} - B_{12}^T A_{11}^{-1} A_{12} - A_{12}^T A_{11}^{-T} (B_{12} - B_{11} A_{11}^{-1} A_{12}),$$

A_{11}^{-T} 是 A_{11}^{-1} 的转置, 其左上角的 2×2 分块矩阵可表示为

$$\begin{pmatrix} A_{11}^{-T} & O \\ -A_{12}^T A_{11}^{-T} & I \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{12}^T & B_{22} \end{pmatrix} \begin{pmatrix} A_{11}^{-1} & -A_{11}^{-1} A_{12} \\ O & I \end{pmatrix},$$

所以是对 B 作的一种合同变换。

1.2 向量与矩阵的范数及其性质

1.2.1 向量与矩阵的范数

设 x 是 n 维向量空间 R^n 的一个向量. x 的范数是一个非负实数, 记为 $\|x\|$, 它满足以下三个条件:

- (i) $\|x\| \geq 0$, 等号仅当 $x=0$ 时成立;
- (ii) 对于任意实数 $\alpha \in R$, $\|\alpha x\| = |\alpha| \|x\|$;
- (iii) 对于任意向量 $x, y \in R^n$, 有 $\|x+y\| \leq \|x\| + \|y\|$.

容易证明

$$\|x\| - \|y\| \leq \|x+y\|, \quad \|y\| - \|x\| \leq \|x+y\|.$$

所以

$$|\|x\| - \|y\|| \leq \|x+y\| \leq \|x\| + \|y\|,$$

此式称为三角不等式.

设 $x = (x_1, x_2, \dots, x_n)^T$, 常用的向量范数有以下几种:

$$\|x\|_1 = \sum_{i=1}^n |x_i|,$$

$$\|x\|_2 = \sqrt{\sum_{i=1}^n x_i^2},$$

$$\|x\|_\infty = \max_i |x_i|.$$

设 A 是 n 阶方阵. A 的范数是一个非负实数, 记为 $\|A\|$, 它满足以下四个条件:

- (i) $\|A\| \geq 0$, 等号仅当 $A=O$ 时成立;
- (ii) 对于任意实数 $\alpha \in R$, $\|\alpha A\| = |\alpha| \|A\|$;
- (iii) 对于任意 n 阶方阵 A 和 B , 有 $\|A+B\| \leq \|A\| + \|B\|$;

(iv) $\|AB\| \leq \|A\| \|B\|$.

设 $A = (a_{ij})_{n \times n}$, 常用的矩阵范数有以下三种:

$$\|A\|_1 = \max_j \sum_{i=1}^n |a_{ij}|,$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)},$$

其中 $\lambda_{\max}(A^T A)$ 表示 $A^T A$ 的最大特征值,

$$\|A\|_\infty = \max_i \sum_{j=1}^n |a_{ij}|.$$

可以证明, 当 A 是实对称矩阵时, $\|A\|_2 = \max_i |\lambda_i(A)|$, 其中 $\lambda_i(A)$ 表示 A 的第 i 个特征值.

对于任意 n 维向量 x 和 n 阶方阵 A , 如果

$$\|Ax\| \leq \|A\| \|x\|$$

成立, 称矩阵范数 $\|A\|$ 与向量范数 $\|x\|$ 相容.

可以证明, $\|A\|_1, \|A\|_2, \|A\|_\infty$ 分别与 $\|x\|_1, \|x\|_2, \|x\|_\infty$ 相容.

1.2.2 有关公式

向量 $x = (x_1 \ x_2 \ \dots \ x_n)$ 与 $y = (y_1 \ y_2 \ \dots \ y_n)$ 的内积是一个实数, 记为 $x \cdot y$, 规定

$$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

定义了内积的向量空间 R^n 称为 n 维欧氏(向量)空间. 欧氏空间中向量 x 的长度为

$$\|x\|_2 = (x_1^2 + x_2^2 + \dots + x_n^2)^{\frac{1}{2}}.$$

两向量 x 与 y 的距离为

$$\|x - y\|_2 = [(x_1 - y_1)^2 + (x_2 - y_2)^2 + \dots + (x_n - y_n)^2]^{\frac{1}{2}}.$$

如果 $x \cdot y = 0$, 称 x 与 y 直交, 记为 $x \perp y$.

可证明以下公式成立:

公式 1.2.1(平行四边形公式) $\|x+y\|_2^2 + \|x-y\|_2^2 = 2\|x\|_2^2 + 2\|y\|_2^2$;

公式 1.2.2(勾股定理) 当 $x \perp y$ 时, $\|x+y\|_2^2 = \|x\|_2^2 + \|y\|_2^2$;

公式 1.2.3(柯西-许瓦茨不等式) $|x \cdot y| \leq \|x\|_2 \|y\|_2$.

为了简单起见, 以后我们将范数符号 $\|\cdot\|_2$ 记为 $\|\cdot\|$.

1.3 凸锥及 Farkas 引理

1.3.1 向量空间

设 V 是一个向量集, R 是实数集. 如果

(i) 对于任意的 $x \in V, y \in V$, 有 $x + y \in V$;

(ii) 对于任意的 $x \in V, k \in R$, 有 $kx \in V$,

就称 V 为一个向量空间. 也就是说向量空间关于向量的加法运算和数乘向量运算是封闭的. 所有 n 维向量的集合是一个向量空间, 记为 R^n . 通常在 R^n 中引入向量的内积运算, 这时 R^n 称为 n 维欧氏空间.

设 $a_1, a_2, \dots, a_m \in R^n$, 可证明集合

$$Z_m = \{x \in R^n \mid a_i \cdot x = 0, i = 1, 2, \dots, m\}$$

和

$$S_m = \{a = k_1 a_1 + k_2 a_2 + \dots + k_m a_m \mid k_1, k_2, \dots, k_m \in R\}$$

是两个向量空间. 这两个空间分别称为向量组 $a_1, a_2, \dots, a_m$ 的零空间和张成的空间.

设 V 和 S 是两个向量空间, 如果 S 是 V 的子集, 则称 S 为 V 的子空间. 上面定义的两个空间 Z_m 和 S_m 都是 R^n 的子空间.

设集合 V 是 R^n 的子集, 向量 $z \in R^n$. 如果 z 与 V 中的一切向量直交, 就说 z 与 V 直交, 记作 $z \perp V$. 设 V_1 和 V_2 都是 R^n 的子集. 如果对于任何 $x \in V_1$ 和任何 $y \in V_2$, 均有 $x \perp y$, 就说 V_1 与 V_2 直交, 记作 $V_1 \perp V_2$.

考虑向量组 $a_1, a_2, \dots, a_m$ 的零空间 Z_m 和张成的空间 S_m . 对于任何 $x \in Z_m$, 任何 $a \in S_m$, 由于 a 可表示为 $a = k_1 a_1 + k_2 a_2 + \dots + k_m a_m$, 所以 $a \cdot x = k_1 a_1 \cdot x + k_2 a_2 \cdot x + \dots + k_m a_m \cdot x = 0$, 即 Z_m 和 S_m 是 R^n 的两个互相直交的子空间. 如果 x 既属于 Z_m 也属于 S_m , 那么, $x \cdot x = 0$. 所以 Z_m 和 S_m 的交集等于 $\{0\}$.

设 V 是一个向量空间, $a_1, a_2, \dots, a_l$ 是其中的 l 个线性无关的向量. 如果 V 中每个向量可以表示为 $a_1, a_2, \dots, a_l$ 的线性组合, 则称 $a_1, a_2, \dots, a_l$ 为 V 的一个基, 数字 l 称为 V 的维数.

维数为 l 的向量空间称为 l 维向量空间. 显然 R^n 是 n 维向量空间. 向量组 $a_1, a_2, \dots, a_m$ 张成的空间 S_m 的维数等于向量组 $a_1, a_2, \dots, a_m$ 的秩. 还可以证明, 如果 n 维向量组 $a_1, a_2, \dots, a_m$ 的秩为 l , 则此向量组的零空间 Z_m 的维数为 $n - l$.

设 V 是 R^n 的子空间, $x \in V, z \in R^n$. 如果 $z \cdot x$ 与 V 的任何向量直交, 就称

x 是 z 在 V 上的直交投影. 容易证明

(i) $\|x\| \leq \|z\|$ 其中等号仅当 $x = z$ 时成立;

(ii) $z \cdot x = \|x\|^2$.

设 x 和 y 分别是向量 z 在向量组 $a_1, a_2, \dots, a_m$ 的零空间 Z_m 和张成的空间 S_m 上的直交投影, 由于 Z_m 和 S_m 直交, 所以

$$z = x + y.$$

1.3.2 仿射集

设 x 和 y 是 R^n 中两个不同的点, $t \in R$, 由 $(1-t)x + ty$ 表示的点的集合称为通过 x 和 y 的直线. 设 V 是 R^n 的子集, 如果对于任意的 $x, y \in V, t \in R$, 有 $(1-t)x + ty \in V$, 则称 V 为仿射集 (affine set). 也就是说, 仿射集包含通过该集合中任意两点的直线.

规定空集 ϕ 是一个仿射集. 而 R^n 的子空间是过原点的仿射集. 设 $a_1, a_2, \dots, a_m$ 是 n 维向量, $b_1, b_2, \dots, b_m$ 是实数. 则方程组 $a_i \cdot x = b_i, i = 1, 2, \dots, m$ 的所有解的集合

$$A_m = \{x \in R^n \mid a_i \cdot x = b_i, i = 1, 2, \dots, m\}$$

是个仿射集. 作为 A_m 的一个特殊情形, $H = \{x \in R^n \mid a_1 \cdot x = b_1\}$ 称为 R^n 中一个超平面, 通常说成超平面 $a_1 \cdot x = b_1$, a_1 称为法向量.

设 V_1 和 V_2 是两个仿射集. 如果存在 $\bar{x} \in R^n$, 使得对于任何 $x \in V_1$ 有 $\bar{x} + x \in V_2$, 就称 V_1 和 V_2 平行.

设 $\bar{x}$ 属于 $A_m = \{x \in R^n \mid a_i \cdot x = b_i, i = 1, 2, \dots, m\}$. 由于对于任何属于 $Z_m = \{x \in R^n \mid a_i \cdot x = 0, i = 1, 2, \dots, m\}$ 的 x , 均有 $\bar{x} + x \in A_m$, 所以 A_m 与 Z_m 平行.

显然, 包含原点的仿射集是一个向量空间.

一个非空仿射集的维数定义为与其平行的子空间的维数. 仅含一个点的集, 一条直线, R^3 中的一个平面分别是 0, 1, 2 维的仿射集.

设 n 维向量组 $a_1, a_2, \dots, a_m$ 的秩为 l , 则上述定义的仿射集 A_m 的维数等于 $a_1, a_2, \dots, a_m$ 的零空间 Z_m 的维数等于 $n - l$.

1.3.3 凸集

设 $x^{(1)}, x^{(2)}, \dots, x^{(m)} \in R^n, k_1, k_2, \dots, k_m \in R$. 如果 $k_1, k_2, \dots, k_m \geq 0$ 且

$k_1 + k_2 + \dots + k_m = 1$, 则称

$$x = k_1 x^{(1)} + k_2 x^{(2)} + \dots + k_m x^{(m)}$$

为 $x^{(1)}, x^{(2)}, \dots, x^{(m)}$ 的凸组合. 两个不同点 $x^{(1)}$ 和 $x^{(2)}$ 的所有凸组合 $\{(1-t)x^{(1)} + tx^{(2)} \mid 0 \leq t \leq 1\}$ 是一条以 $x^{(1)}$ 和 $x^{(2)}$ 为端点的线段. 设 C 是 R^n 的一个子集, 如果 C 中任意两点的所有凸组合属于 C , 即对于任意的 $x^{(1)} \in C, x^{(2)} \in C, 0 \leq t \leq 1$, 有

$$(1-t)x^{(1)} + tx^{(2)} \in C,$$

则称 C 为凸集. R^n 本身是凸集, 仿射集是凸集, 一个点的集是凸集, 通常还规定空集 $\emptyset$ 是凸集.

设 $a_1 \in R^n, b_1 \in R$. 则不等式 $a_1 \cdot x \geq b_1$ 的所有解的集合 $H_+ = \{x \in R^n \mid a_1 \cdot x \geq b_1\}$ 和不等式 $a_1 \cdot x \leq b_1$ 的所有解的集合 $H_- = \{x \in R^n \mid a_1 \cdot x \leq b_1\}$ 是凸集, 这种凸集称为半空间, 超平面 $a_1 \cdot x = b_1$ 称为其边界.

考虑等式与不等式组

$$\left. \begin{aligned} a_i \cdot x &= b_i, \quad i=1, 2, \dots, l, \\ a_i \cdot x &\geq b_i, \quad i=l+1, l+2, \dots, m. \end{aligned} \right\} \quad (1.3.1)$$

其中, $0 \leq l \leq m, a_1, a_2, \dots, a_m \in R^n, b_1, b_2, \dots, b_m \in R$. 容易证明(1.3.1)的解集是凸集, 它是 l 个超平面和 $m-l$ 个半空间的交集, 称为多面体.

1.3.4 凸锥

设 S 是 R^n 的子集, $\bar{x} \in S$. 如果对于任意的 $x \in S$ 和任何非负实数 t , 有

$$\bar{x} + (x - \bar{x})t \in S,$$

就称 S 为一个锥, $\bar{x}$ 称为这个锥的顶点. 如果 S 还是个凸集, 则称为凸锥.

如果(1.3.1)相应的方程组

$$a_i \cdot x = b_i, \quad i=1, 2, \dots, m$$

有解, 容易证明(1.3.1)的解集是个凸锥, 并且以上方程组的任一解是其顶点. 为叙述方便起见, 有时将(1.3.1)称为一个锥.

如果(1.3.1)的系数向量线性无关, 或者其所有右端常数项为 0, 那么它相应的方程组有解, 因此这时(1.3.1)是个锥. 令

$$S = \{x = k_1 a_1 + \dots + k_l a_l + k_{l+1} a_{l+1} + \dots + k_m a_m \mid k_1, k_2, \dots, k_m \in R \text{ 且 } k_{l+1}, \dots, k_m$$

$\geq 0\}$,

则 S 是个凸锥, $x=0$ 是其顶点.

1.3.5 Farkas 引理

引理 1.3.1 设 $a_1, a_2, \dots, a_m$ 和 c 是 n 维向量, $S = \{x = k_1 a_1 + k_2 a_2 + \dots + k_m a_m \mid k_1, k_2, \dots, k_m \geq 0\}$, c 不属于 S , $\bar{x}$ 是 S 内距离 c 最近的点, 那么, $\bar{x} \cdot (\bar{x} - c) = 0$.

证 若 $\bar{x} = 0$, 此结论成立, 所以下面设 $\bar{x} \neq 0$. 按 $\bar{x}$ 的定义, 对于任何 $x \in S$, 均有

$$\|\bar{x} - c\| \leq \|x - c\|.$$

由于 S 是以原点为顶点的凸锥, 对于任何非负实数 t , 均有 $t\bar{x} \in S$, 所以

$$\|\bar{x} - c\| \leq \|t\bar{x} - c\|. \quad (1.3.2)$$

考虑关于 t 的二次多项式

$$f(t) = \|t\bar{x} - c\|^2 = \|\bar{x}\|^2 t^2 - 2c \cdot \bar{x}t + \|c\|^2,$$

令其导数

$$f'(t) = 2\|\bar{x}\|^2 t - 2c \cdot \bar{x} = 0,$$

得

$$t = \frac{c \cdot \bar{x}}{\|\bar{x}\|^2}.$$

这是 $f(t)$ 在 $(-\infty, +\infty)$ 内的极小值点. 这里的 $c \cdot \bar{x}$ 应该是个正数. 因为如果 $c \cdot \bar{x} \leq 0$, 则从

$$\|\bar{x} - c\|^2 = \|\bar{x}\|^2 + \|c\|^2 - 2c \cdot \bar{x}$$

可以看出 $\|\bar{x} - c\| > \|0 - c\|$, 与 $\bar{x}$ 的定义矛盾. 另外由 (1.3.2) 可见, $f(t)$ 在

区间 $[0, +\infty)$ 内的极小值点为 1. 所以 $\frac{c \cdot \bar{x}}{\|\bar{x}\|^2} = 1$, 即 $\bar{x} \cdot (\bar{x} - c) = 0$.

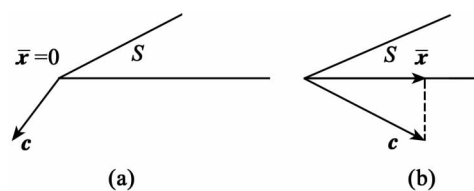


图 1.1 凸锥与其外一点的关系

引理 1.3.1 的几何意义如图 1.1 所示,图 1.1(a)表示 $\bar{x} = 0$ 的情况,图 1.1(b)表示 $\bar{x} \neq 0$ 的情况.

以下定理通常称为 Farkas 引理,它是最优化理论与方法的基础.

定理 1.3.1 设 a_i 和 c 是给定的 n 维行向量, x 是 n 维未知列向量, y_i 是未知标量,则以下两组中有一个且仅有一个有解:

$$\text{组 1: } a_i x \geq 0, i = 1, 2, \dots, m, cx < 0;$$

$$\text{组 2: } y_1 a_1 + y_2 a_2 + \dots + y_m a_m = c, y_1, y_2, \dots, y_m \geq 0.$$

证 我们将证明:若组 2 有解,则组 1 无解;若组 2 无解,则组 1 有解.

(i) 设存在实数 $y_1, y_2, \dots, y_m \geq 0$, 使得 $y_1 a_1 + y_2 a_2 + \dots + y_m a_m = c$, 那么对于组 1 前 m 式的任一解 x 有

$$cx = y_1 a_1 x + y_2 a_2 x + \dots + y_m a_m x \geq 0,$$

所以组 1 无解.

(ii) 记 $S = \{x = k_1 a_1 + k_2 a_2 + \dots + k_m a_m \mid k_1, k_2, \dots, k_m \geq 0\}$. 组 2 无解即 c 不属于 S . 用 $\bar{x}$ 表示 S 内距离 c 最近的点, 则由引理 1.3.1 有

$$\bar{x} \cdot (\bar{x} - c) = 0. \quad (1.3.3)$$

于是

$$c \cdot (\bar{x} - c) = (\bar{x} + c - \bar{x}) \cdot (\bar{x} - c) = \bar{x} \cdot (\bar{x} - c) - \|\bar{x} - c\|^2 = -\|\bar{x} - c\|^2 < 0.$$

这说明 $\bar{x} - c$ 满足组 1 的最后一式. 另外, 由于 S 是凸集, 对于任何 $x \in S$ 和 $t \in (0, 1)$ 有 $\bar{x} + t(x - \bar{x}) \in S$. 但 $\bar{x}$ 是 S 中距 c 最近的点, 所以

$$\begin{aligned} \|\bar{x} - c\|^2 &\leq \|\bar{x} + t(x - \bar{x}) - c\|^2 \\ &= t^2 \|x - \bar{x}\|^2 + 2t(x - \bar{x}) \cdot (\bar{x} - c) + \|\bar{x} - c\|^2. \end{aligned}$$

从上式两边去掉 $\|\bar{x} - c\|^2$, 然后除以 t , 得

$$t \|x - \bar{x}\|^2 + 2(x - \bar{x}) \cdot (\bar{x} - c) \geq 0.$$

令 t 趋于 0 得 $(x - \bar{x}) \cdot (\bar{x} - c) \geq 0$, 再由式(1.3.3)有

$$x \cdot (\bar{x} - c) \geq 0, x \in S.$$

因为 $a_i \in S (i = 1, 2, \dots, m)$, 所以

$$a_i \cdot (\bar{x} - c) \geq 0, i = 1, 2, \dots, m,$$

即 $\bar{x} - c$ 是组 1 的解. ■

此定理的几何意义见图 1.2, 假设向量 $a_2, \dots, a_{m-1}$ 都夹在 a_1 和 a_m 之间. 图 1.2(a)表示 c 是 $a_1, a_2, \dots, a_m$ 的非负线性组合, 组 1 的解集是空集; 图 1.2(b)表示 c 不是 $a_1, a_2, \dots, a_m$ 的非负线性组合, 组 1 的解集如图 1.2(b)阴影区域所示.

推论 1.3.1 设 a_i 和 c 是 n 维行向量, x 是 n 维未知列向量, b_i 和 y_i (未知) 是标量. 设 $\bar{x}$ 是方程组 $a_i x = b_i, i = 1, 2, \dots, m$ 的一个解, 则以下两组中有一个且

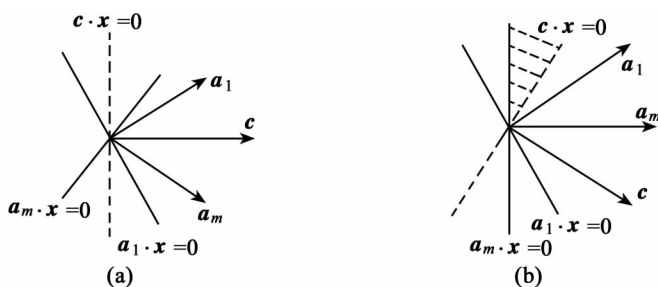


图 1.2 Farkas 引理的几何意义

仅有一个有解：

$$\begin{aligned} \text{组 1 : } & a_i x = b_i, i = 1, 2, \dots, l, \\ & a_i x \geq b_i, i = l+1, l+2, \dots, m, \\ & cx < c\bar{x}; \end{aligned}$$

$$\begin{aligned} \text{组 2 : } & y_1 a_1 + y_2 a_2 + \dots + y_m a_m = c, \\ & y_{l+1} \leq 0, y_{l+2} \leq 0, \dots, y_m \leq 0. \end{aligned}$$

或者等价地讲，等式与不等式组

$$\begin{aligned} & a_i x = b_i, i = 1, 2, \dots, l, \\ & a_i x \geq b_i, i = l+1, l+2, \dots, m \end{aligned}$$

的所有解满足 $cx \geq c\bar{x}$ 的必要充分条件是组 2 有解。

证 只需将组 1 化为以下形式，再应用定理 1.3.1：

$$\begin{aligned} & a_i(x - \bar{x}) \geq 0, -a_i(x - \bar{x}) \geq 0, i = 1, 2, \dots, l, \\ & a_i(x - \bar{x}) \geq 0, i = l+1, l+2, \dots, m, \\ & c(x - \bar{x}) < 0. \end{aligned}$$

1.4 约束最优化及 Kuhn-Tucker 条件

1.4.1 基础知识

n 元实值函数 $f(x) = f(x_1, x_2, \dots, x_n)$ 在 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ 处的 n 个一阶偏导数构成的向量称为 $f(x)$ 在 $\bar{x}$ 处的梯度，记为 $\nabla_x f(\bar{x})$ 或 $\nabla f(\bar{x})$ ，即

$$\nabla f(\bar{x}) = \left(\frac{\partial f(\bar{x})}{\partial x_1}, \frac{\partial f(\bar{x})}{\partial x_2}, \dots, \frac{\partial f(\bar{x})}{\partial x_n} \right)^T.$$

函数 $f(x)$ 在 $\bar{x}$ 处的海色 (Hesse) 矩阵 $\nabla^2 f(\bar{x})$ 是指其二阶偏导数构成的下述对称

矩阵：

$$\nabla^2 f(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f(\bar{x})}{\partial x_1^2} & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\bar{x})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_1} & \frac{\partial^2 f(\bar{x})}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(\bar{x})}{\partial x_n^2} \end{pmatrix}.$$

记 $f_1(x) = cx$, $f_2(x) = \frac{1}{2}x^T Ax$, 其中 c 是 n 维行向量, x 是 n 维列向量, A 是 n 阶对称矩阵. 容易证明, $\nabla f_1(x) = c^T$, $\nabla f_2(x) = Ax$, $\nabla^2 f_2(x) = A$.

对于更为一般的非线性函数, 通常用到以下两种形式的泰勒 (Taylor) 展开式进行近似计算:

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + o(\|x - \bar{x}\|)$$

和

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})^T (x - \bar{x}) + \frac{1}{2}(x - \bar{x})^T \nabla^2 f(\bar{x})(x - \bar{x}) + o(\|x - \bar{x}\|^2).$$

设 $f(x)$ 是定义在 n 维欧氏空间 R^n 的某个凸集 D 上的函数. 如果对于 D 中的任意两点 $x^{(1)}$ 和 $x^{(2)}$, 以及任何实数 $\alpha (0 < \alpha < 1)$, 有

$$f(\alpha x^{(1)} + (1 - \alpha)x^{(2)}) \leq \alpha f(x^{(1)}) + (1 - \alpha)f(x^{(2)}), \quad (1.4.1)$$

就称 $f(x)$ 是定义在 D 上的凸函数. 如果 $x^{(1)} \neq x^{(2)}$ 时总有

$$f(\alpha x^{(1)} + (1 - \alpha)x^{(2)}) < \alpha f(x^{(1)}) + (1 - \alpha)f(x^{(2)}), \quad (1.4.2)$$

则称 $f(x)$ 是定义在 D 上的严格凸函数.

将 (1.4.1) 和 (1.4.2) 两式中的不等号反向, 则得到凹函数和严格凹函数的定义. 显然, 若 $f(x)$ 是凸函数 (严格凸函数), 则 $-f(x)$ 是凹函数 (严格凹函数).

容易证明, 线性函数既是凸函数也是凹函数; 若 A 是 n 阶 (对称) 正半定矩阵, c 是 n 维行向量, x 是 n 维未知列向量, 则二次函数 $f(x) = x^T Ax + cx$ 是凸函数.

关于凸 (凹) 函数有以下结论:

结论 1.4.1 若干个凸函数的非负线性组合是凸函数.

结论 1.4.2 若 $f(x)$ 是凹函数, 则 $f(x) \geq 0$ 的解集是凸集.

结论 1.4.3 设 $f(x)$ 在 R^n 的某个开凸集 D 上具有一阶连续偏导数, 则 $f(x)$ 为 D 上凸函数的必要充分条件是: 对于 D 中的任何两个不同点 $x^{(1)}$ 和 $x^{(2)}$

有

$$f(x^{(2)}) - f(x^{(1)}) \geq \nabla f(x^{(1)})^T (x^{(2)} - x^{(1)}).$$

结论 1.4.4 设 $f(x)$ 在 R^n 的某个开凸集 D 上具有二阶连续偏导数, 则 $f(x)$ 为 D 上凸函数的必要充分条件是, $f(x)$ 的海色矩阵 $\nabla^2 f(x)$ 在 D 上处处正半定.

设 $f_i(x)$ 是 n 元实值函数 $i=1, 2, \dots, m$. $f(x) = (f_1(x), f_2(x), \dots, f_m(x))^T$ 称为 m 维向量值函数. 由 $f_1(x), f_2(x), \dots, f_m(x)$ 的一阶偏导数构成的 $m \times n$ 矩阵称为雅可比 (Jacobi) 矩阵, 记为 $Df(x)$, 即

$$Df(x) = \begin{pmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \frac{\partial f_2(x)}{\partial x_1} & \frac{\partial f_2(x)}{\partial x_2} & \cdots & \frac{\partial f_2(x)}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m(x)}{\partial x_1} & \frac{\partial f_m(x)}{\partial x_2} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{pmatrix} = \begin{pmatrix} \nabla f_1(x)^T \\ \nabla f_2(x)^T \\ \vdots \\ \nabla f_m(x)^T \end{pmatrix}.$$

由 $f_1(x), f_2(x), \dots, f_m(x)$ 的 m 个海色矩阵构成的 $mn \times n$ 矩阵称为 $f(x)$ 的二阶导数矩阵, 记为 $D^2 f(x)$, 即

$$D^2 f(x) = \begin{pmatrix} \nabla^2 f_1(x) \\ \nabla^2 f_2(x) \\ \vdots \\ \nabla^2 f_m(x) \end{pmatrix}.$$

定理 1.4.1 设 $f(x) = (f_1(x), f_2(x), \dots, f_m(x))^T$, 其中 $f_i(x)$ 在 n 维空间的一个开凸集 D 上连续可微, 那么, 对于 D 中任何两点 x 和 $x+s$ 都有

$$\|f(x+s) - f(x)\| \leq \sup_{0 \leq \xi \leq 1} \|Df(x+\xi s)\| \cdot \|s\|.$$

定理 1.4.2 设 $f(x) = (f_1(x), f_2(x), \dots, f_m(x))^T$, 其中 $f_i(x)$ 在 n 维空间的一个开凸集 D 上连续可微, 那么, 对于 D 中任何三点 $x, x+s$ 和 y 都有

$$\|f(x+s) - f(x) - Df(y)s\| \leq \sup_{0 \leq \xi \leq 1} \|Df(x+\xi s) - Df(y)\| \cdot \|s\|.$$

定理 1.4.3 设 $f(x) = (f_1(x), f_2(x), \dots, f_m(x))^T$, 其中 $f_i(x)$ 在 n 维空间的一个开凸集 D 上二阶连续可微, 那么, 对于 D 中任何两点 x 和 $x+s$ 都有

$$\|f(x+s) - f(x) - Df(x)s\| \leq \sup_{0 \leq \xi \leq 1} \|D^2 f(x+\xi s)\| \cdot \|s\|^2.$$

1.4.2 约束最优化问题的 Kuhn-Tucker 条件

设 $f(x)$ 为定义在 n 维欧氏空间 R^n 的某个子集 D 上的 n 元实值函数, $x^* \in D$. 若存在正数 δ , 使得 D 中所有满足 $\|x - x^*\| < \delta$ 的 x 均满足 $f(x) \geq f(x^*)$, 则称 x^* 为 $f(x)$ 在 D 上的局部极小点, $f(x^*)$ 为局部极小值. 若以上等号仅当 $x = x^*$ 时成立, 则称 x^* 为 $f(x)$ 在 D 上的严格局部极小点, $f(x^*)$ 为严格局部极小值. 若对于所有 $x \in D$ 都有 $f(x) \geq f(x^*)$, 则称 x^* 为 $f(x)$ 在 D 上的全局极小点, $f(x^*)$ 为全局极小值. 若以上等号仅当 $x = x^*$ 时成立, 则称 x^* 为 $f(x)$ 在 D 上的严格全局极小点, $f(x^*)$ 为严格全局极小值.

将上述的不等号反向, 即可得到相应的极大点和极大值的定义. 函数的极小点和极大点统称为极值点. 显然, 凸函数的极值点一定是极小点. 由于函数 $f(x)$ 的极大点(值)相当 $-f(x)$ 的极小点(值), 下面仅讨论极小问题. 全局极小点又称为最优解, 全局极小值称为最优目标函数值.

设 D 是 R^n 的一个给定的子集, $f(x)$ 为 D 上的实值函数. 在 D 中找一点使 $f(x)$ 取得局部或全局极小值, 这个问题称为约束最优化问题, 是数学规划的重要课题之一. D 称为可行域, D 中的点称为可行解(点), 使 $f(x)$ 取得最小值的可行解称为最优解. 约束最优化问题通常表示为

$$\left. \begin{aligned} \min f(x), \\ \text{s. t. } g_i(x) = 0, i = 1, 2, \dots, l, \\ g_i(x) \geq 0, i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (1.4.3)$$

其中 $\min$ 和 s. t. 分别是 minimize 和 subject to 的缩写, $f(x)$ 称为目标函数, 每个等式和不等式称为约束条件.

如果模型(1.4.3)中的 $f(x)$ 和所有 $g_i(x)$ 是线性函数, 称模型(1.4.3)为线性规划, 否则称为非线性规划. 若目标函数为二次函数, 约束条件为线性等式或不等式, 则称为二次规划. 二次规划是最简单的非线性规划.

对于某个 $x^* \in R^n$ 和 $i \in \{1, 2, \dots, l\}$, 如果 $g_i(x^*) = 0$, 则称 $g_i(x) = 0$ 为关于 x^* 的紧约束, 否则称为松约束. 同样, 对于某个 $i \in \{l+1, l+2, \dots, m\}$, 如果 $g_i(x^*) = 0$, 则称 $g_i(x) \geq 0$ 为关于 x^* 的紧约束, 否则称为松约束. 若 x^* 是可行解, 则所有等式约束是紧约束, 不等式约束可能是紧约束, 也可能不是.

设模型(1.4.3)中的所有函数一阶连续可微, x^* 是其可行解, 并记不等式约束中关于 x^* 的所有紧约束的指标集为 I^* , 即 $I^* = \{i \mid g_i(x^*) = 0, i \in \{l+1, l+2, \dots, m\}\}$, 则在一定条件下可以证明, x^* 是 $f(x)$ 的局部极小点的必要条

件是存在标量 λ_i 使得

$$\nabla f(x^*) = \sum_{i=1}^l \lambda_i \nabla g_i(x^*) + \sum_{i \in I^*} \lambda_i \nabla g_i(x^*), \lambda_i \geq 0, i \in I^*.$$

但 x^* 通常未知,所以不知道哪些不等式约束是紧约束哪些不是.为把所有约束考虑在内,通常将上式写为

$$\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla g_i(x^*) \quad \lambda_i \geq 0, \lambda_i g_i(x^*) = 0, i = 1 + 1 \quad 1 + 2 \quad \dots \quad \mu.$$

这组式子便是有名的 Kuhn-Tucker 条件(简称 KT 条件).由于 Karush(1939)也给出了类似的结论,有的文献也称其为 Karush-Kuhn-Tucker 条件,或 KKT 条件.上式中的 $\lambda_i g_i(x^*) = 0 (i = 1 + 1 \quad 1 + 2 \quad \dots \quad \mu)$ 称为互补松弛条件.为了明显反映 x^* 是可行解,有时将式子

$$g_i(x^*) = 0, i = 1 \quad 2 \quad \dots \quad l; g_i(x^*) \geq 0, i = 1 + 1 \quad 1 + 2 \quad \dots \quad \mu$$

也写在一起,统称为 KT 条件.满足 KT 条件的 (x^*, λ^*) 称为 KT 对, x^* 称为 KT 点,其中 $x^* = (x_1^* \quad x_2^* \quad \dots \quad x_n^*)^T$, $\lambda^* = (\lambda_1^* \quad \lambda_2^* \quad \dots \quad \lambda_m^*)^T$.

下面仅就线性约束证明这一结果.

定理 1.4.4 考虑以下数学规划

$$\begin{aligned} \min & f(x), \\ \text{s.t.} & a_i x = b_i \quad i = 1 \quad 2 \quad \dots \quad l, \\ & a_i x \geq b_i, \quad i = 1 + 1 \quad 1 + 2 \quad \dots \quad \mu. \end{aligned}$$

其中 $f(x)$ 在包含可行域的某个开集上一阶连续可微, x 是 n 维未知列向量, a_i 是 n 维行向量, b_i 是标量, $i = 1 \quad 2 \quad \dots \quad \mu$.设 x^* 是可行解,则 x^* 是局部极小点的必要条件是存在标量 λ_i 使得

$$\nabla f(x^*) = \sum_{i=1}^l \lambda_i a_i^T + \sum_{i \in I^*} \lambda_i a_i^T, \lambda_i \geq 0 (i \in I^*).$$

其中, $I^* = \{i \mid a_i x^* = b_i \quad i \in \{1 + 1 \quad 1 + 2 \quad \dots \quad \mu\}\}$.

证 由于 x^* 是可行解,它满足 $a_i x = b_i \quad i = 1 \quad 2 \quad \dots \quad l$,所以 x^* 是方程组

$$a_i x = b_i \quad i \in \{1 \quad 2 \quad \dots \quad l\} \cup I^*$$

的解,因而集合 $C = \{x \mid a_i x = b_i \quad i = 1 \quad 2 \quad \dots \quad l; a_i x \geq b_i \quad i \in I^*\}$ 是个凸锥, x^* 是其顶点.设 x 是 C 内任一点,则对于任何正数 t ,有 $x^* + t(x - x^*) \in C$.当 t 足够小时, $x^* + t(x - x^*)$ 满足所有关于 x^* 的松约束,因而是可行解,这时 $f(x^* + t(x - x^*)) \geq f(x^*)$.记

$$\varphi(t) = f(x^* + t(x - x^*)).$$

由于 $\varphi(0) = f(x^*)$, 所以 $\frac{\varphi(t) - \varphi(0)}{t} \geq 0$. 令 t 趋于 0 得

$$\varphi'(0) = \nabla f(x^*)^T (x - x^*) \geq 0,$$

即 C 内任何点 x 满足

$$\nabla f(x^*)^T x \geq \nabla f(x^*)^T x^*.$$

根据推论 1.3.1, 存在实数 λ_i 使得

$$\nabla f(x^*) = \sum_{i=1}^l \lambda_i a_i^T + \sum_{i \in I^*} \lambda_i a_i^T, \lambda_i \geq 0 (i \in I^*).$$

1.5 约束最优化的其他一些定理

1.5.1 基本概念和定理

考虑约束优化问题模型(1.4.3). 设 $x^*, s \in R^n, s \neq 0$. 如果存在一个正数 δ , 使得

$$f(x^* + ts) < f(x^*), 0 < t < \delta,$$

称 s 为 $f(x)$ 在 x^* 处的下降方向. 设 t 是充分小的正数, 由泰勒展开式

$$f(x^* + ts) = f(x^*) + t \nabla f(x^*)^T s + o(t)$$

可见, 如果向量 s 使得 $\nabla f(x^*)^T s < 0$, 即 s 与 $f(x)$ 在 x^* 处梯度 $\nabla f(x^*)$ 的夹角大于 90° , 则 s 是 x^* 处的下降方向.

设 x^* 是可行域 D 内一点, s 是一个非零的 n 维向量. 如果存在一个正数 δ , 使得

$$x^* + ts \in D, 0 \leq t \leq \delta,$$

则称 s 为 x^* 处的可行方向. 如果 s 既是下降方向又是可行方向, 则称 s 为可行下降方向.

引理 1.5.1 考虑模型(1.4.3), 设 x^* 是可行域 D 内一点, $f(x)$ 和 $g_i(x)$ ($i=1, 2, \dots, m$) 在 x^* 处可微. 如果 x^* 不是严格局部极小点, 那么存在一个非零向量 s 满足

$$\left. \begin{aligned} \nabla g_i(x^*)^T s &= 0, \quad i \in E = \{1, 2, \dots, l\}, \\ \nabla g_i(x^*)^T s &\geq 0, \quad i \in I^*. \end{aligned} \right\} \quad (1.5.1)$$

并且使得

$$\nabla f(x^*)^T s \leq 0.$$

其中 I^* 是关于 x^* 的紧不等式约束的指标集, 即

$$I^* = \{i \mid g_i(x^*) = 0, i \in \{1, 2, \dots, m\}\}.$$

证 由于 x^* 不是严格局部极小点, 存在一个收敛于 x^* 的可行点列 $\{x^{(k)} \mid k=1, 2, \dots\}$ ($x^{(k)} \neq x^*$) 使

$$f(x^{(k)}) \leq f(x^*), \quad k=1, 2, \dots$$

因为 $\frac{x^{(k)} - x^*}{\|x^{(k)} - x^*\|}$ ($k=1, 2, \dots$) 在以 x^* 为中心的单位球面上, 必收敛于其内一点 s , 即

$$\lim_{k \rightarrow \infty} \frac{x^{(k)} - x^*}{\|x^{(k)} - x^*\|} = s.$$

由于 $x^{(k)}$ 可行, $g_i(x^*) = 0, i \in E \cup I^*$, 所以

$$g_i(x^{(k)}) - g_i(x^*) = 0, \quad i \in E,$$

$$g_i(x^{(k)}) - g_i(x^*) \geq 0, \quad i \in I^*.$$

又因

$$\lim_{k \rightarrow \infty} \frac{g_i(x^{(k)}) - g_i(x^*)}{\|x^{(k)} - x^*\|} = \lim_{k \rightarrow \infty} \frac{\nabla g_i(x^*)^T (x^{(k)} - x^*)}{\|x^{(k)} - x^*\|} = \nabla g_i(x^*)^T s,$$

所以 s 满足

$$\nabla g_i(x^*)^T s = 0, \quad i \in E,$$

$$\nabla g_i(x^*)^T s \geq 0, \quad i \in I^*.$$

类似可证明

$$\nabla f(x^*)^T s \leq 0. \quad \blacksquare$$

定理 1.5.1 考虑模型(1.4.3), 设 x^* 是可行域 D 内一点, $f(x)$ 和 $g_i(x)$ ($i=1, 2, \dots, m$) 在 x^* 处可微. 如果齐次线性等式与不等式组(1.5.1)的任何非零解 s 都使得

$$\nabla f(x^*)^T s > 0,$$

则 x^* 是严格局部极小点.

证 用反证法. 设 x^* 不是严格局部极小点, 那么按照引理 1.5.1, 存在非零向量 s 满足(1.5.1)并使 $\nabla f(x^*)^T s \leq 0$. 这与定理的假设矛盾. $\blacksquare$

模型(1.4.3)的拉格朗日(Lagrange)函数是指目标函数和约束函数的如下线性组合:

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i g_i(x).$$

$L(x, \lambda)$ 关于 x 的梯度记为 $\nabla_x L(x, \lambda)$, 关于 x 的海色矩阵记为 $\nabla_x^2 L(x, \lambda)$. 在

(x^*, λ^*) 对处, KT 条件的第一个式子可写为 $\nabla_x L(x^*, \lambda^*) = 0$.

定理 1.5.2 考虑模型(1.4.3), 设 $f(x)$ 和 $g_i(x) (i=1, 2, \dots, m)$ 在 x^* 处二阶连续可微, (x^*, λ^*) 是 KT 对. 如果齐次线性等式与不等式组

$$\left. \begin{aligned} \nabla g_i(x^*)^T s &= 0, \quad i \in E, \text{ 或者 } i \in I^* \text{ 且 } \lambda_i^* > 0, \\ \nabla g_i(x^*)^T s &\geq 0, \quad i \in I^* \text{ 且 } \lambda_i^* = 0 \end{aligned} \right\} \quad (1.5.2)$$

的任何非零解 s 都使得

$$s^T \nabla_x^2 L(x^*, \lambda^*) s > 0,$$

则 x^* 是严格局部极小点.

证 用反证法. 设 x^* 不是严格局部极小点, 那么存在一个收敛于 x^* 的可行点列 $\{x^{(k)} \mid k=1, 2, \dots\} (x^{(k)} \neq x^*)$ 使

$$f(x^{(k)}) \leq f(x^*), \quad k=1, 2, \dots,$$

而 $\frac{x^{(k)} - x^*}{\|x^{(k)} - x^*\|} (k=1, 2, \dots)$ 的极限 s 满足齐次等式与不等式组(1.5.1) 并使

$$\nabla f(x^*)^T s \leq 0.$$

(1.5.2) 的解集是(1.5.1)解集的子集. 由于

$$\nabla f(x^*) = \sum_{i \in E \cup I^*} \lambda_i^* \nabla g_i(x^*),$$

所以

$$0 \geq \nabla f(x^*)^T s = \sum_{i \in E \cup I^*} \lambda_i^* \nabla g_i(x^*)^T s = \sum_{i \in I^*} \lambda_i^* \nabla g_i(x^*)^T s. \quad (1.5.3)$$

上式右边求和符号后的各项非负. 若存在 $i \in I^*$ 使得

$$\lambda_i^* > 0 \text{ 且 } \nabla g_i(x^*)^T s > 0,$$

式(1.5.3)的右边将是正数, 矛盾. 所以 s 一定满足(1.5.2). 另外,

$$\begin{aligned} & L(x^{(k)}, \lambda^*) - L(x^*, \lambda^*) \\ &= \left(f(x^{(k)}) - \sum_{i=1}^m \lambda_i^* g_i(x^{(k)}) \right) - \left(f(x^*) - \sum_{i=1}^m \lambda_i^* g_i(x^*) \right) \\ &= f(x^{(k)}) - f(x^*) - \sum_{i \in I^*} \lambda_i^* (g_i(x^{(k)}) - g_i(x^*)) \\ &\leq f(x^{(k)}) - f(x^*) \leq 0. \end{aligned}$$

由于 $\nabla_x L(x^*, \lambda^*) = 0$ 利用泰勒展开式有

$$0 \geq \frac{L(x^{(k)}, \lambda^*) - L(x^*, \lambda^*)}{\|x^{(k)} - x^*\|^2}$$

$$= \frac{1}{2} \frac{(\mathbf{x}^{(k)} - \mathbf{x}^*)^T}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|} \nabla_x^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*) \frac{\mathbf{x}^{(k)} - \mathbf{x}^*}{\|\mathbf{x}^{(k)} - \mathbf{x}^*\|} + o(\|\mathbf{x}^{(k)} - \mathbf{x}^*\|^2).$$

令 $k \rightarrow \infty$ 得 $\mathbf{s}^T \nabla_x^2 L(\mathbf{x}^*, \boldsymbol{\lambda}^*) \mathbf{s} \leq 0$ 与定理假设相矛盾. 所以 $\mathbf{x}^*$ 一定是严格局部极小点. ■

1.5.2 较为特殊的优化问题

下面考虑模型(1.4.3)的一些特殊情形.

二次规划可以表示为以下形式:

$$\left. \begin{aligned} \min f(\mathbf{x}) &= \frac{1}{2} \mathbf{x}^T \mathbf{H} \mathbf{x} + \mathbf{c}^T \mathbf{x}, \\ \text{s. t. } a_i \mathbf{x} &= b_i, \quad i = 1, 2, \dots, l, \\ a_i \mathbf{x} &\geq b_i, \quad i = l+1, l+2, \dots, m, \end{aligned} \right\} \quad (1.5.4)$$

其中, $\mathbf{H}$ 是 n 阶实对称矩阵, $\mathbf{c}$ 是 n 维行向量, $\mathbf{x}$ 是 n 维未知列向量, a_i 是 n 维行向量, b_i 是标量, $i = 1, 2, \dots, m$.

定理 1.5.3 考虑二次规划模型(1.5.4), 设 $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ 是其 KT 对. 那么, $\mathbf{x}^*$ 是局部极小点的必要充分条件是: 齐次线性等式与不等式组

$$\left. \begin{aligned} a_i \mathbf{s} &= 0, \quad i \in E, \text{ 或者 } i \in I^* \text{ 且 } \lambda_i^* > 0, \\ a_i \mathbf{s} &\geq 0, \quad i \in I^* \text{ 且 } \lambda_i^* = 0 \end{aligned} \right\} \quad (1.5.5)$$

的任何解 $\mathbf{s}$ 都使

$$\mathbf{s}^T \mathbf{H} \mathbf{s} \geq 0.$$

其中 $E = \{1, 2, \dots, l\}$, I^* 是关于 $\mathbf{x}^*$ 的紧不等式约束的指标集, 即

$$I^* = \{i \mid a_i \mathbf{x}^* = 0, i \in \{l+1, l+2, \dots, m\}\}.$$

证 (i) 设 $\mathbf{x}^*$ 是局部极小点, $\mathbf{s}$ 是(1.5.5)的任一解. 那么对于充分小的一切正数 t , $\mathbf{x}^* + t\mathbf{s}$ 是可行解, 而且

$$f(\mathbf{x}^* + t\mathbf{s}) \geq f(\mathbf{x}^*), \quad \sum_{i \in E \cup I^*} \lambda_i^* a_i \mathbf{s} = 0.$$

所以

$$\begin{aligned} 0 &\leq f(\mathbf{x}^* + t\mathbf{s}) - f(\mathbf{x}^*) = t \nabla f(\mathbf{x}^*)^T \mathbf{s} + \frac{1}{2} t^2 \mathbf{s}^T \mathbf{H} \mathbf{s} \\ &= t \sum_{i \in E \cup I^*} \lambda_i^* a_i \mathbf{s} + \frac{1}{2} t^2 \mathbf{s}^T \mathbf{H} \mathbf{s} = \frac{1}{2} t^2 \mathbf{s}^T \mathbf{H} \mathbf{s}, \end{aligned}$$

也即 $\mathbf{s}^T \mathbf{H} \mathbf{s} \geq 0$.

(ii)用反证法. 设 x^* 不是局部极小点, 那么 x^* 附近有一个可行的 $\bar{x}$ 使

$$f(\bar{x}) < f(x^*).$$

令 $s = \bar{x} - x^*$. 由于 $\varphi(t) = f(x^* + ts)$ 是 t 的一元二次函数, 对于充分小的一切正数 t , 有

$$f(x^* + ts) < f(x^*),$$

因而

$$\nabla f(x^*)^T s \leq 0.$$

于是

$$0 \geq \nabla f(x^*)^T s = \sum_{i \in E \cup I^*} \lambda_i^* a_i s = \sum_{i \in I^*} \lambda_i^* a_i s. \quad (1.5.6)$$

上式右边求和符号后的各项非负. 若存在 $i \in I^*$ 使得

$$\lambda_i > 0 \text{ 且 } \nabla g_i(x^*)^T s > 0,$$

式(1.5.6)的右边将是正数, 出现矛盾. 所以 s 一定满足(1.5.5), 因而

$$\nabla f(x^*)^T s = \sum_{i \in E \cup I^*} \lambda_i^* a_i s = 0.$$

由此又有

$$0 > f(x^* + ts) - f(x^*) = \frac{1}{2} t^2 s^T H s,$$

即 $s^T H s \geq 0$ 不能成立. 所以 x^* 一定是局部极小点. ■

设 D 是凸集, $f(x)$ 为 D 上的凸函数, 在 D 中找一点使得 $f(x)$ 取得最小值, 这个问题称为凸规划. 在模型(1.4.3)中, 若 $g_i(x)$ ($i=1, 2, \dots, l$) 是线性函数, $g_i(x)$ ($i=l+1, l+2, \dots, m$) 是凹函数, $f(x)$ 为可行域上的凸函数, 则它是凸规划.

定理 1.5.4 凸规划的局部极小点是全局极小点.

证 用反证法, 设凸规划的目标函数为 $f(x)$, x^* 是局部极小点, $\bar{x}$ 是可行解且 $f(\bar{x}) < f(x^*)$. 由于可行域是凸集, 以 x^* 和 $\bar{x}$ 为端点的直线段上的所有点 $x = \alpha x^* + (1 - \alpha)\bar{x}$ ($0 \leq \alpha \leq 1$) 在可行域内. 由于 $f(x)$ 是凸函数,

$$\begin{aligned} f(x) &= f(\alpha x^* + (1 - \alpha)\bar{x}) \leq \alpha f(x^*) + (1 - \alpha)f(\bar{x}) \\ &< \alpha f(x^*) + (1 - \alpha)f(x^*) = f(x^*). \end{aligned}$$

其中 $\alpha < 1$, 但当 α 足够接近 1 时, x 足够接近 x^* , 上式与 x^* 是局部极小点相矛盾. ■

定理 1.5.5 设模型(1.4.3)中的 $f(x)$ 和 $g_i(x)$ ($i=l+1, l+2, \dots, m$) 分别

是包含可行域的开集上的可微凸函数和可微凹函数, $g_i(x) (i=1, 2, \dots, l)$ 是线性函数. 若 (x^*, λ^*) 是 KT 对, 那么 x^* 是最优解.

证 由于 (x^*, λ^*) 是 KT 对,

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - \sum_{i \in E \cup I^*} \lambda_i^* \nabla g_i(x^*) = 0, \lambda_i \geq 0 (i \in I^*),$$

其中 $E = \{1, 2, \dots, l\}$, $I^* = \{i \mid g_i(x^*) = 0, i \in \{l+1, l+2, \dots, m\}\}$, 并且

$$L(x, \lambda^*) = f(x) - \sum_{i \in E \cup I^*} \lambda_i^* g_i(x)$$

是可行域上的凸函数.

所以, 对于任何可行解 x 有

$$L(x, \lambda^*) \geq L(x^*, \lambda^*) + (x - x^*)^T \nabla_x L(x^*, \lambda^*) = f(x^*).$$

另一方面, $L(x, \lambda^*) \leq f(x)$, 因而 $f(x) \geq f(x^*)$, 即 x^* 是最优解. ■

第二章 线性不等式组

2.1 基本概念及有关定理

一个线性等式与不等式组,简称不等式组,可以表示为以下形式:

$$\left. \begin{aligned} a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (2.1.1)$$

其中 x 是 n 维未知列向量, a_i 是 n 维行向量, b_i 是标量, $i = 1, 2, \dots, m$. 本章介绍的解法称为旋转算法,它用到以下一些基本概念: $a_1, a_2, \dots, a_m$ 中一个最大线性无关组称为(2.1.1)的基,基中的向量称为基向量,其余向量称为非基向量. 基向量相应的(不)等式称为基(不)等式,否则称为非基(不)等式. 由基等式和基不等式构成的不等式组称为基本不等式组,其解集是个锥,称为基锥. 基本不等式组相应方程组的解即基锥顶点称为基本解.

定理 2.1.1 给定不等式组(2.1.1)的一个基,用 I_0, I_1, I_2 分别表示基等式、基不等式、非基不等式的编号集. 设 $x^{(1)}$ 是基本解,即 $a_j x^{(1)} = b_j, j \in I_0 \cup I_1$. 再设对于某个 $r \in I_2$, 有 $a_r x^{(1)} < b_r$, 并且

$$a_r = \sum_{j \in I_0 \cup I_1} w_{rj} a_j. \quad (2.1.2)$$

(i) 如果对于任何 $j \in I_1$ 有 $w_{rj} \leq 0$, 则不等式组(2.1.1)无解.

(ii) 如果存在某个 $s \in I_1$ 使得 $w_{rs} > 0$, 则 $a_s x^{(2)} > b_s$, 其中 $x^{(2)}$ 是方程组

$$a_j x = b_j, \quad j \in \{r\} \cup I_0 \cup I_1 \setminus \{s\}$$

的解.

(iii) 如果仅有一个 $s \in I_1$ 使得 $w_{rs} > 0$, 但是对任何 $j \in I_1 \setminus \{s\}$ 有 $w_{rj} \leq 0$, 则 $a_s x \geq b_s$ 是不等式组(2.1.1)的一个冗余不等式.

证 (i) 由于 $-a_r = -\sum_{j \in I_0 \cup I_1} w_{rj} a_j$ 其中 $-w_{rj} \geq 0, j \in I_1$ 根据推论 1.3.1, 不等式组

$$a_j x = b_j, \quad j \in I_0,$$

$$a_j x \geq b_j, \quad j \in I_1$$

的任何解满足 $-a_r x \geq -a_r x^{(1)}$, 即满足 $a_r x \leq a_r x^{(1)} < b_r$ 而不满足 $a_r x \geq b_r$, 因此不等式组(2.1.1) 没有解.

(ii) 由(2.1.2) 解出 a_s 得

$$a_s = \frac{1}{w_{rs}} a_r + \sum_{j \in I_0 \cup I_1 \setminus \{s\}} \left(-\frac{w_{rj}}{w_{rs}} \right) a_j, \quad (2.1.3)$$

两边右乘 $x^{(2)}$ 得

$$\begin{aligned} a_s x^{(2)} &= \frac{1}{w_{rs}} a_r x^{(2)} + \sum_{j \in I_0 \cup I_1 \setminus \{s\}} \left(-\frac{w_{rj}}{w_{rs}} \right) a_j x^{(2)} \\ &= \frac{1}{w_{rs}} b_r + \sum_{j \in I_0 \cup I_1 \setminus \{s\}} \left(-\frac{w_{rj}}{w_{rs}} \right) b_j \\ &= \frac{1}{w_{rs}} \left(b_r - \sum_{j \in I_0 \cup I_1} w_{rj} b_j \right) + b_s \\ &= \frac{1}{w_{rs}} (b_r - a_r x^{(1)}) + b_s > b_s. \end{aligned}$$

(iii) 由于(2.1.3) 的 $\frac{1}{w_{rs}} > 0$ 并且对于任何 $j \in I_1 \setminus \{s\}$ 有 $-\frac{w_{rj}}{w_{rs}} \geq 0$ 根据

推论 1.3.1 不等式组

$$a_j x = b_j, \quad j \in I_0,$$

$$a_j x \geq b_j, \quad j \in \{r\} \cup I_1 \setminus \{s\}$$

的所有解满足 $a_s x \geq a_s x^{(2)}$, 因而满足 $a_s x > b_s$, 故 $a_s x \geq b_s$ 是不等式组(2.1.1) 的一个冗余不等式. ■

以上定理的几何意义很简单. 考虑不等式组

$$a_1 x \geq b_1, \quad a_2 x \geq b_2, \quad a_r x \geq b_r.$$

设 a_1 和 a_2 是 a_1, a_2, a_r 的一个最大线性无关组 $a_r = w_1 a_1 + w_2 a_2$, $x^{(1)}$ 是基本解(即 $a_1 x^{(1)} = b_1$ 且 $a_2 x^{(1)} = b_2$) 且满足 $a_r x^{(1)} < b_r$. 图 2.1 的(a), (b), (c) 分别表示 ① $w_1 < 0, w_2 < 0$; ② $w_1 > 0, w_2 > 0$; ③ $w_1 > 0, w_2 < 0$ 这三种情况. 在情况 ①, 以上不等式组构成一个不相容的系统. 在情况 ② 和 ③, 新的基本解 $x^{(2)}$ 满足出基不等式 $a_1 x \geq b_1$ 在后一情况出基不等式 $a_1 x \geq b_1$ 还可去掉.

定理 2.1.2 给定不等式组(2.1.1) 的一个基, 用 I_0, I_1, I_3 分别表示基等式、基不等式、非基等式的编号集. 设 $x^{(1)}$ 是基本解, 并且对于某个 $r \in I_3$ 有

$$a_r = \sum_{j \in I_0 \cup I_1} w_{rj} a_j.$$

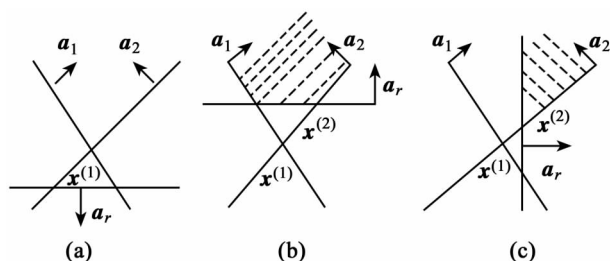


图 2.1 定理 2.1.1 的几何表示

(i) 如果 $a_r x^{(1)} < b_r$ 并且对任何 $j \in I_1$ 有 $w_{rj} \leq 0$; 或者 $a_r x^{(1)} > b_r$ 并且对任何 $j \in I_1$ 有 $w_{rj} \geq 0$, 则不等式组(2.1.1) 无解.

(ii) 如果 $a_r x^{(1)} \leq b_r$ 并且存在某个 $s \in I_1$ 使得 $w_{rs} > 0$ 或者 $a_r x^{(1)} \geq b_r$ 并且存在某个 $s \in I_1$ 使得 $w_{rs} < 0$, 则 $a_s x^{(2)} \geq b_s$ 其中 $x^{(2)}$ 是方程组 $a_j x = b_j$, $j \in \{r\} \cup I_0 \cup I_1 \setminus \{s\}$ 的解.

(iii) 如果 $a_r x^{(1)} = b_r$ 并且对任何 $j \in I_1$ 有 $w_{rj} = 0$, 则 $a_r x = b_r$ 是不等式组(2.1.1) 的一个冗余等式.

证 (i) 和(ii) 的证明只需注意 $a_r x = b_r$ 相当于 $a_r x \geq b_r$ 和 $-a_r x \geq -b_r$ 两个不等式. (iii) 的前提条件说明 a_r 是 $a_i (i \in I_0)$ 的线性组合且 b_r 是 $b_i (i \in I_0)$ 的同一线性组合, 所以基等式构成的方程组的所有解满足 $a_r x = b_r$, 因而 $a_r x = b_r$ 冗余. ■

定理 2.1.3 对于不等式组(2.1.1) 的一个基, 用 I_0 和 I_1 分别表示基等式和基不等式的编号集, x 是基本解, $a_r x \geq b_r$ 是其一个非基不等式. 设 $a_r x = b_r$, 并且

$$a_r = \sum_{j \in I_0 \cup I_1} w_{rj} a_j. \quad (2.1.4)$$

其中 $w_{rj} < 0$, $j \in I_1' \subset I_1$; $w_{rj} = 0$, $j \in I_1 \setminus I_1'$. 那么不等式组(2.1.1) 与以下不等式组等价:

$$\left. \begin{aligned} a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &= b_i, \quad i \in I_1' \cup \{r\}, \\ a_i x &\geq b_i, \quad i \in \{l+1, l+2, \dots, m\} \setminus I_1' \setminus \{r\}. \end{aligned} \right\} \quad (2.1.5)$$

证 设 x 是不等式组(2.1.1) 的任意一个解. 在(2.1.4) 两边右乘 x , 考虑到 $a_i x = b_i$, $i \in I_0$ 得

$$a_r x = \sum_{j \in I_0} w_{rj} b_j + \sum_{j \in I_1} w_{rj} a_j x.$$

在(2.1.4)两边右乘 x 得

$$a_r x = \sum_{j \in I_0} w_{rj} b_j + \sum_{j \in I_1} w_{rj} b_j.$$

以上两式相减, 并将 $a_r x = b_r$ 代入得

$$a_r x - b_r = \sum_{j \in I_1} w_{rj} (a_j x - b_j).$$

由于 $a_r x - b_r \geq 0$, $a_j x - b_j \geq 0$, $w_{rj} < 0$, $j \in I_1'$, 所以 $a_r x - b_r = 0$, $a_j x - b_j = 0$, $j \in I_1'$. 即不等式组(2.1.1)的任意一个解是不等式组(2.1.5)的解. 另外, 不等式组(2.1.5)的解显然是(2.1.1)的解. 所以两不等式组等价. ■

2.2 线性不等式组的旋转算法

2.2.1 旋转运算

从几何上看, 旋转算法开始于一个基锥的顶点, 譬如 $x^{(1)}$. 如果 $x^{(1)}$ 落在由其他不等式确定的半空间或等式确定的超平面上, 则它是整个不等式组的解; 否则存在一个不等式(或等式), 如 $a_r x \geq b_r$ 使 $a_r x^{(1)} < b_r$, 则这样的不等式称为违反不等式. 如果半空间 $a_r x \geq b_r$ 与基锥的交集是空集, 则不等式组没有解; 否则将 $x^{(1)}$ 沿着基锥的一条棱转移到该半空间的边界即平面 $a_r x = b_r$ 上. 这时 $a_r x \geq b_r$ 取代基锥中的一个不等式, 形成一个新的基锥, 其顶点满足出基不等式. 然后重复以上过程.

从代数上看, 用一个不等式或等式取代基锥中的不等式是通过以下旋转运算实现的.

给定不等式组(2.1.1)的一个基, 用 I_0, I_1, I_2, I_3 分别表示基等式、基不等式、非基不等式、非基等式的编号集. 设 $x^{(1)}$ 是基本解, 即 $a_j x^{(1)} = b_j, j \in I_0 \cup I_1$, 并设

$$a_i = \sum_{j \in I_0 \cup I_1} w_{ij} a_j, i \in I_3 \cup I_2. \quad (2.2.1)$$

如果对于某个 $r \in I_3 \cup I_2$ 和 $s \in I_1$ 有 $w_{rs} \neq 0$, 则从(2.2.1)的第 r 式可解得

$$a_s = \frac{1}{w_{rs}} a_r + \sum_{j \in I_0 \cup I_1 \setminus \{s\}} \left(-\frac{w_{rj}}{w_{rs}} \right) a_j, \quad (2.2.2)$$

代入(2.2.1)的其余各式得

$$a_i = \frac{w_{is}}{w_{rs}} a_r + \sum_{j \in I_0 \cup I_1 \setminus \{r\}} \left(w_{ij} - \frac{w_{is}}{w_{rs}} w_{rj} \right) a_j, i \in I_3 \cup I_2 \setminus \{r\} \quad (2.2.3)$$

现在 $\{r\} \cup I_0 \cup I_1 \setminus \{r\}$ 是新基的编号集 相应的基本解记为 $x^{(2)}$. 在(2.2.3) 两边右乘 $x^{(2)}$ 得

$$\begin{aligned} a_i x^{(2)} &= \frac{w_{is}}{w_{rs}} a_r x^{(2)} + \sum_{j \in I_0 \cup I_1 \setminus \{r\}} \left(w_{ij} - \frac{w_{is}}{w_{rs}} w_{rj} \right) a_j x^{(2)} \\ &= \frac{w_{is}}{w_{rs}} b_r + \sum_{j \in I_0 \cup I_1 \setminus \{r\}} \left(w_{ij} - \frac{w_{is}}{w_{rs}} w_{rj} \right) b_j \\ &= \frac{w_{is}}{w_{rs}} b_r + \sum_{j \in I_0 \cup I_1} \left(w_{ij} - \frac{w_{is}}{w_{rs}} w_{rj} \right) b_j. \end{aligned}$$

在(2.2.1) 两边右乘 $x^{(1)}$ 得

$$a_i x^{(1)} = \sum_{j \in I_0 \cup I_1} w_{ij} b_j.$$

以上两式相减得

$$\begin{aligned} a_i x^{(2)} - a_i x^{(1)} &= \frac{w_{is}}{w_{rs}} b_r - \sum_{j \in I_0 \cup I_1} \frac{w_{is}}{w_{rs}} w_{rj} b_j \\ &= \frac{w_{is}}{w_{rs}} \left(b_r - \sum_{j \in I_0 \cup I_1} w_{rj} b_j \right) \\ &= \frac{w_{is}}{w_{rs}} \left(b_r - \sum_{j \in I_0 \cup I_1} w_{rj} a_j x^{(1)} \right) \\ &= \frac{w_{is}}{w_{rs}} (b_r - a_r x^{(1)}). \end{aligned}$$

将上式改写为 $a_i x^{(2)} - b_i = a_i x^{(1)} - b_i - \frac{w_{is}}{w_{rs}} (a_r x^{(1)} - b_r)$, 令 $\sigma_i = a_i x^{(1)} - b_i$, $\sigma_r = a_r x^{(1)} - b_r$, $\sigma'_i = a_i x^{(2)} - b_i$ 得

$$\sigma'_i = \sigma_i - \frac{w_{is}}{w_{rs}} \sigma_r, i \in I_3 \cup I_2 \setminus \{r\}.$$

同样由(2.2.2) 可得 $a_s x^{(2)} - b_s = -\frac{1}{w_{rs}} (a_r x^{(2)} - b_r)$ 或

$$\sigma'_s = -\frac{\sigma_r}{w_{rs}},$$

其中 $\sigma'_s = a_s x^{(2)} - b_s$.

以上运算过程称为旋转运算, w_{rs} 称为枢轴(元素), 其所在行称为枢轴行, 所在列称为枢轴列. 我们说 a_r 入基, a_s 出基, 两者位置的交换记为 $a_r \leftrightarrow a_s$. 旋转

运算前后的数据可用表 2.1、表 2.2 表示.

表 2.1 初始表

	a_s	a_j	
a_r	w_{rs}	w_{rj}	σ_r
a_i	w_{is}	w_{ij}	σ_i

表 2.2 旋转运算的结果

	a_r	a_j	
a_s	$\frac{1}{w_{rs}}$	$-\frac{w_{rj}}{w_{rs}}$	$-\frac{\sigma_r}{w_{rs}}$
a_i	$\frac{w_{is}}{w_{rs}}$	w'_{ij}	σ'_i

其中, $w'_{ij} = w_{ij} - \frac{w_{is}}{w_{rs}}w_{rj}$, $i \in I_3 \cup I_2 \setminus \{r\}$, $j \in I_0 \cup I_1 \setminus \{s\}$;

$$\sigma'_i = \sigma_i - \frac{w_{is}}{w_{rs}}\sigma_r, \quad i \in I_3 \cup I_2 \setminus \{r\}.$$

对表 2.1 来说 $\sigma_i = a_i x^{(1)} - b_i$ 称为 a_i 或相应(不)等式关于 $x^{(1)}$ 的偏差. 如果

$$\sigma_i = a_i x^{(1)} - b_i = 0, \quad i \in I_3,$$

$$\sigma_i = a_i x^{(1)} - b_i \geq 0, \quad i \in I_2,$$

则 $x^{(1)}$ 是不等式组(2.1.1)的解. 如果对某个 $r \in I_2$ 有 $\sigma_r = a_r x^{(1)} - b_r < 0$ 则

称 $a_r x \geq b_r$ 为关于 $x^{(1)}$ 的违反不等式, 并称 $\frac{|a_r x^{(1)} - b_r|}{\|a_r\|}$ 为 $x^{(1)}$ 到违反不等式

$a_r x \geq b_r$ 的距离. 从几何上看, $\frac{|a_r x^{(1)} - b_r|}{\|a_r\|}$ 是 $x^{(1)}$ 到超平面 $a_r x = b_r$ 的距离.

2.2.2 线性不等式组的计算步骤

许多线性不等式组的变量有下界限制, 其形式为

$$\left. \begin{aligned} a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i \in I_Z \subset \{1, 2, \dots, n\}. \end{aligned} \right\} \quad (2.2.4)$$

其中 $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$, $i = 1, 2, \dots, m$.

不等式组(2.2.4)的旋转算法分为构造初始表、预处理和主迭代三个阶段.

为了方便地构造算法的初始表, 初始基向量最好是单位向量, 为此引入不等式

$$x_i \geq -M, \quad i \in \{1, 2, \dots, n\} \setminus I_Z.$$

其中 M 是一个充分大的数. 我们将以上不等式称为人工不等式. $x_i \geq l_i$ 或 $x_i \geq -M$ 的系数向量是 n 阶单位矩阵的第 i 行, 用 e_i 表示 $i = 1, 2, \dots, n$. 初始基向量即为 $e_1, e_2, \dots, e_n$, 初始非基向量为 $a_1, a_2, \dots, a_m$. 预处理是利用旋转运算将非基等式尽可能换为基等式, 或者说将非基等式的系数向量尽可能换为基向量的运算过程. 一个等式一旦成为基本的就不再出基. 这是因为, 如果一个不等式有解, 其每个解必须满足所有等式. 为了使基等式不出基, 只须不在基等式所在列寻找枢轴元素. 因此, 非基等式一旦成为基本的, 其所在列可以删掉. 主迭代是非基不等式与基不等式交换的运算过程, 一直到所有非基向量偏差非负, 即基本解成为不等式组 (2.2.4) 的解, 或者发现其没有解.

算法 2.2.1 不等式组 (2.2.4) 的旋转算法.

步骤 1 构造初始表.

以

$$\begin{aligned} x_i &\geq l_i, \quad i \in I_Z, \\ x_i &\geq -M, \quad i \in \{1, 2, \dots, n\} \setminus I_Z \end{aligned}$$

为初始基本不等式组. 也就是说以 $e_1, e_2, \dots, e_n$ 为初始基向量, 初始基本解为 $x_i^{(0)} = l_i (i \in I_Z), x_i^{(0)} = -M (i \in \{1, 2, \dots, n\} \setminus I_Z)$. 非基向量 a_i 的偏差 $\sigma_i = a_i x^{(0)} - b_i, i = 1, 2, \dots, m$, 其中 $x^{(0)}$ 是以 $x_i^{(0)}$ 为分量的列向量. 初始表如表 2.3 所示.

表 2.3 初始旋转算法表

	e_1	e_2	$\dots$	e_n	
a_1	a_{11}	a_{12}	$\dots$	a_{1n}	σ_1
a_2	a_{21}	a_{22}	$\dots$	a_{2n}	σ_2
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
a_m	a_{m1}	a_{m2}	$\dots$	a_{mn}	σ_m

步骤 2 预处理. 记 $I_3 = \{1, 2, \dots, l\}$.

(i) 如果 $I_3 = \emptyset$, 转步骤 3; 否则, 对于任何 $r \in I_3$, 当 a_r 的偏差 σ_r 是负数、正数或 0 时, 分别转 (ii), (iii), (iv).

(ii) 若 a_r 行没有与基不等式对应的正元素, 则不等式组 (2.2.4) 无解, 否则以其中一个正元素为枢轴进行一次旋转运算, 令 $I_3 = I_3 \setminus \{r\}$, 转 (i).

(iii) 若 a_r 行没有与基不等式对应的负元素, 则不等式组 (2.2.4) 无解; 否

则以其中一个负元素为枢轴进行一次旋转运算,令 $I_3 := I_3 h\{r\}$ 转(i).

(iv) 若 a_r 行与基不等式对应的元素均为零,则 $a_r x = b_r$ 是不等式组(2.2.4)的一个冗余等式,令 $I_3 := I_3 h\{r\}$ 转(i);否则以其中一个非零元素为枢轴进行一次旋转运算,令 $I_3 := I_3 h\{r\}$ 转(i).

步骤3 主迭代.

(i) 若所有非基向量偏差非负,则当前基本解是不等式组(2.2.4)的解,停止;否则转(ii).

(ii) 以其中一个偏差为负的非基向量入基.如果同行没有与基不等式对应的正元素,则不等式组(2.1.1)无解;否则以其中一个正元素为枢轴进行一次旋转运算,转(i).

以上算法的步骤1介绍得比较简单.事实上,如果不等式组(2.2.4)没有形如 $x_i \geq l_i$ 的不等式,但是有形如 $x_i \leq u_i$ 的不等式(u_i 是有限实数),也可以以 $-x_i \geq -u_i$ 为初始基本不等式,无须添加人工不等式.

有时一个线性不等式组含有形如 $b'_i \geq a_i x \geq b_i$ 的不等式,特别是 $u_i \geq x_i \geq l_i$. 按规范方式表示应为 $a_i x \geq b_i$ 和 $-a_i x \geq -b'_i$. 设基本解为 $\bar{x}$, 则 a_i 和 $-a_i$ 的偏差分别为 $a_i \bar{x} - b_i$ 和 $-a_i \bar{x} + b'_i$, 两者的和等于 $b'_i - b_i$. 由于 a_i 和 $-a_i$ 线性相关,它们不可能都是基向量.若其中有一个是基向量,另一个的偏差为 $b'_i - b_i > 0$, 因而可不考虑.如果 a_i 和 $-a_i$ 都是非基向量,由于它们关于基向量表达式系数的符号相反,在表中可共用一行.

值得注意的是,虽然基本解没有明显地出现在旋转运算过程中,但其每个分量很容易利用单位向量计算出来.设基本解 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$, 其获取方法为:如果 e_i 是基向量,则 $x_i \geq l_i$ 或 $-M$ 是基不等式,因而 $\bar{x}_i = l_i$ 或 $-M$;若 e_i 是非基向量,由于 e_i 的偏差等于 $\bar{x}_i - l_i$ 或 $\bar{x}_i + M$, $\bar{x}_i$ 等于 e_i 的偏差加上 x_i 的下界, $i = 1, 2, \dots, n$. 同样,如果不等式 $-x_i \geq -u_i$ 或 $-M$ 的系数向量 $-e_i$ 是基向量,则 $\bar{x}_i$ 等于 x_i 的上界;若 $-e_i$ 是非基向量,则 $\bar{x}_i$ 等于 x_i 的上界减去 $-e_i$ 的偏差.

下面利用两个简单例题说明线性不等式组的计算过程.

例 2.2.1 求以下不等式组的一个解:

$$-2x_1 + x_3 = 1,$$

$$-4x_1 + x_2 + 2x_3 \geq 3,$$

$$3x_1 + 2x_2 + x_3 \geq 12,$$

$$x_1 \geq 0, \quad x_2 \geq 2.$$

解 令 $a_1 = (-2, 0, 1), a_2 = (-4, 1, 2), a_3 = (3, 2, 1), e_1 = (1, 0, 0), e_2 = (0, 1, 0)$. 引入人工不等式 $x_3 \geq -M$ 并令 $e_3 = (0, 0, 1)$.

以 $x_1 \geq 0, x_2 \geq 2, x_3 \geq -M$ 为初始基本不等式组, 初始基本解为 $x^{(0)} = (0, 2, -M)^T$, 初始基向量为 e_1, e_2, e_3 . 非基向量为 a_1, a_2, a_3 , 它们的偏差分别为 $-M-1, 2-2M-3 = -2M-1, 2 \times 2 - M - 12 = -M-8$, 初始表见表 2.4.

表 2.4 初始表

	e_1	e_2	e_3	
a_1	-2	0	1*	$-M-1$
a_2	-4	1	2	$-2M-1$
a_3	3	2	1	$-M-8$

表 2.5 第一次旋转运算结果

	e_1	e_2	
e_3	2	0	$M+1$
a_2	0	1	1
a_3	5	2*	-7

在表 2.4 中, 由于非基向量 a_1 是等式的系数向量, 首先应换入基中. 由于其偏差是负数, 以该行的正元素 1 (带星号者) 为枢轴进行旋转运算, 结果见表 2.5. 其中基向量 a_1 所在列已删掉. 在表 2.5 中, 非基向量 a_3 的偏差是 -7, 以 a_3 入基. 该行两个元素均为正, 不妨以 2 为枢轴进行旋转运算, 结果见表 2.6.

表 2.6 第二次旋转运算结果

	e_1	a_3	
e_3	2	0	$M+1$
a_2	$-\frac{5}{2}$	$\frac{1}{2}$	$\frac{9}{2}$
e_2	$-\frac{5}{2}$	$\frac{1}{2}$	$\frac{7}{2}$

至此, 所有非基向量偏差非负, 当前基本解是不等式组的一个解. 由于 e_1 是基向量, $x_1 \geq 0$ 是基不等式, 所以 $x_1 = 0$. 由于 e_2 和 e_3 是非基向量, $x_2 \geq 2$ 和 $x_3 \geq -M$ 是非基不等式, 所以

$$x_2 = \frac{7}{2} + 2 = \frac{11}{2}, \quad x_3 = M + 1 + (-M) = 1.$$

例 2.2.2 在例 2.2.1 的不等式组中, 如果还要求 $x_2 \leq 4$, 求一个解.

解 将 $x_2 \leq 4$ 写为 $-x_2 \geq -4$, 其系数向量为 $-e_2$. 此问题的初始表及前两次旋转运算见表 2.4 至表 2.6. 虽然表 2.6 列出的非基向量偏差是正数, 但不等

式 $-x_2 \geq -4$ 还未考虑. 由于其偏差 $4 - 2 - \frac{7}{2} = -\frac{3}{2}$ 是负数, 当前基本解不满足此不等式. 将表 2.6 的 e_2 换为 $-e_2$ 将 $\frac{7}{2}$ 换为 $-\frac{3}{2}$, 并将同行其他元素反号, 得表 2.7, 其中 $\bar{\sigma}_i$ 列的 $\frac{7}{2}$ 是 e_2 的偏差.

表 2.7 调整后的表格

	e_1	a_3	σ_i	$\bar{\sigma}_i$
e_3	2	0	$M + 1$	—
a_2	$-\frac{5}{2}$	$\frac{1}{2}$	$\frac{9}{2}$	—
$-e_2$	$\frac{5}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{7}{2}$

表 2.8 第三次旋转运算结果

	$-e_2$	a_3	σ_i
e_3	$\frac{4}{5}$	$\frac{2}{5}$	$M + \frac{11}{5}$
a_2	-1	0	3
e_1	$\frac{2}{5}$	$\frac{1}{5}$	$\frac{3}{5}$

在表 2.7 中以 $-e_2$ 入基, 以 e_1 出基, 结果见表 2.8. 可以看出此时所有非基向量偏差非负, 因此当前基本解是不等式组的一个解. 由于 $-e_2$ 是基向量, $-x_2 \geq -4$ 是基不等式, 所以 $x_2 = 4$. 由于 e_1 和 e_3 是非基向量 $x_1 \geq 0$ 和 $x_3 \geq -M$ 是非基不等式, 所以

$$x_1 = \frac{3}{5} + 0 = \frac{3}{5}, \quad x_3 = M + \frac{11}{5} - M = \frac{11}{5}.$$

2.2.3 旋转算法表的形式不变性及预处理的作用

为了简单起见, 设不等式组(2.2.4) 每个变量 x_i 的下界为 l_i . 这时初始基本不等式组为

$$x_i \geq l_i, \quad i = 1, 2, \dots, n.$$

初始基本解为

$$x_i^{(0)} = l_i, \quad i = 1, 2, \dots, n,$$

或写为 $x^{(0)} = l$, 其中 $l = (l_1, l_2, \dots, l_n)^T$. 于是非基向量 a_i 的偏差为

$$a_i x^{(0)} - b_i = a_i l - b_i, \quad i = 1, 2, \dots, m.$$

不等式组(2.2.4) 是关于 n 维空间一个直角坐标系的表达形式. 现在将坐标系平移, 使原点与 n 维空间的一个特定点 $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$ 重合. 为得到模型(2.2.4) 在新坐标系的表达形式, 将 $x = y + x^*$ 代入, 整理后得

$$\left. \begin{aligned} a_i y &= -a_i x^* + b_i, \quad i = 1, 2, \dots, l, \\ a_i y &\geq -a_i x^* + b_i, \quad i = l+1, l+2, \dots, m, \\ y_i &\geq -x_i^* + l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (2.2.5)$$

同样,以后 n 个不等式为初始基本不等式组,初始基本解为

$$y_i^{(0)} = -x_i^* + l_i, \quad i = 1, 2, \dots, n,$$

或写为 $y^{(0)} = -x^* + l$. 这时,非基向量 a_i 的偏差为

$$\begin{aligned} a_i y^{(0)} - (-a_i x^* + b_i) &= a_i (-x^* + l) + a_i x^* - b_i \\ &= a_i l - b_i, \quad i = 1, 2, \dots, m. \end{aligned}$$

可见与不等式组(2.2.4)初始非基向量的偏差一样.

由于两不等式组的系数向量和初始非基向量的偏差一样,它们的初始表相同,所以只要迭代规则相同,以后的各张旋转算法表也都相同.我们将坐标系平移不改变旋转算法表的这种性质称为旋转算法表的形式不变性.以后将会看到,用旋转算法表解非凸二次规划时,将KT点 x^* 看做“原点”,对判定其是不是局部极小点是非常方便的.

初始表 2.3 的矩阵形式如表 2.9 所示.

表 2.9 初始表的矩阵形式

	e_I	e_{II}	
a_I	A_{11}	A_{12}	σ_I
a_{II}	A_{21}	A_{22}	σ_{II}

其中 (A_{11}, A_{12}) 是前 l 个等式的系数构成的 1×2 分块矩阵, (A_{21}, A_{22}) 是后 $m-l$ 个不等式的系数构成的 1×2 分块矩阵, σ_I 和 σ_{II} 分别是等式和不等式的偏差构成的向量, a_I, a_{II} 和 e_I, e_{II} 分别表示各分块矩阵对应的非基向量和基向量.相应地,将变量 x 划分为 x_I 和 x_{II} .

不失一般性,设 l 个等式的系数向量线性无关.在算法 2.2.1 的第二个步骤预处理中,将 l 个等式换为基等式是以 (A_{11}, A_{12}) 中一个 l 阶非奇异方阵的元素为枢轴进行的 l 次旋转运算.假设这个非奇异方阵为 A_{11} ,那么根据第一章的定理 1.1.1,在表 2.9 中求 A_{11} 逆的结果如表 2.10 所示.

表 2.10 预处理结果

	a_I	e_{II}	
e_I	A_{11}^{-1}	$-A_{11}^{-1}A_{12}$	$-A_{11}^{-1}\sigma_I$
a_{II}	$A_{21}A_{11}^{-1}$	$A_{22} - A_{21}A_{11}^{-1}A_{12}$	$\sigma_{II} - A_{21}A_{11}^{-1}\sigma_I$

其中 a_I 列和 e_I 行可以删掉.

另一方面,由表 2.9 可见,不等式组(2.2.4)等价于以下不等式组:

$$A_{11}x_I + A_{12}x_{II} = -\sigma_I,$$

$$A_{21}x_I + A_{22}x_{II} \geq -\sigma_{II},$$

$$x_I \geq 0, x_{II} \geq 0.$$

由第一个式子解出 x_I 得

$$x_I = -A_{11}^{-1}A_{12}x_{II} - A_{11}^{-1}\sigma_I.$$

代入上面的第二个式子,整理后得

$$(A_{22} - A_{21}A_{11}^{-1}A_{12})x_{II} \geq -(\sigma_{II} - A_{21}A_{11}^{-1}\sigma_I).$$

将此式与表 2.10 对照,可见预处理实际上是将不等式组(2.2.4)转变为一个仅含 $n-1$ 个变量和 $m-1$ 个线性不等式的线性不等式组.

2.3 循环与反循环

在解线性不等式组时,有多种方法选择入基和出基向量,每种方法都在迭代过程中产生一个基(或基本解)的序列.在此序列中如果前后有两个基相同,可想而知,这种情况将重复出现,此现象称为循环.下面是一个循环的例子.

例 2.3.1 解以下线性不等式组:

$$9x_1 + x_2 + 3x_3 \geq 4,$$

$$2x_1 - x_2 - x_3 \geq -3,$$

$$16x_1 - 3x_2 - 2x_3 \geq -4,$$

$$-24x_1 + 12x_2 + x_3 \geq 5,$$

$$x_1 \geq 0, x_2 \geq 1, x_3 \geq -1.$$

解 以 $x_1 \geq 0, x_2 \geq 1, x_3 \geq -1$ 为初始基本不等式组,初始基本解为 $x^{(0)} = (0, 1, -1)^T$,初始基向量为 $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$;

其他向量是非基向量,分别记为 a_1, a_2, a_3, a_4 . 非基向量关于基向量的表达式以及它们关于 $x^{(0)}$ 的偏差如表 2.11 所示.

表 2.11 初始表

	e_1	e_2	e_3	
a_1	9	1*	3	-6
a_2	2	-1	-1	3
a_3	16	-3	-2	3
a_4	-24	12	1	6

表 2.12 第一次旋转运算结果

	e_1	a_2	e_3	
e_2	-9	1	-3	6
a_2	11	-1	2*	-3
a_3	43	-3	7	-15
a_4	-132	12	-35	78

在表 2.11 中以 a_1 入基 e_2 出基,结果如表 2.12 所示. 其他各次旋转运算的结果见表 2.13 至表 2.16,表中带星号的数字是枢轴元素.

表 2.13 第二次旋转运算结果

	e_1	a_2	a_3	
e_2	$\frac{15}{2}$	$-\frac{1}{2}$	$-\frac{3}{2}$	$\frac{3}{2}$
e_3	$-\frac{11}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
a_3	$\frac{9}{2}$	$\frac{1}{2}$ *	$\frac{7}{2}$	$-\frac{9}{2}$
a_4	$\frac{121}{2}$	$-\frac{11}{2}$	$-\frac{35}{2}$	$\frac{51}{2}$

表 2.14 第三次旋转运算结果

	e_1	a_3	a_2	
e_2	12	-1	2	-3
e_3	-10	1	-3	6
a_1	-9	2	-7	9
a_4	110	-11	21*	-24

表 2.15 第四次旋转运算结果

	e_1	a_3	a_4	
e_2	$\frac{32}{21}$	$\frac{1}{21}$ *	$\frac{2}{21}$	$-\frac{5}{7}$
e_3	$\frac{40}{7}$	$-\frac{4}{7}$	$-\frac{1}{7}$	$\frac{18}{7}$
a_1	$\frac{83}{3}$	$-\frac{5}{3}$	$-\frac{1}{3}$	1
a_2	$-\frac{110}{21}$	$\frac{11}{21}$	$\frac{1}{21}$	$\frac{8}{7}$

表 2.16 第五次旋转运算结果

	e_1	e_2	a_4	
a_3	-32	21	-2	15
e_3	24	-12	1*	-6
a_1	81	-35	3	-24
a_2	-22	11	-1	9

在表 2.16 中, 以 e_3 入基 a_4 出基, 则回复到表 2.11, 即发生循环.

解一个线性不等式组需要进行多少次迭代或者是否会发生循环, 与确定入基和出基向量的规则有关. 设 I_0, I_1, I_2 分别是基等式、基不等式、非基不等式的指标集, $x^{(1)}$ 是当前基本解, 并设

$$a_i = \sum_{j \in I_0 \cup I_1} w_{ij} a_j, \sigma_i = a_i x^{(1)} - b_i, \quad i \in I_2.$$

下面是一些可供选择的确定入基和出基向量的规则.

规则 1(最小偏差规则) 在所有偏差为负数的非基不等式中, 以偏差最小者入基. 即若 $\sigma_r = \min\{\sigma_i \mid \sigma_i < 0, i \in I_2\}$, 则以 a_r 入基.

规则 2(最大距离规则) 在所有偏差为负数的非基不等式中, 以距离当前基本解最远的违反不等式入基. 即若 $\frac{\sigma_r}{\|a_r\|} = \min\left\{\frac{\sigma_i}{\|a_i\|} \mid \sigma_i < 0, i \in I_2\right\}$, 则以 a_r 入基.

规则 3(棱上最远规则) 在基锥的某条棱上选择一个距离当前基本解最远的违反不等式入基. 即对于某个 $s \in I_1$, 如果 $\frac{\sigma_r}{w_{rs}} = \min\left\{\frac{\sigma_i}{w_{is}} \mid \sigma_i < 0 \text{ 且 } w_{is} > 0, i \in I_2\right\}$, 则以 a_r 入基.

规则 4(最小下标规则) 在所有偏差为负数的非基不等式中, 以下标(即编号)最小者入基; 在所有可能的出基不等式中, 以下标最小者出基. 即若 $r = \min\{i \in I_2 \mid \sigma_i < 0\}$, 则以 a_r 入基. 若 $s = \min\{j \in I_1 \mid w_{rj} > 0\}$, 则以 a_s 出基.

下面按本文的概念证明, 利用最小下标规则选择入基、出基不等式可以避免循环. 这一证明与 R. G. Bland 的结果本质上是一样的, 但形式大不相同.

定理 2.3.1 在解线性不等式组(2.1.1)的旋转算法中应用最小下标规则可以防止循环.

证 用反证法. 设将最小下标规则用于不等式组(2.1.1)出现循环. 发生循环时, 必定有些向量不断地入基和出基, 我们将这些向量称为游动向量. 用 s 表示游动向量的最大下标, 用 B_1 和 B_2 表示迭代过程中出现的两个基或它们编号的集合, 并设 $B_1 = I_0 \cup I_1$, 其中 I_0 和 I_1 分别对应基等式和基不等式. B_1 相应的基本解记为 $x^{(1)}$, B_2 相应的基本解记为 $x^{(2)}$. 假设对于 B_1 来说 a_r 入基, a_s 出基, 对于 B_2 来说 a_s 入基, 并设 $a_r = \sum_{j \in I_0 \cup I_1} w_{rj} a_j$. 在此式两边分别右乘 $x^{(1)}$ 和 $x^{(2)}$ 得

$$a_r x^{(1)} = \sum_{j \in I_0 \cup I_1} w_{rj} a_j x^{(1)} = \sum_{j \in I_0 \cup I_1} w_{rj} b_j$$

和

$$a_r x^{(2)} = \sum_{j \in I_0} w_{ij} b_j + \sum_{j \in I_1} w_{ij} a_j x^{(2)}.$$

后式减前式得

$$a_r x^{(2)} - a_r x^{(1)} = \sum_{j \in I_1} w_{ij} (a_j x^{(2)} - b_j).$$

由于 a_r 是 B_1 的入基向量 $a_r x^{(1)} < b_r$. 由于 a_s (而不是 a_r) 是 B_2 的入基向量, 必有 $a_r x^{(2)} \geq b_r$. 因此

$$\sum_{j \in I_1} w_{ij} (a_j x^{(2)} - b_j) > 0.$$

上式意味着存在 $t \in I_1$ 使得

$$w_{it} (a_t x^{(2)} - b_t) > 0, \quad (2.3.1)$$

而此式意味着 $a_t x^{(2)} \neq b_t$, 即 t 不属于 B_2 . 但是 $t \in B_1$, 因此 a_t 是游动向量并且 $t \leq s$.

由于 a_r 要进入 B_1 而 a_s 要离开 B_1 , 枢轴元素 $w_{rs} > 0$. 由于 a_s 要进入 B_2 , $a_s x^{(2)} < b_s$, 因此

$$w_{rs} (a_s x^{(2)} - b_s) < 0. \quad (2.3.2)$$

(2.3.1) 和 (2.3.2) 两式意味着 $t \neq s$, 因而 $t < s$. a_s 是 B_2 的入基向量意味着 $a_t x^{(2)} - b_t \geq 0$. 而 (2.3.1) 进一步说明 $a_t x^{(2)} - b_t > 0$ 并且 $w_{it} > 0$. 按照最小下标规则, 这时应选 a_t 离开 B_1 (因为 $w_{it} > 0$), 这与 a_s 是 B_1 的出基向量相矛盾.

■

2.4 参数化技术

前面介绍了计算线性不等式组

$$\left. \begin{aligned} a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m \end{aligned} \right\} \quad (2.4.1)$$

的一个解的方法. 这个解是向量组 $\{a_1, a_2, \dots, a_m\}$ 的一个最大无关组相应方程组的解, 称为基本解. 本节讨论当基等式或基不等式右端的常数项发生变化时, 如何以较少的计算量获取新的解. 这种方法称为参数化技术 (或方法).

对于线性不等式组 (2.4.1) 的一个基, 用 I_B 和 I_N 分别表示基向量和非基向量的指标集, 并设

$$a_i = \sum_{j \in I_B} w_{ij} a_j, \quad i \in I_N.$$

用 x 表示当前基本解. 在上式两边右乘 x , 由于 $a_j x = b_j, j \in I_B$, 有

$$a_i x = \sum_{j \in I_B} w_{ij} b_j.$$

所以非基向量 a_i 的偏差

$$\sigma_i = a_i x - b_i = \sum_{j \in I_B} w_{ij} b_j - b_i, i \in I_N.$$

如果 b_j 变为 $b_j + \delta_j, j \in I_B$, a_i 的偏差变为

$$\sigma'_i = \sum_{j \in I_B} w_{ij} (b_j + \delta_j) - b_i = \sigma_i + \sum_{j \in I_B} w_{ij} \delta_j, i \in I_N. \quad (2.4.2)$$

以上操作记为 $a_j^{+\delta_j}, j \in I_B$.

通常我们一次仅变动一个基等式或基不等式的右端项, 例如将 b_s 变为 $b_s + \Delta, s \in I_B$, 这时非基向量 a_i 的偏差变为

$$\sigma'_i = \sigma_i + w_{is} \Delta, i \in I_N. \quad (2.4.3)$$

(2.4.3) 表明, 当某个基等式或基不等式右端的常数项的改变量为 Δ 时, 将旋转算法表中该基向量对应的列乘以 Δ , 然后加到最后一列(偏差列), 便得到所有非基向量关于新基本解的偏差. 以上操作记为 $a_s^{+\Delta}$.

现设 $I_B = I_0 \cup I_1$, 其中 I_0 和 I_1 分别是基等式和基不等式的指标集. 依定义, 基本不等式组为

$$\begin{aligned} a_j x &= b_j, j \in I_0, \\ a_j x &\geq b_j, j \in I_1. \end{aligned}$$

其解集是个凸锥, 记为 S , 当前基本解 x 是其顶点.

对于某个 $s \in I_0 \cup I_1$, 将 b_s 换为 $b_s + \Delta$ 即将平面 $a_s x = b_s$ 平移到 $a_s x = b_s + \Delta$. $\Delta > 0$ 时, 移动方向与 a_s 的方向相同; $\Delta < 0$ 时, 移动方向与 a_s 的方向相反. 这时, 不等式组(2.4.1)变为另一个不等式组, 其解集记为 $S(\Delta)$. 如果 b_s 是基不等式的右端项, 即 $s \in I_1$, 通常让 $\Delta > 0$. 由于

$$\{x \in \mathbb{R}^n \mid a_s x \geq b_s + \Delta, \Delta > 0\} \subset \{x \in \mathbb{R}^n \mid a_s x \geq b_s\},$$

$S(\Delta) \subset S$, 即新的不等式组的解集是原不等式组解集的子集.

另外, 对于某个 $s \in I_0 \cup I_1$, 将 b_s 换为 $b_s + \Delta$ 时, S 的顶点 x 在集合

$$E = \{x \in \mathbb{R}^n \mid a_i x = b_i, i \in I_0 \cup I_1 \setminus \{s\}\}$$

内移动到另一点, 记为 $x(\Delta)$. 各非基向量的关于 $x(\Delta)$ 的偏差由(2.4.3)确定. 若基向量的数目等于 n , E 是 n 维空间中的一条直线.

在解线性不等式的过程中, 我们总是首先将所有非基等式尽可能换为基等式, 所以下面假设目前没有非基等式, 只有非基不等式, 其指标集记为 I_2 . 设

$$\sigma_i = a_i x - b_i \geq 0, i \in I_2,$$

即当前基本解 x 是不等式组(2.4.1)的解. 将 b_s 换为 $b_s + \Delta$ 时(其中 $s \in I_1, \Delta \geq$

0) 如果非基向量 a_i 的偏差

$$\sigma_i' = \sigma_i + w_{is}\Delta \geq 0, \quad i \in I_2,$$

则 $x(\Delta)$ 不仅是新的不等式组的解, 而且还是原来的模型(2.4.1)的解, 因为这时 $S(\Delta) \subset S$.

下面用一个例子说明参数公式(2.4.3)的一种应用.

例 2.4.1 解不等式组

$$\begin{aligned} x_1 - 2x_2 + x_3 + x_4 &= b_1, \\ 4x_1 - x_2 + x_3 - x_4 &\geq -2, \\ -6x_1 + 3x_2 - 2x_3 + x_4 &\geq 1, \\ x_1 &\geq -1, x_2 \geq 0, x_3 \geq 1. \end{aligned}$$

其中 $b_1 = 1, 3$ 或 5 .

解 首先考虑 $b_1 = 1$ 的情形. 令 $a_1 = (1, -2, 0, 1), a_2 = (4, -1, 1, -1), a_3 = (-6, 3, -2, 1), e_1 = (1, 0, 0, 0), e_2 = (0, 1, 0, 0), e_3 = (0, 0, 1, 0)$. 引入人工不等式 $x_4 \geq -M$ 并令 $e_4 = (0, 0, 0, 1)$.

以 $x_1 \geq -1, x_2 \geq 0, x_3 \geq 1, x_4 \geq -M$ 为初始基本不等式组, 初始基本解为 $x^{(0)} = (-1, 0, 1, -M)^T$, 初始基向量为 e_1, e_2, e_3, e_4 , 非基向量为 a_1, a_2, a_3 , 它们的偏差分别为 $-M-1, M-1, -M+3$, 初始表见表 2.17.

表 2.17 初始表

	e_1	e_2	e_3	e_4	
a_1	1*	-2	1	1	$-M-1$
a_2	4	-1	1	-1	$M-1$
a_3	-6	3	-2	1	$-M+3$

由于非基向量 a_1 是等式的系数向量, 首先应换入基中. 由于其偏差是负数, 以该行的正元素 1(带星号者)为枢轴进行旋转运算, 结果见表 2.18.

表 2.18 第一次旋转运算结果

	a_1	e_2	e_3	e_4	
e_1	1	2	-1	-1	$M+1$
a_2	4	7	-3	-5	$5M+3$
a_3	-6	-9	4	7*	$-7M-3$

在表 2.18 中 a_3 的偏差是最小负数 a_3 入基. 以该行的最大正元素 7 为枢轴作旋转运算, 结果见表 2.19.

表 2.19 第二次旋转运算结果

	a_1	e_2	e_3	a_3	$b_1 = 1$	$b_1 = 3$	$b_1 = 5$
e_1	$\frac{1}{7}$	$\frac{5}{7}$	$-\frac{3}{7}$	$-\frac{1}{7}$	$\frac{4}{7}$	$\frac{6}{7}$	$\frac{8}{7}$
a_2	$-\frac{2}{7}$	$\frac{4}{7}$	$-\frac{1}{7}$	$-\frac{5}{7}$	$\frac{6}{7}$	$\frac{2}{7}$	$-\frac{2}{7}$
e_4	$\frac{6}{7}$	$\frac{9}{7}$	$-\frac{4}{7}$	$\frac{1}{7}$	$M + \frac{3}{7}$	$M + \frac{15}{7}$	$M + \frac{27}{7}$

由表 2.19 可见 $b_1 = 1$ 列的所有非基向量偏差为正, 所以, 当前基本解是不等式组的解. 由于单位向量 e_2 和 e_3 是基向量, x_2 和 x_3 分别等于它们的下界, 即

$$x_2 = 0, x_3 = 1.$$

由于 e_1 和 e_4 是非基向量, x_1 和 x_4 分别等于它们的下界加上 e_1 和 e_4 的偏差, 即

$$x_1 = -1 + \frac{4}{7} = -\frac{3}{7}, x_4 = -M + M + \frac{3}{7} = \frac{3}{7}.$$

当 $b_1 = 3$ 时, 改变量为 $\Delta = 3 - 1 = 2$. 将表 2.19 中 a_1 列的数字乘以 2 加到 $b_1 = 1$ 列:

$$\frac{1}{7} \times 2 + \frac{4}{7} = \frac{6}{7},$$

$$-\frac{2}{7} \times 2 + \frac{6}{7} = \frac{2}{7},$$

$$\frac{6}{7} \times 2 + M + \frac{3}{7} = M + \frac{15}{7}.$$

这些数字在表 2.19 的 $b_1 = 3$ 那一列. 由于所有非基向量偏差为正, $b_1 = 3$ 时的解为

$$x_2 = 0, x_3 = 1,$$

$$x_1 = -1 + \frac{6}{7} = -\frac{1}{7}, x_4 = -M + M + \frac{15}{7} = \frac{15}{7}.$$

当 $b_1 = 5$ 时, 改变量为 $\Delta = 5 - 1 = 4$. 将表 2.19 中 a_1 列的数字乘以 4 加到 $b_1 = 1$ 列:

$$\frac{1}{7} \times 4 + \frac{4}{7} = \frac{8}{7},$$

$$-\frac{2}{7} \times 4 + \frac{6}{7} = -\frac{2}{7},$$

$$\frac{6}{7} \times 4 + M + \frac{3}{7} = M + \frac{27}{7}.$$

这些数字在表 2.19 的最后一列. 由于 a_2 的偏差是负数 a_2 入基. 该行仅有 e_2 列上的 $\frac{4}{7}$ 是正数, 以其为枢轴作旋转运算, 结果见表 2.20.

表 2.20 第三次旋转运算结果

	a_1	a_2	e_3	a_3	$b_1 = 5$
e_1	$\frac{1}{2}$	$\frac{5}{4}$	$-\frac{1}{4}$	$\frac{3}{4}$	$\frac{3}{2}$
e_2	$\frac{1}{2}$	$\frac{7}{4}$	$\frac{1}{4}$	$\frac{5}{4}$	$\frac{1}{2}$
e_4	$\frac{3}{2}$	$\frac{9}{4}$	$-\frac{1}{4}$	$\frac{7}{4}$	$M + \frac{9}{2}$

此时所有非基向量偏差为正, 当前基本解是 $b_1 = 5$ 时不等式组的解. 由于 e_3 是基向量 $x_3 = 1$. 由于 e_1, e_2 和 e_4 是非基向量,

$$x_1 = -1 + \frac{3}{2} = \frac{1}{2}, x_2 = 0 + \frac{1}{2} = \frac{1}{2}, x_4 = -M + M + \frac{9}{2} = \frac{9}{2}.$$

2.5 线性不等式组的极点解

再考虑不等式组

$$\left. \begin{array}{l} a_i x = b_i, i = 1, 2, \dots, l, \\ a_i x \geq b_i, i = l+1, l+2, \dots, m. \end{array} \right\} \quad (2.5.1)$$

其中 a_i 是 n 维行向量 b_i 是标量 $i = 1, 2, \dots, m$. 不失一般性, 可假设 $a_1, a_2, \dots, a_l$ 线性无关且 $0 \leq l < n$.

作为线性不等式组, 一般来说, 向量组 $\{a_1, a_2, \dots, a_m\}$ 的秩等于向量的维数等于 n , 但也有例外. 例如 p 个发点 q 个收点运输问题数学模型的对偶共有 $p \times q$ 个约束, 这些约束都是线性不等式, 它们的系数向量是 $p + q$ 维向量. 但这些向量中最多只有 $p + q - 1$ 个线性无关, 小于向量的维数.

本节我们假设不等式组(2.5.1)的系数向量组的秩等于 n , 并用 S 表示其解集, 这是一个凸多面体. 设 $x \in S$, 如果在 S 中找不到两个不同点 $x^{(1)}$ 和 $x^{(2)}$ 以及

实数 α 使得

$$x = \alpha x^{(1)} + (1 - \alpha)x^{(2)}, \quad 0 < \alpha < 1,$$

就说 x 是 S 的极点, 是不等式组(2.5.1)的极点解.

定理 2.5.1 用 S 表示不等式组(2.5.1)的解集 $x \in S$. 那么 x 是 S 的极点的必要充分条件是 x 是基本解.

证 我们将证明, 如果 x 是不等式组的基本解, 则 x 是其解集 S 的极点. 如果 x 不是不等式组的基本解, 则它不是 S 的极点.

设 x 是由 n 个基向量 $a_i (i \in I_0 \cup I_1)$ 确定的基本解, 其中 $I_0 = \{1, 2, \dots, l\}$, $I_1 \subset \{l+1, l+2, \dots, m\}$, 即 x 是方程组

$$a_i x = b_i, \quad i \in I_0 \cup I_1$$

的惟一解. 假若 S 中有两个不同点 $x^{(1)}$ 和 $x^{(2)}$ 以及实数 α 使得

$$x = \alpha x^{(1)} + (1 - \alpha)x^{(2)}, \quad 0 < \alpha < 1,$$

那么

$$a_i [\alpha x^{(1)} + (1 - \alpha)x^{(2)}] = b_i, \quad i \in I_0 \cup I_1,$$

即

$$\alpha a_i x^{(1)} + (1 - \alpha)a_i x^{(2)} = b_i, \quad i \in I_0 \cup I_1.$$

由于

$$a_i x^{(1)} = b_i, \quad a_i x^{(2)} = b_i, \quad i \in I_0,$$

$$a_i x^{(1)} \geq b_i, \quad a_i x^{(2)} \geq b_i, \quad i \in I_1,$$

可见必有

$$a_i x^{(1)} = b_i, \quad a_i x^{(2)} = b_i, \quad i \in I_0 \cup I_1.$$

所以 $x^{(1)} = x^{(2)} = x$, 即 x 一定是 S 的极点.

设 x 不是不等式组(2.5.1)的基本解, 而仅仅是以下不足 n 个方程的解:

$$a_i x = b_i, \quad i \in I_0 \cup I_1,$$

其中 $I_0 = \{1, 2, \dots, l\}$, $I_1 \subset \{l+1, l+2, \dots, m\}$, $I_0 \cup I_1$ 中的向量线性无关, 但向量数目 $|I_0 \cup I_1| < n$. 另外

$$a_i x > b_i, \quad i \in I_2,$$

$$a_i x = b_i, \quad i \in I_3,$$

其中 I_2 和 I_3 是编号集 $\{1, 2, \dots, m\} \setminus I_0 \setminus I_1$ 的一个划分, I_3 很可能是空集或者其中的向量与 $I_0 \cup I_1$ 中的向量线性相关. 用 x 表示方程组

$$a_i x = b_i, \quad i \in I_0 \cup I_1$$

的任何一个不同于 x 的解,那么,对于充分小的正数 α ,

$$x^{(1)} = x + \alpha(\bar{x} - x) \text{ 和 } x^{(2)} = x - \alpha(\bar{x} - x)$$

不相同且都属于 S . 可见

$$x = \frac{1}{2}x^{(1)} + \frac{1}{2}x^{(2)},$$

所以 x 不是 S 的极点.

在不等式组的旋转算法中,所有等式和不等式的系数向量分为两部分,一部分是基向量,另一部是非基向量.相应地,所有等式和不等式分为基本的和非基本的两部分.基向量对应的方程组的解称为基本解.如果一个基本解满足所有非基等式和不等式,那么它是整个不等式组的解;否则,通过旋转运算将一个非基(不)等式与基不等式互换得到一个新的基本解.

为了叙述方便起见,将满足所有等式和不等式的基本解称为可行基本解.现在假设已经有了一个可行基本解 x ,根据定理 2.5.1,它是不等式组的解集 S 的极点.如果进行一次旋转运算得到另一个不同的可行基本解 $\bar{x}$,那么 x 和 $\bar{x}$ 一定是 S 的两个相邻的极点,因为旋转运算前后的两组基向量仅有一对不相同.设 x 相应的基向量的编号集为 $I_0 \cup I_1$, I_0 对应基等式, I_1 对应基不等式,非基不等式的编号集为 I_2 ,没有非基等式.设

$$a_i = \sum_{j \in I_0 \cup I_1} w_{ij} a_j, \quad i \in I_2.$$

由于 x 是可行基本解, a_i 的偏差

$$\sigma_i = a_i x - b_i \geq 0, \quad i \in I_2.$$

对于某个基不等式 $a_s x \geq b_s$, 将其右端项 b_s 增加 $\Delta (> 0)$, x 将沿着不等式组解集 S 的棱

$$E_s = \{x \in R^n \mid a_i x = b_i, i \in I_0 \cup I_1 \setminus \{s\}, a_s x \geq b_s\}$$

移动到新的点 $x(\Delta)$. 根据(2.4.3),非基向量 a_i 关于 $x(\Delta)$ 的偏差为

$$\sigma'_i = \sigma_i + w_{is} \Delta, \quad i \in I_2. \quad (2.5.2)$$

如前所述,将基不等式 $a_s x \geq b_s$ 的右端项增加 Δ 的操作记为 $a_s^{+\Delta}$,可对任何基不等式进行,以找出与 x 相邻的所有极点.这时可能有以下几种情况:

(i) 如果对于所有 $i \in I_2$ 均有 $w_{is} \geq 0$,那么,无论 Δ 多大,所有非基向量的偏差非负,因此 $x(\Delta)$ 都在 S 内,即沿棱 E_s 这个方向没有极点.

(ii) 如果存在 $i \in I_2$ 使得 $w_{is} < 0$,可求得一个有限的移动量

$$\Delta = \frac{\sigma_r}{-w_{rs}} = \min \left\{ -\frac{\sigma_i}{w_{is}} \mid w_{is} < 0, i \in I_2 \right\}.$$

(a) 若 $\Delta > 0$ 这时 $x(\Delta)$ 是 S 的一个新极点, 其基向量的编号集为 $\{r\} \cup I_0 \cup I_1 \setminus \{s\}$.

(b) 若 $\Delta = 0$ 并且存在 $w_{rj} > 0 (j \in I_1)$,以 w_{rj} 为枢轴作一次旋转运算 将出现一个新的方向 ,以搜索与 x 相邻的极点.

(c) 若 $\Delta = 0$ 并且对于所有 $j \in I_1$ 均有 $w_{rj} \leq 0$ 根据定理 2.1.3 与所有负的 w_{rj} 对应的基不等式都可看做等式 并且 $a_r x \geq b_r$ 冗余.

下面用一个例子说明如何利用参数化技术获取线性不等式组的所有极点.

例 2.5.1 求以下不等式组解集的所有极点：

$$\begin{aligned} & -3x_1 + x_2 - 2x_3 \geq -5, \\ & 2x_1 - 4x_2 - 5x_3 \geq -6, \\ & -3x_1 + 6x_2 - 8x_3 \geq -1, \\ & x_1 \geq -5, x_2 \geq -4, x_3 \geq -3. \end{aligned}$$

解 令 $a_1 = (-3, 1, -2), a_2 = (2, -4, -5), a_3 = (-3, 6, -8), e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$. 以 $x_1 \geq -5, x_2 \geq -4, x_3 \geq -3$ 为初始基本不等式组, 初始基本解为 $x^{(0)} = (-5, -4, -3)^T$, 初始基向量为 e_1, e_2, e_3 , 非基向量为 a_1, a_2, a_3 , 它们的偏差分别为 22, 27, 16, 初始表见表 2.21.

表 2.21 初 始 表

	e_1	e_2	e_3	σ_1	$e_1^{+16/3}$ $\{a_3, e_2, e_3\}$	$e_2^{+27/4}$ $\{e_1, a_2, e_3\}$	e_3^{+2} $\{e_1, e_2, a_3\}$
a_1	- 3	1	- 2	22	6	$\frac{115}{4}$	18
a_2	2	- 4	- 5	27	$\frac{113}{3}$	0	17
a_3	- 3	6	- 8	16	0	$\frac{113}{2}$	0

由 σ_i 列可见 ,所有非基向量偏差为正 ,所以 ,当前基本解

$$x_1 = -5, x_2 = -4, x_3 = -3$$

是不等式组解集 S 的一个极点, 其相应的基向量为 $\{e_1, e_2, e_3\}$.

表 2.21 的 e_1 列有两个负元素. 由于

$$\min\left\{\frac{22}{3}, \frac{16}{3}\right\} = \frac{16}{3} > 0,$$

作 $e_1^{+16/3}$ 可到达另一个极点,其相应的基向量为 $\{a_3, e_2, e_3\}$. 为求出其具体数值,将 e_1 列的数字乘以 $\frac{16}{3}$ 加到 σ_i 列,结果见表 2.21 的 $e_1^{+16/3}$ 列. 于是该极点的各分量为

$$x_1 = -5 + \frac{16}{3} = \frac{1}{3}, \quad x_2 = -4, \quad x_3 = -3.$$

表 2.21 的 e_2 列仅有一个负元素,由于 $\frac{27}{4} > 0$,作 $e_1^{+27/4}$ 可到达另一个极点,其相应的基向量为 $\{e_1, a_2, e_3\}$. 将 e_2 的数字乘以 $\frac{27}{4}$ 加到 σ_i 列,结果见表 2.21 的 $e_1^{+27/4}$ 列. 于是该极点的各分量为

$$x_1 = -5, \quad x_2 = -4 + \frac{27}{4} = \frac{11}{4}, \quad x_3 = -3.$$

同样在表中 2.21 作 e_3^{+2} 可求得另一个极点

$$x_1 = -5, \quad x_2 = -4, \quad x_3 = -3 + 2 = -1,$$

其相应的基向量为 $\{e_1, e_2, a_3\}$.

为求出其他极点,在表 2.21 中首先作 $a_3 \leftrightarrow e_1$,即以 -3 为枢轴作旋转运算,结果见表 2.22.

表 2.22 在表 2.21 作 $a_3 \leftrightarrow e_1$ 的结果

	a_3	e_2	e_3	σ_i	$e_2^{+6/5}$ $\{a_3, a_1, e_3\}$
a_1	1	- 5	6	6	0
a_2	$-\frac{2}{3}$	0	$-\frac{31}{3}$	$\frac{113}{3}$	$\frac{113}{3}$
e_1	$-\frac{1}{3}$	2	$-\frac{8}{3}$	$\frac{16}{3}$	$\frac{116}{15}$

表 2.22 的 a_3 列有两个负元素. 由于

$$\min\left\{\frac{113/3}{2/3}, \frac{16/3}{1/3}\right\} = \frac{16/3}{1/3},$$

它对应 e_1 行,该极点相应的基向量为 $\{e_1, e_2, e_3\}$,前面已经求得,不再考虑.

表 2.22 的 e_2 列仅有一个负元素,由于 $\frac{6}{5} > 0$,作 $e_2^{+6/5}$ 可到达另一个极点,其相应的基向量为 $\{a_3, a_1, e_3\}$. 将 e_2 列的数字乘以 $\frac{6}{5}$ 加到 σ_i 列,结果见表

2.22 的 $e_2^{+6/5}$ 列. 于是该极点的各分量为

$$x_1 = -5 + \frac{116}{15} = \frac{41}{15}, \quad x_2 = -4 + \frac{6}{5} = -\frac{14}{5}, \quad x_3 = -3.$$

在表 2.22 中作 $e_3^{+\Delta}$ 得不到新的极点.

在表 2.21 中作 $a_2 \leftrightarrow e_2$ 结果见表 2.23.

表 2.23 在表 2.21 中作 $a_2 \leftrightarrow e_2$ 的结果

	e_1	a_2	e_3	σ_i	$e_1^{+23/2}$ $\{a_1, a_2, e_3\}$	$e_3^{+113/31}$ $\{e_1, a_2, a_3\}$
a_1	$-\frac{5}{2}$	$-\frac{1}{4}$	$-\frac{13}{4}$	$\frac{115}{4}$	0	$\frac{524}{31}$
e_2	$\frac{1}{2}$	$-\frac{1}{4}$	$-\frac{5}{4}$	$\frac{27}{4}$	$\frac{25}{2}$	$\frac{68}{31}$
a_3	0	$-\frac{3}{2}$	$-\frac{31}{2}$	$\frac{113}{2}$	$\frac{113}{2}$	0

在表 2.23 中作 $e_1^{+\Delta}$ 结果见 $e_1^{+23/2}$ 列. 所以新的极点为

$$x_1 = -5 + \frac{23}{2} = \frac{13}{2}, \quad x_2 = -4 + \frac{25}{2} = \frac{17}{2}, \quad x_3 = -3,$$

其相应的基向量为 $\{a_1, a_2, e_3\}$.

在表 2.23 中作 $a_2^{+\Delta}$ 得不到新的极点.

在表 2.23 中作 $e_3^{+113/31}$ 结果见 $e_3^{+113/31}$ 列. 所以新的极点为

$$x_1 = -5, \quad x_2 = -4 + \frac{68}{31} = -\frac{56}{31}, \quad x_3 = -3 + \frac{113}{31} = \frac{20}{31},$$

其相应的基向量为 $\{e_1, a_2, a_3\}$.

在表 2.21 中作 $a_3 \leftrightarrow e_3$ 结果见表 2.24.

表 2.24 在表 2.21 中作 $a_3 \leftrightarrow e_3$ 的结果

	e_1	e_2	a_3	σ_i
a_1	$-\frac{9}{4}$	$-\frac{1}{2}$	$\frac{1}{4}$	18
a_2	$\frac{31}{8}$	$-\frac{31}{4}$	$\frac{5}{8}$	17
e_3	$-\frac{3}{8}$	$\frac{3}{4}$	$-\frac{1}{8}$	2

在此表中作 $e_1^{+\Delta}$, $e_2^{+\Delta}$, $a_3^{+\Delta}$ 均找不到新的极点.

在表 2.22 中作 $a_1 \leftrightarrow e_2$ 结果见表 2.25.

表 2.25		在表 2.22 中作 $a_1 \leftrightarrow e_2$ 的结果			
	a_3	a_1	e_3	σ_i	$e_3^{+113/31}$ $\{a_3, a_1, a_2\}$
e_2	$\frac{1}{5}$	$-\frac{1}{5}$	$\frac{6}{5}$	$\frac{6}{5}$	$\frac{864}{155}$
a_2	$-\frac{2}{3}$	0	$-\frac{31}{3}$	$\frac{113}{3}$	0
e_1	$\frac{1}{15}$	$-\frac{2}{5}$	$-\frac{4}{15}$	$\frac{116}{15}$	$\frac{1048}{155}$

在此表中作 $e_3^{+113/31}$ 可得一个新的极点：

$$\begin{aligned}x_1 &= -5 + \frac{1048}{155} = \frac{273}{155}, \\x_2 &= -4 + \frac{864}{155} = \frac{244}{155}, \\x_3 &= -3 + \frac{113}{31} = \frac{20}{31},\end{aligned}$$

其相应的基向量为 $\{a_3, a_1, a_2\}$.

在表 2.23 中作 $a_1 \leftrightarrow e_1$ 的结果见表 2.26.

表 2.26		在表 2.23 中作 $a_1 \leftrightarrow e_1$ 的结果			
	a_1	a_2	e_3	σ_i	
e_1	$-\frac{2}{5}$	$-\frac{1}{10}$	$-\frac{13}{10}$	$\frac{23}{2}$	
e_2	$-\frac{1}{5}$	$-\frac{3}{10}$	$-\frac{19}{10}$	$\frac{25}{2}$	
a_3	0	$-\frac{3}{2}$	$-\frac{31}{2}$	$\frac{113}{2}$	

在表 2.23 中作 $a_3 \leftrightarrow e_3$ 的结果见表 2.27.

表 2.27 在表 2.23 中作 $a_3 \leftrightarrow e_3$ 的结果

	e_1	a_2	a_3	σ_i
a_1	$-\frac{5}{2}$	$\frac{2}{31}$	$\frac{13}{62}$	$\frac{524}{31}$
e_2	$\frac{1}{2}$	$-\frac{4}{31}$	$\frac{5}{62}$	$\frac{68}{31}$
e_3	0	$-\frac{3}{31}$	$-\frac{2}{31}$	$\frac{113}{31}$

在表 2.25 中作 $a_2 \leftrightarrow e_3$ 的结果见表 2.28.

表 2.28 在表 2.25 中作 $a_2 \leftrightarrow e_3$ 的结果

	a_3	a_1	a_2	σ_i
e_2	$\frac{19}{155}$	$-\frac{1}{5}$	$-\frac{18}{155}$	$\frac{864}{155}$
e_3	$-\frac{2}{31}$	0	$-\frac{3}{31}$	$\frac{113}{31}$
e_1	$\frac{13}{155}$	$-\frac{2}{5}$	$\frac{4}{155}$	$\frac{1048}{155}$

以上三张表都没有新的相邻极点 ,所以此不等式组的解集共有 8 个极点 ,现列于表 2.29.

表 2.29 不等式组解集的极点及对应的基向量

编号	极点(x_1, x_2, x_3)	基向量
1	$(-5, -4, -3)$	$\{e_1, e_2, e_3\}$
2	$(\frac{1}{3}, -4, -3)$	$\{a_3, e_2, e_3\}$
3	$(-5, \frac{11}{4}, -3)$	$\{e_1, a_2, e_3\}$
4	$(-5, -4, -1)$	$\{e_1, e_2, a_3\}$
5	$(\frac{41}{15}, -\frac{14}{5}, -3)$	$\{a_3, a_1, e_3\}$
6	$(\frac{13}{2}, \frac{17}{2}, -3)$	$\{a_1, a_2, e_3\}$
7	$(-5, -\frac{56}{31}, \frac{20}{31})$	$\{e_1, a_2, a_3\}$
8	$(\frac{273}{155}, \frac{244}{155}, \frac{20}{31})$	$\{a_3, a_1, a_2\}$

这些极点之间的关系如图 2.2 所示 ,有连线的节点表示两个极点相邻.

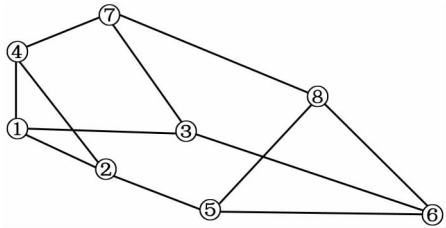


图 2.2 极点及相互关系

第三章 凸二次规划

3.1 二次规划的旋转算法

3.1.1 二次规划的 KT 条件

二次规划的 KT 条件是一个不等式组,除了互补松弛条件外,其他式子都是线性等式或线性不等式.本章和下一章介绍如何利用第二章介绍的旋转算法解其中的线性部分,并使互补松弛条件得以满足.

考虑以下形式的二次规划:

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx + cx \\ \text{s. t. } a_i x &= b_i, i = 1, 2, \dots, l, \\ a_i x &\geq b_i, i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, i = 1, 2, \dots, n. \end{aligned} \right\} \quad (3.1.1)$$

其中 $x = (x_1, x_2, \dots, x_n)^T$, $H = (h_{ij})$ 是 n 阶实对称矩阵, $c = (c_1, c_2, \dots, c_n)$, $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$, b_i 和 l_i 是有限实数.

用 λ_i 表示 $a_i x = b_i$ 或 $a_i x \geq b_i$ 对应的拉格朗日乘子, μ_i 表示 $x_i \geq l_i$ 对应的拉格朗日乘子,并用 e_i 表示 n 阶单位矩阵的第 i 行,模型(3.1.1)的 KT 条件为

$$\begin{aligned} Hx + c^T &= \sum_{i=1}^m \lambda_i a_i^T + \sum_{i=1}^n \mu_i e_i^T, \\ \lambda_i &\geq 0, \lambda_i(a_i x - b_i) = 0, \quad i = l+1, l+2, \dots, m, \\ \mu_i &\geq 0, \mu_i(x_i - l_i) = 0, \quad i = 1, 2, \dots, n, \\ a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n. \end{aligned}$$

由于

$$Hx - \sum_{i=1}^m \lambda_i a_i^T + c^T = \sum_{i=1}^n \mu_i e_i^T = (\mu_1, \mu_2, \dots, \mu_n)^T \geq 0,$$

消去所有 $\mu_i (i = 1, 2, \dots, n)$ 整理后得

$$\left. \begin{aligned} h_{i1}x_1 + \dots + h_{in}x_n - a_{i1}\lambda_1 - \dots - a_{mi}\lambda_m &\geq -c_i, \quad i = 1, 2, \dots, n, \\ a_{i1}x_1 + \dots + a_{in}x_n &= b_i, \quad i = 1, 2, \dots, l, \\ a_{i1}x_1 + \dots + a_{in}x_n &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n, \\ \lambda_i &\geq 0, \quad i = l+1, l+2, \dots, m, \\ (h_{i1}x_1 + \dots + h_{in}x_n - a_{i1}\lambda_1 - \dots - a_{mi}\lambda_m + c_i)(x_i - l_i) &= 0, \quad i = 1, 2, \dots, n, \\ (a_{i1}x_1 + \dots + a_{in}x_n - b_i)\lambda_i &= 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (3.1.2)$$

注意, 消去的

$$\mu_i = h_{i1}x_1 + \dots + h_{in}x_n - a_{i1}\lambda_1 - \dots - a_{mi}\lambda_m + c_i, \quad i = 1, 2, \dots, n.$$

3.1.2 基本概念及旋转算法表的特点

我们将用旋转算法解不等式组(3.1.2)的线性部分, 并且在计算过程中使互补松弛条件得以满足. 此不等式组共有 $n + m$ 个变量, 其中 $\lambda_1, \lambda_2, \dots, \lambda_l$ 是自由变量. 为了启动计算程序, 引入人工不等式

$$\lambda_1 \geq -M, \lambda_2 \geq -M, \dots, \lambda_l \geq -M,$$

其中 M 是充分大的数. 加上这 l 个不等式后, 共有 $2n + 2m$ 个线性(不)等式, 前 $n + m$ 个的系数向量为

$$\begin{aligned} h_i &= (h_{i1}, \dots, h_{in}, -a_{i1}, \dots, -a_{mi}), \quad i = 1, 2, \dots, n, \\ a_i &= (a_{i1}, \dots, a_{in}, 0, \dots, 0), \quad i = 1, 2, \dots, m. \end{aligned}$$

后 $n + m$ 个不等式 $x_i \geq l_i (i = 1, 2, \dots, n)$ 和 $\lambda_i \geq -M$ 或 $0 (i = 1, 2, \dots, m)$ 的系数向量用 e_i 表示, 但这里的 e_i 是 $n + m$ 阶单位矩阵的第 i 行 $i = 1, 2, \dots, n + m$.

在不等式组(3.1.2)中,

$$h_{i1}x_1 + \dots + h_{in}x_n - a_{i1}\lambda_1 - \dots - a_{mi}\lambda_m \geq -c_i \text{ 和 } x_i \geq l_i$$

称为互补不等式; 同样,

$$\lambda_i \geq 0 \text{ 和 } a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n \geq b_i$$

也称为互补不等式, 它们的系数向量称为互补向量. 其中

$$h_1, h_2, \dots, h_n, e_{n+1}, e_{n+2}, \dots, e_{n+m}$$

称为拉格朗日向量,

$$e_1, e_2, \dots, e_n, a_1, a_2, \dots, a_m$$

称为约束向量.

按照线性不等式组基本解的定义,每一对互补不等式中只要有一个是基本的,基本解就满足相应的互补松弛条件.为此,只需在计算过程中使每一对互补不等式(或互补向量)中的一个为基本的另一个为非基本的即可.

与线性不等式组的旋转算法类似,不等式组(3.1.2)的计算过程分为构造初始表、预处理和主迭代三个阶段.

初始基本不等式组为

$$x_1 \geq l_1, \dots, x_n \geq l_n, \lambda_1 \geq -M, \dots, \lambda_1 \geq -M, \lambda_{1+1} \geq 0, \dots, \lambda_m \geq 0,$$

即 $e_1, \dots, e_n, e_{n+1}, \dots, e_{n+m}$ 为初始基向量,初始基本解为

$$y^{(0)} = (l_1, \dots, l_n, -M, \dots, -M, 0, \dots, 0)^T,$$

其余线性(不)等式的系数向量 h_i, a_i 等是非基向量. h_i 的偏差为

$$\sigma_i = h_i y^{(0)} + c_i = h_{i1}l_1 + \dots + h_{in}l_n + (a_{i1} + \dots + a_{im})M + c_i, \quad i=1, 2, \dots, n,$$

a_i 的偏差为

$$\sigma_{n+i} = a_i y^{(0)} - b_i = a_{i1}l_1 + a_{i2}l_2 + \dots + a_{in}l_n - b_i, \quad i=1, 2, \dots, m.$$

各非基向量关于基向量表达式的系数及其偏差如表 3.1 所示.

表 3.1 初 始 表

	e_1	...	e_n	e_{n+1}	...	e_{n+m}	
h_1	h_{11}	...	h_{1n}	$-a_{11}$	...	$-a_{m1}$	σ_1
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
h_n	h_{n1}	...	h_{nn}	$-a_{n1}$	...	$-a_{mn}$	σ_n
a_1	a_{11}	...	a_{1n}	0	...	0	σ_{n+1}
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
a_m	a_{m1}	...	a_{mn}	0	...	0	σ_{n+m}

由于 $a_1, \dots, a_l$ 是等式的系数向量,首先尽可能将它们换入基中.不失一般

性,可设 $a_1, \dots, a_l$ 线性无关,并假设可沿子矩阵 $\begin{pmatrix} a_{11} & \dots & a_{1l} \\ \vdots & & \vdots \\ a_{l1} & \dots & a_{ll} \end{pmatrix}$ 的主对角线进行

1 次旋转运算.为了维持互补松弛条件,沿以上子矩阵主对角线进行旋转运算将

1 个非基等式换为基等式后 ,再沿子矩阵 $\begin{pmatrix} -a_{11} & \cdots & -a_{1l} \\ \vdots & & \vdots \\ -a_{l1} & \cdots & -a_{ll} \end{pmatrix}$ 的主对角线进行 1

次旋转运算 ,将 1 个基人工不等式换为非基不等式. 以上运算过程称为预处理. 由于基本解必须满足所有等式约束 ,等式约束入基后不再出基 ,所以基等式所在列及其对应的人工不等式所在的行都可以删掉. 这意味着人工不等式右端的 M 可以是任何实数 ,为计算方便起见 ,可令 $M = 0$.

为分析以上 2l 次旋转运算的结果以及随后的计算过程 ,将表 3.1 的数据划分为 5 行 5 列 ,如表 3.2 所示.

表 3.2 初始表的矩阵形式

	e_I	e_{II}	e_{III}	e_{IV}	e_V	
h_I	H_{11}	H_{12}	H_{13}	$-A_{11}^T$	$-A_{21}^T$	σ_I
h_{II}	H_{12}^T	H_{22}	H_{23}	$-A_{12}^T$	$-A_{22}^T$	σ_{II}
h_{III}	H_{13}^T	H_{23}	H_{33}	$-A_{13}^T$	$-A_{23}^T$	σ_{III}
a_I	A_{11}	A_{12}	A_{13}	O	O	σ_{IV}
a_{II}	A_{21}	A_{22}	A_{23}	O	O	σ_V

其中目标函数的海色矩阵 H 划分为 3×3 分块矩阵 , (A_{11}, A_{12}, A_{13}) 是等式约束的系数划分而成的 1×3 分块矩阵 , (A_{21}, A_{22}, A_{23}) 是不等式约束的系数划分而成的 1×3 分块矩阵. $h_I, h_{II}, h_{III}, a_I, a_{II}$ 和 $e_I, e_{II}, e_{III}, e_{IV}, e_V$ 分别表示各个矩阵对应的非基向量和基向量. $\sigma_I, \sigma_{II}, \sigma_{III}, \sigma_{IV}, \sigma_V$ 则是非基向量的偏差构成的向量.

上面提到的 2l 次旋转运算相当于先沿 A_{11} 的主对角线进行 1 次旋转运算 ,再沿 $-A_{11}^T$ 的主对角线进行 1 次旋转运算 ,两次应用第一章的定理 1.1.1 ,结果见表 3.3.

表 3.3 预处理的结果

	a_I	e_{II}	e_{III}	h_I	e_V	
e_{IV}	G_{11}	G_{12}	G_{13}	$-B_{11}^T$	$-B_{21}^T$	σ'_I
h_{II}	G_{12}^T	G_{22}	G_{23}	$-B_{12}^T$	$-B_{22}^T$	σ'_{II}
h_{III}	G_{13}^T	G_{23}^T	G_{33}	$-B_{13}^T$	$-B_{23}^T$	σ'_{III}
e_I	B_{11}	B_{12}	B_{13}	O	O	σ'_{IV}
e_{II}	B_{21}	B_{22}	B_{23}	O	O	σ'_V

其中

$$\begin{aligned}
 G_{11} &= A_{11}^{-T} H_{11} A_{11}^{-1}, \quad G_{12} = A_{11}^{-T} (H_{12} - H_{11} A_{11}^{-1} A_{12}), \\
 G_{13} &= A_{11}^{-T} (H_{13} - H_{11} A_{11}^{-1} A_{13}), \\
 G_{22} &= H_{22} - H_{12}^T A_{11}^{-1} A_{12} - A_{12}^T A_{11}^{-T} (H_{12} - H_{11} A_{11}^{-1} A_{12}) \\
 &= (-A_{12}^T A_{11}^{-T} \quad I) \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} -A_{11}^{-1} A_{12} \\ I \end{pmatrix}, \\
 G_{23} &= H_{23} - H_{12}^T A_{11}^{-1} A_{13} - A_{12}^T A_{11}^{-T} (H_{13} - H_{11} A_{11}^{-1} A_{13}), \\
 G_{33} &= H_{33} - H_{13}^T A_{11}^{-1} A_{13} - A_{13}^T A_{11}^{-T} (H_{13} - H_{11} A_{11}^{-1} A_{13}), \\
 B_{11} &= A_{11}^{-1}, \quad B_{12} = -A_{11}^{-1} A_{12}, \quad B_{13} = -A_{11}^{-1} A_{13}, \\
 B_{21} &= A_{21} A_{11}^{-1}, \quad B_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12}, \quad B_{23} = A_{23} - A_{21} A_{11}^{-1} A_{13}, \\
 \sigma'_I &= A_{11}^{-T} (\sigma_I - H_{11} A_{11}^{-1} \sigma_{IV}), \\
 \sigma'_{II} &= \sigma_{II} - H_{12}^T A_{11}^{-1} \sigma_{IV} - A_{12}^T A_{11}^{-T} (\sigma_I - H_{11} A_{11}^{-1} \sigma_{IV}), \\
 \sigma'_{III} &= \sigma_{III} - H_{13}^T A_{11}^{-1} \sigma_{IV} - A_{13}^T A_{11}^{-T} (\sigma_I - H_{11} A_{11}^{-1} \sigma_{IV}), \\
 \sigma'_{IV} &= -A_{11}^{-1} \sigma_{IV}, \quad \sigma'_V = \sigma_V - A_{21} A_{11}^{-1} \sigma_{IV}.
 \end{aligned}$$

其中 A_{11}^{-T} 表示 A_{11} 的逆矩阵的转置, I 为单位矩阵. 根据第一章的定理 1.1.5, 预处理阶段的旋转运算对 H 来说是合同变换, 也就是说表 3.3 左上角的 3×3 分块矩阵与 H 的正、负和零特征值的数目相同.

预处理后的计算过程称为主迭代, 其中的每张旋转算法表称为一般表. 主迭代包括主旋转和双枢轴运算两种运算形式. 主旋转是以对角线上的正元素为枢轴进行的旋转运算. 双枢轴运算是当对角元素为零时, 以该行的一个非零元素及其对称元素为枢轴先后进行的两次旋转运算. 预处理阶段进行的 21 次旋转运算便是 1 次双枢轴运算. 主旋转又分为两种, 一种是以非基拉格朗日向量入基以其对应的基约束向量出基的运算, 称为前向旋转; 另一种是以非基约束向量入基以其对应的基拉格朗日向量出基的运算, 称为后向旋转. 每作一次前向旋转, 基约束的数目减少一个; 每作一次后向旋转, 基约束的数目增加一个. 双枢轴运算则不改变基约束的数目.

现在假设表 3.3 中非基拉格朗日向量 h_{III} 行的主子矩阵 G_{33} 是一个 k 阶正定矩阵, 以其对角元素为枢轴进行 k 次前向旋转, 结果如表 3.4 所示.

表 3.4 主旋转的结果

	a_I	e_{II}	h_{III}	h_I	e_V	
e_{IV}	L_{11}	L_{12}	$-E_{11}^T$	$-E_{21}^T$	$-E_{31}^T$	σ''_I
h_{II}	L_{12}^T	L_{22}	$-E_{12}^T$	$-E_{22}^T$	$-E_{32}^T$	σ''_{II}
e_{III}	E_{11}	E_{12}	D_{11}	D_{12}	D_{13}	σ''_{III}
e_I	E_{21}	E_{22}	D_{12}^T	D_{22}	D_{23}	σ''_{IV}
a_{II}	E_{31}	E_{32}	D_{13}^T	D_{23}^T	D_{33}	σ''_V

其中

$$\begin{aligned} L_{11} &= G_{11} - G_{13} G_{33}^{-1} G_{13}^T, \quad L_{12} = G_{12} - G_{13} G_{33}^{-1} G_{23}^T, \\ L_{22} &= G_{22} - G_{23} G_{33}^{-1} G_{23}^T, \\ D_{11} &= G_{33}^{-1}, \quad D_{12} = G_{33}^{-1} B_{23}^T, \quad D_{13} = G_{33}^{-1} B_{23}^T, \quad D_{22} = B_{13} G_{33}^{-1} B_{13}^T, \\ D_{23} &= B_{13} G_{33}^{-1} B_{23}^T, \quad D_{33} = B_{23} G_{33}^{-1} B_{23}^T, \quad E_{11} = -G_{33}^{-1} G_{13}^T, \quad E_{12} = -G_{33}^{-1} G_{23}^T, \\ E_{21} &= B_{11} - B_{13} G_{33}^{-1} G_{13}^T, \quad E_{22} = B_{12} - B_{13} G_{33}^{-1} G_{13}^T, \\ E_{31} &= B_{21} - B_{23} G_{33}^{-1} G_{13}^T, \quad E_{32} = B_{22} - B_{23} G_{33}^{-1} G_{23}^T, \\ \sigma''_I &= \sigma'_I - G_{13} G_{33}^{-1} \sigma'_{III}, \quad \sigma''_{II} = \sigma'_{II} - G_{23} G_{33}^{-1} \sigma'_{III}, \\ \sigma''_{III} &= -G_{33}^{-1} \sigma'_{III}, \quad \sigma''_{IV} = \sigma'_{IV} - B_{13} G_{33}^{-1} \sigma'_{III}, \quad \sigma''_V = \sigma'_V - B_{23} G_{33}^{-1} \sigma'_{III}. \end{aligned}$$

根据第一章的定理 1.1.4, 非基约束向量行的主子矩阵 $\begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix}$ 和右下

角的 3×3 分块矩阵都是秩为 k 的正半定矩阵.

将表 3.4 的 e_{III} 和 e_I 两行合并为一行, 将 h_{III} 和 h_I 两列合并为一列, 设结果如表 3.5 所示.

表 3.5 一般表的矩阵形式

	e_B	a_B	h_B	e_{n+B}	
h_N	L_{11}	L_{12}	$-E_{11}^T$	$-E_{21}^T$	σ_h
e_{n+N}	L_{12}^T	L_{22}	$-E_{12}^T$	$-E_{22}^T$	σ_λ
e_N	E_{11}	E_{12}	D_{11}	D_{12}	σ_e
a_N	E_{21}	E_{22}	D_{12}^T	D_{22}	σ_a

为了简单起见, 表 3.5 的分块矩阵没有用新的符号. 其中的 h_N 和 h_B 是拉格朗日向量 $\{h_1, h_2, \dots, h_n\}$ 的一个划分, e_{n+N} 和 e_{n+B} 是拉格朗日向量 $\{e_{n+1},$

$e_{n+2}, \dots, e_{n+m}$ 的一个划分 e_N 和 e_B 以及 a_N 和 a_B 则分别是约束向量 $\{e_1, e_2, \dots, e_n\}$ 和 $\{a_1, a_2, \dots, a_m\}$ 的划分. 表中最后一列的 $\sigma_h, \sigma_\lambda, \sigma_e, \sigma_a$ 分别是各非基向量的偏差构成的向量.

为叙述方便起见, 我们将任意一张旋转算法表中非基拉格朗日向量行的主子矩阵称为拉格朗日矩阵, 记为 L ; 非基约束向量行的主子矩阵称为约束矩阵,

记为 Γ . 对于初始表, $L = H, \Gamma = O$; 对于表 3.5, $L = \begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix}, \Gamma =$

$\begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix}$. 表 3.5 是将初始表中的非基向量 a_B 和 h_B 与基向量 e_{n+N} 和 e_N 交

换的结果. 初始表中与 $\begin{pmatrix} L_{22} & -E_{12}^T \\ E_{12} & D_{11} \end{pmatrix}$ 对应的部分称为基矩阵.

从一般表中删掉基等式约束列及其对应的非基拉格朗日向量行, 得到的表称为简约表. 简约表中非基拉格朗日向量行的主子矩阵称为简约拉格朗日矩阵. 一般说来, 等式约束的数目不超过 n , 因而简约表有非基拉格朗日向量. 但是, 如果有个非基约束向量的偏差为 0, 同行与基不等式对应的元素都小于或等于零, 根据定理 2.1.3, 与其中的负元素对应的基不等式都可看做等式. 删掉这些基不等式列及对应的非基拉格朗日向量行, 如果表中不存在非基拉格朗日向量, 就说拉格朗日矩阵为空.

与线性不等式组的旋转算法一样, 从旋转算法表可以得到基本解. 以表 3.5 为例,

$$x_B = l_B, x_N = \sigma_e + l_N, \lambda_B = 0, \lambda_N = \sigma_\lambda.$$

另外还可得到 $\mu_B = 0, \mu_N = \sigma_h$. 其中 l_B 和 l_N 分别是基不等式和非基不等式 $x_i \geq l_i$ 的右端项 l_i 构成的向量; λ_B 和 λ_N 是拉格朗日向量 $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$ 的划分, μ_B 和 μ_N 则是 $\mu = (\mu_1, \mu_2, \dots, \mu_n)^T$ 的划分.

如果简约表的所有非基向量偏差非负, 则当前基本解 $x_B = l_B, x_N = \sigma_e + l_N$ 为二次规划(3.1.1)的 KT 点. 若 H 正半定, (3.1.1) 是凸规划, KT 点是其全局极小点, 或最优解; 否则可能连局部极小点都不是. 例如, $f(x_1, x_2) = (0, 0)$ 是二次规划

$$\begin{aligned} \min f(x) &= x_1^2 - 20x_1x_2 + 99x_2^2, \\ \text{s. t. } x_1 + x_2 &\leq 1, x_1, x_2 \geq 0 \end{aligned}$$

的 KT 点 $(x_1, x_2) = (1, 1)$ 是

$$\min f(x) = -x_1^2 - x_2^2,$$

$$\text{s. t. } x_1 + x_2 \leq 2, x_1, x_2 \geq 0$$

的 KT 点. 但这两个点都不是局部极小点. 非凸二次规划的解法将在第四章中进行讨论.

3.1.3 基础定理

引理 3.1.1 在旋转算法表中, 约束矩阵正半定, 并且它的秩与基约束的数目之和等于二次规划变量的数目 n .

证 若基约束的数目等于 n , 约束矩阵为零矩阵, 此引理显然成立. 现在假设某张旋转算法表中的约束矩阵正半定, 并且它的秩与基约束的数目之和等于 n . 若此表的拉格朗日矩阵有正对角元素, 每作一次前向旋转, 基约束减少一个, 约束矩阵的阶数和秩都增加 1 且仍然正半定 (见第一章的定理 1.1.4). 反之, 若约束矩阵不是零矩阵, 每作一次后向旋转, 基约束增加一个, 约束矩阵的阶数和秩都减少 1 且仍然正半定. 另外, 若某行的对角元素等于零, 以该行的一个非零元素及其对称元素为枢轴进行双枢轴运算, 约束矩阵不变或者是原约束矩阵合同变换的结果 (见定理 1.1.5), 而且基约束的数目不变. 所以引理成立. ■

定理 3.1.1 在旋转算法表中 $e_1, e_2, \dots, e_n$ 中的非基向量对应的主子矩阵正半定, 并且它的秩与基约束的数目之和等于 n .

证 以下提及的单位向量仅指 $e_1, e_2, \dots, e_n$, 不包括 $e_{n+1}, e_{n+2}, \dots, e_{n+m}$. 设 $a_1, a_2, \dots, a_m$ 中基向量的指标集为 I_0 , $e_1, e_2, \dots, e_n$ 中基向量的指标集为 I_1 . 由于非基单位向量对应的主子矩阵是约束矩阵的一个主子矩阵, 它正半定且其秩与基约束的数目之和小于或等于 n , 本定理则断定恰好等于 n .

显然只需考虑 $|I_0| + |I_1| < n$ 的情况. 这时, 非基单位向量对应的主子矩阵不可能是零矩阵, 否则这些单位向量因而全部 n 个单位向量都是不到 n 个基约束向量的线性组合. 所以非基单位向量对应的主子矩阵至少有一个正的对角元素. 以其为枢轴作后向旋转, 基约束数目变为 $|I_0| + |I_1| + 1$, 非基单位向量对应的主子矩阵的阶数和秩都减少 1 且仍然正半定. 若 $|I_0| + |I_1| + 1 < n$, 同样可利用后向旋转将一个非基单位向量换入基中, 一直到基约束的数目等于 n , 非基单位向量对应的主子矩阵为零矩阵或没有非基单位向量, 所以定理成立. ■

定理 3.1.2 任何一张旋转算法表都可从初始表开始, 先作双枢轴运算再作前向旋转得到.

证 只需说明从任何一张旋转算法表开始, 先进行后向旋转, 再进行双枢轴运算可以回复到初始表. 定理 3.1.1 说明, 如果当前表有非基单位向量并且其对应的主子矩阵的秩不等于零, 则可通过后向旋转将若干非基单位向量换入基中, 直到剩余非基单位向量对应的主子矩阵为零矩阵, 使得它们仅为基约束向量的

线性组合. 这时 $a_1, a_2, \dots, a_n$ 中的基向量与非基单位向量的数目相等, 并且任何一个非基单位向量关于基约束向量的表达式中, 不可能所有基向量 a_i 的系数都等于零. 也就是说, 可以利用双枢轴运算将一个非基单位向量与一个基向量 a_i 交换, 再将 a_i 对应的非基向量 e_{n+i} 换为基向量. 如此继续, 直到所有 $e_1, e_2, \dots, e_{n+m}$ 都是基向量, 即回复到初始表.

3.1.4 预处理的作用

预处理是在初始表构造出来之后, 将等式约束换为基等式的计算过程, 它类似于用代入消元法解线性方程组. 为说明这一点, 将表 3.1 的数据重新划分为表 3.6 的形式.

表 3.6 初始表的矩阵形式

	e_I	e_{II}	e_{III}	e_{IV}	
h_I	H_{11}	H_{12}	$-A_{11}^T$	$-A_{21}^T$	σ_I
h_{II}	H_{12}^T	H_{22}	$-A_{12}^T$	$-A_{22}^T$	σ_{II}
a_I	A_{11}	A_{12}	O	O	σ_{III}
a_{II}	A_{21}	A_{22}	O	O	σ_{IV}

其中 (A_{11}, A_{12}) 是等式约束的系数划分而成的 1×2 分块矩阵. 设 A_{11} 为可逆方阵, 在表 3.6 先后求 A_{11} 和 $-A_{11}^T$ 的逆阵, 结果见表 3.7.

表 3.7 预处理的结果

	a_I	e_{II}	h_I	e_{IV}	
e_{III}	G_{11}	G_{12}	$-B_{11}^T$	$-B_{21}^T$	σ'_I
h_{II}	G_{12}^T	G_{22}	$-B_{12}^T$	$-B_{22}^T$	σ'_{II}
e_I	B_{11}	B_{12}	O	O	σ'_{III}
a_{II}	B_{21}	B_{22}	O	O	σ'_{IV}

其中,

$$G_{22} = H_{22} - H_{12}^T A_{11}^{-1} A_{12} - A_{12}^T A_{11}^{-T} (H_{12} - H_{11} A_{11}^{-1} A_{12}),$$
$$B_{12} = -A_{11}^{-1} A_{12}, \quad B_{22} = A_{22} - A_{21} A_{11}^{-1} A_{12},$$

$$\sigma'_{II} = \sigma_{II} - H_{12}^T A_{11}^{-1} \sigma_{III} - A_{12}^T A_{11}^{-T} (\sigma_I - H_{11} A_{11}^{-1} \sigma_{III}) ,$$
$$\sigma'_{III} = - A_{11}^{-1} \sigma_{III} \quad \sigma'_{IV} = \sigma_{IV} - A_{21} A_{11}^{-1} \sigma_{III} .$$

另外,表 3.6 可看做是以下二次规划的初始表:

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} (x_I^T \quad x_{II}^T) \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} x_I \\ x_{II} \end{pmatrix} + (\sigma_I^T \quad \sigma_{II}^T) \begin{pmatrix} x_I \\ x_{II} \end{pmatrix} \\ \text{s. t. } &A_{11} x_I + A_{12} x_{II} = - \sigma_{III} , \\ &A_{21} x_I + A_{22} x_{II} \geq - \sigma_{IV} , \\ &x_I \geq 0 , x_{II} \geq 0 . \end{aligned} \right\} \quad (3.1.3)$$

以上模型是模型(3.1.1)经过平移变换后得到的系统.具体地说 x_I 应该是 $x_I - l_I$, x_{II} 应该是 $x_{II} - l_{II}$,但是为了简单起见,没有引入新的符号.

由等式约束得

$$x_I = - A_{11}^{-1} A_{12} x_{II} - A_{11}^{-1} \sigma_{III} .$$

将其代入(3.1.3) 整理后可得

$$\left. \begin{aligned} \min &\frac{1}{2} x_{II}^T G_{22} x_{II} + x_{II}^T \sigma'_{II} + \dots , \\ \text{s. t. } &B_{12} x_{II} \geq - \sigma'_{III} , \\ &B_{22} x_{II} \geq - \sigma'_{IV} , \\ &x_{II} \geq 0 . \end{aligned} \right\} \quad (3.1.4)$$

目标函数中的省略部分为常数项 G_{22} , B_{12} , σ'_{II} 等的表达式如前所述. 以上模型对应的旋转算法表如表 3.8 所示.

表 3.8 预处理的结果(简化表)

	e_{II}	h_I	e_{IV}	
h_{II}	G_{22}	$- B_{12}^T$	$- B_{22}^T$	σ'_{II}
e_I	B_{12}	O	O	σ'_{III}
a_{II}	B_{22}	O	O	σ'_{IV}

表 3.8 正是表 3.7 的一部分.所以,预处理相当于将一个含有等式约束的二次规划转化为一个不含等式约束的二次规划.

在此我们进一步说明了为什么基等式所在的列及其互补不等式所在的行可以删掉.

3.2 凸二次规划的计算步骤

现在假定模型(3.1.1)是一个凸二次规划,即 H 正半定. 3.1 节已指出,其计算过程分为构造初始表、预处理和主迭代三个阶段.

由定理 3.1.2、定理 1.1.5、定理 1.1.4 和定理 1.1.2 可见,计算过程中的任何一张旋转算法表具有以下特点:约束矩阵和拉格朗日矩阵正半定,并且它们的秩之和等于 H 的秩.

二次规划(3.1.1)的解法如下,其中 H 正半定.在此,不要求等式的系数向量 $a_1, a_2, \dots, a_l$ 线性无关.

算法 3.2.1 凸二次规划(3.1.1)的计算步骤.

步骤 1 初始步骤. 以 $x_1 \geq l_1, \dots, x_n \geq l_n, \lambda_1 \geq 0, \dots, \lambda_m \geq 0$ 为初始基本不等式组,列出表 3.1 所示初始表.

步骤 2 预处理. 将非基等式尽可能换为基等式,相应地,将基人工不等式换为非基不等式.

对于 $i = 1, 2, \dots, l$, 当 a_i 的偏差小于 0, 大于 0 或等于 0 时,分别转(i), (ii)和(iii).

(i) 若 a_i 行没有与基不等式对应的正元素,则原问题没有可行解,停止;否则以其中一个正元素为枢轴进行一次旋转运算,然后以其对称元素为枢轴进行一次旋转运算.

(ii) 若 a_i 行没有与基不等式对应的负元素,则原问题没有可行解,停止;否则以其中一个负元素为枢轴进行一次旋转运算,然后以其对称元素为枢轴进行一次旋转运算.

(iii) 若 a_i 行与基不等式对应的元素均为 0,则去掉 a_i 行和 e_{n+i} 列($a_i x = b_i$ 是冗余等式);否则,以其中一个非零元素为枢轴进行一次旋转运算,然后以其对称元素为枢轴进行一次旋转运算.

步骤 3 主迭代. 进行非基不等式与基不等式的交换.

(i) 若所有非基向量(不包括基等式对应的非基向量)偏差非负,则当前基本解(与 x_i 对应的分量)是原问题的最优解,停止;否则转(ii).

(ii) 以一个偏差是负数的非基向量入基. 如果入基向量行没有与基不等式对应的正元素,原问题没有最优解,停止;否则转(iii).

(iii) 如果入基向量行的对角元素为正,以其为枢轴进行一次旋转运算,返回(i);否则以该行与基不等式对应的一个正元素为枢轴进行一次旋转运算,然后以其对称元素为枢轴进行一次旋转运算,返回(i).

例 3.2.1 解二次规划

$$\begin{aligned} \min \quad & \frac{1}{2}x^T Hx + cx, \\ \text{s. t.} \quad & -x_1 - 2x_2 + x_3 + 4x_4 = 1, \\ & 2x_1 - 6x_2 + x_3 - x_4 \geq -1, \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq -1, x_4 \geq 1. \end{aligned}$$

$$\text{其中 } H = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 2 & 5 & 0 & 3 \\ -1 & 0 & 5 & 1 \\ 1 & 3 & 1 & 2 \end{pmatrix}, c = (1, -2, 0, 4).$$

解 沿 H 的主对角线进行两次旋转运算可以看到残余矩阵是一个 2×2 零矩阵, 所以 H 是一个秩为 2 的正半定矩阵. 以上二次规划的 KT 条件为

$$\begin{aligned} x_1 + 2x_2 - x_3 + x_4 + \lambda_1 - 2\lambda_2 &\geq -1, \\ 2x_1 + 5x_2 + 3x_4 + 2\lambda_1 + 6\lambda_2 &\geq 2, \\ -x_1 + 5x_3 + x_4 - \lambda_1 - \lambda_2 &\geq 0, \\ x_1 + 3x_2 + x_3 + 2x_4 - 4\lambda_1 + \lambda_2 &\geq -4, \\ -x_1 - 2x_2 + x_3 + 4x_4 &= 1, \\ 2x_1 - 6x_2 + x_3 - x_4 &\geq -1, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq -1, x_4 \geq 1, \lambda_2 &\geq 0, \\ (x_1 + 2x_2 - x_3 + x_4 + \lambda_1 - 2\lambda_2 + 1)x_1 &= 0, \\ (2x_1 + 5x_2 + 3x_4 + 2\lambda_1 + 6\lambda_2 - 2)x_2 &= 0, \\ (-x_1 + 5x_3 + x_4 - \lambda_1 - \lambda_2)(x_3 + 1) &= 0, \\ (x_1 + 3x_2 + x_3 + 2x_4 - 4\lambda_1 + \lambda_2 + 4)(x_4 - 1) &= 0, \\ (2x_1 - 6x_2 + x_3 - x_4 + 1)\lambda_2 &= 0. \end{aligned}$$

记

$$\begin{aligned} h_1 &= (1, 2, -1, 1, 1, -2), \\ h_2 &= (2, 5, 0, 3, 2, 6), \\ h_3 &= (-1, 0, 5, 1, -1, -1), \\ h_4 &= (1, 3, 1, 2, -4, 1), \\ a_1 &= (-1, -2, 1, 4, 0, 0), \\ a_2 &= (2, -6, 1, -1, 0, 0), \\ e_1 &= (1, 0, 0, 0, 0, 0), \\ e_2 &= (0, 1, 0, 0, 0, 0), \\ e_3 &= (0, 0, 1, 0, 0, 0), \end{aligned}$$

$$e_4 = (0, 0, 0, 1, 0, 0),$$

$$e_6 = (0, 0, 0, 0, 0, 1).$$

引入人工不等式 $\lambda_1 \geq 0$ 并令

$$e_5 = (0, 0, 0, 0, 1, 0).$$

以 $x_1 \geq 0, x_2 \geq 0, x_3 \geq -1, x_4 \geq 1, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式组, 初始基本解为 $y^{(0)} = (0, 0, -1, 1, 0, 0)^T$. 初始基向量为 $e_1, e_2, e_3, e_4, e_5, e_6$. 非基向量为 $h_1, h_2, h_3, h_4, a_1, a_2$, 它们的偏差分别为 $3, 1, -4, 5, 2, -1$, 初始表见表 3.9.

表 3.9 初 始 表

	e_1	e_2	e_3	e_4	e_5	e_6	
h_1	1	2	- 1	1	1	- 2	3
h_2	2	5	0	3	2 *	6	1
h_3	- 1	0	5	1	- 1	- 1	- 4
h_4	1	3	1	2	- 4	1	5
a_1	- 1	- 2 *	1	4	0	0	2
a_2	2	- 6	1	- 1	0	0	- 1

由于 a_1 是等式的系数向量, 以 a_1 入基. a_1 的偏差 2 是正数, 以该行的最小负元素 - 2 为枢轴进行一次旋转运算, 然后以其对称元素为枢轴进行一次旋转运算, 结果见表 3.10. 其中 a_1 列及 e_5 行已删掉.

表 3.10 预处理结果

	e_1	e_3	e_4	h_2	e_6	
h_1	0.25	- 1.25	- 1.5	0.5	- 5 *	2
h_3	- 1.25	6.25	7.5	- 0.5	2	- 1
h_4	- 1.5	7.5	34	- 2	13	20
e_2	- 0.5	0.5	2	0	0	1
a_2	5 *	- 2	- 13	0	0	- 7

在表 3.10 中 a_2 的偏差最小, 以该行的正元素 5 及其对称元素 - 5 为枢轴先后进行两次旋转运算, 结果见表 3.11.

表 3.11 双枢轴运算结果

	a ₂	e ₃	e ₄	h ₂	h ₁	
e ₆	0.01	- 0.23	- 0.17	0.1	- 0.2	0.47
h ₃	- 0.23	5.29 *	3.91	- 0.3	- 0.4	- 1.81
h ₄	- 0.17	3.91	27.89	- 0.7	- 2.6	24.01
e ₂	- 0.1	0.3	0.7	0	0	0.3
e ₁	0.2	0.4	2.6	0	0	1.4

在表 3.11 中仅 h₃ 的偏差是负数，以该行的对角元素 5.29 为枢轴作旋转运算 结果见表 3.12.

表 3.12 前向旋转运算结果

	a ₂	h ₃	e ₄	h ₂	h ₁	
e ₆	0	- 0.0435	0	0.087	- 0.2174	0.3913
e ₃	0.0435	0.189	- 0.7391	0.0567	0.0756	0.3422
h ₄	0	0.7391	25	- 0.4783	- 2.3043	25.3478
e ₂	- 0.087	0.0567	0.4783	0.017	0.0227	0.4023
e ₁	0.2174	0.0756	2.3043	0.0227	0.0302	1.5369

表 3.12 的所有非基向量偏差为正 ,最优解为

$x_1 = 1.5369$, $x_2 = 0.4026$, $x_3 = 0.3422 + l_3 = 0.3422 - 1 = - 0.6578$,
 $x_4 = l_4 = 1$.

最优目标值为 12.7354. 由表 3.12 还可以看到

$\lambda_2 = 0.3913$, $\mu_1 = \mu_2 = \mu_3 = 0$, $\mu_4 = 25.3478$.

非基向量 e₅ 行已删掉 ,若保留该行可从其偏差得到 $\lambda_1 = - 4.2174$.

3.3 变量有上界的凸二次规划

下面考虑变量有上界限制的凸二次规划 模型如下：

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx + cx \\ \text{s. t. } a_i x &= b_i, \quad i=1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i=l+1, l+2, \dots, m, \\ u_i &\geq x_i \geq l_i, \quad i=1, 2, \dots, n. \end{aligned} \right\} \quad (3.3.1)$$

当然, 我们可以把变量的上界约束 $u_i \geq x_i (i=1, 2, \dots, n)$ 都看做一般约束, 按上一节的方法计算, 但这样做计算量较大. 由于 $x_i \geq l_i$ 和 $-x_i \geq -u_i$ 两不等式系数向量的符号相反, 可简化计算. 用 λ_i 表示第 i 个约束对应的拉格朗日乘子, $i=1, 2, \dots, m$; μ_i 和 $\bar{\mu}_i$ 分别表示 $x_i \geq l_i$ 和 $-x_i \geq -u_i$ 对应的拉格朗日乘子, $i=1, 2, \dots, n$. 此二次规划的 KT 条件为

$$\left. \begin{aligned} h_{11}x_1 + \dots + h_{1n}x_n - a_{1l}\lambda_1 - \dots - a_{ml}\lambda_m + c_1 - \mu_i + \bar{\mu}_i &= 0, \quad i=1, 2, \dots, n, \\ \lambda_i &\geq 0, \lambda_i(a_i x - b_i) = 0, \quad i=l+1, l+2, \dots, m, \\ \mu_i &\geq 0, \mu_i(x_i - l_i) = 0, \bar{\mu}_i \geq 0, \bar{\mu}_i(-x_i + u_i) = 0, \quad i=1, 2, \dots, n, \\ a_i x &= b_i, \quad i=1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i=l+1, l+2, \dots, m, \\ x_i &\geq l_i, -x_i \geq -u_i, \quad i=1, 2, \dots, n. \end{aligned} \right\} \quad (3.3.2)$$

去掉与不等式 $x_i \geq l_i$ 对应的乘子 $\mu_i (i=1, 2, \dots, n)$ 得到与之等价的不等式组:

$$\left. \begin{aligned} h_{11}x_1 + \dots + h_{1n}x_n - a_{1l}\lambda_1 - \dots - a_{ml}\lambda_m + \bar{\mu}_i &\geq -c_1, \\ (h_{11}x_1 + \dots + h_{1n}x_n - a_{1l}\lambda_1 - \dots - a_{ml}\lambda_m + c_1 + \bar{\mu}_i)(x_i - l_i) &= 0, \quad i=1, 2, \dots, n, \\ \lambda_i &\geq 0, \lambda_i(a_i x - b_i) = 0, \quad i=l+1, l+2, \dots, m, \\ \bar{\mu}_i &\geq 0, \bar{\mu}_i(-x_i + u_i) = 0, \quad i=1, 2, \dots, n, \\ a_i x &= b_i, \quad i=1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i=l+1, l+2, \dots, m, \\ x_i &\geq l_i, -x_i \geq -u_i, \quad i=1, 2, \dots, n. \end{aligned} \right\} \quad (3.3.3)$$

此不等式组可用算法 3.2.1 计算, 即将模型 (3.3.1) 中的所有上界约束 $-x_i \geq -u_i$ 当做一般不等式求解.

不等式组 (3.3.3) 共有 $4n+2m-1$ 个线性(不)等式和 $2n+m$ 个变量, 对应大小为 $(2n+m) \times (2n+m+1)$ 的旋转算法表, 其大部分元素为零. 下面说明如

何进一步消去所有 $\bar{\mu}_i$, 使表格大小减少为 $(n+m) \times (n+m+1)$.

对于任何 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ 和每个 $i \in \{1, 2, \dots, n\}$, $x_i \geq l_i$ 和 $-x_i \geq -u_i$ 不可能都是紧约束, 所以 μ_i 和 $\bar{\mu}_i$ 至少有一个为 0. 记

$$h_i(x, \lambda) = h_{i1}x_1 + \dots + h_{in}x_n - a_{i1}\lambda_1 - \dots - a_{im}\lambda_m + c_i.$$

由不等式组(3.3.2)可见, 若 $\bar{x}_i = l_i$, 则 $\bar{\mu}_i = 0$, 因而 $\mu_i = h_i(x, \lambda) \geq 0$; 若 $\bar{x}_i = u_i$, 则 $\mu_i = 0$, 因而 $\bar{\mu}_i = -h_i(x, \lambda) \geq 0$; 若 $\bar{x}_i$ 既不等于 l_i 也不等于 u_i , 则 $\mu_i = \bar{\mu}_i = 0$, 于是 $h_i(x, \lambda) = 0$. 所以在计算过程中, 可以或者仅使用 $h_i(x, \lambda) \geq 0$, 或者仅使用 $-h_i(x, \lambda) \geq 0$. 不难验证, 如果求得不等式组

$$\left. \begin{aligned} h_i(x, \lambda) &\geq 0, x_i \geq l_i, h_i(x, \lambda)(x_i - l_i) = 0, \quad i \in I_1, \\ -h_i(x, \lambda) &\geq 0, -x_i \geq -u_i, -h_i(x, \lambda)(-x_i + u_i) = 0, \quad i \in I_2, \\ a_i x &\geq b_i, \lambda_i \geq 0, \lambda_i(a_i x - b_i) = 0, \quad i = 1 + 1, 1 + 2, \dots, m, \\ a_i x &= b_i, \quad i = 1, 2, \dots, l \end{aligned} \right\} \quad (3.3.4)$$

的一个解 $\bar{x}$, 并且其每个分量还满足 $-x_i \geq -u_i (i \in I_1)$ 和 $x_i \geq l_i (i \in I_2)$, 则 $\bar{x}$ 是不等式组(3.3.2)的解, 因而是模型(3.3.1)的最优解. 其中 I_1 和 I_2 是 $\{1, 2, \dots, n\}$ 的一个划分, 即 $I_1 \cup I_2 = \{1, 2, \dots, n\}$, $I_1 \cap I_2 = \emptyset$.

不等式组(3.3.4)的解法如下. 首先令 $I_1 = \{1, 2, \dots, n\}$, $I_2 = \emptyset$, 即不考虑变量上界, 解不等式组(3.1.2). 前两个计算步骤与算法 3.2.1 的步骤 1 和步骤 2 相同. 此时 $-e_1, \dots, -e_n, -h_1, \dots, -h_n$ 还未出现于表中. 在以后的计算过程中, 不仅要使基本解满足 $x_i \geq l_i$, 还要使基本解满足 $-x_i \geq -u_i$, 即不仅要使 e_i 的偏差非负, 还要使 $-e_i$ 的偏差非负, $i = 1, 2, \dots, n$. 反过来说, 如果 e_i 或 $-e_i$ 的偏差是负数, 则该向量是一个可能的入基向量. 但是这两个向量仅有一个出现于表中, 下面说明如何处理没有出现于表中的那个单位向量. 设当前基本解为 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$.

(i) 如果 e_i 是基向量, 则 $\bar{x}_i = l_i$, 因而 $-e_i$ 的偏差 $-\bar{x}_i + u_i > 0$. 所以不考虑 $-e_i$. 同样, 如果 $-e_i$ 是基向量, 则可以不考虑 e_i .

(ii) 设 e_i 是出现于当前表中的非基向量, 其偏差 $\sigma_i = \bar{x}_i - l_i > 0$. 但是 $-e_i$ 的偏差 $-\bar{x}_i + u_i = u_i - l_i - \sigma_i < 0$ 并以 $-e_i$ 入基. 为进行以 $-e_i$ 入基的旋转运算, 可先作如下操作: 将非基向量 e_i 换为 $-e_i$, 将 e_i 的偏差 σ_i 换为 $u_i - l_i - \sigma_i$, 将第 e_i 行其他数据全部反号, 然后将基向量 h_i 列所有数据反号, 并将 h_i 换为 $-h_i$. 我们将这样的操作称为对 e_i 作向量替换. 这相当于将非基不等式 $x_i \geq l_i$ 换为 $-x_i \geq -u_i$, 将基不等式 $h_i(x, \lambda) \geq 0$ 换为 $-h_i(x, \lambda) \geq 0$. 由于这两个基

不等式对应的等式均是 $h_i(x, \lambda) = 0$ 基本解不变, 所以各非基向量偏差不变. 同样, 如果 $-e_i$ 是出现于当前表中的非基向量, 其偏差大于 0, 但 e_i 的偏差小于 0 并以 e_i 入基, 则先进行将 $-x_i \geq -u_i$ 换为 $x_i \geq l_i$, 将 $-h_i(x, \lambda) \geq 0$ 换为 $h_i(x, \lambda) \geq 0$ 的操作, 并称之为对 $-e_i$ 作向量替换.

不难证明, 将一个正半定矩阵的第 i 行和第 i 列所有元素反号 (主对角线上第 i 个元素不变), 得到的矩阵仍然正半定. 所以, 向量替换后的拉格朗日矩阵和约束矩阵仍然正半定.

按照线性不等式组的旋转算法, 入基向量确定后就要在入基向量行与基不等式对应的元素中找一个正元素进行旋转运算. 如果这些元素都小于或等于 0, 那么入基向量行在约束矩阵或在拉格朗日矩阵中的元素全部等于 0, 因为这两个矩阵正半定. 下面进一步说明“入基向量行与基不等式对应的元素都小于或等于 0”意味着什么以及如何处理.

(i) 设约束向量 $a'_i (= a_i, e_i \text{ 或 } -e_i)$ 入基. 由于此时 a'_i 仅仅是基约束向量的线性组合, 其中与基不等式对应的系数非正. 根据第二章的定理 2.1.1(i), 二次规划没有可行解.

(ii) 设拉格朗日向量 h_i 的偏差 $\sigma_i < 0$ 并企图以 h_i 入基, 但 h_i 行的元素都小于或等于 0. 可先作如下处理: 将非基向量 h_i 换为 $-h_i$, 将 h_i 行所有数据反号, 然后将基向量 e_i 列所有数据反号, 并将 e_i 换为 $-e_i$. 这相当于将非基不等式 $h_i(x, \lambda) \geq 0$ 换为 $-h_i(x, \lambda) \geq 0$, 将基不等式 $x_i \geq l_i$ 换为 $-x_i \geq -u_i$. 这两个基不等式对应的等式为 $x_i = l_i$ 和 $x_i = u_i$, 右端项的改变量为 $u_i - l_i$, 所以基本解发生了变化. 根据参数公式 (2.4.3) 并考虑到 $-e_i$ 列的数据已经反了号, 将 $-e_i$ 列的数据乘以 $l_i - u_i$ 然后加到最后一列, 便得到各非基向量关于新基本解的偏差. 由于 $-e_i$ 列的对角元素等于 0, $-h_i$ 的偏差为 $-\sigma_i > 0$, 所以无需进行以 $-h_i$ 入基的旋转运算. 我们将以上操作称为对 h_i 作向量替换. 同样, 如果拉格朗日向量 $-h_i$ 的偏差小于 0 并企图以其入基, 但 $-h_i$ 行的元素都小于或等于 0, 则进行将 $-h_i(x, \lambda) \geq 0$ 换为 $h_i(x, \lambda) \geq 0$, 将 $-x_i \geq -u_i$ 换为 $x_i \geq l_i$ 的操作, 并称之为对 $-h_i$ 作向量替换.

(iii) 设拉格朗日向量 e_{n+i} 的偏差小于 0 并企图以其入基, 但 e_{n+i} 行的元素都小于或等于 0. 由于 e_{n+i} 是第 $n+i$ 个分量为 1 的 $n+m$ 维单位向量, h_j 和 $-h_j$ 是前 n 个和后 m 个分量都可能不等于 0 的向量, 约束向量则是后 m 个分量都为 0 的 $n+m$ 维向量, 所以 e_{n+i} 关于基向量的表达式中一定有 h_j 或 $-h_j$. 设 e_{n+i} 行的元素 $w_{ij} < 0$, w_{ij} 对应基向量 $h'_j (= h_j \text{ 或 } -h_j)$. 这时, 将 h'_j 换为 $-h'_j$, 将该列所有元素反号, 将非基向量 $e'_j (= e_j \text{ 或 } -e_j)$ 换为 $-e'_j$, 并将 e'_j 的偏差 σ_j 换为 $u_j - l_j - \sigma_j$, 并将该行其余元素反号. 这样的操作称为对 e_{n+i} 作向量替换.

以下算法中提到的非基向量除表中列出的非基向量外, 还包括未列出的单

位向量 e_i 或负单位向量 $-e_i$, $i = 1, 2, \dots, n$ 其中 h'_i 表示 h_i 或 $-h_i$, e'_i 表示 e_i 或 $-e_i$, $i = 1, 2, \dots, n$.

算法 3.3.1 二次规划(3.3.1)的计算步骤.

步骤 1 初始步骤. 同算法 3.2.1 的步骤 1.

步骤 2 预处理. 同算法 3.2.1 的步骤 2.

步骤 3 主迭代.

(i) 若所有非基向量(不包括与基等式对应的非基向量)偏差非负, 则当前基本解(与 x_i 对应的分量)是模型(3.3.1)的最优解, 停止; 否则以偏差最小的非基向量入基.

(ii) 若入基向量为 a_i , 转(a); 若为 e'_i , 转(b); 若为 h'_i , 转(c); 若为 e_{n+i} , 转(d).

(a) 如果 a_i 行与基不等式对应的所有元素均小于或等于 0, 原问题没有可行解, 停止. 如果 a_i 行的对角元素为正, 以其为枢轴进行一次旋转运算, 返回(i); 否则以该行与基不等式对应的一个正元素为枢轴进行一次旋转运算, 然后以其对称元素为枢轴进行一次旋转运算, 返回(i).

(b) 如果 e'_i 未出现于当前表中, 对 e'_i 作向量替换. 若该行与基不等式对应的所有元素均小于或等于 0, 原问题没有可行解, 停止. 如果该行的对角元素为正, 以其为枢轴进行一次旋转运算, 返回(i); 否则以该行与基不等式对应的一个正元素为枢轴进行一次旋转运算, 然后以其对称元素为枢轴进行一次旋转运算, 返回(i).

(c) 若 h'_i 行的对角元素为正, 以其为枢轴进行一次旋转运算, 返回(i). 若 h'_i 行有一个与基不等式对应的正元素, 以其为枢轴进行一次旋转运算, 然后以其对称元素为枢轴进行一次旋转运算, 返回(i); 否则, 对 h'_i 作向量替换, 返回(i).

(d) 若 e_{n+i} 行的对角元素为正, 以其为枢轴进行一次旋转运算, 返回(i). 若 e_{n+i} 行有一个与基不等式对应的正元素, 以其为枢轴进行一次旋转运算, 然后以其对称元素为枢轴进行一次旋转运算, 返回(i); 否则, 对 e_{n+i} 作向量替换, 返回(i).

例 3.3.1 解二次规划

$$\begin{aligned} \min f(x) &= 0.5x_1^2 + 2.5x_2^2 + x_3^2 - 2x_1x_2 - x_1x_3 + 3x_2x_3 + 6x_1 - 2x_2 + x_3, \\ \text{s. t. } x_1 + 4x_2 + 5x_3 &= 4, \\ &\quad -x_1 - 6x_2 + 2x_3 \geq -8, \\ &\quad 2 \geq x_1 \geq 0, 1 \geq x_2 \geq -1, x_3 \geq 0. \end{aligned}$$

解 此问题的目标函数可化为

$$f(x) = \frac{1}{2}[(x_1 - 2x_2 - x_3 + 6)^2 + (x_2 + x_3 + 10)^2 - 6x_3 - 136],$$

所以是凸函数. 此二次规划的 KT 条件为

$$\begin{aligned}
 & x_1 - 2x_2 - x_3 - \lambda_1 + \lambda_2 + 6 - \mu_1 + \bar{\mu}_1 = 0, \\
 & -2x_1 + 5x_2 + 3x_3 - 4\lambda_1 + 6\lambda_2 - 2 - \mu_2 + \bar{\mu}_2 = 0, \\
 & -x_1 + 3x_2 + 2x_3 - 5\lambda_1 - 2\lambda_2 + 1 - \mu_3 = 0, \\
 & x_1 + 4x_2 + 5x_3 = 4, \\
 & -x_1 - 6x_2 + 2x_3 \geq -8, \\
 & 2 \geq x_1 \geq 0, \quad 1 \geq x_2 \geq -1, \quad x_3 \geq 0, \\
 & \lambda_2, \mu_1, \bar{\mu}_1, \mu_2, \bar{\mu}_2, \mu_3 \geq 0, \\
 & (-x_1 - 6x_2 + 2x_3 + 8)\lambda_2 = 0, \\
 & x_1\mu_1 = (-x_1 + 2)\bar{\mu}_1 = (x_2 + 1)\mu_2 = (-x_2 + 1)\bar{\mu}_2 = x_3\mu_3 = 0.
 \end{aligned}$$

按照算法 3.3.1, 首先解以下不等式组:

$$\begin{aligned}
 & x_1 - 2x_2 - x_3 - \lambda_1 + \lambda_2 \geq -6, \\
 & -2x_1 + 5x_2 + 3x_3 - 4\lambda_1 + 6\lambda_2 \geq 2, \\
 & -x_1 + 3x_2 + 2x_3 - 5\lambda_1 - 2\lambda_2 \geq -1, \\
 & x_1 + 4x_2 + 5x_3 = 4, \\
 & -x_1 - 6x_2 + 2x_3 \geq -8, \\
 & x_1 \geq 0, \quad x_2 \geq -1, \quad x_3 \geq 0, \quad \lambda_2 \geq 0, \\
 & (x_1 - 2x_2 - x_3 - \lambda_1 + \lambda_2 + 6)x_1 = 0, \\
 & (-2x_1 + 5x_2 + 3x_3 - 4\lambda_1 + 6\lambda_2 - 2)(x_2 + 1) = 0, \\
 & (-x_1 + 3x_2 + 2x_3 - 5\lambda_1 - 2\lambda_2 + 1)x_3 = 0, \\
 & (-x_1 - 6x_2 + 2x_3 + 8)\lambda_2 = 0.
 \end{aligned}$$

引入人工不等式 $\lambda_1 \geq 0$, 记

$$\begin{aligned}
 h_1 &= (1, -2, -1, -1, 1), \\
 h_2 &= (-2, 5, 3, -4, 6), \\
 h_3 &= (-1, 3, 2, -5, -2), \\
 a_1 &= (1, 4, 5, 0, 0), \\
 a_2 &= (-1, -6, 2, 0, 0),
 \end{aligned}$$

并用 e_i 表示 5 阶单位矩阵的第 i 行 $i = 1, 2, 3, 4, 5$.

以 $x_1 \geq 0, x_2 \geq -1, x_3 \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式组, 初始基本解为 $y^{(0)} = (0, -1, 0, 0, 0)^T$. 初始基向量为 e_1, e_2, e_3, e_4, e_5 . 非基向量为 h_1, h_2, h_3, a_1, a_2 , 它们的偏差分别为 8, -7, -2, -8, 14. 初始表见表 3.13. 由于 a_1 是等式的系数向量 a_1 入基. 该行最大正元素为 5, 先后作旋转运算 $a_1 \leftrightarrow e_3, h_3 \leftrightarrow e_4$ 结果见表 3.14.

表 3.13 初 始 表

	e_1	e_2	e_3	e_4	e_5	
h_1	1	- 2	- 1	- 1	1	8
h_2	- 2	5	3	- 4	6	- 7
h_3	- 1	3	2	- 5 *	- 2	- 2
a_1	1	4	5 *	0	0	- 8
a_2	- 1	- 6	2	0	0	14

表 3.14 预处理结果

	e_1	e_2	a_1	h_3	e_5	
h_1	$\frac{37}{25}$	$-\frac{37}{25}$	$-\frac{7}{25}$	$\frac{1}{5}$	$\frac{7}{5}$	$\frac{154}{25}$
h_2	$-\frac{37}{25}$	$\frac{37}{25}$ *	$\frac{7}{25}$	$\frac{4}{5}$	$\frac{38}{5}$	$-\frac{79}{25}$
e_4	$-\frac{7}{25}$	$\frac{7}{25}$	$\frac{2}{25}$	$-\frac{1}{5}$	$-\frac{2}{5}$	$\frac{6}{25}$
e_3	$-\frac{1}{5}$	$-\frac{4}{5}$	$\frac{1}{5}$	0	0	$\frac{8}{5}$
a_2	$-\frac{7}{5}$	$-\frac{38}{5}$	$\frac{2}{5}$	0	0	$\frac{86}{5}$

在表 3.14 中以 h_2 入基以 e_2 出基 结果见表 3.15.

表 3.15 第一次主旋转结果

	e_1	h_2	a_1	h_3	e_5	σ_i	$\tilde{\sigma}_i$
h_1	0	- 1	0	1	9	3	—
e_2	1	$\frac{25}{37}$	$-\frac{7}{37}$	$-\frac{20}{37}$	$-\frac{190}{37}$	$\frac{79}{37}$	$-\frac{5}{37}$
e_4	0	$\frac{7}{37}$	$\frac{1}{37}$	$-\frac{13}{37}$	$-\frac{68}{37}$	$\frac{31}{37}$	—
e_3	- 1	$-\frac{20}{37}$	$\frac{13}{37}$	$\frac{16}{37}$	$\frac{152}{37}$	$-\frac{4}{37}$	—
a_2	- 9	$-\frac{190}{37}$	$\frac{68}{37}$	$\frac{152}{37}$	$\frac{1\ 444}{37}$	$\frac{36}{37}$	—

由于 e_2 已是非基向量且偏差是正数,需考虑 x_2 是否超过其上界,即 $-e_2$ 的偏差是否为负数. $-e_2$ 的偏差等于 $u_2 - l_2 - \frac{79}{37} = 1 - (-1) - \frac{79}{37} = -\frac{5}{37}$,见表 3.15 的 $\bar{\sigma}_i$ 列. $-e_2$ 的偏差 $-\frac{5}{37}$ 是最小负偏差,应以 $-e_2$ 入基.为此先将表中的 e_2 换为 $-e_2$,将 $\frac{79}{37}$ 换为 $-\frac{5}{37}$,并将该行其余各元素反号;再将 h_2 换为 $-h_2$,并将该列各元素反号,结果见表 3.16.

表 3.16 对 e_2 作向量替换的结果

	e_1	$-h_2$	a_1	h_3	e_5	σ_i	$\bar{\sigma}_i$
h_1	0	1	0	1	9	3	—
$-e_2$	-1	$\frac{25}{37}^*$	$\frac{7}{37}$	$\frac{20}{37}$	$\frac{190}{37}$	$-\frac{5}{37}$	$\frac{79}{37}$
e_4	0	$-\frac{7}{37}$	$\frac{1}{37}$	$-\frac{13}{37}$	$-\frac{68}{37}$	$\frac{31}{37}$	—
e_3	-1	$\frac{20}{37}$	$\frac{13}{37}$	$\frac{16}{37}$	$\frac{152}{37}$	$-\frac{4}{37}$	—
a_2	-9	$\frac{190}{37}$	$\frac{68}{37}$	$\frac{152}{37}$	$\frac{1444}{37}$	$\frac{36}{37}$	—

在表 3.16 中以 $-e_2$ 入基以 $-h_2$ 出基,结果见表 3.17.

表 3.17 第二次主旋转结果

	e_1	$-e_2$	a_1	h_3	e_5	σ_i
h_1	$\frac{37}{25}$	$\frac{37}{25}$	$-\frac{7}{25}$	$\frac{1}{5}$	$\frac{7}{5}$	$\frac{16}{5}$
$-h_2$	$\frac{37}{25}$	$\frac{37}{25}$	$-\frac{7}{25}$	$-\frac{4}{5}$	$-\frac{38}{5}$	$\frac{1}{5}$
e_4	$-\frac{7}{25}$	$-\frac{7}{25}$	$\frac{2}{25}$	$-\frac{1}{5}$	$-\frac{2}{5}$	$\frac{4}{5}$
e_3	$-\frac{1}{5}$	$\frac{4}{5}$	$\frac{1}{5}$	0	0	0
a_2	$-\frac{7}{5}$	$\frac{38}{5}$	$\frac{2}{5}$	0	0	2

至此所有非基向量偏差非负,原问题的最优解为

$$x_1 = l_1 = 0, \quad x_2 = u_2 = 1, \quad x_3 = 0 + l_1 = 0.$$

由表 3.17 还可得到 $\lambda_1 = \frac{4}{5}, \lambda_2 = 0$. 由于 h_1 是非基向量, $\bar{\mu}_1 = 0, \mu_1 = \frac{16}{5}$. 由于 $-h_2$ 是非基向量 $\mu_2 = 0, \bar{\mu}_2 = \frac{1}{5}$. 由于 h_3 是基向量 $\mu_3 = 0$. 可以验证这些解满足二次规划的 KT 条件. 以上计算还可以简化. 因为 a_1 是等式的系数向量, e_4 是人工不等式的系数向量, 表 3.14 至表 3.17 中的 a_1 列和 e_4 行均可以删掉.

例 3.3.2 解二次规划

$$\begin{aligned} \min \quad & \frac{1}{2}x^T Hx + cx, \\ \text{s. t.} \quad & x_1 - 2x_2 \geq -2, \\ & x_1 - x_2 + 2x_3 \geq 5, \\ & 6 \geq x_1 \geq 0, 4 \geq x_2 \geq 0, 4 \geq x_3 \geq 0, \end{aligned}$$

其中 $H = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 2 & 0 \\ -1 & 0 & 2 \end{pmatrix}, c = (-1.25, -1, 3).$

解 先不考虑各变量的上界,其初始表见表 3.18.

表 3.18 初 始 表

	e_1	e_2	e_3	e_4	e_5	
h_1	1	-1	-1	-1	-1*	-1.25
h_2	-1	2	0	2	1	-1
h_3	-1	0	2	0	-2	3
a_1	1	-2	0	0	0	2
a_2	1*	-1	2	0	0	-5

由于非基向量 a_2 的偏差 -5 最小,以 a_2 入基. a_2 行 e_1 列的元素 1 是正数,不妨作旋转运算 $a_2 \leftrightarrow e_1$,再作 $h_1 \leftrightarrow e_5$,结果见表 3.19. 其中 $\bar{\sigma}_1$ 列的 1 是 $-e_1$ 的偏差.

表 3.19 第一次双枢轴运算结果

	a_2	e_2	e_3	e_4	h_1	σ_i	$\bar{\sigma}_i$
e_5	1	0	- 3	- 1	- 1	3. 75	—
h_2	0	1	- 1	1	- 1	- 2. 25	—
h_3	- 3	- 1	10 *	2	2	- 9. 5	—
a_1	1	- 1	- 2	0	0	7	—
e_1	1	1	- 2	0	0	5	1

在表 3.19 中 ,非基向量 h_3 的偏差 - 9.5 最小 ,以 h_3 入基. h_3 行的对角元素是正数 ,作主旋转 $h_3 \leftrightarrow e_3$,结果见表 3.20.

表 3.20 第一次主旋转结果

	a_2	e_2	h_3	e_4	h_1	σ_i	$\bar{\sigma}_i$
e_5	0. 1	- 0. 3	- 0. 3	- 0. 4	- 0. 4	0. 9	—
h_2	- 0. 3	0. 9 *	- 0. 1	1. 2	- 0. 8	- 3. 2	—
e_3	0. 3	0. 1	0. 1	- 0. 2	- 0. 2	0. 95	3. 05
a_1	0. 4	- 1. 2	- 0. 2	0. 4	0. 4	5. 1	—
e_1	0. 4	0. 8	- 0. 2	0. 4	0. 4	3. 1	2. 9

在表 3.20 中再作 $h_2 \leftrightarrow e_2$,结果见表 3.21.

表 3.21 第二次主旋转结果

	a_2	h_2	h_3	e_4	h_1	σ_i	$\bar{\sigma}_i$
e_5	0	- 0. 3333	- 0. 3333	0	- 0. 6667	- 0. 1667	—
e_2	0. 3333	1. 1111	0. 1111	- 1. 3333	0. 8889	3. 5556	0. 4444
e_3	0. 3333	0. 1111	0. 1111	- 0. 3333	- 0. 1111	1. 3056	2. 6944
a_1	0	- 1. 3333	- 0. 3333	2	- 0. 6667	0. 8333	—
e_1	0. 6667	0. 8889	- 0. 1111	- 0. 6667	1. 1111	5. 9444	0. 0556

在表 3.21 中 ,仅 e_5 的偏差是负数 ,但此行所有元素都小于或等于零 ,不妨

取该行的第一个负元素 - 0. 3333. 此元素对应基向量 h_2 , 于是将 h_2 换为 $- h_2$ 并将 e_2 换为 $- e_2$. 为此 将 e_2 的偏差 3. 5556 换为

$$u_2 - l_2 - 3. 5556 = 4 - 3. 5556 = 0. 4444 ,$$

并将此行其他数据反号 然后将第二列所有数据反号 结果为表 3. 22.

表 3. 22 向量替换的结果

	a_2	$- h_2$	h_3	e_4	h_1	σ_i	$\bar{\sigma}_i$
e_5	0	0. 3333 *	- 0. 3333	0	- 0. 6667	- 0. 1667	—
$- e_2$	- 0. 3333 *	1. 1111	- 0. 1111	1. 3333	- 0. 8889	0. 4444	3. 5556
e_3	0. 3333	- 0. 1111	0. 1111	- 0. 3333	- 0. 1111	1. 3056	2. 6944
a_1	0	1. 3333	- 0. 3333	2	- 0. 6667	0. 8333	—
e_1	0. 6667	- 0. 8889	- 0. 1111	- 0. 6667	1. 1111	5. 9444	0. 0556

在表 3. 22 作双枢轴运算($e_5 \leftrightarrow - h_2$, $- e_2 \leftrightarrow a_2$) , 结果为表 3. 23.

表 3. 23 第二次双枢轴运算结果

	$- e_2$	e_5	h_3	e_4	h_1	σ_i	$\bar{\sigma}_i$
$- h_2$	0	3	1	0	2	0. 5	—
a_2	- 3	10	3	4	4	3	—
e_3	- 1	3	1	1	1	2. 25	1. 75
a_1	0	4	1	2	2	1. 5	—
e_1	- 2	4	1	2	2	7. 5	- 1. 5

由表 3. 23 可见仅 $- e_1$ 的偏差 $- 1. 5 < 0$,但 $- e_1$ 未出现于表中. 所以将 e_1 换为 $- e_1$ 将其偏差 7. 5 换为 $- 1. 5$, 并将该行其余数字反号 ;再将 h_1 换为 $- h_1$, 并将该列数字反号. 在得到的表中作主旋转 $- e_1 \leftrightarrow - h_1$, 结果见表 3. 24.

表 3.24 第三次主旋转结果

	$-e_2$	e_5	h_3	e_4	$-e_1$	σ_i	$\bar{\sigma}_i$
$-h_2$	2 *	- 1	0	- 2	- 1	- 1	—
a_2	1	2	1	0	- 2	0	—
e_3	0	1	0.5	0	- 0.5	1.5	2.5
a_1	2	0	0	0	- 1	0	—
$-h_1$	- 1	2	0.5	1	0.5	0.75	0.75

在表 3.24 中作主旋转 $-h_2 \leftrightarrow -e_2$, 结果见表 3.25.

表 3.25 第四次主旋转结果

	$-h_2$	e_5	h_3	e_4	$-e_1$	σ_i	$\bar{\sigma}_i$
$-e_2$	0.5	0.5	0	1	0.5	0.5	3.5
a_2	0.5	2.5	1	1	- 1.5	0.5	—
e_3	0	1	0.5	0	- 0.5	1.5	2.5
a_1	1	1	0	2	0	1	—
$-h_1$	- 0.5	1.5	0.5	0	0	0.25	—

此表所有非基向量偏差非负, 最优解为

$$x_1 = u_1 = 6, \quad x_2 = 4 - 0.5 = 3.5, \quad x_3 = 1.5,$$

对应的目标函数值 $f(x) = -4$. 还可看出 $\bar{\mu}_1 = 0.25$, 其余拉格朗日乘子均等于零.

第四章 非凸二次规划

4.1 二次齐次函数的下降方向

考虑 n 个变量的二次齐次函数 $x^T H x$, 其中 $x = (x_1, x_2, \dots, x_n)^T$,

$$H = \begin{pmatrix} h_{11} & h_{12} & \dots & h_{1n} \\ h_{12} & h_{22} & \dots & h_{2n} \\ \vdots & \vdots & & \vdots \\ h_{1n} & h_{2n} & \dots & h_{nn} \end{pmatrix}$$

是实对称矩阵. 如果存在一个 $d \in R^n$ 使得 $d^T H d < 0$, 称 d 为 $x^T H x$ 或 H 的下降方向.

根据正半定矩阵的定义, 当 H 正半定时, $x^T H x$ 没有下降方向; 当 H 负定 (即 $-H$ 正定) 时, 任何一个非零的 n 维向量都是下降方向. 所以我们一般考虑不定矩阵的下降方向.

容易看出, 如果 d 是 H 的下降方向, 则 $-d$ 也是 H 的下降方向; 如果 d 是 H 的下降方向, 则 d 的充分小的邻域内的一切点也是 H 的下降方向. 另外, H 的下降方向的集合还具有以下引理描述的性质.

引理 4.1.1 如果 x 和 y 是 H 的下降方向, 那么对于任何不全为 0 的非负实数 α 和 β , 或者 $\alpha x + \beta y$ 是 H 的下降方向, 或者 $\alpha x - \beta y$ 是 H 的下降方向.

证 因为

$$(\alpha x^T + \beta y^T) H (\alpha x \pm \beta y) = \alpha^2 x^T H x + \beta^2 y^T H y \pm 2\alpha\beta x^T H y,$$

如果 $x^T H y \leq 0$, 则

$$(\alpha x^T + \beta y^T) H (\alpha x + \beta y) \leq \alpha^2 x^T H x + \beta^2 y^T H y < 0,$$

即 $\alpha x + \beta y$ 是 H 的下降方向; 否则 $(\alpha x^T - \beta y^T) H (\alpha x - \beta y) < 0$, 即 $\alpha x - \beta y$ 是 H 的下降方向. ■

从几何上看, H 的所有下降方向的集合是一个没有边界的“对顶锥”. 例如, H 为二阶矩阵

$$\begin{pmatrix} 1 & -2 \\ -2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

时,下降方向的集合分别如图 4.1(a), (b), (c) 的阴影区域所示,不包括边界线.

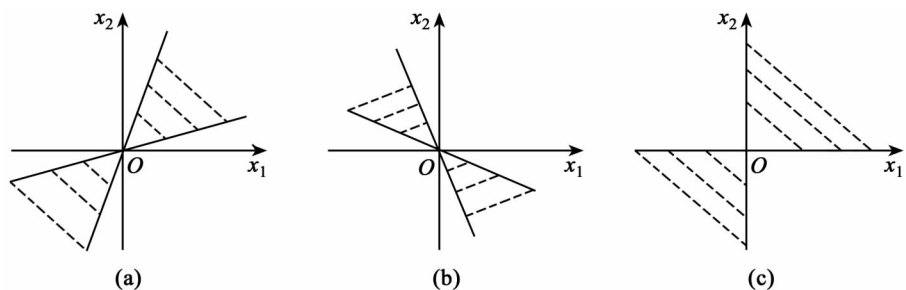


图 4.1 下降方向的集合

4.2 I 型和 II 型非负下降方向

4.2.1 非负下降方向的产生方法

仍然考虑二次齐次函数 $x^T H x$. 如果有一个向量 $d \geq 0$, 使得 $d^T H d < 0$, 称 d 为 $x^T H x$ 或 H 的非负下降方向.

显然, 如果 H 的一个主子矩阵有非负下降方向, 则 H 有非负下降方向. 特殊地讲:

(i) 如果 H 的第 i 个对角元素是负数, 则 H 有非负下降方向. 例如 $h_{ii} < 0$, 令 d 为第 i 个分量等于 Δ 、其他分量等于零的 n 维向量, 则 $d^T H d = h_{ii} \Delta^2 < 0$, 其中 $\Delta > 0$. 我们称仅含一个负对角元素的主子矩阵 (h_{ii}) 为 H 的 I 型矩阵, 并称上面的 d 为 H 的一个(明显的)I 型非负下降方向.

(ii) 如果 H 含有一个 2 阶主子矩阵, 其两个对角元素为零, 非对角元素为负, 则 H 有非负下降方向. 例如, H 第 i, j 两行(列)的主子矩阵为 $\begin{pmatrix} 0 & h_{ij} \\ h_{ij} & 0 \end{pmatrix}$, 其中 $h_{ij} < 0$, 令 d 为第 i, j 两个分量等于某个正数 Δ 、其他分量等于零的 n 维向量, 则 $d^T H d = 2h_{ij}\Delta^2 < 0$. 此时我们说 $\begin{pmatrix} 0 & h_{ij} \\ h_{ij} & 0 \end{pmatrix}$ 是 H 的 II 型矩阵, 并称 d 是 H 的一

个(明显的)II型非负下降方向.

(iii) 设 $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$ 其中 H_{11} 为 k 阶可逆方阵 相应地, 令 $x = \begin{pmatrix} x_N \\ x_B \end{pmatrix}$. 可

以验证

$$x^T H x = (x_N^T + x_B^T H_{12}^T H_{11}^{-1}) H_{11} (x_N + H_{11}^{-1} H_{12} x_B) + x_B^T (H_{22} - H_{12}^T H_{11}^{-1} H_{12}) x_B. \quad (4.2.1)$$

(4.2.1)式称为 $x^T H x$ 的关于 H_{11} 的展开式. 出于算法的需要, 我们总是让 H_{11} 正定. 由此公式可见, 如果 $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 有非负下降方向, 即存在 $d_B \geq 0$ 使得

$$d_B^T (H_{22} - H_{12}^T H_{11}^{-1} H_{12}) d_B < 0$$

并且 $-H_{11}^{-1} H_{12} d_B \geq 0$ 那么

$$\left. \begin{aligned} x_B &= d_B, \\ x_N &= -H_{11}^{-1} H_{12} d_B \end{aligned} \right\} \quad (4.2.2)$$

是 H 的非负下降方向.

设 $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 的第 i 个对角元素 $w_{ii} < 0$, d_B 为第 i 个分量等于某个正数 Δ 、其他分量等于零的 $n-k$ 维向量. 当 $-H_{11}^{-1} H_{12} d_B \geq 0$ 时, 称(4.2.2)为 H 的一个(隐含的)I型非负下降方向.

设 $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 有个主子矩阵 $\begin{pmatrix} 0 & w_{ij} \\ w_{ij} & 0 \end{pmatrix}$, 其中 $w_{ij} < 0$, d_B 为第 i, j 两

个分量等于某个正数 Δ 、其他分量等于零的 $n-k$ 维向量. 当 $-H_{11}^{-1} H_{12} d_B \geq 0$ 时, 称(4.2.2)为 H 的一个(隐含的)II型非负下降方向.

显然 H 满足以下条件之一时, 它没有非负下降方向:

- (i) H 正半定;
- (ii) H 是一个非负矩阵(即 H 的所有元素非负);
- (iii) 在公式(4.2.1)中 H_{11} 正定, $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 为非负矩阵.

从约束最优化的角度看, 本节的问题可表示为

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} x^T H x, \\ \text{s. t. } x &\geq 0. \end{aligned} \right\} \quad (4.2.3)$$

其中 H 不定, $x = (x_1, x_2, \dots, x_n)^T$. 此问题的解有两种情形, 或者 $x = 0$ 是全局极小点, 或者目标函数在可行域(即 n 维空间第一象限)上无下界. 这是一个最简单的非凸非线性规划, 但却是一个 NP 完全问题(见 Murty & Kabadi (1987)), 至

今没有多项式时间的算法. 但是我们将提供一种简便方法, 较好地解决这个问题.

用 μ_i 表示 $x_i \geq 0$ 对应的拉格朗日乘子, 模型(4.2.3)的 KT 条件为

$$\begin{aligned} h_{i1}x_1 + h_{i2}x_2 + \dots + h_{in}x_n &= \mu_i, \quad i=1, 2, \dots, n, \\ \mu_i &\geq 0, \quad \mu_i x_i = 0, \quad i=1, 2, \dots, n, \\ x_i &\geq 0, \quad i=1, 2, \dots, n. \end{aligned}$$

化简得

$$\left. \begin{aligned} h_{i1}x_1 + h_{i2}x_2 + \dots + h_{in}x_n &\geq 0, \quad i=1, 2, \dots, n, \\ (h_{i1}x_1 + h_{i2}x_2 + \dots + h_{in}x_n)x_i &= 0, \quad i=1, 2, \dots, n, \\ x_i &\geq 0, \quad i=1, 2, \dots, n. \end{aligned} \right\} \quad (4.2.4)$$

我们将用旋转算法解此不等式组的线性部分, 并使互补松弛条件得以满足. 初始基本不等式组为 $x_i \geq 0, i=1, 2, \dots, n$, 初始基本解为 $x^{(0)} = 0$, 初始基向量为 e_i , 非基向量为 $h_i, i=1, 2, \dots, n$. 其中 e_i 表示 n 阶单位矩阵的第 i 行, h_i 表示 H 的第 i 行. 非基向量 h_i 关于 $x^{(0)}$ 的偏差为 $h_i x^{(0)} = 0, i=1, 2, \dots, n$. 各非基向量关于基向量的组合系数如表 4.1 所示, 暂时不管其最后两列. 由于所有非基向量的偏差等于 0, 没有在表的右侧列出这些 0.

表 4.1

初 始 表

	e_1	e_2	...	e_n	$e_1^{+\Delta}$	$e_1^{+\Delta} \quad e_2^{+\Delta}$
h_1	h_{11}	h_{12}	...	h_{1n}	$h_{11}\Delta$	$h_{11}\Delta + h_{12}\Delta$
h_2	h_{21}	h_{22}	...	h_{2n}	$h_{21}\Delta$	$h_{21}\Delta + h_{22}\Delta$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
h_n	h_{n1}	h_{n2}	...	h_{nn}	$h_{n1}\Delta$	$h_{n1}\Delta + h_{n2}\Delta$

表 4.1 的基本解为 $x^{(0)} = 0$, 但我们可利用第二章的参数公式(2.4.3)或(2.4.2)改变某些基不等式右端的常数项从而改变基本解, 使其对应的目标函数值小于 0.

(i) 设初始表有一个负对角元素, 例如 $h_{11} < 0$. 作 $e_1^{+\Delta}$, 即将基不等式 $x_1 \geq 0$ 的右端项变为一个正数 Δ . 这时各变量的数值为

$$\left. \begin{aligned} x_1 &= \Delta, \\ x_i &= 0, \quad i = 2, 3, \dots, n. \end{aligned} \right\} \quad (4.2.5)$$

按照第二章的公式(2.4.3), 各非基向量的偏差如表 4.1 的右数第二列所示. 这

时

$$2f(x) = x^T H x = h_{11} \Delta^2 < 0,$$

所以(4.2.5)式是 H 的一个下降方向, 并且是上面提到的 I 型非负下降方向.

(ii) 设初始表有一个 2 阶主子矩阵, 其两个对角元素等于 0, 非对角元素是负数. 例如 $h_{11} = h_{22} = 0, h_{12} < 0$. 作 $e_1^{+\Delta}$ 和 $e_2^{+\Delta}$, 即将基不等式 $x_1 \geq 0$ 和 $x_2 \geq 0$ 的右端项都变为正数 Δ . 这时各变量的数值为

$$\left. \begin{aligned} x_1 = x_2 = \Delta, \\ x_i = 0, \quad i = 3, 4, \dots, n. \end{aligned} \right\} \quad (4.2.6)$$

按照第二章的公式(2.4.2), 各非基向量的偏差如表 4.1 的最后一列所示. 这时

$$2f(x) = x^T H x = 2h_{12} \Delta^2 < 0,$$

所以(4.2.6)式是 H 的一个下降方向, 并且是上面提到的 II 型非负下降方向.

(iii) 现在设初始表左上角的 $k \times k$ 矩阵正定, 并沿主对角线进行 k 次旋转运算, 设结果如表 4.2 所示.

表 4.2 k 次前向旋转的结果

	h_1	...	h_k	e_{k+1}	...	e_n	$e_{k+1}^{+\Delta}$	$e_{k+1}^{+\Delta} e_n^{+\Delta}$
e_1	w_{11}	...	w_{1k}	$w_{1,k+1}$	...	w_{1n}	$w_{1,k+1} \Delta$	$w_{1,k+1} \Delta + w_{1n} \Delta$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
e_k	w_{k1}	...	w_{kk}	$w_{k,k+1}$	...	w_{kn}	$w_{k,k+1} \Delta$	$w_{k,k+1} \Delta + w_{kn} \Delta$
h_{k+1}	$w_{k+1,1}$	...	$w_{k,k+1}$	$w_{k+1,k+1}$	...	$w_{k+1,n}$	$w_{k+1,k+1} \Delta$	$w_{k+1,k+1} \Delta + w_{k+1,n} \Delta$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
h_n	w_{n1}	...	w_{kn}	$w_{n,k+1}$	...	w_{nn}	$w_{n,k+1} \Delta$	$w_{n,k+1} \Delta + w_{nn} \Delta$

设表 4.2 中的对角元素 $w_{k+1,k+1} < 0$, 作 $e_{k+1}^{+\Delta}$, 即将基不等式 $x_{k+1} \geq 0$ 的右端项变为正数 Δ . 按照第二章的公式(2.4.3), 各非基向量的偏差如表 4.2 的右数第二列所示. 这时基本解为

$$\left. \begin{aligned} x_{k+1} = \Delta, \\ x_i = 0, \quad i = k+2, k+3, \dots, n, \\ x_i = w_{i,k+1} \Delta, \quad i = 1, 2, \dots, k, \end{aligned} \right\} \quad (4.2.7)$$

其对应的函数值

$$2f(x) = x^T H x = \sum_{i=1}^n x_i h_i x = x_{k+1} h_{k+1} x = \Delta w_{k+1,k+1} \Delta = w_{k+1,k+1} \Delta^2 < 0.$$

若 $w_{i,k+1} (i=1, 2, \dots, k)$ 均非负, (4.2.7) 就是上面提到的隐含的 I 型非负下降方向.

(iv) 若表 4.2 中的对角元素 $w_{k+1,k+1}$ 和 w_{nn} 等于 0, 非对角元素 $w_{k+1,n} = w_{n,k+1}$ 是负数, 作 $e_{k+1}^{+\Delta}$ 和 $e_n^{+\Delta}$, 即将基不等式 $x_{k+1} \geq 0$ 和 $x_n \geq 0$ 的右端项变为正数 Δ , 按照第二章的公式 (2.4.2), 各非基向量的偏差如表 4.2 的最后一列所示. 这时基本解为

$$\left. \begin{aligned} x_{k+1} &= x_n = \Delta, \\ x_i &= 0, \quad i = k+2, k+3, \dots, n-1, \\ x_i &= w_{i,k+1}\Delta + w_{in}\Delta, \quad i = 1, 2, \dots, k, \end{aligned} \right\} \quad (4.2.8)$$

其对应的函数值

$$2f(x) = x^T H x = \sum_{i=1}^n x_i h_i x = x_{k+1} h_{k+1} x + x_n h_n x = 2w_{k+1,n} \Delta^2 < 0.$$

若 $w_{i,k+1} + w_{in} (i=1, 2, \dots, k)$ 均非负, (4.2.8) 就是上面提到的隐含的 II 型非负下降方向.

为看出表 4.2 与 H 的展开式 (4.2.1) 之间的关系, 将表 4.1 中用标量表示的矩阵 H 划分为 2×2 矩阵 $\begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$, 其中 H_{11} 为 k 阶正定矩阵. 上面提到的 k 次旋转运算结果用矩阵表示即为表 4.3.

表 4.3 k 次前向旋转的结果 (矩阵形式)

	$h_1 \quad \dots \quad h_k$	$e_{k+1} \quad \dots \quad e_n$
e_1 $\vdots$ e_k	H_{11}^{-1}	$-H_{11}^{-1}H_{12}$
h_{k+1} $\vdots$ h_n	$H_{12}^T H_{11}^{-1}$	$H_{22} - H_{12}^T H_{11}^{-1} H_{12}$

可见 (4.2.1) 中的两个矩阵都出现在表 4.3 中.

4.2.2 非负下降方向的有关定理

为了解 H 的 I 型和 II 型非负下降方向的特点, 进一步引入以下概念. 设 d 是 H 的非负下降方向. 若 d 的所有分量为正, 称其为非退化的非负下降方向. 若

d 至少有一个分量等于 0, 则称其为退化的非负下降方向.

引理 4.2.1 如果实对称矩阵 H 有非负下降方向, 则有非退化的非负下降方向.

证 设 d 是 H 的非负下降方向. 如果 d 的所有分量为正, 则 d 是非退化的非负下降方向. 否则, 不妨设其前 k 个分量为正, 这些分量构成的向量记为 d_N ; 后 $n-k$ 个分量均等于 0, 这些分量构成的向量记为 d_B . 相应地将矩阵 H 划分为

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \text{ 则}$$

$$(d_N^T, d_B^T) \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} d_N \\ d_B \end{pmatrix} = d_N^T H_{11} d_N + 2d_N^T H_{12} d_B + d_B^T H_{22} d_B < 0.$$

显然, 将零向量 d_B 的每个分量换为充分小的正数, 上式仍然成立, 于是得到 H 的一个非退化的非负下降方向. ■

定理 4.2.1 如果 n 阶实对称矩阵 H 有一个非负下降方向, 也存在一个下降方向, 其至少有一个分量为正一个分量为负, 则 H 存在一个退化的非负下降方向, 或者说, H 的一个 k 阶主子矩阵有非负下降方向, $1 \leq k < n$.

证 由于 H 有非负下降方向, 根据引理 4.2.1, H 有非退化的非负下降方向, 设其为 $x = (x_1, x_2, \dots, x_n)^T$, 其所有分量为正. 设 $y = (y_1, y_2, \dots, y_n)^T$ 是 H 的另一个下降方向, 其中至少有一个分量为正一个分量为负. 根据引理 4.1.1, 对于任何非负实数 θ , 或者

$$x + \theta y = (x_1 + \theta y_1, x_2 + \theta y_2, \dots, x_n + \theta y_n)^T$$

是 H 的下降方向, 或者

$$x - \theta y = (x_1 - \theta y_1, x_2 - \theta y_2, \dots, x_n - \theta y_n)^T$$

是 H 的下降方向.

若 $x + \theta y$ 是 H 的下降方向, 让 θ 由 0 逐渐增加, 直至 $x + \theta y$ 有一个分量为 0, 其余分量非负, 则 $x + \theta y$ 成为一个退化的非负下降方向. 反之, 若 $x - \theta y$ 是下降方向, 则存在一个正的 θ 使 $x - \theta y$ 成为 H 的退化非负下降方向. 也就是说 H 的一个主子矩阵有非负下降方向. ■

定理 4.2.2 设 $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$, 其中 H_{11} 正定. 如果 $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 没

有非负下降方向, 那么 H_{22} 也没有非负下降方向.

证 $H_{22} - H_{12}^T H_{11}^{-1} H_{12}$ 没有非负下降方向即对于任何非负向量 y 均有

$$y^T (H_{22} - H_{12}^T H_{11}^{-1} H_{12}) y \geq 0.$$

所以

$$y^T H_{22} y \geq y^T H_{12}^T H_{11}^{-1} H_{12} y \geq 0,$$

即 H_{22} 没有非负下降方向. ■

定理 4.2.3 用 $\tilde{H}$ 表示去掉矩阵 H 的非负行和列后得到的矩阵, 则 H 有非负下降方向的必要充分条件是 $\tilde{H}$ 有非负下降方向.

证 记 $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$ 并且设 H_{12}^T 和 H_{22} 的所有元素非负, 于是 $\tilde{H} = H_{11}$.

相应地, 将 x 划分为 x_N 和 x_B , 则

$$x^T H x = \begin{pmatrix} x_N^T & x_B^T \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} x_N \\ x_B \end{pmatrix} = x_N^T \tilde{H} x_N + 2x_N^T H_{12} x_B + x_B^T H_{22} x_B.$$

若 x_N 是 $\tilde{H}$ 的非负下降方向, 即 $x_N^T \tilde{H} x_N < 0$, 令 $x_B = 0$, 则得到 H 的非负下降方向. 反之, 若 $x = \begin{pmatrix} x_N \\ x_B \end{pmatrix}$ 是 H 的非负下降方向, 由于 $x_N^T H_{12} x_B$ 和 $x_B^T H_{22} x_B$ 非

负, $x_N^T \tilde{H} x_N \leq x^T H x < 0$, 即 x_N 是 $\tilde{H}$ 的非负下降方向. ■

根据定理 4.2.3, 下面假设 H 的每行(列)至少有一个负元素. 容易看出, 对于每行(列)都有负元素的二阶矩阵 $H = \begin{pmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{pmatrix}$, 仅当满足以下条件之一时, 它才有非负下降方向:

- (i) H 有个负对角元素;
- (ii) H 的两个对角元素均为 0, 非对角元素为负;
- (iii) H 的一个对角元素为 0, 另一个为正, 非对角元素是负数;
- (iv) H 的两个对角元素均是正数, 非对角元素是负数, 且 $h_{11} h_{22} - h_{12}^2 < 0$.

在情况(iii)和(iv), 以对角线上任何一个正元素为枢轴作旋转运算便可得到 I 型非负下降方向.

对于 $n \geq 3$ 的情况, 设 H 的所有对角元素非负, 第 i 个对角元素 $h_{ii} = 0$, 非对角元素 $h_{ij} < 0$. 那么, 如果 $h_{ij} = 0$, H 有一个 II 型非负下降方向; 如果 $h_{ij} > 0$, 以其为枢轴作旋转运算便可得到 I 型非负下降方向, 此情况与上面(iii)的情况类似. 所以, 下面只需讨论 H 的所有对角元素是正数且每行(列)至少有一个负的非对角元素的情况.

模型(4.2.3)可以看做是一般二次规划模型(3.1.1)的特殊情况. 所以 $h_1, h_2, \dots, h_n$ 是拉格朗日向量, 旋转算法表中非基拉格朗日向量行和基单位向量列

上的元素构成的矩阵是拉格朗日矩阵,非基单位向量行和基拉格朗日向量列上的元素构成的矩阵是约束矩阵.

定理 4.2.4 考虑 n 阶实对称矩阵 $H = (h_{ij})$, 设其所有对角元素为正且每行(列)至少有一个负的非对角元素. 如果 H 有非负下降方向, 那么它有 I 型或 II 型非负下降方向.

证 下面归纳地证明 H 有 I 型或 II 型非负下降方向.

(i) 上面已说明了 $n = 2$ 的情况, 让我们从 $n = 3$ 开始, 设

$$H = \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{12} & h_{22} & h_{23} \\ h_{13} & h_{23} & h_{33} \end{pmatrix}.$$

用 h_1, h_2, h_3 分别表示 H 的三行, 用 e_1, e_2, e_3 分别表示三阶单位矩阵的 3 行. 以 e_1, e_2, e_3 为初始基向量, 将 h_1, h_2, h_3 表示它们的线性组合, 其表格形式如表 4.4 所示.

表 4.4 初 始 表

	e_1	e_2	e_3
h_1	h_{11}	h_{12}	h_{13}
h_2	h_{12}	h_{22}	h_{23}
h_3	h_{13}	h_{23}	h_{33}

不妨以 h_{11} 为枢轴进行旋转运算, 结果如表 4.5 所示.

表 4.5 第一次旋转运算结果

	h_1	e_2	e_3
e_1	$\frac{1}{h_{11}}$	$-\frac{h_{12}}{h_{11}}$	$-\frac{h_{13}}{h_{11}}$
h_2	$\frac{h_{12}}{h_{11}}$	$h_{22} - \frac{h_{12}^2}{h_{11}}$	$h_{23} - \frac{h_{12}h_{13}}{h_{11}}$
h_3	$\frac{h_{13}}{h_{11}}$	$h_{23} - \frac{h_{12}h_{13}}{h_{11}}$	$h_{33} - \frac{h_{13}^2}{h_{11}}$

(a) 若表 4.5 中的 $h_{22} - \frac{h_{12}^2}{h_{11}}$ 和 $h_{33} - \frac{h_{13}^2}{h_{11}}$ 均非正但至少有一个为负, 不妨设前

者为负. 作 $e_2^{+\Delta}$ 得基本解 $x_2 = \Delta, x_3 = 0, x_1 = -\frac{h_{12}\Delta}{h_{11}}$, 且 $x^T H x = (h_{22} - \frac{h_{12}^2}{h_{11}})\Delta^2 < 0$, 其中 $\Delta > 0$. 若表 4.5 中的 $-\frac{h_{12}}{h_{11}} \geq 0$, x 是 I 型非负下降方向; 否则 x 既有正分量也有负分量. 根据定理 4.2.1, H 的一个二阶主子矩阵有非负下降方向, 因而 H 有 I 型非负下降方向.

(b) 若表 4.5 中的 $h_{22} - \frac{h_{12}^2}{h_{11}}$ 和 $h_{33} - \frac{h_{13}^2}{h_{11}}$ 均为 0, 这时一定有 $h_{23} - \frac{h_{12}h_{13}}{h_{11}} < 0$.

因为若 $h_{23} - \frac{h_{12}h_{13}}{h_{11}} \geq 0$, 根据 (4.2.1), 对于任何非负的 x 均有 $x^T H x \geq 0$. 令 $x_2 = \Delta, x_3 = \Delta, x_1 = (-\frac{h_{12}}{h_{11}})\Delta + (-\frac{h_{13}}{h_{11}})\Delta$, 则 $x^T H x = 2(h_{23} - \frac{h_{12}h_{13}}{h_{11}})\Delta^2 < 0$. 所以, 若 $(-\frac{h_{12}}{h_{11}}) + (-\frac{h_{13}}{h_{11}}) \geq 0$, 此解是 II 型非负下降方向; 否则 x 既有正分量也有负分量, 根据定理 4.2.1, H 有个二阶主子矩阵有非负下降方向, 因而 H 有 I 型非负下降方向.

(c) 若 $h_{22} - \frac{h_{12}^2}{h_{11}}$ 和 $h_{33} - \frac{h_{13}^2}{h_{11}}$ 均非负但至少有一个为正, 不妨设前者为正并以其为枢轴进行一次旋转运算, 设结果如表 4.6 所示.

表 4.6

第二次旋转运算结果

	h_1	h_2	$e_3^{+\Delta}$
e_1	w_{11}	w_{12}	w_{13}
e_2	w_{12}	w_{22}	w_{23}
h_3	$-w_{13}$	$-w_{23}$	w_{33}

由于 H 不是正半定的, 因此 $w_{33} < 0$. 这时作 $e_3^{+\Delta}$ 得基本解 $x_3 = \Delta, x_1 = w_{13}\Delta, x_2 = w_{23}\Delta$, 且 $x^T H x = w_{33}\Delta^2 < 0$, 其中 $\Delta > 0$. 若 $w_{13}, w_{23} \geq 0$, x 是 H 的 I 型非负下降方向; 否则 x 既有正分量也有负分量. 根据定理 4.2.1, H 有退化的非负下降方向, 即 H 的一个二阶主子矩阵有非负下降方向. H 的 3 个二阶主子矩阵都是对角元素为正的矩阵. 上面已经说明, 如果这样的矩阵有非负下降方向, 那么它有 I 型非负下降方向, 因而 H 有 I 型非负下降方向.

(ii) 设此定理对 2 至 $n-1$ 阶实对称矩阵成立, 下面证明它对 n 阶实对称矩阵 H 成立. 以 H 对角线上的正元素为枢轴进行若干次旋转运算, 直到拉格朗日矩阵的所有对角元素均小于或等于 0. 不妨设旋转运算结果为表 4.2, 这时有以

下两种情况：

(a) 拉格朗日矩阵有个对角元素如 $w_{ss} < 0, s \in \{k+1, k+2, \dots, n\}$. 令 $x_s = \Delta, x_i = 0 (i \in \{k+1, k+2, \dots, n\} \setminus \{s\}), x_i = w_{is}\Delta (i = 1, 2, \dots, k)$, 则 $x^T H x = w_{ss}\Delta^2 < 0$. 若 $w_{1s}, w_{2s}, \dots, w_{ks}$ 均非负, 此解为 H 的 I 型非负下降方向; 否则此解既有正分量也有负分量. 根据定理 4.2.1, H 有退化的非负下降方向, 即有一个阶数小于 n 的主子矩阵有非负下降方向. 由归纳假设, 这样的矩阵有 I 型或 II 型非负下降方向, 因而 H 有 I 型或 II 型非负下降方向.

(b) 拉格朗日矩阵所有对角元素为 0. 这时拉格朗日矩阵至少有一个非对角元素为负, 否则根据 (4.2.1), 对于任何非负的 x 均有 $x^T H x \geq 0$. 设 $w_{sr} < 0, s, r \in \{k+1, k+2, \dots, n\}$. 令 $x_s = \Delta, x_r = \Delta, x_i = 0 (i \in \{k+1, k+2, \dots, n\} \setminus \{s, r\}), x_i = w_{is}\Delta + w_{ir}\Delta (i = 1, 2, \dots, k)$, 此时 $x^T H x = 2w_{sr}\Delta^2 < 0$. 若此解的所有分量非负, 则为 H 的一个 II 型非负下降方向; 否则有一个阶数小于 n 的主子矩阵有非负下降方向, 因而 H 有 I 型或 II 型非负下降方向. ■

4.2.3 搜索非负方向时无需考虑的情况

严格地说, “方向”是一个不考虑长度的非零向量. 也就是说, 如果 $d^{(1)}$ 和 $d^{(2)}$ 都是矩阵 H 的非负下降方向并且存在一个正数 α 使 $d^{(2)} = \alpha d^{(1)}$, 那么 $d^{(1)}$ 和 $d^{(2)}$ 标示同一个方向. 定理 4.2.4 的意义在于将搜索非负下降方向的范围由一个不可数集缩小为一个有限集. 尽管如此, 当 H 的阶数较大时, 可能有很多正定主子矩阵. 把所有这些正定主子矩阵都找出来, 并将 H 表示为 (4.2.1) 的形式, 计算量很大. 事实上仅进行若干次旋转运算便可判定 H 的许多主子矩阵是否有非负下降方向. 例如, H 的所有对角元素为正时, 以第一个对角元素 h_{11} 为枢轴作旋转运算, 可以判定 H 的包括 h_{11} 在内的所有 $n-1$ 个二阶主子矩阵是否有非负下降方向. 下面进一步考虑在什么场合, 关于某些正定主子矩阵的展开式可以不考虑.

引理 4.2.2 设 (A, B) 是一个所有元素非负的分块矩阵, 其中 A 为可逆方阵, 矩阵 B 的第 s 列和第 r 列分别记为 b_s 和 b_r . 那么

(i) 若 b_s 不是零向量, 则 $-A^{-1}b_s$ 至少有一个分量为负;

(ii) 若 b_s 和 b_r 至少有一个不是零向量, 则 $-A^{-1}(b_s + b_r)$ 至少有一个分量为负.

证 (i) 令 $-A^{-1}b_s = p_s$, 则 $-b_s = Ap_s$. 由于 b_s 的所有分量非负且至少有一个为正, $-b_s$ 至少有一个分量为负. 由于 A 的所有元素非负, p_s 即 $-A^{-1}b_s$ 至少有一个分量为负.

(ii) 令 $-A^{-1}b_s = p_s$, $-A^{-1}b_r = p_r$, 则 $-(b_s + b_r) = A(p_s + p_r)$. 由于 $-(b_s + b_r)$ 至少有一个分量为负, A 的所有元素非负, $p_s + p_r$ 即 $-A^{-1}(b_s + b_r)$ 至少有一个分量为负. ■

定理 4.2.5 用 L 表示以 n 阶实对称矩阵 H 对角线上若干正元素为枢轴进行旋转运算得到的拉格朗日矩阵, L 每一行的分量或者全都非负或者在 L 对角线上的分量非正. 如果当前表未呈现 I 型或 II 型非负下降方向, 那么以 L 对角线上的正元素为枢轴进行旋转运算, 也不会出现 I 型或 II 型非负下降方向.

证 设当前表为表 4.7.

表 4.7 当前表

	h_I	e_{II}	e_{III}	e_{IV}
e_I	H_{11}^{-1}	E_1	E_2	E_3
h_{II}	$-E_1^T$	L_{11}	L_{12}	L_{13}
h_{III}	$-E_2^T$	L_{12}^T	L_{22}	L_{23}
h_{IV}	$-E_3^T$	L_{13}^T	L_{23}^T	L_{33}

其中右下角的 3×3 分块矩阵即为拉格朗日矩阵 L . 不妨设 $L_{11}, L_{12}, L_{13}, L_{22}, L_{23}$ 的所有元素非负且 L_{11} 正定, L_{33} 的所有对角元素小于或等于 0. 表中的 e_I 和 e_{II} 等表示相应的单位向量, h_I 和 h_{II} 等表示相应的拉格朗日向量. 此表未出现 I 型和 II 型非负下降方向意味着: 若 L_{33} 的第 s 个对角元素为负, 则 E_3 的第 s 列至少有一个分量为负; 若 L_{33} 含一个二阶主子矩阵, 其两个对角元素为 0, 非对角元素为负, 设其在 L_{33} 的第 s 行第 r 列, 则 $(E_3)_s + (E_3)_r$ 至少有一个分量为负. 其中 $(E_3)_s$ 和 $(E_3)_r$ 分别表示 E_3 的第 s 列和第 r 列.

沿 L_{11} 的对角线进行旋转运算, 根据定理 1.1.1, 结果见表 4.8.

表 4.8 旋转运算的结果

	h_I	h_{II}	e_{III}	e_{IV}
e_I	$H_{11}^{-1} + E_1 L_{11}^{-1} E_1^T$	$E_1 L_{11}^{-1}$	$E_2 - E_1 L_{11}^{-1} L_{12}$	$E_3 - E_1 L_{11}^{-1} L_{13}$
e_{II}	$L_{11}^{-1} E_1^T$	L_{11}^{-1}	$-L_{11}^{-1} L_{12}$	$-L_{11}^{-1} L_{13}$
h_{III}	$-E_2^T + L_{12}^T L_{11}^{-1} E_1^T$	$L_{12}^T L_{11}^{-1}$	$L_{22} - L_{12}^T L_{11}^{-1} L_{12}$	$L_{23} - L_{12}^T L_{11}^{-1} L_{13}$
h_{IV}	$-E_3^T + L_{13}^T L_{11}^{-1} E_1^T$	$L_{13}^T L_{11}^{-1}$	$L_{23}^T - L_{13}^T L_{11}^{-1} L_{12}$	$L_{33} - L_{13}^T L_{11}^{-1} L_{13}$

由于 $\begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix}$ 的所有元素非负, 它没有非负下降方向. 也就是说, 即使

$L_{22} - L_{12}^T L_{11}^{-1} L_{12}$ 有 I 型或 II 型矩阵, 也不能利用这两种形式的矩阵产生非负下降方向. 因此下面仅考虑能否利用 $L_{33} - L_{13}^T L_{11}^{-1} L_{13}$ 中的 I 型或 II 型矩阵产生非负下降方向.

设 $L_{33} - L_{13}^T L_{11}^{-1} L_{13}$ 的第 s 个对角元素是负数. 根据引理 4.2.2 的(i), 若 L_{13} 的第 s 列不是零向量, 则 $-L_{11}^{-1} L_{13}$ 的第 s 列至少有一个分量为负. 所以作 e_s^{+A} 不会产生非负的 I 型下降方向. 如果 L_{13} 的第 s 列是零向量, 则 $L_{33} - L_{13}^T L_{11}^{-1} L_{13}$ 的第 s 个对角元素就是 L_{33} 的第 s 个对角元素, $E_3 - E_1 L_{11}^{-1} L_{13}$ 的第 s 列就是 E_3 的第 s 列. 这正是表 4.7 的情况. 按假定该表没出现 I 型非负下降方向. 现在假设 $L_{33} - L_{13}^T L_{11}^{-1} L_{13}$ 的第 s 列和第 r 列上有个 2 阶主子矩阵, 其两个对角元素等于 0, 非对角元素小于 0. 据引理 4.2.2 的(ii), 若 $(L_{13})_s$ 和 $(L_{13})_r$ 至少有一个不是零向量, 则 $(-L_{11}^{-1} L_{13})_s + (-L_{11}^{-1} L_{13})_r$ 至少有一个分量为负, 因而不会出现 II 型非负下降方向. 若 $(L_{13})_s$ 和 $(L_{13})_r$ 都是零向量, 则 $L_{33} - L_{13}^T L_{11}^{-1} L_{13}$ 的第 s 列和第 r 列上的 II 型矩阵就是 L_{33} 的相同位置上的 II 型矩阵, $E_3 - E_1 L_{11}^{-1} L_{13}$ 的第 s 列和第 r 列就是 E_3 的第 s 列和第 r 列. 这也是表 4.7 的情况, 按假定该表没出现 II 型非负下降方向. 其中 $(L_{13})_s$ 和 $(L_{13})_r$ 分别表示 L_{13} 的第 s 列和第 r 列, $(-L_{11}^{-1} L_{13})_s$ 和 $(-L_{11}^{-1} L_{13})_r$ 分别表示 $-L_{11}^{-1} L_{13}$ 的第 s 列和第 r 列. ■

以上定理说明, 如果拉格朗日矩阵的某个对角元素为正, 但同行没有负的非对角元素, 无需以该对角元素为枢轴作旋转运算.

现在设 L 是以 H 的某个正对角元素, 如 h_{11} , 为枢轴进行旋转运算得到的拉格朗日矩阵. L 每一行的分量或者全都非负或者在 L 对角线上的分量非正. 并设当前表(以 $\{e_1, h_2, \dots, h_n\}$ 为非基向量)未呈现 I 型或 II 型非负下降方向. 定理 4.2.5 说明, 以 $\{e_1, a_2, \dots, a_n\}$ 为非基向量的任何一张表不会出现 I 型或 II 型非负下降方向, 其中 $a_i = e_i$ 或 h_i , $i = 2, 3, \dots, n$ (而且 H 的与 $\{e_1, a_2, \dots, a_n\}$ 中的单位向量对应的主子矩阵可能不是正定的). 因此, 为考察 H 是否有非负下降方向, 只需考察非基向量为 $\{h_1, a_2, \dots, a_n\}$ 的情况. 也就是说, 应该永远以 e_1 作为基向量, 或者说 H 的第一行和第一列可以删掉.

4.2.4 非负下降方向的计算要点及例

为获取二次函数 $x^T H x$ 的一个非负下降方向, 根据以上分析, 我们只需考虑

H 的所有对角元素是正数且每行(列)至少有一个负的非对角元素的情况. 定理 4.2.4 还说明, 如果 H 有非负下降方向, 那么一定有 I 型或 II 型非负下降方向.

为获取这样一个矩阵的 I 型或 II 型非负下降方向, 往往需要反复进行旋转运算. 与一般二次规划的旋转算法一样, 以拉格朗日矩阵的正对角元素为枢轴进行的旋转运算称为前向旋转, 以约束矩阵的正对角元素为枢轴进行的旋转运算称为后向旋转. 但是在这里作前向旋转时, 拉格朗日矩阵的枢轴元素行应该有负的非对角元素.

根据笔者的经验, 一旦进行前向旋转就不停地进行下去; 同样, 一旦进行后向旋转也不停地进行下去, 两者交替地进行. 每进行一次旋转运算, 检查是否出现 I 型或 II 型非负下降方向, 或者发现 H 根本没有非负下降方向. 枢轴元素按循环的次序搜索. 例如, 上一次旋转的枢轴在对角线上的第 s 个位置, 则按 $s+1, \dots, n, 1, \dots, s-1$ 的次序搜索下一次旋转运算的枢轴元素. 只要 H 不是正定的, 最后将出现以下情况:

(i) 拉格朗日矩阵的所有元素非负, 这时 H 无非负下降方向.

(ii) 出现 I 型或 II 型非负下降方向.

(iii) 拉格朗日矩阵有负对角元素(即 I 型矩阵)或 II 型矩阵但 H 没有非负下降方向.

(a) 如果拉格朗日矩阵有负对角元素, 设 w_{ss} 是最前面的一个负对角元素, w_{rs} 是第 s 列非基单位行上第一个负元素, 以 w_{rr} 为枢轴进行后向旋转. 如果没有出现非负下降方向并且对角线上第 s 个元素仍然是负数, 以类似的方式再进行后向旋转, 直到对角线上第 s 个元素为正. 然后再考虑作前向旋转.

(b) II 型矩阵的情况. 设拉格朗日矩阵的对角元素 $w_{ss} = w_{rr} = 0$, 非对角元素 $w_{sr} < 0$, w_{ks} 是第 s 列非基单位行上第一个负元素, 以 w_{kk} 为枢轴进行后向旋转. 然后再考虑作前向旋转.

下面用几个例子说明计算细节.

例 4.2.1 求一个非负向量 x 使 $2f(x) = x^T H x < 0$, 其中

$$H = \begin{pmatrix} 1 & -1 & 3 & 3 \\ -1 & 2 & -1 & 4 \\ 3 & -1 & 10 & -5 \\ 3 & 4 & -5 & 2 \end{pmatrix}$$

解 用 h_1, h_2, h_3, h_4 分别表示 H 的 4 行, e_1, e_2, e_3, e_4 分别表示 4 阶单位矩阵的 4 行, 以单位向量为初始基向量, 将 h_1, h_2, h_3, h_4 分别表示为基向量的线性组合, 如表 4.9 所示.

由表 4.9 可见, 其右下角的 2 阶主子矩阵有非负下降方向, 具体地说, 只要

以其中任何一个对角元素为枢轴进行旋转运算便可得到非负下降方向. 下面不妨以表 4.9 的第一个对角元素为枢轴进行旋转运算 结果如表 4.10 所示.

表 4.9 初 始 表

	e_1	e_2	e_3	e_4
h_1	1*	- 1	3	3
h_2	- 1	2	- 1	4
h_3	3	- 1	10	- 5
h_4	3	4	- 5	2

表 4.10 第一次前向旋转结果

	h_1	e_2	e_3	e_4
e_1	1	1	- 3	- 3
h_2	- 1	1	2	7
h_3	3	2	1*	- 14
h_4	3	7	- 14	- 7

在表 4.10 中 最后一列出现一个负对角元素 - 7 作 $e_4^{+\Delta}$ 便得到一个下降方向. 但该列非基向量 e_1 行的元素 - 3 < 0 ,该下降方向的第一个分量是负数. 表中的拉格朗日矩阵是右下角的 3×3 矩阵 其第一行所有元素为正 忽略此行. 第二行的对角元素 1 是正数且同行有负元素 以该对角元素为枢轴作旋转运算 结果如表 4.11 所示.

表 4.11 第二次前向旋转结果

	h_1	e_2	h_3	e_4
e_1	10*	7	- 3	- 45
h_2	- 7	- 3	2	35
e_3	- 3	- 2	1	14
h_4	45	35	- 14	- 203

表 4.12 第一次后向旋转结果

	e_1	e_2	h_3	$e_4^{+\Delta}$
h_1	$\frac{1}{10}$	$-\frac{7}{10}$	$\frac{3}{10}$	$\frac{9}{2}$
h_2	$-\frac{7}{10}$	$\frac{19}{10}$	$-\frac{1}{10}$	$\frac{7}{2}$
e_3	$-\frac{3}{10}$	$\frac{1}{10}$	$\frac{1}{10}$	$\frac{1}{2}$
h_4	$\frac{9}{2}$	$\frac{7}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$

表 4.11 中的拉格朗日矩阵是由第 2 4 两行和第 2 4 两列上的 4 个元素构成的矩阵 其两个对角元素均是负数 前向旋转暂时停止. 由于第 2 和第 4 两列都有负元素在非基单位向量行 两个 I 型下降方向都不是非负的. 由于上一次旋转运算在对角线的第 3 个位置进行 从第 4 个位置开始在拉格朗日矩阵中搜索负对角元素进行后向旋转 这正是 - 203. 该列非基向量 e_1 行的元素 - 45 < 0 , 以对角线上的元素 10 作后向旋转 结果见表 4.12. 此表出现非负下降方向 $x_4 =$

$\Delta, x_1 = x_2 = 0, x_3 = \frac{\Delta}{2}$ 相应的函数值 $f(x) = -\frac{\Delta^2}{4}$.

例 4.2.2 求一个非负向量 x 使 $2f(x) = x^T H x < 0$, 其中

$$H = \begin{pmatrix} 1 & -1 & 3 & 3 \\ -1 & 2 & -1 & 4 \\ 3 & -1 & 10 & -3 \\ 3 & 4 & -3 & 6 \end{pmatrix}.$$

解 用 h_1, h_2, h_3, h_4 分别表示 H 的 4 行, e_1, e_2, e_3, e_4 分别表示 4 阶单位矩阵的 4 行, 以单位向量为初始基向量, 将 h_1, h_2, h_3, h_4 表示为基向量的线性组合, 如表 4.13 所示. 首先以 h_1 入基以 e_1 出基作前向旋转, 结果见表 4.14. 此表 h_4 行的对角元素 $-3 < 0$, 且 e_4 列 e_1 行上的元素 $-3 < 0$, 作 $e_4^{+\Delta}$ 将产生一个有的分量为正有的分量为负的下降方向. 根据定理 4.2.1, 如果 H 有非负下降方向, 则 H 有个 2 阶或 3 阶主子矩阵有非负下降方向. 反过来说, 如果 H 的所有 2 阶或 3 阶主子矩阵没有非负下降方向, 那么 H 没有非负下降方向.

表 4.13 初始表

	e_1	e_2	e_3	e_4
h_1	1*	-1	3	3
h_2	-1	2	-1	4
h_3	3	-1	10	-3
h_4	3	4	-3	6

表 4.14 第一次前向旋转结果

	h_1	e_2	e_3	e_4
e_1	1	1	-3	-3
h_2	-1	1	2	7
h_3	3	2	1*	-12
h_4	3	7	-12	-3

在表 4.14 中再以 h_3 入基以 e_3 出基作前向旋转, 结果见表 4.15. 此表的 2 阶拉格朗日矩阵的两个对角元素均为负数, 前向旋转暂时停止. 可以看出, 利用这两个对角元素产生的下降方向都不是非负的. 由于上一次旋转运算在对角线上的第 3 个位置进行, 从第 4 个位置开始搜索负对角元素进行后向旋转, 这正是 -147. 该列非基向量 e_1 行上的元素 $-39 < 0$, 以 e_1 入基以 h_1 出基作后向旋转, 结果见表 4.16. 这时 -147 已经变为正数 5.1, 后向旋转暂时停止.

表 4.15 第二次前向旋转结果

	h_1	e_2	h_3	e_4
e_1	10*	7	-3	-39
h_2	-7	-3	2	31
e_3	-3	-2	1	12
h_4	39	31	-12	-147

表 4.16 第一次后向旋转结果

	e_1	e_2	h_3	e_4
h_1	0.1	-0.7	0.3	3.9
h_2	-0.7	1.9*	-0.1	3.7
e_3	-0.3	0.1	0.1	0.3
h_4	3.9	3.7	-0.3	5.1

再考虑作前向旋转. 表 4.16 中的拉格朗日矩阵是由第 1, 2, 4 三行和第 1, 2, 4 三列的元素构成的 3×3 矩阵. 由于上一次旋转运算在对角线上的第一个位置进行, 从第 2 个位置开始搜索正对角元素进行前向旋转, 这正是拉格朗日矩阵的对角元素 1.9. 由于该行有负的非对角元素, 以 1.9 为枢轴作前向旋转, 结果如表 4.17 所示.

表 4.17 第三次前向旋转运算结果

	e_1	h_2	h_3	e_4
h_1	-0.158	-0.368	0.263	5.263
e_2	0.368	0.526	0.053	-1.947
e_3	-0.263	0.053	0.105*	0.105
h_4	5.263	1.947	-0.105	-2.105

表 4.18 第二次后向旋转运算结果

	e_1	h_2	e_3	e_4
h_1	0.5	-0.5	2.5	5
e_2	0.5	0.5	0.5	-2
h_3	2.5	-0.5	9.5*	-1
h_4	5	2	-1	-2

表 4.17 中, 二阶拉格朗日矩阵的两个对角元素均为负数, 前向旋转暂时停止. 可以看出, 利用这两个对角元素产生的下降方向都不是非负的. 由于上一次旋转运算在对角线上的第 2 个位置进行, 从第 3 个位置开始在拉格朗日矩阵中搜索负对角元素进行后向旋转, 这正是 -2.105. 该列非基向量 e_2 行上的元素 $-1.947 < 0$, 应以对角线上的元素 0.526 作后向旋转. 但在这一位置刚刚进行了一次旋转运算, 让我们在拉格朗日矩阵中另找一个负对角元素 -0.158 作后向旋转. 该元素所在列和非基向量 e_3 行上的元素 $-0.263 < 0$, 以 e_3 行对角线上的元素 0.105 作后向旋转, 结果如表 4.18 所示. 可见 -0.158 变成了 0.5, 后向旋转暂时停止. 再考虑前向旋转. 表 4.18 中的拉格朗日矩阵是由第 1, 3, 4 三行和第 1, 3, 4 三列上的元素构成的 3×3 矩阵. 由于上一次旋转运算在对角线上的第 3 个位置进行, 从第 4 个位置开始搜索正对角元素进行前向旋转. 第 4 个位置上的对角元素是负数. 再看其后的对角元素 0.5, 这是表 4.18 对角线上的第 1 个元素. 但第一行拉格朗日矩阵的所有元素都是正数. 拉格朗日矩阵对角线上的下一个元素是表中对角线上的元素 9.5. 如果以其为枢轴作前向旋转又回到表 4.17, 于是停下来检查 H 是否没有非负下降方向.

为叙述方便起见, 用 $(1; 2)$ 表示矩阵 H 左上角的二阶主子矩阵, $(1; 3; 4)$ 表示 H 的第 1, 3, 4 行的 3 阶主子矩阵, 其余情况类似表示. 由表 4.14 可见, H 的主子矩阵 $(1; 2; 3)$ 没有非负下降方向, 因为这个主子矩阵的拉格朗日矩阵 $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$ 的所有元素是正数. 由表 4.16 可见, H 的主子矩阵 $(1; 3; 4)$ 和 $(2; 3; 4)$

没有非负下降方向, 因为这两个主子矩阵各自的拉格朗日矩阵的所有元素都是正数. 由表 4.14 可见, H 的主子矩阵 $(1; 2; 4)$ 有下降方向但有的分量是负数, $(1; 2)$ 正定, $(1; 4)$ 没有非负下降方向. 另外, 由表 4.18 可见, $(2; 4)$ 没有非负下降方向, 所以 $(1; 2; 4)$ 没有非负下降方向. 由于 H 的 4 个三阶主子矩阵都没有非负下降方向, 所以 H 没有非负下降方向.

4.3 二次齐次函数在锥内的下降方向

4.3.1 问题的模型及求解思路

本节考虑在一般形式的齐次线性不等式组的解集中找一个解, 使一个二次齐次函数小于零, 称为锥二次规划, 其模型如下:

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx, \\ \text{s. t. } a_i x &= 0, \quad i = 1, 2, \dots, l, \\ a_i x &\geq 0, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq 0, \quad i \in I_Z \subset \{1, 2, \dots, n\}. \end{aligned} \right\} \quad (4.3.1)$$

其中 $x = (x_1, x_2, \dots, x_n)^T$, $H = (h_{ij})$ 是一个 n 阶实对称矩阵, $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$, $i = 1, 2, \dots, m$. 模型(4.3.1)的解可分为以下两种情况:

情况 1, $x=0$ 是全局极小点;

情况 2, 目标函数在可行域上无下界.

用 D 表示模型(4.3.1)的可行域, 则 D 是一个以原点为顶点的凸锥. 如果存在 $d \in D$, 使 $f(d) = \frac{1}{2}d^T H d < 0$, 则称 d 为 $f(x)$ (在 $x=0$ 处)的可行下降方向.

显然, 情况 1 的必要充分条件是 $f(x)$ 没有可行下降方向, 或者说情况 2 的必要充分条件是 $f(x)$ 有可行下降方向.

用 λ_i 表示 $a_i x = 0$ 或 $a_i x \geq 0$ 对应的拉格朗日乘子, 模型(4.3.1)的 KT 条件为

入基中. 为了维持互补松弛条件, 每将一个 $a_i (i = 1, 2, \dots, l)$ 换入基中时, 再将其对应的人工不等式的系数向量 e_{n+i} 换出基. 但是与模型 (3.1.1) 不同的是, (4.3.1) 中有的变量无下界, 也就是说人工不等式 $x_i \geq 0 (i \in \{1, 2, \dots, n\} \cap H_Z)$ 不是其本身固有的. 为叙述简单起见, 将这些不等式称为人工约束不等式. 在将非基等式约束换为基等式时, 应尽量使其与基人工约束不等式交换. 但是, 当等式约束较少, 人工约束不等式较多时, 可能还会剩余一些基人工约束不等式. 这时还可采用两种方式将基人工约束不等式换出基外. 一种方式是利用双枢轴运算将非基约束不等式与基人工约束不等式互换. 另一种方式是当基人工约束不等式所在列上的对角元素为正时, 以其为枢轴作主旋转, 直接将基人工约束不等式与其互补等式交换, 一直到不能进行时为止. 以上运算过程称为模型 (4.3.1) 的预处理.

下面假设 $\alpha_1, \alpha_2, \dots, \alpha_l$ 线性无关, 先分析这些向量入基的情形. 与 3.1 节类似, 将表 4.19 的数据划分为 4×4 部分如表 4.20 所示.

表 4.20 初始表的矩阵形式

	$\mathbf{e}_I$	$\mathbf{e}_{II}$	$\mathbf{e}_{III}$	$\mathbf{e}_{IV}$
$\mathbf{h}_I$	$\mathbf{H}_{11}$	$\mathbf{H}_{12}$	$-\mathbf{A}_{11}^T$	$-\mathbf{A}_{21}^T$
$\mathbf{h}_{II}$	$\mathbf{H}_{12}^T$	$\mathbf{H}_{22}$	$-\mathbf{A}_{12}^T$	$-\mathbf{A}_{22}^T$
$\mathbf{a}_I$	$\mathbf{A}_{11}$	$\mathbf{A}_{12}$	$\mathbf{O}$	$\mathbf{O}$
$\mathbf{a}_{II}$	$\mathbf{A}_{21}$	$\mathbf{A}_{22}$	$\mathbf{O}$	$\mathbf{O}$

其中 (A_{11}, A_{12}) 是等式约束的系数划分而成的 1×2 分块矩阵. 设 A_{11} 为可逆方阵, 在表 4.20 先后求 A_{11} 和 $-A_{11}^T$ 的逆阵, 结果见表 4.21.

表 4.21 预处理的结果

	$\mathbf{a}_I$	$\mathbf{e}_{II}$	$\mathbf{h}_I$	$\mathbf{e}_{IV}$
$\mathbf{e}_{III}$	$\mathbf{G}_{11}$	$\mathbf{G}_{12}$	$-\mathbf{B}_{11}^T$	$-\mathbf{B}_{21}^T$
$\mathbf{h}_{II}$	$\mathbf{G}_{12}^T$	$\mathbf{G}_{22}$	$-\mathbf{B}_{12}^T$	$-\mathbf{B}_{22}^T$
$\mathbf{e}_I$	$\mathbf{B}_{11}$	$\mathbf{B}_{12}$	$\mathbf{O}$	$\mathbf{O}$
$\mathbf{a}_{II}$	$\mathbf{B}_{21}$	$\mathbf{B}_{22}$	$\mathbf{O}$	$\mathbf{O}$

其中，

$$G_{22} = H_{22} - H_{12}^T A_{11}^{-1} A_{12} - A_{12}^T A_{11}^{-T} (H_{12} - H_{11} A_{11}^{-1} A_{12}) ,$$

$$B_{12} = -A_{11}^{-1}A_{12}, B_{22} = A_{22} - A_{21}A_{11}^{-1}A_{12}.$$

定理 4.3.1 在等式约束成为基约束之后,如果简约拉格朗日矩阵正定,则模型(4.3.1)的任何非零可行解都使目标函数为正;如果简约拉格朗日矩阵非负并且其对应的基约束都是模型(4.3.1)的非负约束,那么任何可行解都使目标函数非负.

证 考虑表 4.20,与之对应,模型(4.3.1)可写为

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}(x_I^T H_{11} x_I + 2x_I^T H_{12} x_{II} + x_{II}^T H_{22} x_{II}) , \\ \text{s. t. } A_{11} x_I + A_{12} x_{II} &= 0 , \\ A_{21} x_I + A_{22} x_{II} &\geq 0 , \\ x_i &\geq 0 , \quad i \in I_Z . \end{aligned} \right\} \quad (4.3.3)$$

其中 x_I 和 x_{II} 是 x 的划分,分别对应 A_{11} 和 A_{12} . 由等式约束得

$$x_I = -A_{11}^{-1} A_{12} x_{II} .$$

将其代入模型(4.3.3)的目标函数,可得

$$\begin{aligned} f(x) &= \frac{1}{2} x_{II}^T (H_{22} - H_{12}^T A_{11}^{-1} A_{12} - A_{12}^T A_{11}^{-1} (H_{12} - H_{11} A_{11}^{-1} A_{12})) x_{II} \\ &= \frac{1}{2} x_{II}^T G_{22} x_{II} . \end{aligned}$$

其中 G_{22} 便是表 4.21 中的简约拉格朗日矩阵. 若 G_{22} 正定,任何非零向量 $x = \begin{pmatrix} x_I \\ x_{II} \end{pmatrix}$, 包括模型(4.3.1)的所有非零可行解都使 $f(x) > 0$. 若 G_{22} 非负且 $x_{II} \geq 0$ 是模型(4.3.1)的约束条件,则任何可行解都使 $f(x) \geq 0$. ■

在模型(4.3.1)中,如果约束条件系数向量组的秩等于变量数 n ,在预处理过程中利用双枢轴运算可将所有基人工不等式换为非基不等式并删掉,否则将剩余一些基人工约束不等式. 另一方面,基等式所在列及其相应的行也可删掉. 所以,我们只需考虑仅含不等式约束的锥二次规划. 当约束条件系数向量组的秩等于 n 时,称其为标准锥二次规划,否则称其为非标准锥二次规划.

4.3.2 标准锥二次规划

标准锥二次规划可表示为以下形式:

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} x^T H x , \\ \text{s. t. } a_i x &\geq 0 , \quad i = 1, 2, \dots, m , \\ x_i &\geq 0 , \quad i = 1, 2, \dots, n . \end{aligned} \right\} \quad (4.3.4)$$

上一节的模型(4.2.3)可看做以上模型的特殊情况,即 $m = 0$ 时的情况. 算

法的初始表为表 4.22 ,但不包括最后两列.

表 4.22 初 始 表

	e_1	...	e_n	e_{n+1}	...	e_{n+m}	$e_1^{+\Delta}$	$e_1^{+\Delta} \quad e_2^{+\Delta}$
h_1	h_{11}	...	h_{1n}	$-a_{11}$	...	$-a_{m1}$	$h_{11}\Delta$	$h_{11}\Delta + h_{12}\Delta$
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
h_n	h_{n1}	...	h_{nn}	$-a_{n1}$	...	$-a_{mn}$	$h_{n1}\Delta$	$h_{n1}\Delta + h_{n2}\Delta$
a_1	a_{11}	...	a_{1n}	0	...	0	$a_{11}\Delta$	$a_{11}\Delta + a_{12}\Delta$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$
a_m	a_{m1}	...	a_{mn}	0	...	0	$a_{m1}\Delta$	$a_{m1}\Delta + a_{m2}\Delta$

以下两种情况很容易确定 $2f(x) = x^T Hx$ 的一个可行下降方向.

(i) 初始表有一个负对角元素且同列非基约束向量行上的所有元素非负. 例如 $h_{11} < 0$ 且 $a_{11}, \dots, a_{m1} \geq 0$. 作 $e_1^{+\Delta}$, 即将基不等式 $x_1 \geq 0$ 的右端项变为一个正数 Δ , 这时各变量的数值为

$$\begin{aligned} x_1 &= \Delta, \\ x_i &= 0, \quad i = 2, 3, \dots, n. \end{aligned}$$

按照 2.4 节的公式(2.4.3), 各非基向量的偏差如表 4.22 的右数第二列所示, 其中所有非基约束向量的偏差非负, 因而以上解是一个可行解, 并且

$$2f(x) = x^T Hx = \Delta h_{11} \Delta = h_{11} \Delta^2 < 0.$$

(ii) 初始表有个 II 型矩阵且同列非基约束向量行上的两个元素之和均非负. 例如 $h_{11} = h_{22} = 0$, $h_{12} < 0$, 并且 $a_{11} + a_{12}, \dots, a_{m1} + a_{m2}$ 均非负. 作 $e_1^{+\Delta}$ 和 $e_2^{+\Delta}$, 即将基不等式 $x_1 \geq 0$ 和 $x_2 \geq 0$ 的右端项都变为正数 Δ . 这时各变量的数值为

$$\begin{aligned} x_1 &= x_2 = \Delta, \\ x_i &= 0, \quad i = 3, 4, \dots, n. \end{aligned}$$

按照 2.4 节的公式(2.4.2), 各非基向量的偏差如表 4.22 的最后一列所示. 其中所有非基约束向量的偏差非负, 因而以上解是一个可行解, 并且

$$2f(x) = x^T Hx = 2h_{12} \Delta^2 < 0.$$

现在假设初始表中有负的对角元素或 II 型矩阵, 但是不能按照上面的方法得到一个可行下降方向. 例如, 基向量 e_s 列的对角元素 $h_{ss} < 0$, 但同列非基向量 a_r 行上的元素 $a_{rs} < 0$. 这时可能出现以下两种情况之一:

- (i) a_r 行上的元素全都小于或等于 0;
- (ii) a_r 行有正元素.

如果是情况(i) 根据第二章的定理 2.1.3, a_r 行上的负元素对应的基不等式(包括 $x_s \geq 0$)都可换为等式. 也就是说, 这些负元素对应的列以及相应的行都可以删掉. 如果简约拉格朗日矩阵的所有元素非负或者为空, 则没有可行下降方向. 这种情况使计算变得非常简单, 但很少见. 在情况(ii), 以 a_r 行的一个正元素及其对称元素为枢轴先后进行两次旋转运算. 如果还不能发现可行下降方向, 再以类似方式进行旋转运算.

以上我们仅考虑了双枢轴运算. 如果简约拉格朗日矩阵有正的对角元素, 也可考虑作前向旋转, 一直到出现新的负对角元素或 II 型矩阵. 此时, 如果还不能参数化方法产生可行下降方向, 再进行前向或后向旋转. 例如, 基约束向量 a'_s ($= e_s$ 或 a_s) 列的对角元素 $w_{ss} < 0$, 但同列非基约束向量 a'_r ($= e_r$ 或 a_r) 行上的元素 $w_{rs} < 0$ 时, 如果 a'_r 行的对角元素为正, 以其为枢轴作后向旋转; 否则, 如果 a'_r 行有正的非对角元素, 以此元素及其对称元素为枢轴先后进行两次旋转运算; 否则, 删掉 a'_r 行的负元素所在列及相应的行.

基约束的数目等于 n 时, 拉格朗日矩阵都是 $n \times n$ 矩阵, 其正、负和零特征值的数目与 H 的相同. 本章 4.2 节介绍了二次齐次函数 $2f(x) = x^T H x$ 的非负下降方向的计算方法. 非负下降方向就是 $f(x)$ 在 n 维空间直角坐标系的第一象限内的下降方向. 同样可用于求经过合同变换的 H 在以 n 个基约束向量为基的斜坐标系的第一象限内的下降方向, 或“非负”下降方向. 如果非基拉格朗日向量的数目小于 n , 可以看出, 表中任何一个包括所有非基拉格朗日向量行的元素在内的 $n \times n$ 主子矩阵都可看做是对 H 进行合同变换, 然后以对角线上的正元素为枢轴进行若干次前向旋转的结果.

现在设当前表为表 4.23, 其中, a_i' 表示约束向量, 即 $a_i' = e_i$ 或 a_i, h_i' 表示 a_i' 对应的拉格朗日向量.

表 4.23

一般表

	a'_1	...	a'_k	h'_{k+1}	...	h'_{n+m}
h'_1	w_{11}	...	w_{1k}	$-w_{k+1,1}$	...	$-w_{n+m,1}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
h'_k	w_{k1}	...	w_{kk}	$-w_{k+1,k}$	...	$-w_{n+m,k}$
a'_{k+1}	$w_{k+1,1}$	...	$w_{k+1,k}$	$w_{k+1,k+1}$	...	$w_{k+1,n+m}$
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$
a'_{n+m}	$w_{n+m,1}$	...	$w_{n+m,k}$	$w_{n+m,k+1}$	...	$w_{n+m,n+m}$

(i) 设拉格朗日矩阵有个对角元素 $w_{ss} < 0, s \in \{1, 2, \dots, k\}$. 用 $\bar{x}$ 表示作

$a_s' + \Delta$ 得到的基本解, 那么

$$f(\bar{x}) = w_{ss} \frac{\Delta^2}{2} < 0.$$

如果 $w_{is} \geq 0, i = k+1, k+2, \dots, n+m$, 称 $\bar{x}$ 为 H 的一个 I 型可行下降方向.

(ii) 设拉格朗日矩阵有个二阶主子矩阵 $\begin{pmatrix} 0 & w_{sr} \\ w_{sr} & 0 \end{pmatrix}$, 其中 $w_{sr} < 0, s, r \in$

$\{1, 2, \dots, k\}$. 用 $\bar{x}$ 表示作 $a_s' + \Delta$ 和 $a_r' + \Delta$ 得到的基本解, 那么

$$f(\bar{x}) = w_{sr} \Delta^2 < 0.$$

如果 $w_{is} + w_{ir} \geq 0, i = k+1, k+2, \dots, n+m$, 称 $\bar{x}$ 为 H 的一个 II 型可行下降方向.

设 $d = (d_1, d_2, \dots, d_n)^T$ 是模型(4.3.4)的一个可行下降方向. 如果

$$a_i d > 0, \quad i = 1, 2, \dots, m,$$

$$d_i > 0, \quad i = 1, 2, \dots, n,$$

称 d 为非退化的可行下降方向; 如果以上式子中至少有一个是等式, 称 d 为退化的可行下降方向.

定理 4.3.2 如果模型(4.3.4)有可行下降方向, 则有 I 型或 II 型可行下降方向.

证 下面归纳地证明此定理. 用 a_i' 表示约束向量, h_i' 表示其对应的拉格朗日向量.

(i) $n=2$ 的情况. 设模型(4.3.4)的可行域为

$$\{x \mid a_s' x \geq 0, a_r' x \geq 0\}.$$

如果 a_s' 和 a_r' 线性无关, 那么在以 a_s' 和 a_r' 为基的坐标系中有“非负”下降方向, 因而模型(4.3.4)有 I 型或 II 型可行下降方向. 如果 a_s' 和 a_r' 线性相关, 任取一个与 a_s' (或 a_r') 线性无关的约束向量 a_k' , 在以 a_k' 和 a_s' 为基向量的旋转算法表中, 作 $a_k' + \Delta$, 便可得到模型(4.3.4)的一个 I 型可行下降方向.

(ii) 假设定理对于 2 到 $n-1$ 个变量的问题都成立, 下面证明对于 n 个变量的问题也成立.

(a) 设模型(4.3.4)存在一个退化的可行下降方向 $d = (d_1, d_2, \dots, d_n)^T$, 那么或者有个 $d_s = 0 (s \in \{1, 2, \dots, n\})$ 或者有个 $a_s d = 0 (s \in \{1, 2, \dots, m\})$. 在前一情况, 应将初始表中的 e_s 永远作为基向量; 在下一情况, 先利用双枢轴运算将非基向量 a_s 换为基向量, 然后使 a_s 永远作为基向量. 也就是说以后的计算实际上是在解 $n-1$ 个变量的问题. 由归纳假设, 这样的问题有 I 型或 II 型可行下降方向, 因而 n 个变量的问题有 I 型或 II 型可行下降方向.

(b) 设 $d = (d_1, d_2, \dots, d_n)^T$ 是模型(4.3.4)的一个非退化的可行下降方向. 任何一张旋转算法表中的拉格朗日矩阵或者含有 I 型或 II 型矩阵, 或者进行前向旋转可得到这样的矩阵. 如果利用 I 型或 II 型矩阵得到的基本解 $y = (y_1, y_2, \dots, y_n)^T$ 可行, 那么 y 就是 I 型或 II 型可行下降方向. 否则 $\{y_1, \dots, y_n, a_1 y, \dots, a_m y\}$ 中至少有一个是正数且至少有一个是负数. 根据引理 4.1.1, 对于任何非负的 θ , 或者

$$d + \theta y = (d_1 + \theta y_1, d_2 + \theta y_2, \dots, d_n + \theta y_n)^T$$

或者

$$d - \theta y = (d_1 - \theta y_1, d_2 - \theta y_2, \dots, d_n - \theta y_n)^T$$

是 H 的下降方向. 在前一情况让 θ 由 0 逐渐增加, 直到 $d_i + \theta y_i (i=1, 2, \dots, n)$ 和 $a_i(d + \theta y) = a_i d + \theta a_i y (i=1, 2, \dots, m)$ 至少有一个等于 0, 其他的非负, 则 $d + \theta y$ 成为退化的可行下降方向. 同样, 若 $d - \theta y$ 是 H 的下降方向, 可找到一个正数 θ , 使 $d - \theta y$ 成为退化的可行下降方向. 根据前述(a), 模型(4.3.4)有 I 型或 II 型可行下降方向. ■

显然, 如果 $2f(x) = x^T H x$ 有可行下降方向, 除了 H 要有非负下降方向外, 经过合同变换后的 H 在每一个斜坐标系中都必需有“非负”下降方向. 反过来说, 如果发现在某一坐标系中没有非负下降方向, 则 $f(x)$ 没有可行下降方向. 具体地讲, 如果某张旋转算法表中的拉格朗日矩阵正定或非负, 则模型(4.3.4)没有可行下降方向, 所以有以下定理.

定理 4.3.3 用旋转算法解模型(4.3.4), 如果旋转算法表中的拉格朗日矩阵正定或者非负, 则没有可行下降方向.

与 4.1 节类似, 为寻找可行下降方向, 拉格朗日矩阵中必须有负的对角元素或者出现 II 型矩阵. 但在本节, 下降方向受到的约束要多一些. 下面通过几个例子说明计算细节.

例 4.3.1 在不等式组

$$\begin{aligned} -x_1 + 2x_2 + x_3 - 4x_4 &\geq 0, \\ 3x_1 + 4x_2 - 4x_3 + 2x_4 &\geq 0, \\ x_1, x_2, x_3, x_4 &\geq 0 \end{aligned}$$

的解集中找一个 x 使二次函数 $f(x) = \frac{1}{2}x^T H x < 0$, 其中

$$H = \begin{pmatrix} 1 & -4 & -2 & -1 \\ -4 & 1 & 1 & -2 \\ -2 & 1 & -1 & -3 \\ -1 & -2 & -3 & -1 \end{pmatrix}.$$

解 记

$$\begin{aligned}h_1 &= (1, -4, -2, -1, 1, -3), \\h_2 &= (-4, 1, 1, -2, -2, -4), \\h_3 &= (-2, 1, -1, -3, -1, 4), \\h_4 &= (-1, -2, -3, -1, 4, -2), \\a_1 &= (-1, 2, 1, -4, 0, 0), \\a_2 &= (3, 4, -4, 2, 0, 0),\end{aligned}$$

用 $e_1, e_2, e_3, e_4, e_5, e_6$ 分别表示 6 阶单位矩阵的各行.

以单位向量为初始基向量 将 $h_1, h_2, h_3, h_4, a_1, a_2$ 分别表示为基向量的线性组合 如表 4.24 所示.

表 4.24 初 始 表

	e_1	e_2	e_3	e_4	e_5	e_6
h_1	1	- 4	- 2	- 1	1	- 3
h_2	- 4	1	1	- 2	- 2	- 4 *
h_3	- 2	1	- 1	- 3	- 1	4
h_4	- 1	- 2	- 3	- 1	4	- 2
a_1	- 1	2	1	- 4	0	0
a_2	3	4 *	- 4	2	0	0

初始表基向量 e_3 和 e_4 两列上的对角元素是负数 ,但这两列非基约束向量行上都有负元素无论作 $e_3^{+\Delta}$ 或是 $e_4^{+\Delta}$,下降方向都不可行. 不妨取 e_3 列. 此列非基约束向量 a_2 行上的元素 $- 4 < 0$. 以 a_2 行上的最大正元素 4 及其对称元素 $- 4$ 为枢轴先后进行的两次旋转运算 结果见表 4.25.

表 4.25 第一次双枢轴运算结果

	e_1	a_2	e_3	e_4	e_5	h_2
h_1	7.5625 *	- 1.1875	- 7.5	2.875	2.5	0.75
e_6	- 1.1875	0.0625	0.5	- 0.625	- 0.5	- 0.25
h_3	- 7.5	0.5	2	- 6	- 3	- 1
h_4	2.875	- 0.625	- 6	1.25	5	0.5
a_1	- 2.5	0.5	3	- 5	0	0
e_2	- 0.75	0.25	1	- 0.5	0	0

在表 4.25 中 ,拉格朗日矩阵的 4 个对角元素都是正数 ,不妨以第一个对角元素 7.5625 为枢轴作前向旋转 结果见表 4.26.

表 4.26 第一次前向旋转运算结果

	h_1	$a_2^{+\Delta}$	e_3	e_4	e_5	h_2
e_1	0.1322	0.157	0.9917	- 0.3802	- 0.3306	- 0.0992
e_6	- 0.157	- 0.124	- 0.6777	- 0.1736	- 0.1074	- 0.1322
h_3	- 0.9917	- 0.6777	- 5.438	- 3.1488	- 0.5207	- 0.2562
h_4	0.3802	- 0.1736	- 3.1488	0.157	4.0496	0.2149
a_1	- 0.3306	0.1074	0.5207	- 4.0496	0.8264	0.2479
e_2	- 0.0992	0.1322	0.2562	- 0.2149	0.2479	0.0744

在表 4.26 中,基向量 a_2 列上的对角元素是负数,并且同列 3 个非基约束向量行上的元素都是正数,作 $a_2^{+\Delta}$ 得可行下降方向

$$x_1 = 0.157\Delta, \quad x_2 = 0.1332\Delta, \quad x_3 = x_4 = 0,$$

其中 Δ 是任意正数,相应的函数值为 $2f(x) = -0.124\Delta^2$. 同样作 $e_3^{+\Delta}$ 也能得到一个可行下降方向

$$x_1 = 0.9917\Delta, \quad x_2 = 0.2562\Delta, \quad x_3 = \Delta, \quad x_4 = 0,$$

相应的函数值为 $2f(x) = -5.438\Delta^2$.

例 4.3.2 在不等式组

$$\begin{aligned} -2x_1 + x_2 + x_3 - 2x_4 &\geq 0, \\ x_1 - 4x_2 - 4x_3 + 8x_4 &\geq 0, \\ x_1, x_2, x_3, x_4 &\geq 0 \end{aligned}$$

的解集中找一个 x 使二次函数 $f(x) = \frac{1}{2}x^T Hx < 0$, 其中

$$H = \begin{pmatrix} -6 & 2 & 0 & 1 \\ 2 & -2 & -2 & 3 \\ 0 & -2 & 1 & -1 \\ 1 & 3 & -1 & 1 \end{pmatrix}.$$

解 与前面类似,初始表为表 4.27.

表 4.27 初 始 表

	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆
h ₁	- 6	2	0	1	2	- 1
h ₂	2	- 2	- 2	3	- 1 *	4
h ₃	0	- 2	1	- 1	- 1	4
h ₄	1	3	- 1	1	2	- 8
a ₁	- 2	1 *	1	- 2	0	0
a ₂	1	- 4	- 4	8	0	0

初始表基向量 e_1 和 e_2 两列上的对角元素是负数 ,但这两列非基约束向量行上都有负元素无论作 $e_1^{+\Delta}$ 或是 $e_2^{+\Delta}$,下降方向都不可行. 不妨取 e_1 列. 此列非基约束向量 a_1 行上的元素 $- 2 < 0$. 以 a_1 行上的正元素 1 及其对称元素 - 1 为枢轴先后进行的两次旋转运算 结果见表 4.28.

表 4.28 第一次双枢轴运算结果

	e ₁	a ₁	e ₃	e ₄	h ₂	e ₆
h ₁	- 6	- 2	- 2	3	- 2	7
e ₅	- 2	- 2	0	- 1	- 1	4
h ₃	- 2	0	3 *	- 4	1	0
h ₄	3	- 1	- 4	5	- 2	0
e ₂	2	1	- 1	2	0	0
a ₂	- 7	- 4	0	0	0	0

非基向量 a_2 行上的元素都小于或等于 0 ,其中与基向量 e_1 和 a_1 对应元素是负数. 根据定理 2.1.3 ,这两个基向量对应的不等式可换为等式 ,于是删掉第 1 ,2 两列和第 1 ,2 两行. 由于 a_2 行只剩下 0 元素 ,再删掉 a_2 行和 e_6 列. 这时的拉格朗日矩阵仅为 h_3 和 h_4 两行的主子矩阵 $\begin{pmatrix} 3 & - 4 \\ - 4 & 5 \end{pmatrix}$. 以其中的第一个对角元素 3 为枢轴作前向旋转 结果见表 4.29.

表 4.29 第一次前向旋转结果

	h ₃	$e_4^{+\Delta}$	h ₂
e ₃	$\frac{1}{3}$	$\frac{4}{3}$	$-\frac{1}{3}$
h ₄	$-\frac{4}{3}$	$-\frac{1}{3}$	$-\frac{2}{3}$
e ₂	$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$

在表 4.29 中,基向量 e_4 列上的对角元素是负数,并且同列两个非基约束向量行上的元素都是正数,作 $e_4^{+\Delta}$ 得可行下降方向

$$x_1 = 0, x_2 = \frac{2}{3}\Delta, x_3 = \frac{4}{3}\Delta, x_4 = \Delta,$$

相应的函数值为 $2f(x) = -\frac{1}{3}\Delta^2$.

例 4.3.3 在不等式组

$$\begin{aligned} x_1 + x_2 - x_3 &\geq 0, \\ -x_1 - x_2 + 2x_3 &\geq 0, \\ x_1, x_2, x_3 &\geq 0 \end{aligned}$$

的解集中找一个 x 使二次函数 $f(x) = \frac{1}{2}x^THx < 0$, 其中 $H = \begin{pmatrix} 1 & -1 & -1 \\ -1 & 1 & -1 \\ -1 & -1 & 1 \end{pmatrix}$.

解 初始表为表 4.30.

表 4.30 初 始 表

	e_1	e_2	e_3	e_4	e_5
h_1	1*	-1	-1	-1	1
h_2	-1	1	-1	-1	1
h_3	-1	-1	1	1	-2
a_1	1	1	-1	0	0
a_2	-1	-1	2	0	0

表 4.30 拉格朗日矩阵的 3 个对角元素都是正数,不妨以第一个对角元素 1 为枢轴作前向旋转,结果见表 4.31.

表 4.31 第一次前向旋转结果

	h_1	e_2	e_3	e_4	e_5
e_1	1	1	1	1	-1
h_2	-1	0	-2	-2	2
h_3	-1	-2	0	0	-1
a_1	1	2	0	1	-1
a_2	-1	-2	1	-1	1*

表 4.31 的拉格朗日矩阵是一个 II 型矩阵,但其所在列非基约束向量 a_2 行上的两个元素之和 $-2 + 1 = -1 < 0$,II 型下降方向不可行. a_2 行上的对角元素是正数,以其为枢轴作后向旋转,结果见表 4.32.

表 4.32 第一次后向旋转结果

	h_1	e_2	e_3	e_4	a_2
e_1	0	- 1	2	0	- 1
h_2	1	4^*	- 4	0	2
h_3	- 2	- 4	1	- 1	- 1
a_1	0	0	1	0	- 1
e_5	1	2	- 1	1	1

表 4.32 的拉格朗日矩阵是一个对角元素都是正数的 3×3 矩阵. 以其第一个对角元素 4 为枢轴作前向旋转, 结果见表 4.33.

表 4.33 第二次前向旋转结果

	h_1	h_2	$e_3^{+\Delta}$	e_4	a_2
e_1	$\frac{1}{4}$	$-\frac{1}{4}$	1	0	$-\frac{1}{2}$
e_2	$-\frac{1}{4}$	$\frac{1}{4}$	1	0	$-\frac{1}{2}$
h_3	- 1	- 1	- 3	- 1	- 1
a_1	0	0	1	0	- 1
e_5	$\frac{1}{2}$	$\frac{1}{2}$	1	1	0

表 4.33 的拉格朗日矩阵是一个 2×2 矩阵, 其在 e_3 列上有个负的对角元素 - 3, 并且此列非基约束向量行上的元素都是正数. 于是, 作 $e_3^{+\Delta}$ 得可行下降方向

$$x_1 = \Delta, x_2 = \Delta, x_3 = \Delta,$$

其中 Δ 是任意正数, 相应的函数值为 $2f(x) = -3\Delta^2$.

例 4.3.4 在不等式组

$$-x_1 + x_2 - x_3 - x_4 \geq 0,$$

$$x_1, x_2, x_3, x_4 \geq 0$$

的解集中找一个 x 使二次函数 $f(x) = \frac{1}{2}x^T Hx < 0$, 其中 $H = \begin{pmatrix} 1 & -2 & 0 & -1 \\ -2 & 5 & -1 & 1 \\ 0 & -1 & 1 & 1 \\ -1 & 1 & 1 & 1 \end{pmatrix}$.

解 与前面类似, 初始表为表 4.34.

表 4.34 初 始 表

	e_1	e_2	e_3	e_4	e_5
h_1	1 *	- 2	0	- 1	1
h_2	- 2	5	- 1	1	- 1
h_3	0	- 1	1	1	1
h_4	- 1	1	1	1	1
a_1	- 1	1	- 1	- 1	0

表 4.34 中拉格朗日矩阵的 4 个对角元素都是正数, 不妨以第一个对角元素 1 为枢轴作前向旋转, 结果见表 4.35.

表 4.35 第一次前向旋转结果

	h_1	e_2	e_3	e_4	e_5
e_1	1	2	0	1	- 1
h_2	- 2	1 *	- 1	- 1	1
h_3	0	- 1	1	1	1
h_4	- 1	- 1	1	0	2
a_1	- 1	- 1	- 1	- 2	1

表 4.35 中的 3×3 拉格朗日矩阵没有负对角元素和 II 型矩阵, 再以其中的第一个对角元素 1 为枢轴作前向旋转, 结果见表 4.36.

表 4.36 第二次前向旋转结果

	h_1	h_2	e_3	e_4	e_5
e_1	5	2	2	3	- 3
e_2	2	1	1	1	- 1
h_3	- 2	- 1	0	0	2
h_4	- 3	- 1	0	- 1	3
a_1	- 3	- 1	- 2	- 3	2^*

表 4.36 中 2×2 拉格朗日矩阵有一个负对角元素 - 1, 但同列非基向量 a_1 行上的元素 - 3 < 0, 下降方向不可行. a_1 行上的对角元素为 2, 以其为枢轴作后向旋转, 结果见表 4.37.

表 4.37 第一次后向旋转结果

	h_1	h_2	e_3	e_4	a_1
e_1	$\frac{1}{2}$	$\frac{1}{2}$	- 1	$-\frac{3}{2}$	$-\frac{3}{2}$
e_2	$\frac{1}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	$-\frac{1}{2}$
h_3	1	0	2	3	1
h_4	$\frac{3}{2}$	$\frac{1}{2}$	3	$\frac{7}{2}$	$\frac{3}{2}$
e_5	$\frac{3}{2}$	$\frac{1}{2}$	1	$\frac{3}{2}$	$\frac{1}{2}$

由于右下角 3×3 拉格朗日矩阵的所有元素是正数, 没有可行下降方向.

例 4.3.5 在不等式组

$$\begin{aligned}
 & -2x_2 + x_3 + x_4 - x_5 = 0, \\
 & -2x_1 + x_2 + 4x_3 - 6x_4 + x_5 = 0, \\
 & x_1 + 8x_2 - 5x_3 - 2x_4 \geq 0, \\
 & x_1, x_2, x_3, x_4 \geq 0
 \end{aligned}$$

的解集中找一个 x 使 $f(x) = \frac{1}{2}x^THx < 0$, 其中 $H = \begin{pmatrix} -1 & -8 & 5 & -2 & 1 \\ -8 & 4 & 2 & 5 & 8 \\ 5 & 2 & -1 & 1 & 2 \\ -2 & 5 & 1 & 4 & 1 \\ 1 & 8 & 2 & 1 & 5 \end{pmatrix}$.

解 可以看出不等式组系数向量组的秩为 5 , 等于变量的数目 , 所以此问题相当于一个标准锥二次规划. 引入人工不等式 $x_5 \geq 0$, 初始表为表 4. 38. 其中 a_1 和 a_2 对应等式 , e_5 对应的变量 x_5 无非负要求.

表 4. 38 初 始 表

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
h_1	- 1	- 8	5	- 2	1	0	2	- 1
h_2	- 8	4	2	5	8	2	- 1	- 8
h_3	5	2	- 1	1	2	- 1	- 4	5
h_4	- 2	5	1	4	1	- 1	6	2
h_5	1	8	2	1	5	1 *	- 1	0
a_1	0	- 2	1	1	- 1 *	0	0	0
a_2	- 2	1	4	- 6	1	0	0	0
a_3	1	8	- 5	- 2	0	0	0	0

首先考虑将等式约束的系数向量 a_1 换入基中. a_1 行 e_5 列上的元素 $- 1 \neq 0$, 以其为枢轴作旋转运算不仅可把第一个等式换入基中 , 还可把人工不等式 $x_5 \geq 0$ 换出基. 以 $- 1$ 及其对称元素 1 为枢轴先后进行两次旋转运算 , 结果见表 4. 39. 其中 a_1 行及其对应的 e_6 列已删掉. 在给定的不等式组中没有不等式 $x_5 \geq 0$. 现在 e_5 成为非基向量 , e_5 行及其对应的 h_5 列本可以删掉 , 但考虑到要用 e_5 计算基本解的分量 x_5 , 特保留在表 4. 39 及以后各表内.

表 4.39 第一次双枢轴运算结果

	e_1	e_2	e_3	e_4	h_5	e_7	e_8
h_1	- 1	- 10	5	- 1	0	2	- 1
h_2	- 10	- 8	- 4	1	2	1	- 8
h_3	6	- 4	8	9	- 1	- 5	5
h_4	- 1	1	9	11	- 1	5 *	2
e_5	0	- 2	1	1	0	0	0
a_2	- 2	- 1	5	- 5 *	0	0	0
a_3	1	8	- 5	- 2	0	0	0

在表 4.39 中, a_2 也是等式的系数向量, 其所在行约束向量列上的元素都不为 0 不妨以该行上的 - 5 及其对称元素 5 为枢轴先后进行两次旋转运算, 结果见表 4.40 其中基等式 a_2 列以及其对应的非基向量 e_7 行已经删掉.

表 4.40 第二次双枢轴运算结果

	e_1	e_2	e_3	h_5	h_4	e_8
h_1	1.56	- 9.32	- 3	0.4	0.4	- 1.8
h_2	- 9.32	- 7.96	- 7	2.2	0.2	- 8.4
h_3	- 3	- 7	37	- 2	- 1 *	7
e_5	- 0.4	- 2.2	2	0	0	0
e_4	- 0.4	- 0.2	1 *	0	0	0
a_3	1.8	8.4	- 7	0	0	0

在表 4.40 中, e_2 列上的对角元素为 - 7.96 此列 e_5 和 e_4 行上的元素都是负数. e_5 对应的变量 x_5 无非负要求, 但 e_4 对应的变量 x_4 要求非负. 以 e_4 行的正元素 1 及其对称元素 - 1 为枢轴先后进行两次旋转运算, 结果见表 4.41.

表 4.41 第三次双枢轴运算结果

	e_1	$e_2^{+\Delta}$	e_4	h_5	h_3	e_8
h_1	5.08	- 9.76	11.8	- 0.4	- 0.4	1
h_2	- 9.76	- 9.28	0.4	1.8	- 0.2	- 7
h_4	11.8	0.4	37	- 2	- 1	7
e_5	0.4	- 1.8	2	0	0	0
e_3	0.4	0.2	1	0	0	0
a_3	- 1	7	- 7	0	0	0

由表 4.41 可见，基向量 e_2 列对角线上的元素 $- 9.28 < 0$ ，同列非基约束向量 e_3 和 a_3 行上的元素为正，作 $e_2^{+\Delta}$ 得

$$x_1 = 0, \quad x_2 = \Delta, \quad x_3 = 0.2\Delta, \quad x_4 = 0, \quad x_5 = - 1.8\Delta,$$

其中 Δ 是任意正数，对应的函数值为 $2f(x) = - 9.28\Delta^2$.

4.3.3 非标准锥二次规划

非标准锥二次规划可表示为以下形式：

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx = \frac{1}{2} \begin{pmatrix} x_I^T & x_{II}^T \end{pmatrix} \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} x_I \\ x_{II} \end{pmatrix}, \\ \text{s. t. } Ax_I &\geq 0, \\ x_{II} &\geq 0. \end{aligned} \right\} \quad (4.3.5)$$

引入人工不等式 $x_{II} \geq 0$ ，初始表为表 4.42.

表 4.42 初始表

	e_I	e_{II}	e_{III}
h_I	H_{11}	H_{12}	$- A^T$
h_{II}	H_{12}^T	H_{22}	O
a	A	O	O

其中 e_I 和 e_{II} 分别表示约束不等式 $x_I \geq 0$ 和人工不等式 $x_{II} \geq 0$ 的系数向量，其他符号不再赘述.

以下情况表 4.42 存在明显的可行下降方向，其中 $\Delta > 0$.

(i) 基约束对应的主子矩阵 H_{11} 有个负的对角元素 h_{ss} 且 A 的第 s 列非负。作 $e_s^{+\Delta}$, 即将基约束不等式 $x_s \geq 0$ 的右端项换为 Δ , 将产生可行下降方向

$$x_s = \Delta; \quad x_i = 0, \quad i = 1, 2, \dots, n, i \neq s.$$

(ii) 基人工不等式对应的主子矩阵 H_{22} 有个负的对角元素 h_{ss} . 作 $e_s^{+\Delta}$ 或作 $e_s^{-\Delta}$, 即无论将基人工不等式 $x_s \geq 0$ 的右端项换为 Δ 或 $-\Delta$, 将产生可行下降方向.

下面设 H 有个 2×2 主子矩阵 $\begin{pmatrix} 0 & h_{sr} \\ h_{sr} & 0 \end{pmatrix}$, 其对应的基向量分别为 e_s 和 e_r , 或者说对应的基不等式分别为 $x_s \geq 0$ 和 $x_r \geq 0$.

(iii) $x_s \geq 0$ 和 $x_r \geq 0$ 都是模型 (4.3.5) 的约束, $h_{sr} < 0$, 并且 A 的每一行第 s 列和第 r 列上的两个元素之和非负. 作 $e_s^{+\Delta}$ 和 $e_r^{+\Delta}$ 将产生可行下降方向

$$x_s = x_r = \Delta; \quad x_i = 0, \quad i = 1, 2, \dots, n, i \neq s, i \neq r.$$

(iv) $x_s \geq 0$ 是约束不等式, $x_r \geq 0$ 是人工不等式, A 的第 s 列非负, $h_{sr} \neq 0$. $h_{sr} < 0$ 时, 作 $e_s^{+\Delta}$ 和 $e_r^{+\Delta}$; $h_{sr} > 0$ 时, 作 $e_s^{+\Delta}$ 和 $e_r^{-\Delta}$, 将产生可行下降方向.

(v) $x_s \geq 0$ 和 $x_r \geq 0$ 都是人工不等式, $h_{sr} \neq 0$. $h_{sr} < 0$ 时, 作 $e_s^{+\Delta}$ 和 $e_r^{+\Delta}$; $h_{sr} > 0$ 时, 作 $e_s^{+\Delta}$ 和 $e_r^{-\Delta}$, 将产生可行下降方向.

下面设表 4.42 不存在明显的可行下降方向.

如果 H_{22} 正定, 在表 4.42 中沿 H_{22} 的对角线进行旋转运算, 并删掉枢轴行和枢轴列, 得到的表相当于前面标准锥二次规划的旋转算法表. 如果 H_{22} 不是正定的但有正的对角元素. 以正的对角元素为枢轴进行旋转运算, 并删掉枢轴行和枢轴列, 直到不能进行为止. 那么, 剩余基人工不等式对应的主子矩阵或者有负的对角元素; 或者有一个 2 阶主子矩阵, 其两个对角元素为 0, 非对角元素不等于 0; 或者所有元素等于 0. 不妨设表 4.42 的 H_{22} 是这样一个矩阵, 在前两种情况下很容易确定一个可行下降方向, 见上述 (ii) 和 (v). 若 $H_{22} = 0$ 且 $H_{12} = 0$, 表 4.42 相当于标准锥二次规划的旋转算法表. 若 $H_{22} = 0$ 但 $H_{12} \neq 0$, 表 4.42 成为表 4.43.

表 4.43 初始表 ($H_{12} \neq 0$)

	e_I	e_{II}	e_{III}
h_I	H_{11}	H_{12}	$-A^T$
h_{II}	H_{12}^T	O	O
a	A	O	O

这时可按以下方式寻求可行下降方向. 不考虑 e_{II} 列和 h_{II} 行, 将其作为标准锥二次规划计算可行下降方向. 在计算过程中如果 e_{II} 对应的主子矩阵出现负的对角元素, 或者出现二阶主子矩阵, 其两个对角元素为 0, 非对角元素不等于 0, 利用它们产生的下降方向可能是可行下降方向.

例 4.3.6 在不等式组

$$-x_1 + x_2 \geq 0, \quad x_1 \geq 0, \quad x_2 \geq 0$$

的解集中找一个 x 使 $f(x) = \frac{1}{2}x^T Hx < 0$, 其中 $H = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$.

解 引入人工不等式 $x_3 \geq 0$, 初始表如表 4.44 所示.

表 4.44 不存在明显的可行下降方向, 不妨作前向旋转 $h_1 \leftrightarrow e_1$, 结果见表 4.45.

表 4.44 初 始 表

	e_1	e_2	e_3	e_4
h_1	1^*	1	1	1
h_2	1	1	1	- 1
h_3	1	1	0	0
a_1	- 1	1	0	0

表 4.45 第一次前向旋转结果

	h_1	e_2	$e_3^{-\Delta}$	e_4
e_1	1	- 1	- 1	- 1
h_2	1	0	0	- 2
h_3	1	0	- 1	- 1
a_1	- 1	2	1	1^*

表中人工不等式所在 e_3 列的对角元素变为 - 1, 但该列非基约束向量行上的元素一个为负一个为正. 如果作 $e_3^{+\Delta} \leftrightarrow e_1$ 的偏差 $-\Delta < 0$; 如果作 $e_3^{-\Delta} \leftrightarrow a_1$ 的偏差 $-\Delta < 0$. 在前一情况作后向旋转 $e_1 \leftrightarrow h_1$ 将回复到初始表, 在后一情况作 $a_1 \leftrightarrow e_4$ 结果为表 4.46.

表 4. 46 第一次后向旋转结果

	h_1	e_2	e_3	a_1
e_1	0	1	0	- 1
h_2	- 1	4 *	2	- 2
h_3	0	2	0	- 1
e_4	1	- 2	- 1	1

在表 4. 46 中 ,拉格朗日矩阵的三个对角元素非负 ,第一个对角元素为 4 ,以其为枢轴作前向旋转得表 4. 47.

表 4. 47 第二次前向旋转结果

	h_1	e_2	$e_3^{-\Delta}$	a_1	
e_1	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{\Delta}{2}$
e_2	$\frac{1}{4}$	$\frac{1}{4}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{\Delta}{2}$
h_3	$\frac{1}{2}$	$\frac{1}{2}$	- 1	0	Δ
e_4	$\frac{1}{2}$	$-\frac{1}{2}$	0	0	0

在表 4. 47 作 $e_3^{-\Delta}$,各非基向量偏差见最后一列. 由于非基约束向量 e_1 和 e_2 的偏差为正 , e_3 列的对角元素为负 ,可行下降方向为 $x_1 = \frac{\Delta}{2}$, $x_2 = \frac{\Delta}{2}$, $x_3 = - \Delta$,对应的目标函数值 $f(x) = - \frac{\Delta^2}{2}$.

事实上 ,一开始就在表 4. 44 中作前向旋转 $e_2 \leftrightarrow h_2$ 便可得到一个可行下降方向. 该旋转运算的结果见表 4. 48.

表 4. 48 在初始表中作 $e_2 \leftrightarrow h_2$ 的结果

	e_1	h_2	$e_3^{-\Delta}$	e_4	
h_1	0	1	0	2	0
e_2	- 1	1	- 1	1	Δ
h_3	0	1	- 1	1	Δ
a_1	- 2	1	- 1	1	Δ

在表 4.48 作 $e_3^{-\Delta}$, 各非基向量偏差见最后一列. 此时的可行下降方向为 $x_1 = 0, x_2 = \Delta, x_3 = -\Delta$ 对应的目标函数值 $f(x) = -\frac{\Delta^2}{2}$.

例 4.3.7 在不等式组

$$\begin{aligned} x_1 - x_2 - x_3 &\geq 0, \\ -x_1 + 2x_2 - 4x_3 &\geq 0 \end{aligned}$$

的解集中找一个 x 使 $f(x) = \frac{1}{2}x^THx < 0$, 其中 $H = \begin{pmatrix} 1 & 1 & -1 \\ 1 & 2 & -3 \\ -1 & -3 & 4 \end{pmatrix}$.

解 由于不等式组系数向量组的秩为 2, 小于变量的数目, 此题相当于一个非标准锥二次规划. 引入人工不等式 $x_1, x_2, x_3 \geq 0$ 并以其为初始基本不等式组, 初始表如表 4.49 所示.

表 4.49 初 始 表

	e_1	e_2	e_3	e_4	e_5
h_1	1	1	-1	-1*	1
h_2	1	2	-3	1	-2
h_3	-1	-3	4	1	4
a_1	1*	-1	-1	0	0
a_2	-1	2	-4	0	0

首先利用双枢轴运算尽可能将基人工不等式换为非基不等式. 以非基向量 a_1 行的正元素 1 及其对称元素为枢轴先后进行两次旋转运算, 结果见表 4.50.

表 4.50 第一次双枢轴运算结果

	a_1	e_2	e_3	h_1	e_5
e_4	1	2	0	-1	1
h_2	2	5	-2	-1	-1*
h_3	0	-2	3	-1	5
e_1	1	1	1	0	0
a_2	-1	1*	-5	0	0

我们的二次规划模型为

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx + cx, \\ \text{s. t. } a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (4.4.1)$$

其中 $c = (c_1, c_2, \dots, c_n)$, $x = (x_1, x_2, \dots, x_n)^T$, $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$, $i = 1, 2, \dots, m$, $H = (h_{ij})_{n \times n}$ 不定或负半定. 以上模型的 KT 条件为

$$\left. \begin{aligned} H_i x - p_i^T \lambda &\geq -c_i, \quad i = 1, 2, \dots, n, \\ a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n, \\ \lambda_i &\geq 0, \quad i = l+1, l+2, \dots, m, \\ (H_i x - p_i^T \lambda + c_i)(x_i - l_i) &= 0, \quad i = 1, 2, \dots, n, \\ \lambda_i(a_i x - b_i) &= 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (4.4.2)$$

其中 H_i 是 H 的第 i 行, p_i 是 $a_1, a_2, \dots, a_m$ 构成的 $m \times n$ 矩阵的第 i 列, $i = 1, 2, \dots, n$, $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_m)^T$.

如 3.1 节所述, 二次规划的计算过程分为构造初始表、预处理和主迭代三个阶段. 非凸二次规划解法的前两个步骤与凸二次规划相同. 首先引入人工不等式 $\lambda_i \geq 0, i = 1, 2, \dots, l$, 并以

$$x_1 \geq l_1, \dots, x_n \geq l_n, \lambda_1 \geq 0, \dots, \lambda_m \geq 0$$

为初始基本不等式组构造初始表, 其矩阵形式如表 4.53 所示.

表 4.53 初始表的矩阵形式

	e_I	e_{II}	e_{III}	e_{IV}	
h_I	H_{11}	H_{12}	$-A_{11}^T$	$-A_{21}^T$	σ_I
h_{II}	H_{12}^T	H_{22}	$-A_{12}^T$	$-A_{22}^T$	σ_{II}
a_I	A_{11}	A_{12}	O	O	σ_{III}
a_{II}	A_{21}	A_{22}	O	O	σ_{IV}

预处理的结果见表 4. 54.

表 4. 54 预处理的结果

	a_I	e_{II}	h_I	e_{IV}	
e_{III}	G_{11}	G_{12}	$-B_{11}^T$	$-B_{21}^T$	σ'_I
h_{II}	G_{12}^T	G_{22}	$-B_{12}^T$	$-B_{22}^T$	σ'_{II}
e_I	B_{11}	B_{12}	O	O	σ'_{III}
a_{II}	B_{21}	B_{22}	O	O	σ'_{IV}

一般来说，旋转算法表用表 4. 55 表示.

表 4. 55 一般表的矩阵形式

	e_B	a_B	h_B	e_{n+B}	
h_N	L_{11}	L_{12}	$-E_{11}^T$	$-E_{21}^T$	σ_h
e_{n+N}	L_{12}^T	L_{22}	$-E_{12}^T$	$-E_{22}^T$	σ_λ
e_N	E_{11}	E_{12}	D_{11}	D_{12}	σ_e
a_N	E_{21}	E_{22}	D_{12}^T	D_{22}	σ_a

其中 e_B 和 e_N 分别是以 $\{e_1, e_2, \dots, e_n\}$ 中的基向量和非基向量为分量的列向量， a_B 和 a_N 分别是以 $\{a_1, a_2, \dots, a_m\}$ 中的基向量和非基向量为分量的列向量， h_B 和 h_N 分别是以 $\{h_1, h_2, \dots, h_n\}$ 中的基向量和非基向量为分量的列向量， e_{n+B} 和 e_{n+N} 分别是以 $\{e_{n+1}, e_{n+2}, \dots, e_{n+m}\}$ 中的基向量和非基向量为分量的列向量。 $\sigma_h, \sigma_\lambda, \sigma_e, \sigma_a$ 是由非基向量的偏差构成的列向量。

如前所述，表 4. 55 中非基拉格朗日向量行和基约束向量列的元素构成的方阵 $\begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix}$ 称为拉格朗日矩阵，其阶数等于基约束的数目。非基约束向量行和基拉格朗日向量列的元素构成的方阵 $\begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix}$ 称为约束矩阵，这是一个正半定矩阵。

主迭代的运算方式有两种：一种是以对角线上的正元素为枢轴所作的旋转运算，称为主旋转。其中，以拉格朗日矩阵的正对角元素为枢轴进行的旋转运

算称为前向旋转,以约束矩阵的正对角元素为枢轴进行的旋转运算称为后向旋转.另一种运算形式称为双枢轴运算,它是当约束矩阵的对角元素等于零时,以同行的一个非零元素及其对称元素为枢轴先后进行的两次旋转运算.

但是在主迭代的计算细节上,非凸二次规划比凸二次规划要复杂很多.对于凸二次规划,只要通过旋转运算使所有非基向量(不包括基等式约束对应的非基向量)偏差非负,即使基本解为KT点,就是全局极小点.对于非凸二次规划,KT点可能连局部极小点都不是.对于非凸二次规划,应尽量利用拉格朗日矩阵中的负对角元素,使目标函数值有足够的改进,不能简单地以偏差最负的非基向量入基.

4.4.2 参数公式及其意义

下面考虑基不等式右端的常数项发生变化时,基本解及对应的目标函数值如何变化.有关计算公式称为参数公式.利用这些公式可判断KT点处是否有可行下降方向,以寻求更好的解.

由表4.55可见,非基向量 h_N 和 e_{n+N} 关于基向量的表达式为

$$h_N = L_{11}e_B + L_{12}a_B - E_{11}^T h_B - E_{21}^T e_{n+B}, \quad (4.4.3)$$

$$e_{n+N} = L_{12}^T e_B + L_{22}a_B - E_{12}^T h_B - E_{22}^T e_{n+B}. \quad (4.4.4)$$

设当前基本解为 $y = \begin{pmatrix} x \\ \lambda \end{pmatrix}$,那么

$$e_B x = l_B, \quad a_B x = b_B, \quad h_B y = -c_N^T, \quad e_{n+B} y = 0.$$

其中, l_B 是由 $x_i \geq l_i (i=1, 2, \dots, n)$ 中的基不等式的右端项 l_i 构成的列向量, b_B 是由 $a_i x = b_i$ 和 $a_i x \geq b_i (i=1, 2, \dots, m)$ 中的基(不)等式的右端项 b_i 构成的列向量, c_N 是目标函数线性部分的系数向量 c 中与非基变量 x_N 对应的部分.

在(4.4.3)和(4.4.4)两式两边右乘 $y = \begin{pmatrix} x \\ \lambda \end{pmatrix}$,由于 $\sigma_h = h_N y + c_B^T \sigma_\lambda = e_{n+N} y$,非基向量 h_N 和 e_{n+N} 的偏差可表示为

$$\sigma_h = L_{11}l_B + L_{12}b_B + E_{11}^T c_N^T + c_B^T, \quad (4.4.5)$$

$$\sigma_\lambda = L_{12}^T l_B + L_{22}b_B + E_{12}^T c_N^T. \quad (4.4.6)$$

同样,由表4.55有

$$e_N = E_{11}e_B + E_{12}a_B + D_{11}h_B + D_{12}e_{n+B},$$

$$a_N = E_{21}e_B + E_{22}a_B + D_{12}^T h_B + D_{12}e_{n+B}.$$

在以上两式两边右乘 $y = \begin{pmatrix} x \\ \lambda \end{pmatrix}$,由于 $\sigma_e = e_N x - l_N \sigma_a = a_N x - b_N$,非基向量 e_N 和 a_N 的偏差可表示为

$$\sigma_e = E_{11}l_B + E_{12}b_B - D_{11}c_N^T - l_N, \quad (4.4.7)$$

$$\sigma_a = E_{21}l_B + E_{22}b_B - D_{12}c_N^T - b_N. \quad (4.4.8)$$

由于

$$\begin{aligned} 2f(x) &= x^T H x + 2cx = \sum_{i=1}^n x_i H_i x + 2 \sum_{i=1}^n c_i x_i \\ &= \sum_{i=1}^n x_i (H_i x - p_i^T \lambda + c_i) + \sum_{i=1}^n x_i p_i^T \lambda + \sum_{i=1}^n c_i x_i \\ &= \sum_{i=1}^n x_i (h_i y + c_i) + \sum_{i=1}^m \lambda_i a_i x + \sum_{i=1}^n c_i x_i. \end{aligned} \quad (4.4.9)$$

其中 $y = \begin{pmatrix} x \\ \lambda \end{pmatrix}$. 将基本解 $y = \begin{pmatrix} x \\ \lambda \end{pmatrix}$ 代入上式, 考虑到

$$\begin{aligned} x_i &= l_i, i \in B_e; & \sigma_i &= h_i y + c_i, i \in N_h, \\ h_i y + c_i &= 0, i \in B_h, \\ a_i x &= b_i, i \in B_a; & \lambda_i &= \sigma_i, i \in N_\lambda, \\ x_i &= l_i + \sigma_i, i \in N_e, \end{aligned}$$

其中, B_e 和 N_e , B_a 和 N_a , B_h 和 N_h , B_λ 和 N_λ 分别为 $\{e_1, e_2, \dots, e_n\}, \{a_1, a_2, \dots, a_m\}, \{h_1, h_2, \dots, h_n\}, \{e_{n+1}, e_{n+2}, \dots, e_{n+m}\}$ 中基向量和非基向量的指标集. 所以

$$\begin{aligned} 2f(x) &= \sum_{i \in B_e} l_i \sigma_i + \sum_{i \in B_a} \sigma_i b_i + \sum_{i \in B_e} c_i l_i + \sum_{i \in N_e} c_i (l_i + \sigma_i) \\ &= l_B^T \sigma_h + b_B^T \sigma_\lambda + c_B l_B + c_N (l_N + \sigma_e). \end{aligned} \quad (4.4.10)$$

现在, 让基不等式右端的常数项变化. 设基不等式 $x_B \geq l_B$, $a_B x \geq b_B$, $h_B y \geq -c_N^T e_{n+B} y \geq 0$ 的右端项分别变为 $l_B + \delta_e$, $b_B + \delta_a$, $-c_N^T + \delta_h$, δ_λ , 其中“ $\geq$ ”表示“ $=$ ”或“ $\geq$ ”, 分别对应等式和不等式. 根据第二章的参数公式(2.4.2), 表 4.55 中非基向量 h_N, e_{n+N}, e_N, a_N 的偏差分别为

$$\begin{aligned} \sigma'_h &= \sigma_h + L_{11} \delta_e + L_{12} \delta_a - E_{11}^T \delta_h - E_{21}^T \delta_\lambda, \\ \sigma'_\lambda &= \sigma_\lambda + L_{12}^T \delta_e + L_{22} \delta_a - E_{12}^T \delta_h - E_{22}^T \delta_\lambda, \\ \sigma'_e &= \sigma_e + E_{11} \delta_e + E_{12} \delta_a + D_{11} \delta_h + D_{12} \delta_\lambda, \\ \sigma'_a &= \sigma_a + E_{21} \delta_e + E_{22} \delta_a + D_{12}^T \delta_h + D_{22} \delta_\lambda. \end{aligned}$$

设此时的基本解为 $\tilde{y} = \begin{pmatrix} \tilde{x} \\ \tilde{\lambda} \end{pmatrix}$, 依定义它满足

$$\begin{aligned}x_B &= l_B + \delta_e, & h_N y + c_B^T &= \sigma'_h, \\h_B y + c_N^T &= \delta_h, & x_N &= l_N + \sigma'_e, \\a_B x &= b_B + \delta_a, & \lambda_N &= \sigma'_\lambda, \\\lambda_B &= \delta_\lambda, & a_N x &= b_N + \sigma'_a.\end{aligned}$$

于是

$$\begin{aligned}2f(\tilde{x}) &= \sum_{i=1}^n x_i (h_i y + c_i) + \sum_{i=1}^m \lambda_i a_i x + \sum_{i=1}^n c_i x_i \\&= (l_B^T + \delta_e^T) \sigma'_h + \delta_h^T (l_N + \sigma'_e) + (b_B^T + \delta_a^T) \sigma'_\lambda + \delta_\lambda^T (b_N + \sigma'_a) \\&\quad + c_B (l_B + \delta_e) + c_N (l_N + \sigma'_e) \\&= (l_B^T + \delta_e^T) (\sigma_h + L_{11} \delta_e + L_{12} \delta_a - E_{11}^T \delta_h - E_{21}^T \delta_\lambda) \\&\quad + \delta_h^T (l_N + \sigma_e + E_{11} \delta_e + E_{12} \delta_a + D_{11} \delta_h + D_{12} \delta_\lambda) \\&\quad + (b_B^T + \delta_a^T) (\sigma_\lambda + L_{12}^T \delta_e + L_{22} \delta_a - E_{12}^T \delta_h - E_{22}^T \delta_\lambda) \\&\quad + \delta_\lambda^T (b_N + \sigma_a + E_{21} \delta_e + E_{22} \delta_a + D_{12}^T \delta_h + D_{22} \delta_\lambda) \\&\quad + c_B (l_B + \delta_e) + c_N (l_N + \sigma_e + E_{11} \delta_e + E_{12} \delta_a + D_{11} \delta_h + D_{12} \delta_\lambda).\end{aligned}$$

所以

$$\begin{aligned}2f(\tilde{x}) - 2f(x) &= (l_B^T + \delta_e^T) (L_{11} \delta_e + L_{12} \delta_a - E_{11}^T \delta_h - E_{21}^T \delta_\lambda) + \delta_e^T \sigma_h \\&\quad + \delta_h^T (l_N + \sigma_e + E_{11} \delta_e + E_{12} \delta_a + D_{11} \delta_h + D_{12} \delta_\lambda) \\&\quad + (b_B^T + \delta_a^T) (L_{12}^T \delta_e + L_{22} \delta_a - E_{12}^T \delta_h - E_{22}^T \delta_\lambda) + \delta_a^T \sigma_\lambda \\&\quad + \delta_\lambda^T (b_N + \sigma_a + E_{21} \delta_e + E_{22} \delta_a + D_{12}^T \delta_h + D_{22} \delta_\lambda) \\&\quad + c_B \delta_e + c_N (E_{11} \delta_e + E_{12} \delta_a + D_{11} \delta_h + D_{12} \delta_\lambda) \\&= \delta_e^T (L_{11} \delta_e + L_{12} \delta_a) + \delta_a^T (L_{12}^T \delta_e + L_{22} \delta_a) \\&\quad + \delta_h^T (D_{11} \delta_h + D_{12} \delta_\lambda) + \delta_\lambda^T (D_{12}^T \delta_h + D_{22} \delta_\lambda) \\&\quad + (l_B^T L_{11} + b_B^T L_{12}^T + c_N E_{11} + c_B) \delta_e + \delta_e^T \sigma_h \\&\quad + (l_B^T L_{12} + b_B^T L_{22}^T + c_N E_{12}) \delta_a + \delta_a^T \sigma_\lambda \\&\quad + (-l_B^T E_{11}^T - b_B^T E_{12}^T + c_N D_{11} + l_N^T) \delta_h + \delta_h^T \sigma_e \\&\quad + (-l_B^T E_{21}^T - b_B^T E_{22}^T + c_N D_{12} + b_N^T) \delta_\lambda + \delta_\lambda^T \sigma_a.\end{aligned}$$

将(4.4.5)至(4.4.8)各式代入上式,化简后可得到

$$\begin{aligned}2f(\tilde{x}) - 2f(x) &= 2(\sigma_h^T, \sigma_\lambda^T) \begin{pmatrix} \delta_e \\ \delta_a \end{pmatrix} + (\delta_e^T, \delta_a^T) \begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix} \begin{pmatrix} \delta_e \\ \delta_a \end{pmatrix} + (\delta_h^T, \delta_\lambda^T) \begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix} \begin{pmatrix} \delta_h \\ \delta_\lambda \end{pmatrix}.\end{aligned}\tag{4.4.11}$$

下面考虑公式(4.4.11)的几种较为特殊的情况. 用 $x^{(I)}$ 表示仅仅将基约束 $x_B \geq l_B$ 和 $a_B x \geq b_B$ 的右端项分别变为 $l_B + \delta_e$ 和 $b_B + \delta_a$ 得到的基本解, 那么

$$2f(x^{(I)}) - 2f(x) = 2(\sigma_h^T \sigma_\lambda^T) \begin{pmatrix} \delta_e \\ \delta_a \end{pmatrix} + (\delta_e^T, \delta_a^T) \begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix} \begin{pmatrix} \delta_e \\ \delta_a \end{pmatrix}. \quad (4.4.12)$$

用 $x^{(II)}$ 表示仅仅将基拉格朗日不等式 $h_B y \geq -c_N^T$ 和 $e_{n+B} y \geq 0$ 的右端项分别变为 $-c_N^T + \delta_h$ 和 δ_λ 得到的基本解, 那么

$$2f(x^{(II)}) - 2f(x) = (\delta_h^T \delta_\lambda^T) \begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix} \begin{pmatrix} \delta_h \\ \delta_\lambda \end{pmatrix}. \quad (4.4.13)$$

特殊地讲, 如果仅有一个基约束 $x_s \geq l_s$ (或 $a_s x \geq b_s$) 右端项的改变量为 Δ , 其他基约束的右端项不变, 则目标函数的改变量为

$$f(\bar{x}) - f(x) = \sigma_s \Delta + w_{ss} \frac{\Delta^2}{2}. \quad (4.4.14)$$

其中, σ_s 是变化之前基向量 e_s (或 a_s) 对应的非基拉格朗日向量 h_s (或 e_{n+s}) 的偏差, w_{ss} 是该行的对角元素, $\bar{x}$ 则是基约束不等式的右端项发生变化后得到的基本解.

在搜索 x 处的下降方向时, 公式(4.4.14)非常管用. 本章 4.2 节和 4.3 节主要用这个公式确定二次齐次函数在 $x=0$ 处的下降方向, 只不过非基向量的偏差 $\sigma_s=0$. (4.4.14) 式是一个关于 Δ 的二次多项式, 可以看出

(i) 如果 $w_{ss} < 0$ 且 $\sigma_s \leq 0$, 或者 $w_{ss} = 0$ 且 $\sigma_s < 0$, 则对于任何 $\Delta > 0$ 都有 $f(\bar{x}) < f(x)$;

(ii) 如果 $w_{ss} = 0$ 且 $\sigma_s = 0$, 则对于任何 $\Delta \geq 0$ 都有 $f(\bar{x}) = f(x)$;

(iii) 如果 $w_{ss} > 0$ 且 $\sigma_s < 0$, 则当 $0 < \Delta < \frac{-2\sigma_s}{w_{ss}}$ 时 $f(\bar{x}) < f(x)$, 当 $\Delta > \frac{-2\sigma_s}{w_{ss}}$ 时 $f(\bar{x}) > f(x)$.

4.4.3 非凸二次规划的有关定理

引理 4.4.1 在旋转算法表 4.55 中, 适当改变基拉格朗日不等式 $h_B y \geq -c_N^T$ 的右端项, 可得到基约束相应方程组的任何一个解.

证 表 4.55 的基约束相应的方程组为

$$e_B x = l_B,$$

$$a_B x = b_B.$$

设基约束的数目为 k . 如果 $k < n$, 那么此方程组有无穷多个解, 包括当前基本解 x . 若基拉格朗日不等式 $h_{By} \geq -c_N^T$ 的右端项变为 $-c_N^T + \delta_h$, 则非基单位向量 e_N 的偏差变为 $\sigma_e + D_{11}\delta_h$. 所以新的基本解为

$$\begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} l_B \\ l_N + \sigma_e + D_{11}\delta_h \end{pmatrix} = \begin{pmatrix} l_B \\ l_N + \sigma_e \end{pmatrix} + \begin{pmatrix} O & O \\ O & D_{11} \end{pmatrix} \begin{pmatrix} 0 \\ \delta_h \end{pmatrix}.$$

根据第三章的定理 3.1.1, D_{11} 正半定且其秩等于 $n - k$. 所以适当选择 δ_h 可得到基约束相应方程组的任何一个解. ■

定理 4.4.1 用旋转算法解二次规划, 基本解是目标函数在基约束相应方程组解集上的全局极小点.

证 如果基约束有 n 个, 它们相应的方程组只有惟一解, 此定理当然成立. 下面设基约束的数目 $k < n$. 不妨考虑表 4.55, 基本解为 x . 设 $\bar{x}$ 是基约束相应方程组的任意一个解, 根据引理 4.4.1, 适当改变基拉格朗日不等式 $h_{By} \geq -c_N^T$ 的右端项可得到 $\bar{x}$. 在 (4.4.13) 式中令 $\delta_\lambda = 0$, 由于 D_{11} 正半定, 可见 $f(\bar{x}) \geq f(x)$. ■

引理 4.4.2 用 S 表示基约束构成的不等式组的解集, 那么适当改变基约束不等式和基拉格朗日不等式 $h_{By} \geq -c_N^T$ 的右端项可得到 S 内任一点.

证 考虑表 4.55, 基约束构成的不等式组为

$$e_B x \geq l_B,$$

$$a_B x \geq b_B.$$

其中“ $\geq$ ”表示等式对应的式子应取等号, 设 $\tilde{x}$ 是其解集 S 内任意一个点. 如果基约束的数目为 n , 令 l_B 和 b_B 分别变为

$$l_B + \delta_e = l_B + (e_B \tilde{x} - l_B) \text{ 和 } b_B + \delta_a = b_B + (a_B \tilde{x} - b_B),$$

则表 4.55 将产生 $\tilde{x}$. 其中 δ_e 和 δ_a 的各分量非负且 δ_a 的与基等式对应的分量为 0.

如果基约束的数目小于 n , 令 $d = \tilde{x} - x$, 其中 x 表示当前基本解. 用 d_s 表示 d 在 e_B 和 a_B 张成的空间上的直交投影, d_z 表示 d 在 e_B 和 a_B 的零空间上的直交投影. 那么, $\tilde{x} = x + d_s + d_z$. 与上面一样, 适当改变基约束不等式右端的常数项, 可得到 $x + d_s$. 然后适当改变基拉格朗日不等式 $h_{By} \geq -c_N^T$ 的右端项, 便可得到 $\tilde{x} = x + d_s + d_z$. ■

第三章的 3.1 节已说明, 将等式约束都换为基约束后的旋转算法表称为一般表. 删掉一般表中基等式约束所在列及其对应的非基拉格朗日向量行, 得到的表称为简约表. 现在设简约表的所有非基向量偏差非负, 即其对应的基本解是 KT 点. 如果有些非基拉格朗日向量的偏差等于 0, 用 T_0 表示偏差为 0 的所有

非基向量对应的主子矩阵,包括其相应的非基向量和基向量. 根据旋转算法表的形式不变性, T_0 相当于模型(4.3.4)的旋转算法表. 通过 T_0 求一个可行下降方向的问题称为锥问题.

定理 4.4.2(局部极小点的必要充分条件) 考虑二次规划(4.4.1), 设简约表的所有非基向量偏差非负, 并且至少有一个非基拉格朗日向量的偏差等于 0. 那么, 当前基本解是局部极小点的必要充分条件是锥问题没有可行下降方向.

证 不失一般性, 假设没有等式约束, 并设表 4.56 为简约表.

表 4.56 简 约 表

	a'_{B1}	a'_{B2}	h'_{B1}	h'_{B2}	
h'_{N1}	L_{11}	L_{12}	$-E_{11}^T$	$-E_{21}^T$	σ_{N1}
h'_{N2}	L_{12}^T	L_{22}	$-E_{12}^T$	$-E_{22}^T$	0
a'_{N1}	E_{11}	E_{12}	D_{11}	D_{12}	0
a'_{N2}	E_{21}	E_{22}	D_{12}^T	D_{22}	σ_{N2}

其左上角的 2×2 分块矩阵为拉格朗日矩阵, 右下角的 2×2 分块矩阵为约束矩阵, h'_{N1} 和 a'_{B1} 分别表示 L_{11} 对应的非基拉格朗日向量和基约束向量, 其他子块对应的非基向量和基向量的符号类似. 非基向量 h'_{N1} 的偏差 σ_{N1} 和 a'_{N2} 的偏差 σ_{N2} 的所有分量为正 h'_{N2} 和 a'_{N1} 的偏差为零向量. 所以锥问题的旋转算法表为表 4.57.

表 4.57 锥问题的旋转算法表

	a'_{B2}	h'_{B1}
h'_{N2}	L_{22}	$-E_{12}^T$
a'_{N1}	E_{12}	D_{11}

设当前基本解为 x . 我们将证明, 如果 x 不是局部极小点, 则锥问题有可行下降方向; 如果锥问题有可行下降方向, 那么 x 不是局部极小点.

(i) 设 x 不是局部极小点, 那么在 x 附近有一个可行解 $\tilde{x}$, 使得 $f(\tilde{x}) < f(x)$. 根据引理 4.4.2, 适当改变基不等式的右端项可得到 $\tilde{x}$, 设基不等式右端的改变量分别为 $\delta_{a1}, \delta_{a2}, \delta_{h1}, \delta_{h2}$, 其中 δ_{a1} 和 δ_{a2} 的分量是充分小的正数或 0, δ_{h1} 和 δ_{h2} 的分量充分接近 0. 根据参数公式(4.4.11),

$$2f(\tilde{x}) - 2f(x)$$

$$= 2(\sigma_{N1}^T, 0^T) \begin{pmatrix} \delta_{a1} \\ \delta_{a2} \end{pmatrix} + (\delta_{a1}^T, \delta_{a2}^T) \begin{pmatrix} L_{11} & L_{12} \\ L_{12}^T & L_{22} \end{pmatrix} \begin{pmatrix} \delta_{a1} \\ \delta_{a2} \end{pmatrix} + (\delta_{h1}^T, \delta_{h2}^T) \begin{pmatrix} D_{11} & D_{12} \\ D_{12}^T & D_{22} \end{pmatrix} \begin{pmatrix} \delta_{h1} \\ \delta_{h2} \end{pmatrix}$$

$$= (2\sigma_{N1}^T + \delta_{a1}^T L_{11} + 2\delta_{a1}^T L_{12}) \delta_{a1} + \delta_{a2}^T L_{22} \delta_{a2} + \delta_{h1}^T D_{11} \delta_{h1} + 2\delta_{h1}^T D_{12} \delta_{h2} + \delta_{h2}^T D_{22} \delta_{h2}.$$

由于 $2f(\tilde{x}) - 2f(x) < 0$, 所以

$$\delta_{a2}^T L_{22} \delta_{a2} + \delta_{h1}^T D_{11} \delta_{h1}$$

$$< - (2\sigma_{N1}^T + \delta_{a1}^T L_{11} + 2\delta_{a1}^T L_{12}) \delta_{a1} - 2\delta_{h1}^T D_{12} \delta_{h2} - \delta_{h2}^T D_{22} \delta_{h2} < 0,$$

即锥问题目标函数的改变量是负数.

由于 $\tilde{x}$ 是二次规划(4.4.1)的可行解, 对于 $\tilde{x}$ 来说, a'_{N1} 的偏差非负, 即 $\tilde{x} - x$ 是锥问题的可行下降方向.

(ii) 设锥问题有可行下降方向, 那么存在向量 δ_{a2} 和 δ_{h1} , 使得

$$\delta_{a2}^T L_{22} \delta_{a2} + \delta_{h1}^T D_{11} \delta_{h1} < 0,$$

并且非基向量 a'_{N1} 的偏差

$$E_{12} \delta_{a2} + D_{11} \delta_{h1} \geq 0,$$

其中 δ_{a2} 的分量是充分小的正数或 0, δ_{h1} 的分量充分接近 0. 令 $\delta_{a1} = 0, \delta_{h2} = 0$, 即基向量 a'_{B1} 和 h'_{B2} 的右端项不改变. 设新的基本解为 $\bar{x}$. 那么

$$2f(\bar{x}) - 2f(x) = \delta_{a2}^T L_{22} \delta_{a2} + \delta_{h1}^T D_{11} \delta_{h1} < 0.$$

由于 σ_{N2} 的所有分量是正数, 对于 $\bar{x}$ 来说, a'_{N2} 的偏差

$$\sigma_{N2} + E_{22} \delta_{a2} + D_{12}^T \delta_{h1}$$

仍然是正数. 即 $\bar{x}$ 是二次规划(4.4.1)的可行解, 所以 x 不是局部极小点. ■

定理 4.4.3 (局部极小点的充分条件) 考虑二次规划(4.4.1), 设简约表的所有非基向量偏差非负.

(i) 如果所有非基拉格朗日向量的偏差大于 0, 则当前基本解是二次规划的局部极小点;

(ii) 如果偏差等于 0 的非基拉格朗日向量行的主子矩阵正半定或非负, 则当前基本解是局部极小点;

(iii) 如果没有非基拉格朗日向量, 则当前基本解是全局极小点.

证 (i) 由参数公式(4.4.11)可以看出当前基本解是二次规划的局部极

小点.

(ii) 由于锥问题没有可行下降方向, 根据定理 4.4.2, 当前基本解是局部极小点.

(iii) 简约表没有非基拉格朗日向量也就是 3.1 节所提到的拉格朗日矩阵为空. 这时所有基约束都是等式. 根据定理 4.4.1, 当前基本解是基约束相应方程组解集上的全局极小点, 因而是二次规划的全局极小点. ■

4.4.4 旋转算法的要点

下面介绍非凸二次规划主迭代过程中的关键技术, 称为 I 型平推和 II 型平推, 一般来说在当前基本解是可行解时使用.

为叙述简单起见, 下面将约束不等式 $a_i x \geq b_i$ 和 $x_i \geq l_i$ 的系数向量统一地记为 a'_i . 将它们对应的拉格朗日不等式 $\lambda_i \geq 0$ 和 $H_i x - p_i^T \lambda \geq -c_i$ 的系数向量统一地记为 h'_i . 用 B_C 和 B_L 分别表示基约束向量和基拉格朗日向量的指标集, N_C 和 N_L 分别表示非基约束向量和非基拉格朗日向量的指标集. 根据前面的说明, 非基约束向量行和基拉格朗日向量列上的元素构成的矩阵, 即约束矩阵, 正半定.

(i) I 型平推. 设当前表的部分数据如表 4.58 所示, 其中 $w_{ss} \leq 0$ 且 $\sigma_s < 0$, 或者 $w_{ss} = 0$ 且 $\sigma_s < 0, \sigma_i \geq 0, i \in N_C$, 当前基本解为 x .

表 4.58 I 型平推的初始表

	$a'_s + \Delta$	h'_r	
h'_s	w_{ss}	$-w_{rs}$	σ_s
a'_r	w_{rs}	w_{rr}	σ_r

在 a'_s 处作平推, 记为 $a'_s + \Delta$, 它是进行下述情况 (a), (b), (c) 或 (d) 中陈述的操作之一.

(a) 若 a'_s 列非基约束向量行上的所有元素非负, 二次规划无下界, 停止. 因为将基约束 a'_s 的右端项不断增加, 目标函数值越来越小且对应的基本解仍然可行.

下面假设 a'_s 列非基约束向量行至少有一个负元素, 并设

$$\theta = \frac{\sigma_r}{-w_{rs}} = \min \left\{ \frac{\sigma_i}{-w_{is}} \mid w_{is} < 0, i \in N_C \right\},$$

上式确定的 θ 称为平推量.

(b) $\theta = 0$ 的情况. 若 $w_{rr} > 0$, 以其为枢轴进行一次旋转运算; 否则, 若 a'_r 行有正的非对角元素, 以该元素及其对称元素为枢轴先后进行两次旋转运算; 否则, 将 a'_r 行上负元素对应的基不等式都当做等式 (见第二章的定理 2.1.3).

(c) $\theta > 0$ 且 $w_{rr} > 0$ 的情况. 以 w_{rr} 为枢轴进行一次旋转运算, 结果见表 4.59.

表 4.59 后向旋转的结果

	a'_s	a'_r	
h'_s	$w_{ss} + \frac{w_{rs}^2}{w_{rr}}$	$\frac{-w_{rs}}{w_{rr}}$	$\sigma_s + \frac{w_{rs}\sigma_r}{w_{rr}}$
h'_r	$\frac{-w_{rs}}{w_{rr}}$	$\frac{1}{w_{rr}}$	$-\frac{\sigma_r}{w_{rr}}$

若 $w_{ss} + \frac{w_{rs}^2}{w_{rr}} > 0$, 以其为枢轴再进行一次旋转运算, 结果见表 4.60.

表 4.60 前向旋转的结果

	h'_s	a'_r	
a'_s	$\frac{w_{rr}}{w_{ss}w_{rr} + w_{rs}^2}$	$\frac{w_{rs}}{w_{ss}w_{rr} + w_{rs}^2}$	$-\frac{w_{rr}\sigma_s + w_{rs}\sigma_r}{w_{ss}w_{rr} + w_{rs}^2}$
h'_r	$\frac{-w_{rs}}{w_{ss}w_{rr} + w_{rs}^2}$	$\frac{w_{ss}}{w_{ss}w_{rr} + w_{rs}^2}$	$\frac{w_{rs}\sigma_s - w_{ss}\sigma_r}{w_{ss}w_{rr} + w_{rs}^2}$

可见非基向量 a'_s 和 h'_r 的偏差为正. 用 $\bar{x}$ 表示此表相应的基本解, 可以证明

$$2f(\bar{x}) - 2f(x) = \frac{w_{ss}\sigma_r^2 - w_{rr}\sigma_s^2 - 2w_{rs}\sigma_s\sigma_r}{w_{ss}w_{rr} + w_{rs}^2}. \tag{4.4.15}$$

可以看出上式结果是一个负数, 即目标函数值减少.

(d) $\theta > 0$ 且 $w_{rr} = 0$ 的情况. 以 w_{rs} 和 $-w_{rs}$ 为枢轴先后进行两次旋转运算, 结果如表 4.61 所示.

表 4. 61 双枢轴运算的结果

	a'_r	h'_s	
h'_r	$\frac{w_{ss}}{w_{rs}^2}$	$-\frac{1}{w_{rs}}$	$\frac{w_{rs}\sigma_s - w_{ss}\sigma_r}{w_{rs}^2}$
a'_s	$\frac{1}{w_{ks}}$	0	$-\frac{\sigma_r}{w_{rs}}$

可见 a'_s 和 h'_r 的偏差为正. 用 $\bar{x}$ 表示此表相应的基本解, 可以证明

$$2f(\bar{x}) - 2f(x) = \frac{w_{ss}\sigma_r^2 - 2w_{rs}\sigma_s\sigma_r}{w_{rs}^2}. \tag{4.4.16}$$

可以看出上式结果是一个负数, 即目标函数值减少.

(ii) II 型平推. 设当前表的部分数据如表 4. 62 所示, 其中 $w_{sk} < 0, \sigma_i \geq 0, i \in N_C$, 当前基本解为 x .

表 4. 62 II 型平推的初始表

	a'_s	a'_k	h'_r	
h'_s	0	w_{sk}	$-w_{rs}$	0
h'_k	w_{sk}	0	$-w_{rk}$	0
a'_r	w_{rs}	w_{rk}	w_{rr}	σ_r

当我们说作平推 $a'_s + \Delta$ 时, 意味着进行下述情况 (a), (b), (c) 或 (d) 中陈述的操作之一.

(a) 若 a'_s 和 a'_k 两列非基约束行上的两个元素之和均非负, 二次规划无下界, 停止.

下面设 a'_s 列非基约束行上至少有一个负元素, 并设

$$\theta = \frac{\sigma_r}{-w_{rs}} = \min \left\{ \frac{\sigma_i}{-w_{is}} \mid w_{is} < 0, i \in N_C \right\}.$$

(b) $\theta = 0$ 的情况. 若 $w_{rr} > 0$, 以其为枢轴进行一次旋转运算. 否则, 若 a'_r 行有正的非对角元素, 以该元素及其对称元素为枢轴先后进行两次旋转运算; 否则, 将 a'_r 行上负元素对应的基不等式都当做等式.

(c) $\theta > 0$ 且 $w_{rr} > 0$ 的情况. 以 w_{rr} 为枢轴作后向旋转, 然后再以 h'_s 入基以 a'_s 出基作前向旋转, 结果见表 4. 63.

表 4. 63 两次旋转运算的结果

	h'_s	a'_k	a'_r	
a'_s	$\frac{w_{rr}}{w_{rs}^2}$	$-\frac{w_{rr}w_{sk} + w_{rs}w_{rk}}{w_{rs}^2}$	$\frac{1}{w_{rs}}$	$-\frac{\sigma_r}{w_{rs}}$
h'_k	$\frac{w_{rr}w_{sk} + w_{rs}w_{rk}}{w_{rs}^2}$	$-\frac{w_{rr}^2w_{sk}^2 + 2w_{rs}w_{sk}w_{rk}}{w_{rs}^2}$	$\frac{w_{sk}}{w_{rs}}$	$-\frac{w_{sk}\sigma_r}{w_{rs}}$
h'_r	$-\frac{1}{w_{rs}}$	$\frac{w_{sk}}{w_{rs}}$	0	0

可见非基约束向量 a'_s 的偏差为正.

(d) $\theta > 0$ 且 $w_{rr} = 0$ 的情况. 以 w_{rs} 和 $-w_{rs}$ 为枢轴先后进行两次旋转运算, 结果如表 4. 64 所示.

表 4. 64 双枢轴运算的结果

	a'_r	a'_k	h'_s	
h'_r	0	$\frac{w_{sk}}{w_{rs}}$	$-\frac{1}{w_{rs}}$	0
h'_k	$\frac{w_{sk}}{w_{rs}}$	$-\frac{2w_{sk}w_{rk}}{w_{rs}}$	$\frac{w_{rk}}{w_{rs}}$	$-\frac{w_{sk}\sigma_r}{w_{rr}}$
a'_s	$\frac{1}{w_{rs}}$	$-\frac{w_{rk}}{w_{rs}}$	0	$-\frac{\sigma_r}{w_{rs}}$

可见 h'_k 和 a'_s 的偏差为正.

4. 4. 5 旋转算法的计算步骤及例子

现将非凸二次规划局部极小点的计算步骤简述如下, 其中 L 表示旋转算法表中的拉格朗日矩阵, L 中与基约束不等式和非正偏差对应的主子矩阵记为 L^* .

算法 4. 4. 1 非凸二次规划(4. 4. 1)局部极小点的计算步骤.

步骤 1 初始步骤. 同算法 3. 2. 1 的步骤 1.

步骤 2 预处理. 同算法 3. 2. 1 的步骤 2.

步骤 3 主迭代.

(i) 求基本可行解.

(a) 如果所有非基约束的偏差非负, 转(ii); 否则以偏差最小者入基. 若该行与基不等式对应的所有元素均小于或等于 0, 停止; 否则转(b).

(b) 若该行对角元素为正, 以其为枢轴进行一次旋转运算, 转(i); 否则以该行最大正元素及其对称元素为枢轴先后进行两次旋转运算, 转(i).

(ii) 寻找可行下降方向.

如果 L^* 矩阵为空, 停止. 否则转(a).

(a) 如果 L^* 有一个对角元素为负, 同行偏差小于或等于 0, 作平推, 转(i); 否则转(b).

(b) 如果 L^* 有一个对角元素为 0, 同行偏差小于 0, 作平推, 转(i); 否则转(c).

(c) 如果 L^* 有一个 II 型主子矩阵, 同行两个偏差为 0, 作平推, 转(i); 否则转(d).

(d) 如果 L^* 有一个正的对角元素, 该行偏差为负, 以该对角元素枢轴进行一次旋转运算, 转(i); 否则转(e).

(e) 如果 L^* 有一个 2×2 主子矩阵, 其非对角元素为负且至少有一个正对角元素, 相应的两个偏差均为 0, 以一个正对角元素为枢轴进行一次旋转运算, 转(i).

例 4.4.1 求以下二次规划的一个局部极小点:

$$\min f(x) = \frac{-x_1^2 + x_2^2}{2},$$

$$\text{s. t. } -x_1 + x_2 \geq -2,$$

$$-x_2 \geq -1,$$

$$x_1 \geq 0.1, x_2 \geq -1.$$

解 其 KT 条件为

$$-x_1 + \lambda_1 \geq 0,$$

$$x_2 - \lambda_1 + \lambda_2 \geq 0,$$

$$-x_1 + x_2 \geq -2,$$

$$-x_2 \geq -1,$$

$$x_1 \geq 0.1, x_2 \geq -1, \lambda_1, \lambda_2 \geq 0,$$

$$(-x_1 + \lambda_1)(x_1 - 0.1) = 0,$$

$$(x_2 - \lambda_1 + \lambda_2)(x_2 + 1) = 0,$$

$$(-x_1 + x_2 + 2)\lambda_1 = 0,$$

$$(-x_2 + 1)\lambda_2 = 0.$$

记

$$\begin{aligned}h_1 &= (-1, 0, 1, 0), \\h_2 &= (0, 1, -1, 1), \\a_1 &= (-1, 1, 0, 0), \\a_2 &= (0, -1, 0, 0),\end{aligned}$$

并用 e_1, e_2, e_3, e_4 分别表示 4 阶单位矩阵的各行.

以 $x_1 \geq 0.1, x_2 \geq -1, \lambda_1, \lambda_2 \geq 0$ 为初始基本不等式组, 即 $y^{(0)} = (0.1, -1, 0, 0)^T$ 是初始基本解. e_1, e_2, e_3, e_4 是初始基向量, h_1, h_2, a_1, a_2 是非基向量, 它们的偏差分别为 $h_1 y^{(0)} = -0.1, h_2 y^{(0)} = -1, a_1 y^{(0)} + 2 = 0.9, a_2 y^{(0)} + 1 = 2$. 非基向量关于基向量的线性组合及其偏差表如 4.65 所示.

由于第一个对角元素为负且同行偏差为负, 作平推 $e_1^{+\Delta}$. 第一列非基向量 a_1 行上的元素 $-1 < 0$, a_1 入基. 由于 a_1 行的对角元素为 0, 以 -1 及其对称元素 1 为枢轴先后进行两次旋转运算, 结果见表 4.66.

表 4.65 初 始 表

	$e_1^{+\Delta}$	e_2	e_3	e_4	
h_1	-1	0	1^{**}	0	-0.1
h_2	0	1	-1	1	-1
a_1	-1^{*}	1	0	0	0.9
a_2	0	-1	0	0	2

表 4.66 第一次双枢轴运算的结果

	a_1	$e_2^{+\Delta}$	h_1	e_4	
e_3	-1	1	1	0	1
h_2	1	0	-1	1^{**}	-2
e_1	-1	1	0	0	0.9
a_2	0	-1^{*}	0	0	2

在表 4.66 中, 第二个对角元素为 0 但同行偏差为负, 作平推 $e_2^{+\Delta}$. 第二列非基向量 a_2 行上的元素 $-1 < 0$, a_2 入基. 由于 a_2 行的对角元素为 0, 以 -1 及其对称元素 1 为枢轴先后进行两次旋转运算, 结果见表 4.67.

表 4.67 最 终 表

	a ₁	a ₂	h ₁	h ₂	
e ₃	- 1	- 1	1	0	3
e ₄	- 1	0	1	1	2
e ₁	- 1	- 1	0	0	2.9
e ₂	0	- 1	0	0	2

由于非基向量偏差均为正，局部极小点为 $x_1 = 2.9 + 0.1 = 3$, $x_2 = 2 - 1 = 1$, 可以算得 $f(x) = - 4$.

例 4.4.2 解二次规划

$$\begin{aligned} \min f(x) &= \frac{1}{2}(x_1, x_2, x_3) \begin{pmatrix} 1 & -2 & 0 \\ -2 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \\ \text{s. t. } &- x_1 - x_2 + x_3 \geqslant -1, \\ &- x_3 \geqslant -2, \\ &x_1, x_2, x_3 \geqslant 0. \end{aligned}$$

解 初始表为表 4.68. 由于 H 的所有对角元素为正且偏差为 0 ,以第一个对角元素为枢轴进行旋转运算，结果见表 4.69.

表 4.68 初 始 表

	e ₁	e ₂	e ₃	e ₄	e ₅	
h ₁	1 *	- 2	0	1	0	0
h ₂	- 2	2	- 1	1	0	0
h ₃	0	- 1	1	- 1	1	0
a ₁	- 1	- 1	1	0	0	1
a ₂	0	0	- 1	0	0	2

表 4. 69 第一次前向旋转结果

	h_1	$e_2^{+\Delta}$	e_3	e_4	e_5	
e_1	1	2	0	- 1	0	0
h_2	- 2	- 2	- 1	3	0	0
h_3	0	- 1	1	- 1	1	0
a_1	- 1	- 3	1	1 *	0	1
a_2	0	0	- 1	0	0	2

在表 4. 69 中,第二个对角元素为负且同行偏差为 0,作平推 $e_2^{+\Delta}$. 第二列非基向量 a_1 行上的元素 $- 3 < 0$, a_1 入基. 由于 a_1 行的对角元素为正,以该对角元素为枢轴进行旋转运算,结果见表 4. 70. 这时原来的枢轴元素 $- 2$ 变为正数 7,并且该行偏差为负,以 7 为枢轴进行一次前向旋转,结果见表 4. 71.

表 4. 70 平推 $e_2^{+\Delta}$ 中的后向旋转结果

	h_1	e_2	e_3	a_1	e_5	
e_1	0	- 1	1	- 1	0	1
h_2	1	7 *	- 4	3	0	- 3
h_3	- 1	- 4	2	- 1	1	1
e_4	1	3	- 1	1	0	- 1
a_2	0	0	- 1	0	0	2

表 4. 71 平推 $e_2^{+\Delta}$ 中的前向旋转结果

	h_1	h_2	$e_3^{+\Delta}$	a_1	e_5	
e_1	0. 143	- 0. 143	0. 429	- 0. 571	0	0. 571
e_2	- 0. 143	0. 143	0. 571	- 0. 429	0	0. 429
h_3	- 0. 429	- 0. 571	- 0. 286	0. 714	1 **	- 0. 714
a_1	0. 571	0. 429	0. 714	- 0. 286	0	0. 286
a_2	0	0	- 1 *	0	0	2

在表 4. 71 中,第三个对角元素为负且同行偏差为负,作平推 $e_3^{+\Delta}$. 第三列

非基向量 a_2 行上的元素为 - 1 , a_2 行的对角元素为 0 , 以 - 1 及其对称元素 1 为枢轴先后进行两次旋转运算 , 结果见表 4. 72.

表 4. 72 平推 $e_3^{+\Delta}$ 的后向旋转结果

	h_1	h_2	a_2	a_1	h_3	
e_1	0. 143	- 0. 143	- 0. 429	- 0. 571	0	1. 429
e_2	- 0. 143	0. 143	- 0. 571	- 0. 429	0	1. 571
e_5	0. 429	0. 571	- 0. 286	- 0. 714	1	1. 286
e_4	0. 571	0. 429	- 0. 714	- 0. 286	0	1. 714
e_3	0	0	- 1	0	0	2

由于表 4. 72 的所有偏差为正 , 局部极小点为 $x_1 = 1. 429$, $x_2 = 1. 571$, $x_3 = 2$, $f(x) = - 2. 143$.

例 4. 4. 3 解二次规划

$$\begin{aligned} \min f(x) &= - x_1 x_2 + x_1 x_3 + x_2 x_3 , \\ \text{s. t. } &- x_1 - 2x_2 - 4x_3 \geq - 2 , \\ &x_1 + x_2 + x_3 \geq 1 , \\ &x_1 , x_2 , x_3 \geq 0 . \end{aligned}$$

解 初始表为表 4. 73. 由于非基向量 a_2 的偏差为 - 1 , 当前基本解不可行. 不妨以 a_1 行的第一个正元素 1 为枢轴进行一次旋转运算 , 然后以其对称元素 - 1 为枢轴进行旋转运算 , 结果见表 4. 74.

表 4. 73 初 始 表

	e_1	e_2	e_3	e_4	e_5	
h_1	0	- 1	1	1	- 1 **	0
h_2	- 1	0	1	2	- 1	0
h_3	1	1	0	4	- 1	0
a_1	- 1	- 2	- 4	0	0	2
a_2	1 *	1	1	0	0	- 1

表 4. 74 第一次双枢轴运算结果

	a_2	e_2	e_3	e_4	h_1	
e_5	0	- 1	1	1	- 1	0
h_2	- 1	2 *	1	1	1	- 1
h_3	1	1	- 2	3	1	1
a_1	- 1	- 1	- 3	0	0	1
e_1	1	- 1	- 1	0	0	1

在表 4. 74 中, 非基拉格朗日向量 h_2 的偏差为负, 以对角线上的元素为枢轴进行旋转运算, 结果见表 4. 75. 由于拉格朗日向量 e_5 的偏差和该行的对角元素都是负数, 作平推 $a_2^{+\Delta}$. 第一列非基向量 a_1 行上的元素 $-\frac{3}{2} < 0$ 且该行对角元素为正, 以其为枢轴进行旋转运算, 结果见表 4. 76.

表 4. 75 第一次前向旋转结果

	$a_2^{+\Delta}$	h_2	e_3	e_4	h_1	
e_5	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{3}{2}$	$\frac{3}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$
e_2	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$
h_3	$\frac{3}{2}$	$\frac{1}{2}$	$-\frac{5}{2}$	$\frac{5}{2}$	$\frac{1}{2}$	$\frac{3}{2}$
a_1	$-\frac{3}{2}$	$-\frac{1}{2}$	$-\frac{5}{2}$	$\frac{1}{2}^*$	$\frac{1}{2}$	$\frac{1}{2}$
e_1	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

表 4. 76 平推 $a_2^{+\Delta}$ 中的后向旋转结果

	a_2	h_2	e_3	a_1	h_1	
e_5	4 *	1	9	3	- 2	- 2
e_2	- 1	0	- 3	- 1	0	1
h_3	9	3	10	5	- 2	- 1
e_4	3	1	5	2	- 1	- 1
e_1	2	0	2	1	0	0

在表 4.76 中以 e_5 入基以 a_2 出基 结果如表 4.77 所示.

表 4.77 最 终 表

	e_5	h_2	e_3	a_1	h_1	
a_2	$\frac{1}{4}$	$-\frac{1}{4}$	$-\frac{9}{4}$	$-\frac{3}{4}$	$\frac{1}{2}$	$\frac{1}{2}$
e_2	$-\frac{1}{4}$	$\frac{1}{4}$	$-\frac{3}{4}$	$-\frac{1}{4}$	$-\frac{1}{2}$	$\frac{1}{2}$
h_3	$\frac{9}{4}$	$\frac{3}{4}$	$-\frac{41}{4}$	$-\frac{7}{4}$	$\frac{5}{2}$	$\frac{7}{2}$
e_4	$\frac{3}{4}$	$\frac{1}{4}$	$-\frac{7}{4}$	$-\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{2}$
e_1	$\frac{1}{2}$	$-\frac{1}{2}$	$-\frac{5}{2}$	$-\frac{1}{2}$	1	1

由表 4.77 可见 , 局部极小点为 $x_1 = 1 , x_2 = \frac{1}{2} , x_3 = 0 , f(x) = -\frac{1}{2}$.

例 4.4.4 解二次规划

$$\min f(x) = \frac{1}{2}(x_1, x_2, x_3, x_4) \begin{pmatrix} -1 & 6 & 2 & -4 \\ 6 & -2 & -8 & 1 \\ 2 & -8 & 0 & -2 \\ -4 & 1 & -2 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + (-4, 2, -3, 0) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

s. t. $5x_1 - x_2 + x_3 - 2x_4 = 3,$
 $-x_1 + 2x_2 - x_3 - x_4 = -6,$
 $-3x_1 - 3x_2 + 6x_3 + x_4 \geq -1,$
 $-5x_2 + 2x_3 - 7x_4 \geq 1,$
 $x_1, x_2, x_4 \geq 0, \quad x_3 \geq 1.$

解 以 $x_1, x_2 \geq 0, x_3 \geq 1, x_4 \geq 0, \lambda_1, \lambda_2, \lambda_3, \lambda_4 \geq 0$ 为初始基本不等式组 , 即 $y^{(0)} = (0, 0, 1, 0, 0, 0, 0, 0)^T$ 是初始基本解. 各非基向量的偏差分别为 $h_1 y^{(0)} - 4 = -2, h_2 y^{(0)} + 2 = -6, \dots, a_4 y^{(0)} - 1 = 1$. 非基向量关于基向量的线性组合及其偏差表如表 4.78 所示.

表 4.78 初 始 表

	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	e ₇	e ₈	
h ₁	- 1	6	2	- 4	- 5 *	1	3	0	- 2
h ₂	6	- 2	- 8	1	1	- 2	3	5	- 6
h ₃	2	- 8	0	- 2	- 1	1	- 6	- 2	- 3
h ₄	- 4	1	- 2	5	2	1	- 1	7	- 2
a ₁	5 *	- 1	1	- 2	0	0	0	0	- 2
a ₂	- 1	2	- 1	- 1	0	0	0	0	5
a ₃	- 3	- 3	6	1	0	0	0	0	7
a ₄	0	- 5	2	- 7	0	0	0	0	1

在表 4.78 中,非基约束向量 a₁和 a₂对应等式,首先将 a₁换入基中. a₁的偏差是负数,以该行的最大正元素为 5 及其对称元素 - 5 为枢轴先后进行两次旋转运算,结果见表 4.79 其中 a₁列和 e₅行已删去.

表 4.79 第一次双枢轴运算结果

	e ₂	e ₃	e ₄	h ₁	e ₆	e ₇	e ₈	
h ₂	0.36	- 8.76	2.52	- 0.2	- 1.8	3.6	5	- 4.08
h ₃	- 8.76	- 0.84	- 0.32	0.2	0.8	- 6.6	- 2	- 1.72
h ₄	2.52	- 0.32	1.64	- 0.4	1.4 *	0.2	7	- 4.56
e ₁	0.2	- 0.2	0.4	0	0	0	0	0.4
a ₂	1.8	- 0.8	- 1.4 *	0	0	0	0	4.6
a ₃	- 3.6	6.6	- 0.2	0	0	0	0	5.8
a ₄	- 5	2	- 7	0	0	0	0	1

在表 4.79 中, a₂的偏差是正数,以该行的最小负元素 - 1.4 及其对称元素

1.4为枢轴先后进行两次旋转运算，结果见表 4.80 其中 a_2 列和 e_6 行已删去.

表 4.80 第二次双枢轴运算结果

	e_2	e_3	h_1	h_4	e_7	e_8	
h_2	9.551	- 11.816	- 0.714	- 1.286	3.857	14	5.265
h_3	- 11.816	0.061	0.429	0.571	- 6.714	- 6 *	- 3.245
e_1	0.714	- 0.429	0	0	0	0	1.714
e_4	1.286	- 0.571	0	0	0	0	3.286
a_3	- 3.857	6.714	0	0	0	0	5.143
a_4	- 14	6 *	0	0	0	0	- 22

在表 4.80 中， a_4 的偏差是负数，当前基本解不可行. 以该行的最大正元素 6 及其对称元素 - 6 为枢轴先后进行两次旋转运算，结果见表 4.81.

表 4.81 第三次双枢轴运算结果

	$e_2^{+\Delta}$	a_4	h_1	h_4	e_7	h_3	
h_2	- 45.259	- 1.946	0.286 *	0.048	- 11.81	- 2.333	- 45.109
e_8	- 1.946	0.002	0.071	0.095	- 1.119	- 0.167	- 0.503
e_1	- 0.286 *	- 0.071	0	0	0	0	0.143
e_4	- 0.048	- 0.095	0	0	0	0	1.19
a_3	11.81	1.119	0	0	0	0	29.762
e_3	2.333	0.167	0	0	0	0	3.667

拉格朗日向量 h_2 行上的对角元素及其偏差都是负数，作 $e_2^{+\Delta}$. 由于

$$\min \left\{ \frac{0.143}{0.286}, \frac{1.19}{0.048} \right\} = \frac{0.143}{0.286},$$

此比值对应 e_1 行，以该行的 - 0.286 及其对称元素 0.286 为枢轴先后进行

两次旋转运算，结果见表 4.82.

表 4.82 第四次双枢轴运算结果

	e_1	a_4	h_2	h_4	e_7	h_3	
h_1	- 554.416	- 32.792	3.5	- 0.167	41.333	8.167	237.083
e_8	- 32.792	- 1.854	0.25	0.083	1.833	0.417	15.458
e_2	- 3.5	- 0.25	0	0	0	0	0.5
e_4	0.167	- 0.083	0	0	0	0	1.167
a_3	- 41.333	- 1.833	0	0	0	0	35.667
e_3	- 8.167	- 0.417	0	0	0	0	4.833

至此，所有非基向量偏差是正数，局部极小点为

$$x_1 = 0, \quad x_2 = 0.5, \quad x_3 = 4.833 + 1 = 5.833, \quad x_4 = 1.167.$$

可以算得，对应的目标函数值为 $f(x) = -49.708$.

4.5 局部极小点的改进

4.5.1 改进局部极小点的思路

一般来说，非凸二次规划局部极小点的数目是变量数目和约束数目的指数函数. 当问题规模较大时，要把局部极小点一一找出来，最后确定全局极小点，计算量很大. 本节提供一些思路以获取比当前局部极小点更好的解.

仍考虑二次规划

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2}x^T Hx + cx \\ \text{s. t. } a_i x &= b_i, \quad i = 1, 2, \dots, l, \\ a_i x &\geq b_i, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (4.5.1)$$

其中 $H = (h_{ij})_{n \times n}$ 是非正半定的 n 阶实对称矩阵， $c = (c_1, c_2, \dots, c_n)$ ， $x = (x_1, x_2, \dots, x_n)^T$ ， $a_i = (a_{i1}, a_{i2}, \dots, a_{in})$ ， $i = 1, 2, \dots, m$.

设 x 是 n 维空间一个点， d 是 x 的变动方向， t 是非负实数. 将 $x + td$ 代入目标函数得

$$\begin{aligned}
 f(x+td) &= \frac{1}{2}(x+td)^T H(x+td) + c(x+td) \\
 &= \frac{1}{2}[t^2 d^T H d + 2t(d^T H x + cd) + x^T H x + c x].
 \end{aligned}$$

可见目标函数是 t 的一个不高于二次的多项式. 当 t 由 0 逐渐增加时, 其变化有以下 5 种情况.

- (i) 目标函数值单调递减;
- (ii) 目标函数值单调递增;
- (iii) 目标函数值不变;
- (iv) 目标函数值先减后增;
- (v) 目标函数值先增后减.

若 x 是局部极小点, 我们只能希望找到一个可行的方向 d 使得情况 (v) 出现, 然后看目标函数值能否在可行域内下降到比 $f(x)$ 还小.

现在就设 x 是利用上一节的方法求出的一个局部极小点, 并设其对应表 4.83.

表 4.83 旋转算法表

	a'_1	...	a'_k	h'_{k+1}	...	h'_{n+m}	
h'_1	w_{11}	...	w_{1k}	$-w_{k+1,1}$	...	$-w_{n+m,1}$	σ_1
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
h'_k	w_{k1}	...	w_{kk}	$-w_{k+1,k}$	...	$-w_{n+m,k}$	σ_k
a'_{k+1}	$w_{k+1,1}$	...	$w_{k+1,k}$	$w_{k+1,k+1}$	...	$w_{k+1,n+m}$	σ_{k+1}
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\vdots$
a'_{n+m}	$w_{n+m,1}$	...	$w_{n+m,k}$	$w_{n+m,k+1}$	...	$w_{n+m,n+m}$	σ_{n+m}

其中, a'_i 表示约束向量, h'_i 表示相应的拉格朗日向量, 最后一列是各非基向量关于当前基本解 x 的偏差. 由于当前基本解是局部极小点, 所有偏差非负. 而且当拉格朗日矩阵的第 s 个对角元素 $w_{ss} < 0$ 时, $\sigma_s > 0$; 当 $w_{ss} = w_{rr} = 0$ 且 $w_{sr} < 0$ 时, σ_s 和 σ_r 至少有一个是正数, 其中 $s, r \in \{1, 2, \dots, k\}$.

下面根据基约束右端常数项数变动的多少, 讨论搜索更好极小点的几点思路.

如果拉格朗日矩阵有负的对角元素, 首先考虑利用负的对角元素改进当前局部极小点. 假设在表 4.83 中存在 $s \in \{1, 2, \dots, k\}$ 使得 $w_{ss} < 0$. 根据公式

(4.4.14), 基约束 a'_s 相应不等式的右端项增加 Δ 时, 目标函数值的改变量为

$$\Delta f = \sigma_s \Delta + w_{ss} \frac{\Delta^2}{2}. \quad (4.5.2)$$

可见当

$$\Delta > \frac{2\sigma_s}{-w_{ss}}$$

时, 目标函数值的改变量 $\Delta f < 0$. 如果对于所有 $i \in \{k+1, k+2, \dots, n+m\}$ 都有 $w_{is} \geq 0$, 目标函数在可行域上无下界; 否则计算

$$\theta = \frac{\sigma_r}{-w_{rs}} = \min \left\{ \frac{\sigma_i}{-w_{is}} \mid w_{is} < 0, i \in \{k+1, k+2, \dots, n+m\} \right\}.$$

如果 $\theta > \frac{2\sigma_s}{-w_{ss}}$, 那么当 a'_s 相应基不等式的右端项增加 θ 时, 可得到一个目标函数值更小的可行解.

如果没有以上情况, 再考虑利用拉格朗日矩阵的二阶主子矩阵改进当前局部极小点. 设表 4.83 中拉格朗日矩阵有一个二阶主子矩阵 $\begin{pmatrix} w_{ss} & w_{rs} \\ w_{rs} & w_{rr} \end{pmatrix}$, 其非对角元素 w_{rs} 是负数. 将基约束 a'_s 和 a'_r 的右端项分别增加 Δ_1 和 Δ_2 , 根据公式 (4.4.12), 目标函数的改变量为

$$\Delta f = 2(\sigma_s, \sigma_r) \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix} + (\Delta_1 \Delta_2) \begin{pmatrix} w_{ss} & w_{rs} \\ w_{rs} & w_{rr} \end{pmatrix} \begin{pmatrix} \Delta_1 \\ \Delta_2 \end{pmatrix}. \quad (4.5.3)$$

当上式右边第二项是足够小的负数时, 目标函数值将减少, 因而可能得到一个目标函数值更小的可行解. 这时, 需将基向量 a'_s 和 a'_r 都换为非基向量.

如果 w_{ss} 和 w_{rr} 至少有一个是负数, 如 $w_{ss} < 0$, 则先作平推 a'_s , 然后在 a'_r 列作平推或作前向旋转. 如果 w_{ss} 和 w_{rr} 都是正数, 必须满足

$$w_{ss}w_{rr} - w_{rs}^2 < 0.$$

这时, 首先以其中一个对角元素, 如 w_{ss} 为枢轴作前向旋转, 然后作平推 a'_r .

值得注意的是, 在某一列进行平推, 如果使用的是后向旋转且对角元素仍然是负数, 还可在同一位置再次作平推. 另外, 只要能求得一个有限的平推量, 不要求所有非基约束的偏差非负.

下面通过 3 个例子说明如何改进非凸二次规划的局部极小点.

例 4.5.1 解二次规划

$$\begin{aligned} \min f(x) &= -x_1^2 + x_2, \\ \text{s. t. } -x_1 + 2x_2 &\geq 0, \end{aligned}$$

$$\begin{aligned} & -x_2 \geq -1, \\ & x_1, x_2 \geq 0. \end{aligned}$$

解 其 KT 条件为

$$\begin{aligned} & -2x_1 + \lambda_1 \geq 0, \\ & -2\lambda_1 + \lambda_2 \geq -1, \\ & -x_1 + 2x_2 \geq 0, \\ & -x_2 \geq -1, \\ & x_1, x_2, \lambda_1, \lambda_2 \geq 0, \\ & (-2x_1 + \lambda_1)x_1 = 0, \\ & (-2\lambda_1 + \lambda_2 + 1)x_2 = 0, \\ & (-x_1 + 2x_2)\lambda_1 = 0, \\ & (-x_2 + 1)\lambda_2 = 0. \end{aligned}$$

记

$$\begin{aligned} h_1 &= (-2, 0, 1, 0), \\ h_2 &= (0, 0, -2, 1), \\ a_1 &= (-1, 2, 0, 0), \\ a_2 &= (0, -1, 0, 0), \end{aligned}$$

并用 e_1, e_2, e_3, e_4 分别表示 4 阶单位矩阵的各行.

以 $x_1, x_2, \lambda_1, \lambda_2 \geq 0$ 为初始基本不等式组, 即 $y^{(0)} = (0, 0, 0, 0)^T$ 是初始基本解, e_1, e_2, e_3, e_4 是初始基向量, h_1, h_2, a_1, a_2 是非基向量, 它们的偏差分别为 $h_1 y^{(0)} = 0, h_2 y^{(0)} + 1 = 1, a_1 y^{(0)} = 0, a_2 y^{(0)} + 1 = 1$. 非基向量关于基向量的线性组合及其偏差如表 4.84 所示.

由于第一个对角元素是负数且同行偏差为 0, 作平推 $e_1^{+\Delta}$. 第一列非基向量 a_1 行上的元素 $-1 < 0$, a_1 入基. 由于 a_1 的偏差为 0, 以该行的正元素 2 为枢轴进行旋转运算, 然后以其对称元素 -2 为枢轴进行旋转运算, 结果见表 4.85.

表 4.84 初 始 表

	$e_1^{+\Delta}$	e_2	e_3	e_4	
h_1	-2	0	1	0	0
h_2	0	0	-2*	1	1
a_1	-1	2*	0	0	0
a_2	0	-1	0	0	1

表 4. 85 第一次双枢轴运算的结果

	$e_1^{+\Delta}$	a_1	h_2	e_4	
h_1	- 2	0	$-\frac{1}{2}$	$\frac{1}{2}^*$	$\frac{1}{2}$
e_3	0	0	$-\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
e_2	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0
a_2	$-\frac{1}{2}^*$	$-\frac{1}{2}$	0	0	1

在表 4. 85 中,所有非基向量偏差非负,当前基本解 $x_1 = 0, x_2 = 0$ 是局部极小点,对应的目标函数值为 0.

由表 4. 85 可见,非基拉格朗日向量 h_1 行的对角元素为 - 2 ,偏差为 $\frac{1}{2}$;基向量 e_1 列和非基向量 a_2 行上的元素为 $-\frac{1}{2}$,偏差为 1. 由于

$$\theta = \frac{1}{-\left(-\frac{1}{2}\right)} = 2 > \frac{2 \times \frac{1}{2}}{-(-2)} = \frac{1}{2},$$

当基不等式 $x_1 \geq 0$ 的右端项增加到 2 时,目标函数值的改变量是负数. 所以以 $-\frac{1}{2}$ 及其对称元素 $\frac{1}{2}$ 为枢轴先后进行两次旋转运算,结果见表 4. 86.

表 4. 86 最 终 表

	a_2	a_1	h_2	h_1	
e_4	- 8	- 4	1	2	7
e_3	- 4	- 2	0	1	4
e_2	- 1	0	0	0	1
e_1	- 2	- 1	0	0	2

由于所有非基向量偏差为正,当前基本解 $x_1 = 2, x_2 = 1$ 是局部极小点,而且也是全局极小点,对应的目标函数值为 $f(x) = - 3$.

例 4. 5. 2 试求以下二次规划的全局极小点:

$$\min \frac{1}{2}(x_1, x_2, x_3, x_4) \begin{pmatrix} 1 & -4 & -2 & -1 \\ -4 & 1 & 1 & -2 \\ -2 & 1 & -1 & -3 \\ -1 & -2 & -3 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} + (1, 1, 1, 1) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix},$$

s. t. $-x_1 + 2x_2 + x_3 - 4x_4 \geq -2,$
 $-4x_1 - 3x_2 - 4x_3 + 2x_4 \geq -8,$
 $x_1, x_2, x_3, x_4 \geq 0.$

解 初始表为表 4.87.

表 4.87 初 始 表

	e ₁	e ₂	e ₃	e ₄	e ₅	e ₆	
h ₁	1	- 4	- 2	- 1	1	4	1
h ₂	- 4	1	1	- 2	- 2	3	1
h ₃	- 2	1	- 1	- 3	- 1	4 ^{∗∗}	1
h ₄	- 1	- 2	- 3	- 1	4	- 2	1
a ₁	- 1	2	1	- 4	0	0	2
a ₂	- 4	- 3	- 4 [∗]	2	0	0	8

表 4.87 中所有非基向量偏差是正数，所以 $x = (0, 0, 0, 0)^T$ 是局部极小点，对应的目标值为 0. 表中基向量 e_3 和 e_4 列有两个负对角元素. 但仅利用其中的一个负元素作平推都不能得到目标函数值更小的可行解. 由于这两列的二阶主子矩阵的非对角元素是负数 将 e_3 和 e_4 都换为非基向量可能得到一个目标函数值更小的可行解. 为此，首先作 $e_3^{+\Delta}$. 由于 e_3 列非基约束向量 a_2 行上的元素 $-4 < 0$ 且 a_2 的偏差是正数，先后以 -4 及其对称元素 4 为枢轴作旋转运算，结果见表 4.88.

表 4.88 第一次双枢轴运算结果

	e ₁	e ₂	a ₂	e ₄	e ₅	h ₃	
h ₁	4	- 4.25	0.25	1.5	2	1	- 2
h ₂	- 4.25	- 1.063	- 0.438	1.125	- 1.25	0.75	3.75
e ₆	0.25	- 0.438	- 0.063	0.875	0.25	0.25	0.25
h ₄	1.5	1.125	0.875	- 4.25	3.5 ^{∗∗}	- 0.5	- 5.5
a ₁	- 2	1.25	- 0.25	- 3.5 [∗]	0	0	4
e ₃	- 1	- 0.75	- 0.25	0.5	0	0	2

表 4.88 中,基向量 e_4 列的对角元素是负数,再作平推 $e_4^{+\Delta}$. 由于 e_4 列非基约束向量 a_1 行上的元素 $-3.5 < 0$ 且 a_1 的偏差是正数,先后以 -3.5 及其对称元素为枢轴作旋转运算,结果见表 4.89. 至此已将表 4.87 中的基向量 e_3 和 e_4 都换成了非基向量.

表 4.89 第二次双枢轴运算结果

	e_1	e_2	a_2	a_1	h_4	h_3	
h_1	0.898	- 3.49	- 5.31	- 1.122	0.571	1.286	5.633
h_2	- 3.49	- 0.801	- 0.097	0.112	- 0.357	0.571 **	1.337
e_6	- 0.531	- 0.091	- 2.09	- 0.337	0.071	0.286	1.990
e_5	- 1.122	0.112	- 0.337	- 0.347	0.286	0.143	2.959
e_4	- 0.571	0.357	- 0.071	- 0.286	0	0	1.143
e_3	- 1.286	- 0.571 *	- 0.286	- 0.143	0	0	2.571

表 4.89 的所有非基向量偏差是正数,所以 $x = (0, 0, 2.571, 1.143)^T$ 是局部极小点,对应的目标值为 -9.061 . 此表拉格朗日矩阵左上角的二阶主子矩阵的非对角元素是负数且第二个对角元素是负数,可考虑作 $e_2^{+\Delta}$. 由于 e_2 列基约束向量行仅有一个负元素 -0.571 且 e_3 的偏差是正数,以此元素及其对称元素为枢轴先后进行两次旋转运算,结果见表 4.90.

表 4.90 第三次双枢轴运算结果

	e_1	e_3	a_2	a_1	h_4	h_2	
h_1	12.547 *	2.953	0.531	- 0.953	1.375	2.25	- 4.969
h_3	2.953	- 2.453	- 0.531	- 0.547	0.625	1.75	3.969
e_6	0.531	- 0.531	- 0.313	- 0.469	0.25	0.5	2.688
e_5	- 0.953	- 0.547	- 0.469	- 0.453	0.375	0.25	4.031
e_4	- 1.375	- 0.625	- 0.25	- 0.375	0	0	2.75
e_2	- 2.25	- 1.75	- 0.5	- 0.25	0	0	4.5

在表 4.90 中仅 h_1 的偏差是负数,以该行的对角元素为枢轴作前向旋转得表 4.91. 至此已将表 4.89 中的基向量 e_1 和 e_2 都换成了非基向量.

表 4.91 第一次前向旋转结果

	h_1	e_3	a_2	a_1	h_4	h_2	
e_1	0.08	- 0.235	- 0.042	0.076	- 0.110	- 0.179	0.396
h_3	0.235	- 3.148	- 0.656	- 0.323	0.301	1.220	5.138
e_6	0.042	- 0.656	- 0.335	- 0.428	0.192	0.405	2.898
e_5	- 0.076	- 0.323	- 0.428	- 0.526	0.479	0.421	3.654
e_4	- 0.110	- 0.301	- 0.192	- 0.479	0.151	0.247	2.205
e_2	- 0.179	- 1.220	- 0.405	- 0.421	0.247	0.403	3.609

表 4.91 的所有非基向量偏差是正数，所以 $x = (0.396, 3.609, 0, 2.205)^T$ 是局部极小点，对应的目标值为 - 12.140. 由于每个约束向量或者入基或者出基过一次，可考虑停止迭代.

例 4.5.3 求以下二次规划的局部极小点使其对应的目标函数值尽可能小.

$$\begin{aligned} \min f(x) &= \frac{1}{2}(x_1, x_2, x_3) \begin{pmatrix} 1 & -2 & -5 \\ -2 & 5 & -1 \\ -5 & -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} + (1, 1, 1) \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \\ \text{s. t. } & -x_1 - 2x_2 + x_3 \geq -8, \\ & -2x_1 + 4x_2 - x_3 \geq -9, \\ & x_1 - x_2 - 2x_3 \geq -1, \\ & x_1, x_2, x_3 \geq -1. \end{aligned}$$

解 以 $x_1, x_2, x_3 \geq -1, \lambda_1, \lambda_2, \lambda_3 \geq 0$ 为初始基本不等式组，初始表为表 4.92.

表 4.92 初 始 表

	e_1	e_2	e_3	e_4	e_5	e_6	
h_1	1	- 2	- 5	1	2	- 1	7
h_2	- 2	5 *	- 1	2	- 4	1	- 1
h_3	- 5	- 1	1	- 1	1	2	6
a_1	- 1	- 2	1	0	0	0	10
a_2	- 2	4	- 1	0	0	0	8
a_3	1	- 1	- 2	0	0	0	3

表 4.92 只有非基拉格朗日向量 h_2 的偏差是负数，以该行的对角元素 5 为枢轴作旋转运算 结果见表 4.93.

表 4.93 第一次前向旋转结果

	e_1	h_2	e_3	e_4	e_5	e_6	
h_1	0.2^*	- 0.4	- 5.4	1.8	0.4	- 0.6	6.6
e_2	0.4	0.2	0.2	- 0.4	0.8	- 0.2	0.2
h_3	- 5.4	- 0.2	0.8	- 0.6	0.2	2.2	5.8
a_1	- 1.8	- 0.4	0.6	0.8	- 1.6	0.4	9.6
a_2	- 0.4	0.8	- 0.2	- 1.6	3.2	- 0.8	8.8
a_3	0.6	- 0.2	- 2.2	0.4	- 0.8	0.2	2.8

表 4.93 的所有非基向量偏差是正数，所以得到一个局部极小点

$$x_1 = -1, \quad x_2 = 0.2 - 1 = -0.8, \quad x_3 = -1,$$

对应的目标值为 - 7.6. 此表的拉格朗日矩阵是第 1,3 两行上的主子矩阵 $\begin{pmatrix} 0.2 & -5.4 \\ -5.4 & 0.8 \end{pmatrix}$ ，其非对角元素是负数，两个对角元素都是正数. 不妨以第一个对角元素为枢轴作前向旋转，结果见表 4.94.

表 4.94 第二次前向旋转结果

	h_1	h_2	$e_3^{+\Delta}$	e_4	e_5	e_6	
e_1	5	2	27	- 9	- 2	3	- 33
e_2	2	1	11	- 4	0	1	- 13
h_3	- 27	- 11	- 145	48	11	- 14	184
a_1	- 9	- 4	- 48	17^*	2	- 5	69
a_2	- 2	0	- 11	2	4	- 2	22
a_3	1	1	14	- 5	- 2	2	- 17

表 4.94 的基向量 e_3 列出现负对角元素 - 145. 虽然非基约束向量 e_1, e_2, a_3 的偏差是负数，但在 e_3 列上的元素都是正数，如果作 $e_3^{+\Delta}$ ，可使这些非基约束向量的偏差变大. e_3 列 a_1 和 a_2 行上的两个元素是负数且 a_1 和 a_2 的偏差是正数，

作平推 $e_3^{+\Delta}$. 由于

$$\min \left\{ \frac{69}{48}, \frac{22}{11} \right\} = \frac{69}{48},$$

此比值对应 a_1 行. 由于 a_1 行的对角元素 $17 > 0$, 以其为枢轴作后向旋转, 结果见表 4.95.

表 4.95 第一次后向旋转结果

	h_1	h_2	$e_3^{+\Delta}$	a_1	e_5	e_6	
e_1	0.235	- 0.118	1.588	- 0.529	- 0.941	0.353	3.529
e_2	- 0.118	0.059	- 0.294	- 0.235	0.471	- 0.176	3.235
h_3	- 1.588	0.294	- 9.471	2.824	5.353	0.118	- 10.824
e_4	0.529	0.235	2.824	0.059	- 0.118	0.294	- 4.059
a_2	- 0.941	0.471	- 5.353	0.118	3.765 *	- 1.412	13.882
a_3	0.353	- 0.176	- 0.118	- 0.294	- 1.412	0.529	3.294

在表 4.95 中, e_3 列上的对角元素仍然是负数. 由于 e_3 列 e_2 , a_2 和 a_3 行上的三个元素是负数且这三个非基向量的偏差是正数, 再作平推 $e_3^{+\Delta}$. 由于

$$\min \left\{ \frac{3.235}{0.294}, \frac{13.882}{5.353}, \frac{3.294}{0.118} \right\} = \frac{13.882}{5.353},$$

此比值对应 a_2 行. 由于 a_2 行的对角元素为正, 以其为枢轴再作后向旋转, 结果见表 4.96.

表 4.96 第二次后向旋转结果

	h_1	h_2	$e_3^{+\Delta}$	a_1	a_2	e_6	
e_1	0	0	0.25	- 0.5	- 0.25	0	7
e_2	0	0	0.375	- 0.25	0.125	0	1.5
h_3	- 2.5	- 0.375	- 1.859	2.656	1.422	2.125 *	- 30.563
e_4	0.5	0.25	2.656	0.063	- 0.031	0.25	- 3.625
e_5	0.25	- 0.125	1.422	- 0.031	0.266	0.375	- 3.688
a_3	0	0	- 2.125 *	- 0.25	- 0.375	0	8.5

在表 4.96 中,基向量 e_3 列的对角元素仍然是负数. 再作平推 $e_3^{+\Delta}$. 为此以 e_3 列 a_3 行的负元素 - 2.125 及其对称元素为枢轴先后进行两次旋转运算, 结果见表 4.97. 至此已将表 4.93 的基向量 e_1 和 e_3 都换成了非基向量.

表 4.97 双枢轴运算结果

	h_1	h_2	a_3	a_1	a_2	h_3	
e_1	0	0	- 0.118	- 0.529	- 0.294	0	8
e_2	0	0	- 0.176	- 0.294	0.059	0	3
e_6	0.118	0.176	- 0.412	- 1.353	- 0.824	0.471	17.882
e_4	0.529	0.294	- 1.353	- 0.588	- 0.706	0.118	11.471
e_5	0.294	- 0.059	- 0.824	- 0.706	- 0.294	0.176	8.706
e_3	0	0	- 0.471	- 0.118	- 1.176	0	4

表 4.97 的所有非基向量偏差是正数, 所以又得到一个局部极小点

$$x_1 = 8 - 1 = 7, \quad x_2 = 3 - 1 = 2, \quad x_3 = 4 - 1 = 3,$$

对应的目标值为 - 88. 由于每个约束向量或者入基或者出基过一次, 可考虑停止迭代.

4.5.2 凹二次规划

在二次规划模型(4.5.1)中, 如果 H 负半定, 即对于任何 n 维向量 x 均有 $x^T H x \leq 0$, 称其为凹二次规划; 如果 H 负定, 即对于任何非零的 n 维向量 x 均有 $x^T H x < 0$, 称其为严格凹二次规划.

定理 4.5.1 设 x^* 是二次规划(4.5.1)的局部极小点. 如果 H 负定, x^* 一定是可行域的极点. 如果 H 负半定, 则存在一个极点, 其相应的目标函数值等于 $f(x^*)$.

证 用 I^* 表示关于 x^* 的紧约束的指标集, I^* 对应的系数向量组的秩记为 $r(I^*)$. 如果 $r(I^*) = n$, x^* 是极点, 所以下面设 $r(I^*) < n$. 在 x^* 的紧约束中找一个系数向量 a_s , 它与 I^* 对应的系数向量组线性无关. 用 p 表示 a_s 在 I^* 相应的齐次方程组的解集上的直交投影, 那么 $p \neq 0$. 显然, 对于任何实数 t , $x^{(t)} = x^* + tp$ 是 I^* 对应的线性方程组的解, 并且当 t 充分接近 0 时, $x^{(t)}$ 还是二次规划的可行解. 记

$$\varphi(t) = f(x^{(t)}) = f(x^* + tp) = \frac{1}{2}(x^* + tp)^T H (x^* + tp) + c(x^* + tp)$$

$$= \frac{1}{2}t^2 p^T H p + t (p^T H x^* + c p) + f(x^*).$$

由于 x^* 是二次规划的局部极小点, $t=0$ 是 $\varphi(t)$ 的极小点.

若 $p^T H p < 0$ $\varphi(t)$ 是 t 的凹函数, 与 $t=0$ 是 $\varphi(t)$ 的极小点矛盾, 所以 H 负定时, x^* 一定是可行域的极点.

若 $p^T H p = 0$, 由于 t 可正可负, 一定有 $\varphi(t) = f(x^{(t)}) = f(x^*)$. 让 t 由 0 逐渐减少, 直到 $x^{(t)} = x^* + tp$ 到达一个松约束的边界, 至少可使其到达 $a_s x \geq b_s$ 的边界. 因为

$$a_s x^{(t)} = a_s (x^* + tp) = a_s x^* + t a_s p = a_s x^* + t \|p\|^2$$

随 t 减少而不断变小. 所以, 关于 $x^{(t)}$ 的紧约束的系数向量组的秩至少增加 1. 如果其秩等于 n $x^{(t)}$ 成为极点, 否则重复同样的过程, 直至到达可行域的一个极点. ■

根据以上定理, 对于凹二次规划, 只需要考察可行域的极点, 或者说应使基向量的数目始终保持为 n 个. 为此仅需要作双枢轴运算, 因为初始表基向量的数目为 n , 而双枢轴运算不改变基向量的数目. 事实上, 当 H 负半定时, 拉格朗日矩阵(开始为 H)没有正对角元素, 不可能作前向旋转.

将凹二次规划按一般二次规划求解, 很容易用参数化方法判断与当前极点相邻的极点是否对应更小的目标函数值. 但总的说来, 计算量较大. 可直接用第二章的方法找出可行域的所有极点, 看哪一个极点对应的目标值最小.

第五章 约束非线性规划

5.1 序列二次规划法简介

5.1.1 序列二次规划的类型

约束非线性规划是求一个 n 元实值函数在一个给定的集合内的局部或全局极小点, 其数学模型通常表示为以下形式:

$$\left. \begin{aligned} \min f(x), \\ \text{s. t. } g_i(x) = 0, \quad i = 1, 2, \dots, l, \\ g_i(x) \geq 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (5.1.1)$$

其中 $f(x)$ 和 $g_i(x)$ 为 n 元实值函数并且至少有一个是非线性的. 用 λ_i 表示第 i 个约束对应的拉格朗日乘子, (5.1.1) 的 KT 条件为

$$\left. \begin{aligned} \nabla f(x) &= \sum_{i=1}^m \lambda_i \nabla g_i(x), \\ \lambda_i &\geq 0, \lambda_i g_i(x) = 0, \quad i = l+1, l+2, \dots, m, \\ g_i(x) &= 0, \quad i = 1, 2, \dots, l, \\ g_i(x) &\geq 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (5.1.2)$$

无论是(5.1.1)还是(5.1.2) 其形式可能十分复杂, 难以计算. 为了使问题易于求解, 常在某一点处用一个近似模型特别是二次规划模型代替(5.1.1), 以一系列二次规划的解逼近(5.1.1)的解. 这种方法称为序列二次规划(sequential quadratic programming, SQP)法. (5.1.1)称为原问题, 作为其近似的二次规划称为子问题.

在序列二次规划法中, 子问题的形式为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} s^T H_k s + \nabla f(x^{(k)})^T s, \\ \text{s. t.} \quad & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\ & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (5.1.3)$$

其中 $s = x - x^{(k)}$, $x^{(k)}$ 是 R^n 中一个特定点, H_k 是一个 n 阶实对称矩阵. 可以看出(5.1.3)的约束函数是 $g_i(x)$ 在 $x^{(k)}$ 处泰勒展开式的不高于一次项的部分. 但是目标函数的海色矩阵 H_k 形式多种多样, 相应的方法难以起个合适的名称. 本书讨论以下几种:

(i) $H_k = O$ 这时(5.1.3)是个线性规划, 称为序列线性规划法.

(ii) $H_k = \nabla^2 f(x^{(k)})$ 这时(5.1.3)的目标函数是 $f(x)$ 在 $x^{(k)}$ 处泰勒展开式的二次和一次项部分, 称为牛顿法.

(iii) $H_k = \nabla_x^2 L(x^{(k)}, \lambda^{(k)})$ 这里的 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)})$ 是拉格朗日函数 $L(x, \lambda)$ 在 $(x^{(k)}, \lambda^{(k)})$ 处关于 x 的海色矩阵, 依定义,

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i g_i(x).$$

我们将这种方法称为拉格朗日法.

(iv) H_k 为一个增广拉格朗日函数在 $(x^{(k)}, \lambda^{(k)})$ 处关于 x 的海色矩阵. 增广拉格朗日函数的形式很多, 我们仅使用 Powell (1978) 提出的增广拉格朗日函数, 但实际使用时略有不同. 该增广拉格朗日函数的形式为

$$P(x, \lambda) = f(x) - \sum_{i \in I^*} \lambda_i g_i(x) + r \sum_{i \in I^*} (g_i(x))^2,$$

其中 I^* 是关于 $x^{(k)}$ 的紧约束的指标集, 即

$$I^* = \{i \mid g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, m\}.$$

我们将这种方法称为增广拉格朗日法.

(v) H_k 是一个包含 $\nabla^2 f(x^{(k)})$ 或 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)})$ 信息的正定矩阵. 其形式很多, 统称为拟牛顿法.

以上序列线性规划或二次规划法都有各自的适用范围. 从 H_k 的形式看, 牛顿法的子问题是原问题在 $x^{(k)}$ 处的最好近似. 但是, 实际计算表明, 不少具有非线性约束的问题, 牛顿法不如拉格朗日法或增广拉格朗日法有效.

子问题(5.1.3)的KT条件为

$$\left. \begin{aligned} H_k s + \nabla f(x^{(k)}) &= \sum_{i=1}^m \lambda_i \nabla g_i(x^{(k)}), \\ \lambda_i &\geq 0, \lambda_i (\nabla g_i(x^{(k)})^T s + g_i(x^{(k)})) = 0, \quad i = 1 + 1, 1 + 2, \dots, m, \\ \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) &= 0, \quad i = 1, 2, \dots, l, \\ \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) &\geq 0, \quad i = 1 + 1, 1 + 2, \dots, m. \end{aligned} \right\} \quad (5.1.4)$$

此不等式组的解记为 $(s^{(k)}, \lambda^{(k+1)})$. 如果 $s^{(k)} = 0$, 可以看出 $(x^{(k)}, \lambda^{(k+1)})$ 是原问题的 KT 条件 (5.1.2) 的解, 因而 $x^{(k)}$ 是原问题的一个可供选择的解. 这大概就是为什么子问题 (5.1.3) 目标函数的海色矩阵虽然有多种形式, 但其线性部分的系数向量总是 $\nabla f(x^{(k)})$ 的原因. 实际计算时, 可选择一正数 ε , 一旦 $\|s^{(k)}\| < \varepsilon$ 便停止计算. 在本章的例子中, $\varepsilon = 10^{-4}$, 多数例子选自余俊和周济 (1989)、赵瑞安和吴方 (1992)、菲赫金哥耳茨 (1964) 的相关著述. 但有些例子的约束条件作了修改.

5.1.2 子问题初始解的选择

求解子问题 (5.1.3) 的目的是希望得到的解 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 对于 $f(x)$ 更好些, 即

$$f(x^{(k+1)}) < f(x^{(k)}).$$

另外还希望, 如果 $x^{(k)}$ 在原问题 (5.1.1) 的可行域 D 内, $x^{(k+1)}$ 也在 D 内; 如果 $x^{(k)}$ 不在 D 内, $x^{(k+1)}$ 最好距 D 更近一点, 即

$$\text{dist}(x^{(k+1)}, D) \leq \text{dist}(x^{(k)}, D),$$

其中 $\text{dist}(x, D)$ 表示 x 到 D 的“距离”.

由于子问题仅仅是原问题的一个近似, $x^{(k+1)}$ 与 $x^{(k)}$ 可能相距较远, 以上两式往往不能同时成立, 所以常常在 $x^{(k)}$ 与 $x^{(k)} + s^{(k)}$ 之间找一个点

$$x^\alpha = x^{(k)} + \alpha s^{(k)}, \quad 0 < \alpha \leq 1,$$

使 x^α 对于原问题 (5.1.1) 较 $x^{(k)}$ 更好一点. 但有时候这个目的也实现不了. 上式中的 α 称为步长, $s^{(k)}$ 称为 $x^{(k)}$ 处的搜索方向.

点 $x^{(k)}$ 到可行域 D 的距离是一个非负实数, 常用距离有

$$\text{dist}(x^{(k)}, D) = \sum_{i=1}^l |g_i(x^{(k)})| - \sum_{i=1+m}^m \min\{0, g_i(x^{(k)})\},$$

或

$$\text{dist}(x^{(k)}, D) = \sum_{i=1}^l (g_i(x^{(k)}))^2 + \sum_{i=1+m}^m (\min\{0, g_i(x^{(k)})\})^2.$$

前者类似向量范数 $\|\cdot\|_1$ 称为 L_1 距离, 计算最方便; 后者类似 $\|\cdot\|_2$ 称为 L_2 距离.

我们用以下两种方式确定下一个二次规划的初始点.

(i) 始终以 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 作为下一个二次规划的初始点, 即步长恒为 1.

(ii) 可行性优先准则. 具体地说, 如果 $x^{(k)} + s^{(k)}$ 是原问题的可行解, 以 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 作为下一个二次规划的初始点; 否则, 依次计算

$$x^\alpha = x^{(k)} + \alpha s^{(k)}, \quad \alpha = 0.1, 0.2, \dots, 0.9, 1.0,$$

到可行域 D 的 L_1 距离, 以最接近 D 的 x^α 作为 $x^{(k+1)}$.

对于许多实际问题, 当前状况可作为初始点 $x^{(0)}$; 对于理论问题, $x^{(0)}$ 必须人为给定. 其位置对迭代次数可能影响很大, 甚至影响收敛性.

设 z^* 是 p 维空间 R^p 内一个点, $\{z^{(k)} \mid k=1, 2, \dots\}$ 是收敛于 z^* 但不等于 z^* 的一个点列. 如果

$$\lim_{k \rightarrow +\infty} \frac{\|z^{(k+1)} - z^*\|}{\|z^{(k)} - z^*\|} = 0,$$

就说 $\{z^{(k)}\}$ 超线性收敛于 z^* . 若

$$\lim_{k \rightarrow +\infty} \frac{\|z^{(k+1)} - z^*\|}{\|z^{(k)} - z^*\|^2} = \text{正的常数},$$

称 $\{z^{(k)}\}$ 二阶收敛于 z^* .

超线性收敛, 特别是二阶收敛, 是非常理想的收敛速度.

5.1.3 实用模型及其子问题

模型(5.1.1)将所有约束函数用统一的形式 $g_i(x)$ 表示, 理论分析十分方便. 但实际使用时, 通常将变量的下界约束单独列出来, 即表示为以下形式.

$$\left. \begin{aligned} \min f(x), \\ \text{s. t. } g_i(x) &= 0, \quad i = 1, 2, \dots, l, \\ g_i(x) &\geq 0, \quad i = l+1, l+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.1.5)$$

将约束条件 $x_i \geq l_i$ 对应的拉格朗日乘子用 $\mu_i (i=1, 2, \dots, n)$ 表示, e_i 表示 n 阶单位矩阵的第 i 行, 以上模型的 KT 条件为

$$\left. \begin{aligned} \nabla f(x) &= \sum_{i=1}^m \lambda_i \nabla g_i(x) + \sum_{i=1}^n \mu_i e_i^T, \\ \lambda_i &\geq 0, \lambda_i g_i(x) = 0, \quad i = 1+1, 1+2, \dots, m, \\ \mu_i &\geq 0, \mu_i (x_i - l_i) = 0, \quad i = 1, 2, \dots, n, \\ g_i(x) &= 0, \quad i = 1, 2, \dots, l, \\ g_i(x) &\geq 0, \quad i = 1+1, 1+2, \dots, m, \\ x_i &\geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.1.6)$$

模型(5.1.5)的子问题为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} s^T H_k s + \nabla f(x^{(k)})^T s, \\ \text{s. t.} \quad & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\ & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) \geq 0, \quad i = 1+1, 1+2, \dots, m, \\ & s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.1.7)$$

其中 $s = x - x^{(k)}$. 子问题(5.1.7)也用 $Q(H_k, x^{(k)})$ 表示, 其 KT 条件为

$$\begin{aligned} H_k s + \nabla f(x^{(k)}) &= \sum_{i=1}^m \lambda_i \nabla g_i(x^{(k)}) + \sum_{i=1}^n \mu_i e_i^T, \\ \lambda_i &\geq 0, \lambda_i (\nabla g_i(x^{(k)})^T s + g_i(x^{(k)})) = 0, \quad i = 1+1, 1+2, \dots, m, \\ \mu_i &\geq 0, \mu_i (s_i - l_i + x_i^{(k)}) = 0, \quad i = 1, 2, \dots, n, \\ \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) &= 0, \quad i = 1, 2, \dots, l, \\ \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) &\geq 0, \quad i = 1+1, 1+2, \dots, m, \\ s_i &\geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned}$$

消去其中的 $\mu_i (i = 1, 2, \dots, n)$ 整理后可用标量形式表示为

$$\left. \begin{aligned}
 & \sum_{j=1}^n h_{ij} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \geq 0, \quad i = 1, 2, \dots, n, \\
 & \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\
 & \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m, \\
 & s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n, \\
 & \lambda_i \geq 0, \quad i = l+1, l+2, \dots, m, \\
 & (s_i - l_i + x_i^{(k)}) \left(\sum_{j=1}^n h_{ij} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \right) = 0, \quad i = 1, 2, \dots, n, \\
 & \lambda_i \left(\sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \right) = 0, \quad i = l+1, l+2, \dots, m.
 \end{aligned} \right\} \quad (5.1.8)$$

其中 $(h_{ij})_{n \times n} = H_k$. 再引入人工不等式

$$\lambda_i \geq 0, \quad i = 1, 2, \dots, l.$$

以上不等式组有 $n+m$ 个变量, 其解记为 $(s^{(k)}, \lambda^{(k+1)})$, 其中 $s^{(k)} = (s_1^{(k)}, s_2^{(k)}, \dots, s_n^{(k)})^T$, $\lambda^{(k+1)} = (\lambda_1^{(k+1)}, \lambda_2^{(k+1)}, \dots, \lambda_m^{(k+1)})^T$. 如果加上消掉的 μ_i ($i=1, 2, \dots, n$), 总共有 $2n+m$ 个变量, 其解记为 $(s^{(k)}, \lambda^{(k+1)}, \mu^{(k+1)})$, 其中 $\mu^{(k+1)} = (\mu_1^{(k+1)}, \mu_2^{(k+1)}, \dots, \mu_n^{(k+1)})^T$, 而

$$\mu_i^{(k+1)} = \sum_{j=1}^n h_{ij} s_j^{(k)} - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j^{(k+1)} + \frac{\partial f(x^{(k)})}{\partial x_i}, \quad i = 1, 2, \dots, n.$$

$s^{(k)} = 0$ 时的解称为极限解, 用 (s^*, λ^*, μ^*) 表示, $x^{(k+1)}$ 用 x^* 表示.

不等式组(5.1.8)线性部分的系数向量依次记为

$$h_i, i = 1, 2, \dots, n; \quad a_i, i = 1, 2, \dots, m; \quad e_i, i = 1, 2, \dots, n+m.$$

我们用旋转算法解以上不等式组的线性部分, 并使互补松弛条件得以满足.

初始基本不等式组为

$$s_i \geq l_i - x_i^{(k)}, i = 1, 2, \dots, n; \quad \lambda_i \geq 0, i = 1, 2, \dots, m,$$

即 e_i ($i=1, 2, \dots, n+m$) 是初始基向量, h_i ($i=1, 2, \dots, n$) 和 a_i ($i=1, 2, \dots, m$) 是非基向量. 初始基本解为 $s_i = l_i - x_i^{(k)}, i=1, 2, \dots, n; \lambda_i = 0, i=1, 2, \dots, m$. 将初始表划分为 4×4 部分, 其形式如表 5.1 所示.

表 5.1 子问题的初始表

	e_I	e_{II}	e_{III}	e_{IV}	
h_I	H_{11}	H_{12}	$-A_{11}^T$	$-A_{21}^T$	σ_I
h_{II}	H_{12}^T	H_{22}	$-A_{12}^T$	$-A_{22}^T$	σ_{II}
a_I	A_{11}	A_{12}	O	O	σ_{III}
a_{II}	A_{21}	A_{22}	O	O	σ_{IV}

其中 $\begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$ 是由 H_k 划分而成的分块矩阵, $\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 是由

$$\begin{pmatrix} \frac{\partial g_1(x^{(k)})}{\partial x_1} & \cdots & \frac{\partial g_1(x^{(k)})}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial g_m(x^{(k)})}{\partial x_1} & \cdots & \frac{\partial g_m(x^{(k)})}{\partial x_n} \end{pmatrix}$$

划分而成的矩阵. h_I h_{II} a_I a_{II} 和 e_I e_{II} e_{III} e_{IV} 分别表示以上矩阵对应的非向量
和基向量. 表 5.1 中非基向量的偏差分别为

$$\left. \begin{aligned} \sigma_I &= H_{11}(l_I - x_I^{(k)}) + H_{12}(l_{II} - x_{II}^{(k)}) + \nabla f(x^{(k)})_I, \\ \sigma_{II} &= H_{12}^T(l_I - x_I^{(k)}) + H_{22}(l_{II} - x_{II}^{(k)}) + \nabla f(x^{(k)})_{II}, \\ \sigma_{III} &= A_{11}(l_I - x_I^{(k)}) + A_{12}(l_{II} - x_{II}^{(k)}) + g_I(x^{(k)}), \\ \sigma_{IV} &= A_{21}(l_I - x_I^{(k)}) + A_{22}(l_{II} - x_{II}^{(k)}) + g_{II}(x^{(k)}). \end{aligned} \right\} \quad (5.1.9)$$

其中 l_I 和 l_{II} 是由原问题变量 x 的下界 $(l_1, l_2, \dots, l_n)^T$ 划分而成的两个向量;
 $x_I^{(k)}$ 和 $x_{II}^{(k)}$ 是由 $x^{(k)}$ 划分而成的两个向量; $\nabla f(x^{(k)})_I$ 和 $\nabla f(x^{(k)})_{II}$ 以及
 $g_I(x^{(k)})$ 和 $g_{II}(x^{(k)})$ 分别是由 $\nabla f(x^{(k)})$ 和 $g(x^{(k)})$ 划分而成的向量, $g(x^{(k)}) =$
 $(g_1(x^{(k)}), g_2(x^{(k)}), \dots, g_m(x^{(k)}))^T$.

5.1.4 拉格朗日法的收敛性

在拉格朗日法中,子问题(5.1.7)目标函数的海色矩阵 $H_k = \nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$. 设初始表为表 5.1, 经过若干次迭代后 h_{II} 和 a_I 成为基向量, e_{II} 和 e_{III} 成为非基向量, 即对于当前表来说, 表 5.1 中的

$$\begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}$$

是基矩阵. 根据第一章的旋转运算公式, 当前表非基向量 e_{II} 和 e_{III} 的偏差为

$$\begin{aligned} \begin{pmatrix} \sigma'_{II} \\ \sigma'_{III} \end{pmatrix} &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{II} \\ \sigma_{III} \end{pmatrix} \\ &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \left(\begin{pmatrix} H_{12}^T & H_{22} \\ A_{11} & A_{12} \end{pmatrix} \begin{pmatrix} l_I - x_I^{(k)} \\ l_{II} - x_{II}^{(k)} \end{pmatrix} + \begin{pmatrix} \nabla f(x^{(k)})_{II} \\ g_I(x^{(k)}) \end{pmatrix} \right). \end{aligned}$$

于是

$$\begin{aligned} \begin{pmatrix} s_{II}^{(k)} \\ \lambda_I^{(k+1)} \end{pmatrix} &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{II} \\ \sigma_{III} \end{pmatrix} + \begin{pmatrix} l_{II} - x_{II}^{(k)} \\ 0 \end{pmatrix} \\ &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \left(\begin{pmatrix} H_{12}^T & H_{22} \\ A_{11} & A_{12} \end{pmatrix} \begin{pmatrix} l_I - x_I^{(k)} \\ l_{II} - x_{II}^{(k)} \end{pmatrix} + \begin{pmatrix} \nabla f(x^{(k)})_{II} \\ g_I(x^{(k)}) \end{pmatrix} \right) + \begin{pmatrix} l_{II} - x_{II}^{(k)} \\ 0 \end{pmatrix} \\ &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \left(\begin{pmatrix} H_{12}^T \\ A_{11} \end{pmatrix} (l_I - x_I^{(k)}) + \begin{pmatrix} \nabla f(x^{(k)})_{II} \\ g_I(x^{(k)}) \end{pmatrix} \right), \end{aligned}$$

其中 $s_{II}^{(k)} = x_{II}^{(k+1)} - x_{II}^{(k)}$. 于是

$$\begin{aligned} \begin{pmatrix} x_{II}^{(k+1)} - x_{II}^* \\ \lambda_I^{(k+1)} - \lambda_I^* \end{pmatrix} &= \begin{pmatrix} x_{II}^{(k)} - x_{II}^* \\ \lambda_I^{(k)} - \lambda_I^* \end{pmatrix} - \begin{pmatrix} 0 \\ \lambda_I^{(k)} \end{pmatrix} + \begin{pmatrix} s_{II}^{(k)} \\ \lambda_I^{(k+1)} \end{pmatrix} \\ &= \begin{pmatrix} x_{II}^{(k)} - x_{II}^* \\ \lambda_I^{(k)} - \lambda_I^* \end{pmatrix} - \begin{pmatrix} 0 \\ \lambda_I^{(k)} \end{pmatrix} - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \left(\begin{pmatrix} H_{12}^T \\ A_{11} \end{pmatrix} (l_I - x_I^{(k)}) + \begin{pmatrix} \nabla f(x^{(k)})_{II} \\ g_I(x^{(k)}) \end{pmatrix} \right) \\ &= - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \left(- \begin{pmatrix} H_{12}^T & H_{22} \\ A_{11} & A_{12} \end{pmatrix} \begin{pmatrix} x_I^{(k)} - l_I \\ x_{II}^{(k)} - x_{II}^* \\ \lambda_I^{(k)} - \lambda_I^* \end{pmatrix} + \begin{pmatrix} \nabla_x L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})_{II} \\ g_I(x^{(k)}) \end{pmatrix} \right). \end{aligned}$$

在此假设子问题产生的点列收敛, 并且当 k 充分大时子问题最终表的基向量不发生变化. 这实际上相当于在解一个约束条件都是等式的非线性规划. 不妨假设这些等式对应的向量为 a_I 和 e_I , 忽略表 5.1 的 a_{II} 行及其互补向量 e_{IV} 所在的列. 由于 e_I 是基向量, k 充分大时 $x_I^{(k+1)} = x_I^{(k)} = x_I^* = l_I$, 即序列 $\{x_I^{(k)}\}$ 不发生变化. 另外

$$\nabla_x L(x^*, \lambda^*, \mu^*)_{\parallel} = 0, \quad g_I(x^*) = 0.$$

所以上式可写为

$$\begin{pmatrix} x_{\parallel}^{(k+1)} - x_{\parallel}^* \\ \lambda_I^{(k+1)} - \lambda_I^* \end{pmatrix} = - \begin{pmatrix} H_{22} & -A_{12}^T \\ A_{12} & O \end{pmatrix}^{-1} \begin{pmatrix} \nabla_x L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})_{\parallel} \\ g_I(x^{(k)}) \end{pmatrix} \\ - \begin{pmatrix} \nabla_x L(x^*, \lambda^*, \mu^*)_{\parallel} \\ g_I(x^*) \end{pmatrix} - \begin{pmatrix} H_{12}^T & H_{22} & -A_{12}^T \\ A_{11} & A_{12} & O \end{pmatrix} \begin{pmatrix} x_I^{(k)} - x_I^* \\ x_{\parallel}^{(k)} - x_{\parallel}^* \\ \lambda_I^{(k)} - \lambda_I^* \end{pmatrix},$$

其中 $\begin{pmatrix} H_{12}^T & H_{22} & -A_{12}^T \\ A_{11} & A_{12} & O \end{pmatrix}$ 是向量值函数 $\begin{pmatrix} \nabla_x L(x, \lambda, \mu)_{\parallel} \\ g_I(x) \end{pmatrix}$ 在 $(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$ 处的雅可比矩阵。

如果我们还假设 $f(x)$ 的三阶偏导数、 $g_i(x)$ 的二阶偏导数在 x^* 处连续, 那么根据定理 1.4.3, 有

$$\left\| \begin{pmatrix} x_{\parallel}^{(k+1)} - x_{\parallel}^* \\ \lambda_I^{(k+1)} - \lambda_I^* \end{pmatrix} \right\| \leq M \left\| \begin{pmatrix} x_{\parallel}^{(k)} - x_{\parallel}^* \\ \lambda_I^{(k)} - \lambda_I^* \end{pmatrix} \right\|^2,$$

其中 M 是一个常数. 此式说明点列 $\left\{ \begin{pmatrix} x_{\parallel}^{(k)} \\ \lambda_I^{(k)} \end{pmatrix} \right\}$ 二阶收敛于 $\left\{ \begin{pmatrix} x_{\parallel}^* \\ \lambda_I^* \end{pmatrix} \right\}$.

由于我们关心的是点列 $\left\{ x^{(k)} = \begin{pmatrix} x_I^{(k)} \\ x_{\parallel}^{(k)} \end{pmatrix} \right\}$ 的收敛性, 不讨论点列 $\{\lambda_{\parallel}^{(k)}\}$ 的收敛性. 综合以上所述, 笼统地得出以下结论:

定理 5.1.1 如果拉格朗日法产生的点列收敛(步长均为 1), 并且当点列充分接近于极限点时所有子问题最终表的基向量不发生变化, 那么点列二阶收敛于其极限点.

以上我们假设 $x^{(k)}$ 充分接近 x^* 时, 子问题最终表的基向量不发生变化, 在大多数情况下应该是合理的. 特别是当子问题最终表中所有非基向量(不包括基等式对应的非基拉格朗日向量)的偏差都是正数时. 因为只要 $f(x)$ 的三阶偏导数、 $g_i(x)$ 的二阶偏导数在 x^* 处连续, 那么当 $x^{(k)}$ 充分接近 x^* 时, 子问题最终表中非基向量的偏差将仍然是正数.

以拉格朗日函数的海色矩阵作为子问题目标函数的海色矩阵是 Wilson 1963 年提出来的, 在袁亚湘和孙文瑜的著作(2003)中称为 Lagrange-Newton 法, 其收敛性在该著作及赵瑞安和吴方的著作(1992)中有详细介绍.

5.2 序列线性规划法

用序列线性规划法解非线性规划模型(5.1.5) ,搜索方向由以下模型确定.

$$\left. \begin{aligned} \min \quad & \nabla f(x^{(k)})^T s, \\ \text{s. t.} \quad & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\ & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m, \\ & s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.2.1)$$

其中 $s = x - x^{(k)}$. 此模型是一个线性规划,但也可看做是二次规划目标函数的海色矩阵为零矩阵的情况,用二次规划的算法解之. 模型(5.2.1)的KT条件为

$$\left. \begin{aligned} - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j &\geq - \frac{\partial f(x^{(k)})}{\partial x_i}, \quad i = 1, 2, \dots, n, \\ \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j &= - g_i(x^{(k)}), \quad i = 1, 2, \dots, l, \\ \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j &\geq - g_i(x^{(k)}), \quad i = l+1, l+2, \dots, m, \\ s_i &\geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n, \\ \lambda_i &\geq 0, \quad i = l+1, l+2, \dots, m, \\ (s_i - l_i + x_i^{(k)}) \left(- \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \right) &= 0, \quad i = 1, 2, \dots, n, \\ \lambda_i \left(\sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \right) &= 0, \quad i = l+1, l+2, \dots, m. \end{aligned} \right\} \quad (5.2.2)$$

如前所述,不等式组(5.2.2)线性部分的系数向量依次记为

$$h_i, i = 1, 2, \dots, n; \quad a_i, i = 1, 2, \dots, m; \quad e_i, i = 1, 2, \dots, n+m.$$

初始基本不等式组为

$$s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n; \quad \lambda_i \geq 0, \quad i = 1, 2, \dots, m,$$

即 $e_i (i = 1, 2, \dots, n+m)$ 是初始基向量, $h_i (i = 1, 2, \dots, n)$ 和 $a_i (i = 1, 2, \dots, m)$ 是非基向量. 初始基本解为

$$s_i = l_i - x_i^{(k)}, i = 1, 2, \dots, n; \quad \lambda_i = 0, i = 1, 2, \dots, m.$$

由(5.1.9), 非基向量 h_i 的偏差为

$$\frac{\partial f(x^{(k)})}{\partial x_i}, \quad i = 1, 2, \dots, n,$$

非基向量 a_i 的偏差为

$$\frac{\partial g_i(x^{(k)})}{\partial x_1}(1 - x_1^{(k)}) + \dots + \frac{\partial g_i(x^{(k)})}{\partial x_n}(1 - x_n^{(k)}) + g_i(x^{(k)}), \quad i = 1, 2, \dots, m.$$

由此可列出初始表并进行计算.

下面用两个例子说明计算细节.

例 5.2.1 用序列线性规划法解非线性规划

$$\min f(x) = -x_1^3 + 4x_1^2 - 2x_1x_2 + x_2^2,$$

$$\text{s. t. } g_1(x) = -x_1 + 5 \geq 0,$$

$$g_2(x) = -x_2 + 1 \geq 0,$$

$$x_1 \geq -5, \quad x_2 \geq -1.$$

设给定的初始解为 $x^{(0)} = (-6, -3)^T$.

解 由于

$$\frac{\partial f(x)}{\partial x_1} = -3x_1^2 + 8x_1 - 2x_2, \quad \frac{\partial f(x)}{\partial x_2} = -2x_1 + 2x_2,$$

$$\frac{\partial f(x^{(0)})}{\partial x_1} = -150, \quad \frac{\partial f(x^{(0)})}{\partial x_2} = 6,$$

$$g_1(x^{(0)}) = 11, \quad g_2(x^{(0)}) = 4, \quad x_1^{(0)} + 5 = -1, \quad x_2^{(0)} + 1 = -2,$$

所以 $x^{(0)} = (-6, -3)^T$ 对应的子问题为

$$\min -150s_1 + 6s_2,$$

$$\text{s. t. } -s_1 \geq -11,$$

$$-s_2 \geq -4,$$

$$s_1 \geq 1, \quad s_2 \geq 2.$$

其初始表为表 5.2. 由于拉格朗日向量 h_1 的偏差是负数, 同行对角元素为 0, 作平推 $e_1^{+\Delta}$. 由于 e_1 列非基约束向量行上仅一个负元素, 作双枢轴运算 ($a_1 \leftrightarrow e_1, h_1 \leftrightarrow e_3$) 结果见表 5.3.

表 5.2 子问题 0 的初始表

	e ₁	e ₂	e ₃	e ₄	
h ₁	0	0	1 *	0	- 150
h ₂	0	0	0	1	6
a ₁	- 1 *	0	0	0	10
a ₂	0	- 1	0	0	2

表 5.3 双枢轴运算结果

	a ₁	e ₂	h ₁	e ₄	
e ₃	0	0	1 *	0	150
h ₂	0	0	0	1	6
e ₁	- 1 *	0	0	0	10
a ₂	0	- 1	0	0	2

由于表 5.3 所有非基向量偏差为正 ,当前基本解

$$s_1^{(0)} = 10 + 1 = 11 , \quad s_2^{(0)} = 2$$

是子问题 0 的最优解. 下一个子问题的初始解为

$$x_1^{(1)} = x_1^{(0)} + s_1^{(0)} = - 6 + 11 = 5 , \quad x_2^{(1)} = x_2^{(0)} + s_2^{(0)} = - 3 + 2 = - 1.$$

模型为

$$\begin{aligned} \min & \quad - 33s_1 - 12s_2 , \\ \text{s. t.} & \quad - s_1 \geq 0 , \\ & \quad - s_2 \geq - 2 , \\ & \quad s_1 \geq - 10 , s_2 \geq 0. \end{aligned}$$

其初始表为表 5.4. 表 5.4 中拉格朗日向量 h₁和 h₂的偏差都是负数 ,先后作双枢轴运算(a₁↔e₁ ,h₁↔e₃) ,(a₂↔e₂ ,h₂↔e₄) ,其结果见表 5.5.

表 5.4 子问题 1 的初始表

	e ₁	e ₂	e ₃	e ₄	
h ₁	0	0	1	0	- 33
h ₂	0	0	0	1	- 12
a ₁	- 1	0	0	0	10
a ₂	0	- 1	0	0	2

表 5.5 两次双枢轴运算结果

	a ₁	a ₂	h ₁	h ₂	
e ₃	0	0	1	0	33
e ₄	0	0	0	1	12
e ₁	- 1	0	0	0	10
e ₂	0	- 1	0	0	2

子问题 1 的解为

$$s_1^{(1)} = 10 - 10 = 0 , \quad s_2^{(1)} = 2 + 0 = 2.$$

下一个子问题的初始解为

$$x_1^{(2)} = 5 + 0 = 5 , \quad x_2^{(2)} = - 1 + 2 = 1.$$

其模型为

$$\begin{aligned} \min & -37s_1 - 8s_2, \\ \text{s. t.} & -s_1 \geq 0, \\ & -s_2 \geq 0, \\ & s_1 \geq -10, \quad s_2 \geq -2. \end{aligned}$$

同样可求得

$$s_1^{(2)} = 10 - 10 = 0, \quad s_2^{(2)} = 2 - 2 = 0.$$

由于搜索方向 s 已是零向量, 算法停止. 原问题的解为

$$x_1^* = x_1^{(2)} = 5, \quad x_2^* = x_2^{(2)} = 1,$$

相应的目标函数值 $f(x^*) = -34$. 目标函数和两个紧约束在 $x^* = (5, 1)^T$ 处的梯度分别为

$$\nabla f(x^*) = (-37, -8)^T, \quad \nabla g_1(x^*) = (-1, 0)^T, \quad \nabla g_2(x^*) = (0, -1)^T.$$

由于 $\nabla f(x^*)$ 与 x^* 处任一可行方向的夹角小于 90° , 根据定理 1.5.1, $x^* = (5, 1)^T$ 是严格局部极小点, 事实上也是全局极小点.

例 5.2.2 解非线性规划

$$\begin{aligned} \min & f(x) = x_2, \\ \text{s. t.} & g_1(x) = -2x_1^2 + x_1^3 + x_2 \geq 0, \\ & g_2(x) = -2(1 - x_1)^2 + (1 - x_1)^3 + x_2 \geq 0, \\ & x_1, x_2 \geq 0. \end{aligned}$$

设给定的初始解为 $x^{(0)} = (0, 0)^T$.

解 由于

$$\begin{aligned} \nabla f(x) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} -4x_1 + 3x_1^2 \\ 1 \end{pmatrix}, \\ \nabla g_2(x) &= \begin{pmatrix} (1 - x_1)(1 + 3x_1) \\ 1 \end{pmatrix}, \\ \nabla g_1(x^{(0)}) &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \nabla g_2(x^{(0)}) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad g_1(x^{(0)}) = 0, \\ g_2(x^{(0)}) &= -1, \quad x_1^{(0)} - 0 = 0, \quad x_2^{(0)} - 0 = 0, \end{aligned}$$

初始解 $x^{(0)} = (0, 0)^T$ 对应的子问题为

$$\begin{aligned} \min & s_2, \\ \text{s. t.} & s_2 \geq 0, \\ & s_1 + s_2 \geq 1, \\ & s_1, s_2 \geq 0. \end{aligned}$$

以上问题的初始表为表 5.6. 由于 a_2 的偏差是负数 ,作双枢轴运算 $a_2 \leftrightarrow e_1$,
 $h_1 \leftrightarrow e_4$,其结果见表 5.7.

表 5.6 子问题 0 的初始表

	e_1	e_2	e_3	e_4	
h_1	0	0	0	- 1 *	0
h_2	0	0	- 1	- 1	1
a_1	0	1	0	0	0
a_2	1 *	1	0	0	- 1

表 5.7 双枢轴运算结果

	a_2	e_2	e_3	h_1	
e_4	0	0	0	- 1	0
h_2	0	0	- 1	1	1
a_1	0	1	0	0	0
e_1	1	- 1	0	0	1

表 5.7 的所有非基向量偏差是正数 ,所以

$$s_1^{(0)} = 1 + 0 = 1, \quad s_2^{(0)} = 0.$$

下面按可行性优先的准则确定下一个子问题的初始点. 由于 $\alpha = 0.5$ 时

$$x^\alpha = x^{(0)} + \alpha s^{(0)} = (\alpha \quad \rho)^T, \quad \alpha = 0.1 \quad 0.2 \quad \dots \quad 1.0,$$

到可行域最近(距离为 0.75) ,下一个问题的初始解为

$$x^{(1)} = (0.5 \quad \rho)^T.$$

子问题 1 为

$$\begin{aligned} \min \quad & s_2, \\ \text{s. t.} \quad & -1.25s_1 + s_2 \geq 0.375, \\ & 1.25s_1 + s_2 \geq 0.375, \\ & s_1 \geq -0.5, \quad s_2 \geq 0. \end{aligned}$$

以上问题的初始表为表 5.8. 由于 a_2 的偏差是负数 ,作双枢轴运算 $a_2 \leftrightarrow e_1$,
 $h_1 \leftrightarrow e_4$,其结果见表 5.9.

表 5.8 子问题 1 的初始表

	e_1	e_2	e_3	e_4	
h_1	0	0	1.25	- 1.25 *	0
h_2	0	0	- 1	- 1	1
a_1	- 1.25	1	0	0	0.25
a_2	1.25 *	1	0	0	- 1

表 5.9 第一次双枢轴运算结果

	a_2	e_2	e_3	h_1	
e_4	0	0	1	- 0.8	0
h_2	0	0	- 2 *	0.8	1
a_1	- 1	2 *	0	0	- 0.75
e_1	0.8	- 0.8	0	0	0.8

在表 5.9 中再作双枢轴运算 $a_1 \leftrightarrow e_2$ $h_2 \leftrightarrow e_3$ 结果见表 5.10.

表 5.10 第二次双枢轴运算结果

	a_2	a_1	h_2	h_1	
e_4	0	0	- 0.5	- 0.4	0.5
e_3	0	0	- 0.5	0.4	0.5
e_2	0.5	0.5	0	0	0.375
e_1	0.4	- 0.4	0	0	0.5

由表 5.10 得

$s_1^{(1)} = 0.5 - 0.5 = 0$, $s_2^{(1)} = 0.375 + 0 = 0.375$.

按可行性优先的准则下一个子问题的初始点为

$x_1^{(2)} = x_1^{(1)} + s_1^{(1)} = 0.5$, $x_2^{(2)} = x_2^{(1)} + s_2^{(1)} = 0.375$.

子问题 2 为

$$\begin{aligned} \min \quad & s_2 , \\ \text{s. t.} \quad & - 1.25s_1 + s_2 \geq 0 , \\ & 1.25s_1 + s_2 \geq 0 , \\ & s_1 \geq - 0.5 , \quad s_2 \geq - 0.375 . \end{aligned}$$

以上问题的初始表为表 5.11. 先后作双枢轴运算($a_2 \leftrightarrow e_1$ $h_1 \leftrightarrow e_4$) ,($a_1 \leftrightarrow e_2$ $h_2 \leftrightarrow e_3$) 其结果见表 5.12.

表 5.11 子问题 2 的初始表

	e_1	e_2	e_3	e_4	
h_1	0	0	1.25	- 1.25	0
h_2	0	0	- 1	- 1	1
a_1	- 1.25	1	0	0	0.25
a_2	1.25	1	0	0	- 1

表 5.12 两次双枢轴运算结果

	a_2	a_1	h_1	h_2	
e_4	0	0	- 0.5	- 0.4	0.5
e_3	0	0	- 0.5	0.4	0.5
e_2	0.5	0.5	0	0	0.375
e_1	0.4	- 0.4	0	0	0.5

由表 5.12 得

$s_1^{(2)} = 0.5 - 0.5 = 0$, $s_2^{(2)} = 0.375 - 0.375 = 0$.

由于 $\parallel s^{(2)} \parallel = 0$ 最终解为

$$\mathbf{x}^* = \mathbf{x}^{(2)} = (0.5 \ 0.375)^T,$$

相应的目标函数值为

$$f(\mathbf{x}^*) = 0.375.$$

目标函数和两个紧约束在 $\mathbf{x}^* = (0.5 \ 0.375)^T$ 处的梯度分别为

$$\nabla f(\mathbf{x}^*) = (0 \ 1)^T, \quad \nabla g_1(\mathbf{x}^*) = (-1.25 \ 1)^T, \quad \nabla g_2(\mathbf{x}^*) = (1.25 \ 1)^T.$$

由于 $\nabla f(\mathbf{x}^*)$ 与 $\mathbf{x}^*$ 处任一可行方向的夹角小于 90° , 根据定理 1.5.1, $\mathbf{x}^* = (0.5 \ 0.375)^T$ 是严格局部极小点, 事实上也是全局极小点.

5.3 牛 顿 法

在牛顿法中, 直接将原问题目标函数 $f(\mathbf{x})$ 的海色矩阵作为子问题目标函数的海色矩阵, 与模型(5.1.5)对应的子问题为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} \mathbf{s}^T \nabla^2 f(\mathbf{x}^{(k)}) \mathbf{s} + \nabla f(\mathbf{x}^{(k)})^T \mathbf{s}, \\ \text{s. t.} \quad & \nabla g_i(\mathbf{x}^{(k)})^T \mathbf{s} + g_i(\mathbf{x}^{(k)}) = 0, \quad i = 1 \ 2 \ \dots \ l, \\ & \nabla g_i(\mathbf{x}^{(k)})^T \mathbf{s} + g_i(\mathbf{x}^{(k)}) \geq 0, \quad i = l+1 \ l+2 \ \dots \ m, \\ & s_i \geq l_i - x_i^{(k)}, \quad i = 1 \ 2 \ \dots \ n. \end{aligned} \right\} \quad (5.3.1)$$

其中 $\mathbf{s} = \mathbf{x} - \mathbf{x}^{(k)}$, 其 KT 条件为

$$\left. \begin{aligned} \sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(\mathbf{x}^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_i} &\geq 0, \quad i = 1 \ 2 \ \dots \ n, \\ \sum_{j=1}^n \frac{\partial g_i(\mathbf{x}^{(k)})}{\partial x_j} s_j + g_i(\mathbf{x}^{(k)}) &= 0, \quad i = 1 \ 2 \ \dots \ l, \\ \sum_{j=1}^n \frac{\partial g_i(\mathbf{x}^{(k)})}{\partial x_j} s_j + g_i(\mathbf{x}^{(k)}) &\geq 0, \quad i = l+1 \ l+2 \ \dots \ m, \\ s_i &\geq l_i - x_i^{(k)}, \quad i = 1 \ 2 \ \dots \ n, \\ \lambda_i &\geq 0, \quad i = l+1 \ l+2 \ \dots \ m, \\ (s_i - l_i + x_i^{(k)}) \left(\sum_{j=1}^n \frac{\partial^2 f(\mathbf{x}^{(k)})}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(\mathbf{x}^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(\mathbf{x}^{(k)})}{\partial x_i} \right) &= 0, \\ &\quad i = 1 \ 2 \ \dots \ n, \\ \lambda_i \left(\sum_{j=1}^n \frac{\partial g_i(\mathbf{x}^{(k)})}{\partial x_j} s_j + g_i(\mathbf{x}^{(k)}) \right) &= 0, \quad i = l+1 \ l+2 \ \dots \ m. \end{aligned} \right\} \quad (5.3.2)$$

以上不等式组的解用 $(\mathbf{s}^{(k)} \ \boldsymbol{\lambda}^{(k+1)})$ 表示. 对于某个 $\mathbf{x}^{(k)}$, 若 $\mathbf{s}^{(k)} = \mathbf{0}$, 可以看出

$(x^{(k)}, \lambda^{(k+1)})$ 是 (5.1.6) 的解, 因而 $x^{(k)}$ 是模型 (5.1.5) 的 KT 点.

例 5.3.1 用牛顿法解非线性规划

$$\begin{aligned} \min f(x) &= -x_1^4 x_2 x_3^2 \\ \text{s. t. } g_1(x) &= -x_1^2 x_2 + 675 \geq 0, \\ g_2(x) &= -x_1^2 x_3^2 + 4190000 \geq 0, \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

初始点为 $x^{(0)} = (22.5, 0.5, 125)^T$.

解 $f(x)$ 的一阶和二阶偏导数 $g_1(x)$ 和 $g_2(x)$ 的一阶偏导数如下:

$$\begin{aligned} \frac{\partial f(x)}{\partial x_1} &= -4x_1^3 x_2 x_3^2, \quad \frac{\partial f(x)}{\partial x_2} = -x_1^4 x_3^2, \quad \frac{\partial f(x)}{\partial x_3} = -2x_1^4 x_2 x_3, \\ \frac{\partial^2 f(x)}{\partial x_1^2} &= -12x_1^2 x_2 x_3^2, \quad \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} = -4x_1^3 x_3^2, \quad \frac{\partial^2 f(x)}{\partial x_1 \partial x_3} = -8x_1^3 x_2 x_3, \\ \frac{\partial^2 f(x)}{\partial x_2^2} &= 0, \quad \frac{\partial^2 f(x)}{\partial x_2 \partial x_3} = -2x_1^4 x_3, \quad \frac{\partial^2 f(x)}{\partial x_3^2} = -2x_1^4 x_2, \\ \frac{\partial g_1(x)}{\partial x_1} &= -2x_1 x_2, \quad \frac{\partial g_1(x)}{\partial x_2} = -x_1^2, \quad \frac{\partial g_1(x)}{\partial x_3} = 0, \\ \frac{\partial g_2(x)}{\partial x_1} &= -2x_1 x_3^2, \quad \frac{\partial g_2(x)}{\partial x_2} = 0, \quad \frac{\partial g_2(x)}{\partial x_3} = -2x_1^2 x_3, \end{aligned}$$

它们在 $x^{(0)} = (22.5, 0.5, 125)^T$ 的数值为

$$\begin{aligned} \frac{\partial f(x^{(0)})}{\partial x_1} &= -3.55957 \times 10^8, \quad \frac{\partial f(x^{(0)})}{\partial x_2} = -4.004517 \times 10^9, \\ \frac{\partial f(x^{(0)})}{\partial x_3} &= -3.203613 \times 10^7, \\ \frac{\partial^2 f(x^{(0)})}{\partial x_1^2} &= -47460936, \quad \frac{\partial^2 f(x^{(0)})}{\partial x_1 \partial x_2} = -7119.1405 \times 10^9, \\ \frac{\partial^2 f(x^{(0)})}{\partial x_1 \partial x_3} &= -5695312.5, \\ \frac{\partial^2 f(x^{(0)})}{\partial x_2^2} &= 0, \quad \frac{\partial^2 f(x^{(0)})}{\partial x_2 \partial x_3} = -64072264, \\ \frac{\partial^2 f(x^{(0)})}{\partial x_3^2} &= -256289.0625, \end{aligned}$$

$$\begin{aligned} \frac{\partial g_1(x^{(0)})}{\partial x_1} &= -22.5, \frac{\partial g_1(x^{(0)})}{\partial x_2} = -506.25, \frac{\partial g_1(x^{(0)})}{\partial x_3} = 0, \\ \frac{\partial g_2(x^{(0)})}{\partial x_1} &= -703125, \frac{\partial g_2(x^{(0)})}{\partial x_2} = 0, \frac{\partial g_2(x^{(0)})}{\partial x_3} = -126562.5, \end{aligned}$$

另外

$$g_1(x^{(0)}) = -421.875, g_2(x^{(0)}) = -3720156, l_1 = -22.5, l_2 = -0.5, l_3 = -125.$$

所以 $x^{(0)} = (22.5, 0.5, 125)^T$ 时的子问题为

$$\begin{aligned} \min \frac{1}{2} (s_1, s_2, s_3) & \begin{pmatrix} -47460936 & -7119.1405 \times 10^9 & -5695312.5 \\ -7119.1405 \times 10^9 & 0 & -64072264 \\ -5695312.5 & -64072264 & -256289.0625 \end{pmatrix} \begin{pmatrix} s_2 \\ s_2 \\ s_3 \end{pmatrix} \\ & + (-3.55957 \times 10^8, -4.004517 \times 10^9, -3.203613 \times 10^7) \begin{pmatrix} s_2 \\ s_2 \\ s_3 \end{pmatrix}, \\ \text{s. t. } & -22.5s_1 - 506.25s_2 + 421.875 \geq 0, \\ & -703125s_1 - 126562.5s_3 - 3720156 \geq 0, \\ & s_1 \geq -22.5, s_2 \geq -0.5, s_3 \geq -125. \end{aligned}$$

此问题的初始表见表 5.13.

表 5.13 子问题 0 的初始表

	e_1	e_2	e_3	e_4	e_5	
h_1	-47460936	-7119.1405×10^9	-5695312.5	22.5	703125^*	1779.7852×10^6
h_2	-7119.1405×10^9	0	-64072264	506.25	0	2002.2583×10^3
h_3	-5695312.5	-64072264	-256289.0625	0	126562.5	160180656
a_1	-22.5	-506.25	0	0	0	1181.25
a_2	-703125^*	0	-126562.5	0	0	27920468

表 5.13 中所有非基向量偏差为正,当前基本解是子问题的局部极小点. 为了得到一个新的搜索方向,我们必须寻找子问题的另一个极小点. 为此,在表 5.13中作旋转运算($a_2 \leftrightarrow e_1, h_1 \leftrightarrow e_5$)结果见表 5.14.

表 5.14 第一次双枢轴运算结果

	a_2	e_2	e_3	e_4	h_1	
e_5	- 0.0001	1012.5	- 4.05	0.0000	0.0000	149.1146
h_2	1012.5	0	64072264	506.25 *	0	- 8246.8905 $\times 10^6$
h_3	- 4.05	64072264	256289.0313	- 4.05	0.18	- 47102816
a_1	0.0000	- 506.25 *	4.05	0	0	287.7950
e_1	0.0000	0.0000	- 0.18	0	0	39.7091

在表 5.14 中再作旋转运算($a_1 \leftrightarrow e_2$ $h_2 \leftrightarrow e_4$) 结果见表 5.15.

表 5.15 第二次双枢轴运算结果

	a_2	a_1	e_3	h_2	h_1	
e_5	0.0000	- 2	8.1	0.0000	0.0000	203.4197
e_4	- 2	0	- 126562.5	0.002	0.0000	16290154
h_3	8.1	- 126562.5	1281445.375 *	- 0.008	0.18	- 76653888
e_2	0.0000	- 0.002	0.008	0	0	0.5685
e_1	0.0000	0.0000	- 0.18	0	0	39.7091

在表 5.15 中再作旋转运算 $h_3 \leftrightarrow e_3$ 结果见表 5.16.

表 5.16 第一次主旋转结果

	a_2	a_1	h_3	h_2	h_1	
e_5	0.0000	- 1.2	0.0000	0.0000	0.0000	687.9479
e_4	- 1.2	- 12500	- 0.0988	0.0012	0.0178	8719400
e_3	0.0000	0.0988	0.0000	0.0000	0.0000	59.8183
e_2	0.0000	- 0.0012	0.0000	0	0	1.0470
e_1	0.0000	- 0.0178	0.0000	0	0	28.9418

此时所有非基向量偏差为正 ,当前基本解为

$$s_1^{(0)} = 28.9418 - 22.5 = 6.441814 ,$$

$$s_2^{(0)} = 1.0470 - 0.5 = 0.5470303 ,$$

$$s_3^{(0)} = 59.8183 - 125 = - 65.1817 .$$

下面按可行性优先的准则确定下一问题的初始点. 计算

$$x^\alpha = x^{(0)} + \alpha s^{(0)}, \quad \alpha = 0.1, 0.2, \dots, 1.0$$

到原问题可行域的 L_1 距离, 可发现 $x^{0.8}$ 距可行域最近. 所以下一个问题的初始点为

$$\begin{aligned} x_1^{(1)} &= x_1^{(0)} + 0.8s_1^{(0)} = 27.6534519, \\ x_2^{(2)} &= x_2^{(0)} + 0.8s_2^{(0)} = 0.93762428, \\ x_3^{(3)} &= x_3^{(0)} + 0.8s_3^{(0)} = 72.8546371. \end{aligned}$$

以后的计算过程类似, 总共解了 4 个二次规划, 进行了 20 次旋转运算. 各子问题的初始点及解如表 5.17 所示, 其中 $x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$, $\alpha_0 = 0.8$, $\alpha_1 = \alpha_2 = \alpha_3 = 1$.

表 5.17 子问题的计算结果

子问题 k	初始点及计算结果
0	$x^{(0)} = (22.5, 0.5, 125),$ $s^{(0)} = (6.441814, 0.5470303, -65.1817)$
1	$x^{(1)} = (27.6534519, 0.93762428, 72.8546371),$ $s^{(1)} = (-0.234787, -0.03901917, 1.794739)$
2	$x^{(2)} = (27.41866, 0.8986051, 74.64938),$ $s^{(2)} = (-4.068375 \times 10^{-3}, -4.728436 \times 10^{-4}, 1.701355 \times 10^{-2})$
3	$x^{(3)} = (27.414597, 0.8981323, 74.66639),$ $s^{(3)} = (7.629395 \times 10^{-6}, -5.960464 \times 10^{-7}, -1.525879 \times 10^{-5})$
4	$x^{(4)} = (27.414597, 0.8981317, 74.66637)$

其中 $\|s^{(3)}\| < 10^{-4}$, 按照我们的精度要求

$$x^* = (27.4146, 0.8981317, 74.66637)^T,$$

相应的目标函数值为

$$f(x^*) = -2.828252 \times 10^9.$$

下面用定理 1.5.2 检验 x^* 是不是原问题的局部极小点. 由于

$$g_1(x^*) = -3.7857358 \times 10^{-4}, \quad g_2(x^*) = 2.2270298 \times 10^{-2},$$

可认为前两个约束是关于 x^* 的紧约束, 即 $I^* = \{1, 2\}$. 它们对应的拉格朗日乘子

$$\lambda_1 = 4190001, \quad \lambda_2 = 675.$$

$$s^T \nabla_x^2 L(x^*, \lambda^*, 0) s > 0.$$

根据定理 1.5.2 ,

$$x^* = (27.4146 \ 0.8981317 \ 74.66637)^T$$

是原问题的严格局部极小点.

例 5.3.2 用牛顿法解非线性规划

$$\min f(x) = (x_1 - 2)^2 + x_2^2 ,$$

$$\text{s. t. } g_1(x) = (x_1 - 1)^2 + (x_2 - 1)^2 - 1 \geq 0 ,$$

$$x_1, x_2 \geq 0.$$

设给定的初始解为 $x^{(0)} = (0 \ 0)^T$.

解 此问题的可行域是一个非凸的曲边三角形, 见图 5.1 的阴影区域.

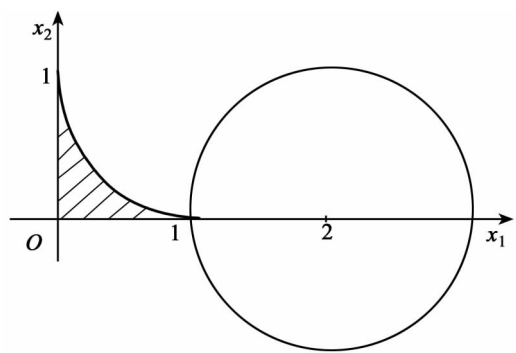


图 5.1 例 5.3.2 图示

可以看出 $x = (1 \ 0)^T$ 是全局极小点. 但在这一点处, 目标函数的梯度与两个紧约束函数的梯度都直交, 不能表示为它们的线性组合, 即 KT 条件不成立.

下面用牛顿法解之. 由于

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 4 \\ 2x_2 \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 2x_1 - 2 \\ 2x_2 - 2 \end{pmatrix},$$

$$\nabla f(x^{(0)}) = \begin{pmatrix} -4 \\ 0 \end{pmatrix}, \quad \nabla g_1(x^{(0)}) = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad g_1(x^{(0)}) = 1, \quad x_1^{(0)} - 0 = 0, \quad x_2^{(0)} - 0 = 0,$$

所以 $x^{(0)} = (0 \ 0)^T$ 的子问题为

$$\min s_1^2 + s_2^2 - 4s_1,$$

$$\text{s. t. } -2s_1 - 2s_2 \geq -1,$$

$$s_1 \geq 0, \quad s_2 \geq 0.$$

其初始表为表 5.20. 由于拉格朗日向量 h_1 的偏差是负数,作旋转运算 $h_1 \leftrightarrow e_1$ 结果见表 5.21.

表 5.20 子问题 0 的初始表

	e_1	e_2	e_3	
h_1	2^*	0	2	- 4
h_2	0	2	2	0
a_1	- 2	- 2	0	1

表 5.21 第一次主旋转结果

	h_1	e_2	e_3	
e_1	0.5	0	- 1	2
h_2	0	2	2	0
a_1	- 1	- 2	2^*	- 3

在表 5.21 中作旋转运算 $a_1 \leftrightarrow e_3$ 结果见表 5.22.

表 5.22 第二次主旋转结果

	h_1	e_2	a_1	
e_1	0	- 1	- 0.5	0.5
h_2	1	4	1	3
e_3	0.5	1	0.5	1.5

由表 5.22 得

$$s_1^{(0)} = 0.5, \quad s_2^{(0)} = 0.$$

下一个子问题的初始解为

$$x_1^{(1)} = 0 + 0.5 = 0.5, \quad x_2^{(1)} = 0 + 0 = 0.$$

以后的计算过程类似. 表 5.23 是子问题的计算结果 其中 $x^{(k+1)} = x^{(k)} + s^{(k)}$.

表 5.23 子问题的计算结果

子问题 k	初始点 $x^{(k)}$	解 $s^{(k)}$
0	(0, 0)	(0.5, 0)
1	(0.5, 0)	(0.25, 0)
2	(0.75, 0)	(0.125, 0)
3	(0.875, 0)	(0.0625, 0)
4	(0.9375, 0)	(0.03125, 0)
5	(0.96875, 0)	(0.015625, 0)
6	(0.984375, 0)	(0.0078125, 0)
7	(0.9921875, 0)	$(3.90625 \times 10^{-3}, 0)$
8	(0.99609375, 0)	$(1.953125 \times 10^{-3}, 0)$
9	(0.99804688, 0)	$(9.765625 \times 10^{-4}, 0)$
10	(0.99902344, 0)	$(4.882813 \times 10^{-4}, 0)$
11	(0.99951172, 0)	$(2.441406 \times 10^{-4}, 0)$
12	(0.99975586, 0)	$(1.220703 \times 10^{-4}, 0)$
13	(0.99987793, 0)	$(6.103516 \times 10^{-5}, 0)$
14	(0.99993896, 0)	

由表 5.23 可见 $x^{(14)} = (0.99993896 \ 0)^T$ 已非常接近原问题的最优解 $x^* = (1 \ 0)^T$.

例 5.3.3 解非线性规划

$$\begin{aligned} \min f(x) &= x_1^2 + x_2^2, \\ \text{s. t. } g_1(x) &= x_1^2 + x_2 - 1 \geq 0, \\ x_1 &\geq -1, \quad x_2 \geq 0. \end{aligned}$$

给定初始解 $x^{(0)} = (0 \ 0)^T$.

解 由于

$$\nabla f(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 1 \end{pmatrix},$$

初始解 $x^{(0)} = (0 \ 0)^T$ 对应的子问题为

$$\begin{aligned} \min s_1^2 + s_2^2, \\ \text{s. t. } s_2 \geq 1, \\ s_1 \geq -1, \quad s_2 \geq 0. \end{aligned}$$

其初始表及计算过程如表 5.24 至表 5.26 所示.

表 5.24 子问题 0 的初始表

	e_1	e_2	e_3	
h_1	2	0	0	- 2
h_2	0	2	- 1 *	0
a_1	0	1 *	0	- 1

表 5.25 第一次双枢轴运算结果

	e_1	a_1	h_2	
h_1	2 *	0	0	- 1
e_3	0	2	- 1	2
e_2	0	1	0	1

表 5.26 第一次主旋转结果

	h_1	a_1	h_2	
e_1	0.5	0	0	1
e_3	0	2	- 1	2
e_2	0	1	0	1

由表 5.26 可见

$$s_1^{(0)} = 1 - 1 = 0, \quad s_2^{(0)} = 1 + 0 = 1.$$

所以步长为 1 时,

$$x_1^{(1)} = 0 + 0 = 0, \quad x_2^{(1)} = 0 + 1 = 1.$$

$x^{(1)} = (0 \ 1)^T$ 对应的子问题为

$$\begin{aligned} \min \quad & s_1^2 + s_2^2 + 2s_2, \\ \text{s. t.} \quad & s_2 \geq 0, \\ & s_1 \geq -1, s_2 \geq -1. \end{aligned}$$

其初始表及计算过程与表 5.24 至表 5.26 相同. 所以

$$s_1^{(1)} = 1 - 1 = 0, \quad s_2^{(1)} = 1 - 1 = 0.$$

由于搜索方向是零向量, 停止. $x^* = (0, 1)^T$ 是原问题的 KT 点, 对应的目标函数值 $f(x^*) = 1$. 但此解不是局部极小点, 例如,

$$x = (\Delta, 1 - \Delta^2)^T, \quad 0 < |\Delta| < 1,$$

是可行解, 其对应的目标函数值 $1 - \Delta^2 + \Delta^4 < 1$.

5.4 拉格朗日法

拉格朗日法以拉格朗日函数关于 x 的海色矩阵作为子问题目标函数的海色矩阵. 模型(5.1.5)的子问题为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} s^T \nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)}) s + \nabla f(x^{(k)})^T s, \\ \text{s. t.} \quad & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\ & \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m, \\ & s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.4.1)$$

其中 $s = x - x^{(k)}$, 拉格朗日函数的形式为

$$L(x, \lambda, \mu) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{i=1}^n \mu_i (x_i - l_i).$$

由于线性函数的二阶偏导数等于 0, 将拉格朗日函数的海色矩阵也记为 $\nabla_x^2 L(x, \lambda, \mu)$.

模型(5.4.1)的 KT 条件为

$$\begin{aligned} \sum_{j=1}^n \frac{\partial^2 L(x^{(k)}, \lambda^{(k)}, 0)}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} &\geq 0, \quad i = 1, 2, \dots, n, \\ \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) &= 0, \quad i = 1, 2, \dots, l, \\ \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) &\geq 0, \quad i = l+1, l+2, \dots, m, \\ s_i &\geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n, \\ \lambda_i &\geq 0, \quad i = l+1, l+2, \dots, m, \end{aligned}$$

$$\begin{aligned}
 (s_i - l_i + x_i^{(k)}) \left(\sum_{j=1}^n \frac{\partial^2 L(x^{(k)}, \lambda^{(k)}, 0)}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \right) &= 0, \\
 i &= 1, 2, \dots, n, \\
 \lambda_i \left(\sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \right) &= 0, \quad i = 1 + 1, 1 + 2, \dots, m.
 \end{aligned}
 \tag{5.4.2}$$

以上不等式组的解用 $(s^{(k)}, \lambda^{(k+1)}, \mu^{(k+1)})$ 表示. 在 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)}, 0)$ 中, $\lambda^{(0)} = 0$, $\lambda^{(k)}$ 是第 $k-1$ 个子问题 KT 条件的解, $k=1, 2, \dots$.

例 5.4.1 用拉格朗日法解例 5.3.3.

解 约束函数 $g_1(x)$ 的海色矩阵 $\nabla^2 g_1(x) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$. 对于 $x^{(0)} = (0, 0)^T$ 和 $\lambda_1^{(0)} = 0$,

$$\nabla_x^2 L(x^{(0)}, \lambda_1^{(0)}, 0) = \nabla^2 f(x^{(0)}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

相应的子问题及计算过程见例 5.3.3. 由表 5.26 可见

$$s_1^{(0)} = 1 - 1 = 0, \quad s_2^{(0)} = 1 + 0 = 1, \quad \lambda_1^{(1)} = 2.$$

令步长为 1 得 $x_1^{(1)} = 0, x_2^{(1)} = 1$.

对于 $x^{(1)} = (0, 1)^T$ 和 $\lambda_1^{(1)} = 2$,

$$\nabla_x^2 L(x^{(1)}, \lambda_1^{(1)}, 0) = \nabla^2 f(x^{(1)}) - \lambda_1^{(1)} \nabla^2 g_1(x^{(1)}) = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix},$$

$$\nabla f(x^{(1)}) = (0, 2)^T, \quad \nabla g_1(x^{(1)}) = (0, 1)^T, \quad g_1(x^{(1)}) = 0.$$

所以相应的子问题为

$$\min -s_1^2 + s_2^2 + 2s_2,$$

$$\text{s. t.} \quad s_2 \geq 0,$$

$$s_1 \geq -1, s_2 \geq -1.$$

初始表及计算过程如表 5.27 至表 5.28 所示.

表 5.27 子问题 0 的初始表

	e_1	e_2	e_3	
h_1	-2	0	0	2
h_2	0	2	-1*	0
a_1	0	1*	0	-1

表 5.28 第一次双枢轴运算结果

	e_1	a_1	h_2	
h_1	-2	0	0	2
e_3	0	2	-1	2
e_2	0	1	0	1

$$\min 10s_1^2 + 10s_2^2 + 15s_1 + 12s_2 ,$$
$$\text{s. t. } 1.6s_1 + 1.2s_2 \geq 0 ,$$
$$s_1 \geq -0.8 , \quad s_2 \geq -0.6 .$$

初始表见表 5.30. 作双枢轴运算 $a_1 \leftrightarrow e_1$ $h_1 \leftrightarrow e_3$ 的结果见表 5.31.

表 5.30 子问题 0 的初始表

	e_1	e_2	e_3	
h_1	20	0	- 1.6 *	- 1
h_2	0	20	- 1	0
a_1	1.6 *	1.2	0	- 2

表 5.31 双枢轴运算结果

	a_1	e_2	h_1	
e_3	7.8125	- 9.375	- 0.625	15
h_2	- 9.375	31.25 *	0.75	- 18
e_1	0.625	- 0.75	0	1.25

在表 5.31 中作主旋转 $h_2 \leftrightarrow e_2$ 的结果见表 5.32.

表 5.32 主旋转结果

	a_1	h_2	h_1	
e_3	5	- 0.3	- 0.4	9.6
e_2	0.3	0.032	- 0.024	0.576
e_1	0.4	- 0.024	0.018	0.818

由表 5.32 可见

$s_1^{(0)} = 0.818 - 0.8 = 0.018 , \quad s_2^{(0)} = 0.576 - 0.6 = -0.024 , \quad \lambda_1^{(1)} = 9.6 .$

令步长为 1 得

$x_1^{(1)} = 0.8 + 0.018 = 0.818 , \quad x_2^{(1)} = 0.6 - 0.024 = 0.576 .$

其他子问题的计算过程类似 ,全部结果见表 5.33 ,其中 $x^{(k+1)} = x^{(k)} + s^{(k)} .$

表 5.33 子问题的计算结果

子问题 k	初始解($x_1^{(k)}$ $x_2^{(k)}$ $\lambda^{(k)}$)	解($s_1^{(k)}$ $s_2^{(k)}$ $\lambda^{(k+1)}$)
0	(0.8 , 0.6 , 0)	(1.799995×10^{-2} , - 2.399999×10^{-2} , 9.6)
1	(0.818 , 0.576 , 9.6)	(0.4050442 , - 0.576 , 9.586819)
2	(1.223044 , 0 , 9.586819)	(- 0.2027061 , 0 , 9.522705)
3	(1.020338 , 0 , 9.522705)	(- 0.0201354 , 0 , 9.500546)
4	(1.000203 , 0 , 9.500546)	(- 2.026558×10^{-4} , 0 , 9.5)
5	(1 , 0 , 9.5)	(0 , 0 , 9.5)
6	(1 , 0 , 9.5)	

容易验证 $x^{(6)} = (1 \ 0)^T$ 是原问题的全局极小点 相应的目标值为 - 1. 此题共进行了 13 次旋转运算.

例 5.4.3 用拉格朗日法解非线性规划

$$\min f(x) = x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4,$$

$$\text{s. t. } g_1(x) = -x_1^2 - x_2^2 - x_3^2 - x_4^2 - x_1 + x_2 - x_3 + x_4 + 8 \geq 0,$$

$$g_2(x) = -x_1^2 - 2x_2^2 - x_3^2 - 2x_4^2 + x_1 + x_4 + 10 \geq 0,$$

$$g_3(x) = -2x_1^2 - x_2^2 - x_3^2 - 2x_1 + x_2 + x_4 + 5 \geq 0,$$

$$x_1, x_2, x_3, x_4 \geq -1.$$

初始点为 $x^{(0)} = (0 \ 0 \ 0 \ 0)^T$.

解 各函数的梯度和海色矩阵如下：

$$\nabla f(x) = \begin{pmatrix} 2x_1 - 5 \\ 2x_2 - 5 \\ 4x_3 - 21 \\ 2x_4 + 7 \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} -2x_1 - 1 \\ -2x_2 + 1 \\ -2x_3 - 1 \\ -2x_4 + 1 \end{pmatrix}, \quad \nabla g_2(x) = \begin{pmatrix} -2x_1 + 1 \\ -4x_2 \\ -2x_3 \\ -4x_4 + 1 \end{pmatrix},$$

$$\nabla g_3(x) = \begin{pmatrix} -4x_1 - 2 \\ -2x_2 + 1 \\ -2x_3 \\ 1 \end{pmatrix},$$

$$\nabla^2 f(x) = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 4 & \\ & & & 2 \end{pmatrix}, \quad \nabla^2 g_1(x) = \begin{pmatrix} -2 & & & \\ & -2 & & \\ & & -2 & \\ & & & -2 \end{pmatrix},$$

$$\nabla^2 g_2(x) = \begin{pmatrix} -2 & & & \\ & -4 & & \\ & & -2 & \\ & & & -4 \end{pmatrix}, \quad \nabla^2 g_3(x) = \begin{pmatrix} -4 & & & \\ & -2 & & \\ & & -2 & \\ & & & 0 \end{pmatrix}.$$

初始点 $x^{(0)} = (0 \ 0 \ 0 \ 0)^T$ 和 $\lambda^{(0)} = (0 \ 0 \ 0)^T$ 时的子问题为

$$\min \frac{2s_1^2 + 2s_2^2 + 4s_3^2 + 2s_4^2}{2} - 5s_1 - 5s_2 - 21s_3 + 7s_4$$

s. t. $-s_1 + s_2 - s_3 + s_4 \geq -8$,
 $s_1 + s_4 \geq -10$,
 $-2s_1 + s_2 + s_4 \geq -5$,
 $s_i \geq -1$, $i=1, 2, 3, 4$.

这是一个凸二次规划 通过其 KT 条件可求得

$s^{(0)} = (2.5, 2.5, 5.25, -1)^T$,

前三个约束对应的拉格朗日乘子 $\lambda_1 = \lambda_2 = \lambda_3 = 0$. 令步长等于 1 得

$x^{(1)} = x^{(0)} + s^{(0)} = (2.5, 2.5, 5.25, -1)^T$.

同样可求得其他子问题的解 如表 5.34 所示 其中 $x^{(k+1)} = x^{(k)} + s^{(k)}$.

表 5.34 子问题的计算结果

子问题 k	子问题的初始点 $x^{(k)}$ $\lambda^{(k)}$ 及计算结果 $s^{(k)}$ $\lambda^{(k+1)}$
0	$x^{(0)} = (0, 0, 0, 0)$, $\lambda^{(0)} = (0, 0, 0)$, $s^{(0)} = (2.5, 2.5, 5.25, -1)$, $\lambda^{(1)} = (0, 0, 0)$
1	$x^{(1)} = (2.5, 2.5, 5.25, -1)$, $s^{(1)} = (-1.996826, -1.331217, -1.913624, 0)$, $\lambda^{(2)} = (0.6656084, 0, 0)$
2	$x^{(2)} = (0.5031744, 1.168783, 3.336376, -1)$, $s^{(2)} = (-0.289081, 0.0441637, -1.160012, 0)$, $\lambda^{(3)} = (1.290554, 0, 0.5899626)$
3	$x^{(3)} = (0.2140934, 1.212947, 2.176364, -1)$, $s^{(3)} = (-0.2155174, -0.2333668, -0.1179011, 0.053167)$, $\lambda^{(4)} = (1.247717, 0, 1.500412)$
4	$x^{(4)} = (-1.423955 \times 10^{-3}, 0.97958, 2.058463, -0.946833)$, $s^{(4)} = (-1.439452 \times 10^{-4}, 2.226841 \times 10^{-2}, -5.720139 \times 10^{-2}, -5.316699 \times 10^{-2})$, $\lambda^{(5)} = (0.973902, 0, 2.022375)$
5	$x^{(5)} = (-0.0015679, 1.001848, 2.001262, -1)$, $s^{(5)} = (1.567245 \times 10^{-3}, -1.847982 \times 10^{-3}, -1.257896 \times 10^{-3}, 4.072179 \times 10^{-6})$, $\lambda^{(6)} = (1.000016, 0, 1.999968)$
6	$x^{(6)} = (-6.556511 \times 10^{-7}, 1, 2.000004, -0.9999959)$, $s^{(6)} = (1.370907 \times 10^{-6}, -2.384186 \times 10^{-7}, -3.099442 \times 10^{-6}, -3.316117 \times 10^{-6})$, $\lambda^{(7)} = (1.000002, 0, 1.999995)$
7	$x^{(7)} = (7.152557 \times 10^{-7}, 1, 2.000001, -0.9999993)$

其中 $\|s^{(6)}\| < 10^{-4}$ 按照我们的误差要求 计算停止. 由于原问题是凸规划 ,

$$x^* = (7.152557 \times 10^{-7}, 1, 2.000001, -0.9999993)^T$$

是全局极小点,对应的目标函数值

$$f(x^*) = -44.$$

此题共进行了 51 次旋转运算.

5.5 增广拉格朗日法

增广拉格朗日法以增广拉格朗日函数的海色矩阵作为子问题目标函数的海色矩阵. 此处的增广拉格朗日函数是在拉格朗日函数中引入某些约束函数的平方得到的函数, 其形式为

$$P(x, \lambda, \mu) = f(x) - \sum_{i=1}^m \lambda_i g_i(x) - \sum_{i=1}^n \mu_i (x_i - l_i) + r \sum_{i \in I^*} (g_i(x))^2. \quad (5.5.1)$$

其中 r 是一个正数, I^* 是某些紧约束的指标集. 由于线性函数的二阶偏导数等于 0, 线性函数不出现在其海色矩阵 $\nabla_x^2 P(x, \lambda, \mu)$ 中. 根据笔者的计算经验, 对于某些具有等式约束的非线性规划, 不一定将所有等式约束引入拉格朗日函数, 只需引入一些变量较多 (系数不等于 0) 的约束函数. 令 $G_i(x) = (g_i(x))^2$, $\nabla^2 G_i(x)$ 的第 i 行第 j 列的元素为

$$\frac{\partial^2 G_i(x)}{\partial x_i \partial x_j} = 2 \left(\frac{\partial g_i(x)}{\partial x_i} \cdot \frac{\partial g_i(x)}{\partial x_j} + g_i(x) \frac{\partial^2 g_i(x)}{\partial x_i \partial x_j} \right), \quad i, j = 1, 2, \dots, n.$$

模型 (5.1.5) 的子问题为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} s^T \nabla_x^2 P(x^{(k)}, \lambda^{(k)}, \mu^{(k)}) s + \nabla f(x^{(k)})^T s, \\ \text{s. t.} \quad & \nabla g_i(x^{(k)})^T s = -g_i(x^{(k)}), \quad i = 1, 2, \dots, l, \\ & \nabla g_i(x^{(k)})^T s \geq -g_i(x^{(k)}), \quad i = l+1, l+2, \dots, m, \\ & s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.5.2)$$

其中 $s = x - x^{(k)}$, 其 KT 条件为

$$\sum_{j=1}^n \frac{\partial^2 P(x^{(k)}, \lambda^{(k)}, \mu)}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \geq 0, \quad i = 1, 2, \dots, n,$$

$$\begin{aligned}
& \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, l, \\
& \sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m, \\
& s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n, \\
& \lambda_i \geq 0, \quad i = l+1, l+2, \dots, m, \\
& (s_i - l_i + x_i^{(k)}) \left(\sum_{j=1}^n \frac{\partial^2 P(x^{(k)}, \lambda^{(k)}, \mu)}{\partial x_i \partial x_j} s_j - \sum_{j=1}^m \frac{\partial g_j(x^{(k)})}{\partial x_i} \lambda_j + \frac{\partial f(x^{(k)})}{\partial x_i} \right) = 0, \\
& \quad i = 1, 2, \dots, n, \\
& \lambda_i \left(\sum_{j=1}^n \frac{\partial g_i(x^{(k)})}{\partial x_j} s_j + g_i(x^{(k)}) \right) = 0, \quad i = l+1, l+2, \dots, m. \quad (5.5.3)
\end{aligned}$$

以上不等式组的解用 $(s^{(k)}, \lambda^{(k+1)}, \mu^{(k+1)})$ 表示. 在 $\nabla_x^2 P(x^{(k)}, \lambda^{(k)}, 0)$ 中 $\lambda^{(0)} = 0$, $\lambda^{(k)}$ 是第 $k-1$ 个子问题 KT 条件的解 $k=1, 2, \dots$.

例 5.5.1 解非线性规划

$$\begin{aligned}
& \min f(x) = \exp(x_1 x_2 x_3 x_4 x_5), \\
& \text{s. t. } g_1(x) = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 - 10 = 0, \\
& \quad g_2(x) = x_2 x_3 - 5 x_4 x_5 = 0, \\
& \quad g_3(x) = x_1^3 + x_2^3 + 1 = 0, \\
& \quad x_1, x_2 \geq -2.3, x_3, x_4, x_5 \geq -3.2.
\end{aligned}$$

初始点为 $x^{(0)} = (-2, 2, 2, -1, -1)^T$.

解 各函数的梯度和海色矩阵如下, 其中 $y = x_1 x_2 x_3 x_4 x_5$.

$$\begin{aligned}
\nabla f(x) &= \begin{pmatrix} x_2 x_3 x_4 x_5 e^y \\ x_1 x_3 x_4 x_5 e^y \\ x_1 x_2 x_4 x_5 e^y \\ x_1 x_2 x_3 x_5 e^y \\ x_1 x_2 x_3 x_4 e^y \end{pmatrix}, \quad \nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \\ 2x_4 \\ 2x_5 \end{pmatrix}, \\
\nabla g_2(x) &= \begin{pmatrix} 0 \\ x_3 \\ x_2 \\ -5x_5 \\ -5x_4 \end{pmatrix}, \quad \nabla g_3(x) = \begin{pmatrix} 3x_1^2 \\ 3x_2^2 \\ 0 \\ 0 \\ 0 \end{pmatrix},
\end{aligned}$$

$\nabla^2 f(x)$ 的第 i 个对角元素为

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \prod_{k=1, k \neq i}^5 x_k^2 e^{y_i} \quad i = 1, 2, 3, 4, 5;$$

第 i 行第 j 列的非对角元素为

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \prod_{k=1, k \neq i, k \neq j}^5 x_k (y_i + 1) e^{y_j}, \quad i, j = 1, 2, 3, 4, 5, i \neq j,$$

$$\nabla^2 g_1(x) = \begin{pmatrix} 2 & & & & \\ & 2 & & & \\ & & 2 & & \\ & & & 2 & \\ & & & & 2 \end{pmatrix}, \quad \nabla^2 g_2(x) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -5 \\ 0 & 0 & 0 & -5 & 0 \end{pmatrix},$$

$$\nabla^2 g_3(x) = \begin{pmatrix} 6 & & & & \\ & 6 & & & \\ & & 0 & & \\ & & & 0 & \\ & & & & 0 \end{pmatrix},$$

此题仅将 $(g_1(x))^2$ 引入拉格朗日函数, 即 (5.5.1) 中的 $I^* = \{1\}$, 并令 $r=1$. 初始点 $x^{(0)} = (-2, 2, 2, -1, -1)^T$ 对应的子问题为

$$\min \frac{1}{2} (s_1 \ s_2 \ s_3 \ s_4) \begin{pmatrix} 48.0054 & -32.0047 & -32.0047 & 16.0094 & 16.0094 \\ -32.0047 & 48.0054 & 32.0047 & -16.0094 & -16.0094 \\ -32.0047 & 32.0047 & 48.0054 & -16.0094 & -16.0094 \\ 16.0094 & -16.0094 & -16.0094 & 24.0215 & 8.0188 \\ 16.0094 & -16.0094 & -16.0094 & 8.0188 & 24.0215 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{pmatrix}$$

$$+ 1.34185 \times 10^{-3} (s_1 - s_2 - s_3) + 2.683701 \times 10^{-3} (s_4 + s_5)$$

$$\text{s. t. } -4s_1 + 4s_2 + 4s_3 - 2s_4 - 2s_5 + 4 = 0,$$

$$2s_2 + 2s_3 + 5s_4 + 5s_5 - 1 = 0,$$

$$12s_1 + 12s_2 + 1 = 0,$$

$$s_1 \geq -0.3, s_2 \geq -4.3, s_3 \geq -5.2, s_4 \geq -2.2, s_5 \geq -2.2.$$

其解为

$$s^{(0)} = (0.2256943, -0.3090272, -0.2528944, 0.2123842, 0.2123842)^T,$$

$$\lambda^{(1)} = (-9.129627, 0.2315068, -0.07485056)^T.$$

同样可求得其他子问题的解, 如表 5.35 所示, 其中 $x^{(k+1)} = x^{(k)} + s^{(k)}$.

表 5.35

子问题的计算结果

子问题 k	子问题的初始点 $x^{(k)}$ $\lambda^{(k)}$ 及计算结果 $s^{(k)}$ $\lambda^{(k+1)}$
0	$x^{(0)} = (-2.2, -1, -1), \lambda^{(0)} = (0, 0, 0),$ $s^{(0)} = (0.2256943, -0.3090272, -0.2528944, 0.2123842, 0.2123842),$ $\lambda^{(1)} = (-9.129627, 0.2315068, -0.07485056)$
1	$x^{(1)} = (-1.774306, 1.690973, 1.747106, -0.7876158, -0.7876158),$ $s^{(1)} = (6.212771 \times 10^{-3}, -3.590822 \times 10^{-2}, -1.752138 \times 10^{-2}, 0.0304389,$ $3.043962 \times 10^{-2}),$ $\lambda^{(2)} = (-0.7886386, 0.1367941, -4.789114 \times 10^{-2})$
2	$x^{(2)} = (-1.768093, 1.655065, 1.729584, -0.7571769, -0.7571762),$ $s^{(2)} = (4.331827 \times 10^{-3}, -5.709171 \times 10^{-3}, 8.805275 \times 10^{-3}, -9.36985 \times$ $10^{-5}, -8.845329 \times 10^{-5}),$ $\lambda^{(3)} = (-4.798394 \times 10^{-2}, 3.919566 \times 10^{-2}, -4.951268 \times 10^{-3})$
3	$x^{(3)} = (-1.763761, 1.649355, 1.738389, -0.7572706, -0.7572646),$ $s^{(3)} = (0.0388326, -4.441404 \times 10^{-2}, 0.0759716, -6.361723 \times 10^{-3},$ $-6.327868 \times 10^{-3}),$ $\lambda^{(4)} = (-4.012999 \times 10^{-2}, 3.753117 \times 10^{-2}, -4.847921 \times 10^{-3})$
4	$x^{(4)} = (-1.724928, 1.604941, 1.814361, -0.7636323, -0.7635925),$ $s^{(4)} = (6.150544 \times 10^{-3}, -7.331848 \times 10^{-3}, 9.876251 \times 10^{-3}, 1.504421 \times 10^{-4},$ $1.189709 \times 10^{-4}),$ $\lambda^{(5)} = (-5.883758 \times 10^{-2}, 3.796491 \times 10^{-2}, -5.16513 \times 10^{-3})$
5	$x^{(5)} = (-1.718778, 1.59761, 1.824237, -0.7634819, -0.7634735),$ $s^{(5)} = (1.275599 \times 10^{-3}, -1.483917 \times 10^{-3}, 2.342224 \times 10^{-3}, -1.111031 \times$ $10^{-4}, -1.416206 \times 10^{-4}),$ $\lambda^{(6)} = (-0.0405543, 3.795841 \times 10^{-2}, -5.210191 \times 10^{-3}),$
6	$x^{(6)} = (-1.717502, 1.596126, 1.82658, -0.763593, -0.7636151),$ $s^{(6)} = (3.212094 \times 10^{-4}, -3.728867 \times 10^{-4}, 5.970001 \times 10^{-4}, -3.576279 \times$ $10^{-5}, -3.480911 \times 10^{-5}),$ $\lambda^{(7)} = (-4.017697 \times 10^{-2}, 3.795754 \times 10^{-2}, -5.220445 \times 10^{-3})$

续表

子问题 k	子问题的初始点 $x^{(k)}$ $\lambda^{(k)}$ 及计算结果 $s^{(k)}$ $\lambda^{(k+1)}$
7	$x^{(7)} = (-1.717181, 1.595753, 1.827177, -0.7636287, -0.7636499),$ $s^{(7)} = (-4.410744 \times 10^{-6}, 5.722046 \times 10^{-6}, -8.583069 \times 10^{-6}, -6.008148 \times 10^{-5}, 6.198883 \times 10^{-5}),$ $\lambda^{(8)} = (-4.016424 \times 10^{-2}, 3.795733 \times 10^{-2}, -5.219713 \times 10^{-3})$
8	$x^{(8)} = (-1.717186, 1.595758, 1.827168, -0.7636888, -0.763588)$

其中 $\|s^{(7)}\| < 10^{-4}$ 最终解为

$x^* = (-1.717186, 1.595758, 1.827168, -0.7636888, -0.763588)^T.$

相应的目标函数值为 $f(x^*) = 0.0539498217$. 此题共进行了 64 次旋转运算.

原问题关于 x^* 的紧约束仅是三个等式约束 因此检查方程组

$\nabla g_1(x^*)^T s = 0, \quad \nabla g_2(x^*)^T s = 0, \quad \nabla g_3(x^*)^T s = 0$

的所有非零解是否使

$s^T \nabla_x^2 L(x^*, \lambda^*, 0)^T s > 0.$

与例 5.3.1 类似,引入人工不等式 $s_1, s_2, s_3, s_4 \geq 0$, 列出初始表, 并将三个等式换为基约束后, 简约拉格朗日矩阵 $L = (0.7597)$ 正定. 根据定理 4.3.1 和定理 1.5.2 x^* 是原问题的严格局部极小点.

例 5.5.2 解非线性规划

$$\begin{aligned} \min f(x) &= x_1 x_4 (x_1 + x_2 + x_3) + x_3, \\ \text{s. t. } g_1(x) &= x_1^2 + x_2^2 + x_3^2 + x_4^2 - 40 = 0, \\ g_2(x) &= x_1 x_2 x_3 x_4 - 25 \geq 0, \\ x_1, x_2, x_3, x_4 &\geq 0. \end{aligned}$$

初始点为 $x^{(0)} = (1, 5, 5, 1)^T.$

解 各函数的梯度和海色矩阵如下：

$$\nabla f(x) = \begin{pmatrix} x_4(2x_1 + x_2 + x_3) \\ x_1 x_4 \\ x_1 x_4 + 1 \\ x_1(x_1 + x_2 + x_3) \end{pmatrix}, \quad \nabla^2 f(x) = \begin{pmatrix} 2x_4 & x_4 & x_4 & 2x_1 + x_2 + x_3 \\ x_4 & 0 & 0 & x_1 \\ x_4 & 0 & 0 & x_1 \\ 2x_1 + x_2 + x_3 & x_1 & x_1 & 0 \end{pmatrix},$$

$$\nabla g_1(x) = \begin{pmatrix} 2x_1 \\ 2x_2 \\ 2x_3 \\ 2x_4 \end{pmatrix}, \quad \nabla^2 g_1(x) = \begin{pmatrix} 2 & & & \\ & 2 & & \\ & & 2 & \\ & & & 2 \end{pmatrix},$$

$$\nabla g_2(x) = \begin{pmatrix} x_2 x_3 x_4 \\ x_1 x_3 x_4 \\ x_1 x_2 x_4 \\ x_1 x_2 x_3 \end{pmatrix}, \quad \nabla^2 g_2(x) = \begin{pmatrix} 0 & x_3 x_4 & x_2 x_4 & x_2 x_3 \\ x_3 x_4 & 0 & x_1 x_4 & x_1 x_3 \\ x_2 x_4 & x_1 x_4 & 0 & x_1 x_2 \\ x_2 x_3 & x_1 x_3 & x_1 x_2 & 0 \end{pmatrix}.$$

此题仅将 $(g_1(x))^2$ 引入拉格朗日函数, 即(5.5.1)式中的 $I^* = 1$, 并令 $r =$

1. 初始点 $x^{(0)} = (1 \ 5 \ 5 \ 1)^T$ 时的子问题为

$$\min \frac{1}{2} (s_1 \ s_2 \ s_3 \ s_4) \begin{pmatrix} 58 & 41 & 41 & 20 \\ 41 & 248 & 200 & 41 \\ 41 & 200 & 248 & 41 \\ 20 & 41 & 41 & 56 \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix} + (12 \ 1 \ 2 \ 11) \begin{pmatrix} s_1 \\ s_2 \\ s_3 \\ s_4 \end{pmatrix},$$

$$\text{s. t. } 2s_1 + 10s_2 + 10s_3 + 2s_4 = -12,$$

$$25s_1 + 5s_2 + 5s_3 + 25s_4 \geq 0,$$

$$s_1 \geq -1, \quad s_2 \geq -5, \quad s_3 \geq -5, \quad s_4 \geq -1.$$

其解为

$$s^{(0)} = \begin{pmatrix} 0.108108 \\ -0.6145844 \\ -0.6354156 \\ 0.141892 \end{pmatrix}, \quad \lambda^{(1)} = \begin{pmatrix} -27.31457 \\ 0.9795051 \end{pmatrix}.$$

下面均按步长等于 1 确定下一个子问题的初始点, 因而

$$x^{(1)} = x^{(0)} + s^{(0)} = \begin{pmatrix} 1.108108 \\ 4.385416 \\ 4.364584 \\ 1.141892 \end{pmatrix}.$$

类似可求得其他子问题的解, 部分实验结果如表 5.36 所示, 其中 $x^{(k+1)} = x^{(k)} + s^{(k)}$.

表 5.36 例 5.5.2 的计算结果

k	子问题的初始点 $x^{(k)}$, $\lambda^{(k)}$ 及计算结果 $s^{(k)}$, $\lambda^{(k+1)}$
0	$x^{(0)} = (1.551)$, $\lambda^{(0)} = (0.0)$, $s^{(0)} = (0.108108, -0.6145844, -0.6354156, 0.141892)$, $\lambda^{(1)} = (-27.31475, 0.9795051)$
1	$x^{(1)} = (1.108108, 4.385416, 4.364584, 1.141892)$, $s^{(1)} = (0.02349126, -0.0467844, -0.06291676, 0.04124546)$, $\lambda^{(2)} = (-2.253339, 0.7053058)$
2	$x^{(2)} = (1.131599, 4.338631, 4.301668, 1.183137)$, $s^{(2)} = (-0.07769799, 0.0853529, -0.08950281, 0.08318424)$, $\lambda^{(3)} = (-0.1865068, 0.5693744)$
⋮	⋮
9	$x^{(9)} = (0.7544838, 4.639303, 3.788626, 1.885193)$, $s^{(9)} = (3.248453 \times 10^{-5}, 3.480911 \times 10^{-5}, -4.053116 \times 10^{-6}, -9.334087 \times 10^{-5})$, $\lambda^{(10)} = (-0.1798344, 0.573539)$
10	$x^{(10)} = (0.7545162, 4.639338, 3.788622, 1.8851)$, $s^{(10)} = (2.920628 \times 10^{-6}, 4.863739 \times 10^{-5}, -6.270409 \times 10^{-5}, 3.695488 \times 10^{-6})$, $\lambda^{(11)} = (-0.1798049, 0.5735412)$
11	$x^{(11)} = (0.7545192, 4.639386, 3.788559, 1.885103)$

由于 $\|s^{(10)}\| < 10^{-4}$,最终解为

$$x^* = (0.7545192, 4.639386, 3.788559, 1.885103) ,$$

相应的目标函数值 $f(x^*) = 16.84921$. 此题共进行了 66 次旋转运算.

原问题的两个一般约束都是关于 x^* 的紧约束 ,其中不等式约束对应的拉格朗日乘子 $\lambda_2 = 0.5735412 > 0$,所以检查方程组

$$\nabla g_1(x^*)^T s = 0 , \quad \nabla g_2(x^*)^T s = 0$$

的所有非零解是否使

$$s^T \nabla_x^2 L(x^*, \lambda^*, 0)^T s > 0.$$

与例 5.3.1 类似 ,引入人工不等式 $s_1 \ s_2 \ s_3 \ s_4 \geq 0$,列出初始表 ,并将两个等式换为基约束后 ,可以看到简约拉格朗日矩阵

$$L = \begin{pmatrix} 2.1104 & 0.3641 \\ 0.3641 & 1.0775 \end{pmatrix}.$$

这是一个正定矩阵, 所以 x^* 是原问题的严格局部极小点.

如果将此题所有变量的下界改为 1, 最后一个子问题的初始点及计算结果

$$\lambda^{(7)} = (-0.1620555, 0.5522918),$$

$$x^{(7)} = (1, 4.743034, 3.821128, 1.3794),$$

$$s^{(7)} = (0, -3.767014 \times 10^{-5}, 2.598763 \times 10^{-5}, 7.748604 \times 10^{-6}),$$

$$\lambda^{(8)} = (-0.1617455, 0.5522937),$$

$$x^{(8)} = x^{(7)} + s^{(7)} = (1, 4.742997, 3.821154, 1.379408).$$

目标函数值 $f(x^{(8)}) = 17.01402$. 此题共进行了 53 次旋转运算.

5.6 拟牛顿法

拟牛顿法是近几十年来研究的最多的一类方法. 在拟牛顿法中, 子问题目标函数的海色矩阵是一个包含 $\nabla^2 f(x^{(k)})$ 或 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)})$ 信息的矩阵, 并且在构造时尽量使其正定. 所以一般来说, 二次规划子问题是凸二次规划.

拟牛顿法的重点在于确定子问题目标函数的海色矩阵和罚函数, 其形式可以说形形色色. 搜索方向由子问题确定, 罚函数则综合考虑目标函数的下降程度和迭代点到可行域的距离以确定步长. 但近期 Fletcher 等资深学者也提出不用罚函数的方法.

5.6.1 无约束极小值问题的牛顿法和拟牛顿法

首先考虑无约束极值问题

$$\min f(x). \quad (5.6.1)$$

将 $f(x)$ 在 $x^{(k)}$ 处作泰勒展开, 并略去高于二次的项, 得

$$f(x) \approx Q(x) = f(x^{(k)}) + \nabla f(x^{(k)})^T (x - x^{(k)}) + \frac{1}{2} (x - x^{(k)})^T \nabla^2 f(x^{(k)}) (x - x^{(k)}). \quad (5.6.2)$$

牛顿法通过求解线性方程组

$$\nabla^2 f(x^{(k)}) (x - x^{(k)}) + \nabla f(x^{(k)}) = 0 \quad (5.6.3)$$

获取(5.6.1)的解. (5.6.3)即是子问题 $\min Q(x)$ 的 KT 条件 $\nabla Q(x) = 0$.

现在假设在 $x^{(k)}$ 的某个邻域内 $\nabla^2 f(x)$ 正定, 于是

$$x = x^{(k)} - \nabla^2 f(x^{(k)})^{-1} \nabla f(x^{(k)}). \quad (5.6.4)$$

当由(5.6.4)确定的 x 等于或充分接近于 $x^{(k)}$ 时, 将其记为 x^* . 这时 x^* 满足

(5.6.1)的KT条件 $\nabla f(x) = 0$. 由于局部极小点的二阶充分条件也成立, x^* 是 $f(x)$ 的局部极小点. 由(5.6.4)确定的 x 与 $x^{(k)}$ 相距较远时, 就将 x 视作 $x^{(k)}$ 并开始下一轮迭代, 这就是牛顿法(步长为1)的实施过程.

由于二次函数 $Q(x)$ 仅仅是原函数 $f(x)$ 的近似, 由(5.6.4)确定的 x 距 $f(x)$ 的局部极小点可能相差较远. 所以常按下式确定下一个迭代点,

$$x^{(k+1)} = x^{(k)} + \alpha_k (x - x^{(k)}), \quad 0 < \alpha_k \leq 1.$$

以上假设 $\nabla^2 f(x^{(k)})$ 正定, 这不仅可以获得一个显式解, 而且可以使子问题的最终解 x^* 不至于是 $f(x)$ 的极大点. 但是 $\nabla^2 f(x^{(k)})$ 可能不是正定的, 而且计算二阶偏导数工作量较大. 所以常用一个接近于 $\nabla^2 f(x^{(k)})$ 的正定矩阵 B_k 代替 $\nabla^2 f(x^{(k)})$ 进行计算, 并在迭代过程中不断更新. 这种方法称为拟牛顿法或变尺度法. 第 $k+1$ 次迭代使用的 B_{k+1} 与第 k 次迭代使用的 B_k 最好用一个公式联系起来. 在(5.6.2)两边求梯度, 得

$$\nabla f(x) \approx \nabla f(x^{(k)}) + \nabla^2 f(x^{(k)})(x - x^{(k)}). \quad (5.6.5)$$

若 $f(x)$ 是二次函数, (5.6.5)式为严格等式, 可看做是二次函数的“特征”方程. 记 $\delta^{(k)} = x - x^{(k)}$, $\gamma^{(k)} = \nabla f(x) - \nabla f(x^{(k)})$, 将 $\nabla^2 f(x^{(k)})$ 换为 B_{k+1} , 然后令两边相等得

$$B_{k+1} \delta^{(k)} = \gamma^{(k)}. \quad (5.6.6)$$

(5.6.6)称为拟牛顿方程. 两边减去 $B_k \delta^{(k)}$, 并令 $C_k = B_{k+1} - B_k$ 得

$$C_k \delta^{(k)} = -B_k \delta^{(k)} + \gamma^{(k)}. \quad (5.6.7)$$

满足(5.6.6)的 C_k 很多, 例如对于适合 $(q^{(k)})^T \delta^{(k)} = (w^{(k)})^T \delta^{(k)} = 1$ 的任何向量 $q^{(k)}$ 和 $w^{(k)}$, 令

$$C_k = -B_k \delta^{(k)} (q^{(k)})^T + \gamma^{(k)} (w^{(k)})^T \quad (5.6.8)$$

便可. 特别是

$$C_k = -\frac{B_k \delta^{(k)} (\delta^{(k)})^T B_k}{(\delta^{(k)})^T B_k \delta^{(k)}} + \frac{\gamma^{(k)} (\gamma^{(k)})^T}{(\delta^{(k)})^T \gamma^{(k)}}.$$

将 $C_k = B_{k+1} - B_k$ 代入上式, 移项得

$$B_{k+1} = B_k - \frac{B_k \delta^{(k)} (\delta^{(k)})^T B_k}{(\delta^{(k)})^T B_k \delta^{(k)}} + \frac{\gamma^{(k)} (\gamma^{(k)})^T}{(\delta^{(k)})^T \gamma^{(k)}}. \quad (5.6.9)$$

(5.6.9)就是著名的 BFGS (Broyden-Fletcher-Goldfarb-Shanno) 公式. 还有一些类似公式, 在此不一一列举.

以上考虑问题的思路主要参考赵瑞安和吴方的著述(1992).

可以证明, 如果 B_k 正定, 那么由 BFGS 公式得到的 B_{k+1} 正定的必要充分条件是 $(\delta^{(k)})^T \gamma^{(k)} > 0$. 此式容易满足. 因为 $x - x^{(k)}$ 应是 $x^{(k)}$ 处的下降方向, 因而

$\nabla f(x^{(k)})^T (x - x^{(k)}) < 0$;而且一般说来可以找到 $x = x^{(k+1)}$ 使 $\nabla f(x^{(k+1)})^T (x^{(k+1)} - x^{(k)}) = 0$ 或

$$\nabla f(x^{(k+1)})^T (x^{(k+1)} - x^{(k)}) > \nabla f(x^{(k)})^T (x^{(k+1)} - x^{(k)}).$$

此式便是 $(\delta^{(k)})^T \gamma^{(k)} > 0$.

5.6.2 约束极小值问题的拟牛顿法

现在考虑约束极小值问题

$$\left. \begin{aligned} \min f(x), \\ \text{s. t. } g_i(x) = 0, \quad i = 1, 2, \dots, l, \\ g_i(x) \geq 0, \quad i = l+1, l+2, \dots, m, \\ x_i \geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.6.10)$$

序列二次规划法是求解约束极值问题最有效的方法. 以拉格朗日函数的海色矩阵作为子问题目标函数的海色矩阵, 子问题的模型为

$$\left. \begin{aligned} \min \frac{1}{2} s^T \nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)}) s + \nabla f(x^{(k)})^T s, \\ \text{s. t. } \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) = 0 \quad i = 1, 2, \dots, l, \\ \nabla g_i(x^{(k)})^T s + g_i(x^{(k)}) \geq 0, \quad i = l+1, l+2, \dots, m, \\ s_i \geq l_i - x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (5.6.11)$$

其中 $s = x - x^{(k)}$. 5.4 节便是通过求解此子问题获取(5.6.10)的解, 并将其称为拉格朗日法.

目标函数中的 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$ 可能不是正定的而且计算复杂, 与无约束极值问题的拟牛顿法一样, 目前流行的做法是将 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$ 换为一个与之接近的正定矩阵 B_k 进行计算.

子问题(5.6.11)KT 条件的第一式为

$$\nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)}) (x - x^{(k)}) + \nabla_x L(x^{(k)}, \lambda, \mu) = 0. \quad (5.6.12)$$

将上式与(5.6.3)对照, 约束极值问题中 BFGS 公式的 $\gamma^{(k)}$ 应该是 $\nabla_x L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$ 的改变量而不再是 $\nabla f(x^{(k)})$ 的改变量.

但是, (5.6.12)与(5.6.3)有个重要区别, 就是 $\nabla_x L(x^{(k)}, \lambda, \mu)$ 是未知的, λ 和 μ 还需满足 KT 条件的其他式子. 这使 $(\delta^{(k)})^T \gamma^{(k)} > 0$ 有时难以成立, 因而由 BFGS 公式确定的 B_{k+1} 可能不是正定的. 为了使得 B_{k+1} 正定, Powell(1978) 建议将 $\gamma^{(k)}$ 换为 $\gamma^{(k)}$ 和 $B_k \delta^{(k)}$ 的线性组合

$$\eta^{(k)} = \theta_k \gamma^{(k)} + (1 - \theta_k) B_k \delta^{(k)}.$$

其中

$$\begin{aligned} \delta^{(k)} &= x^{(k+1)} - x^{(k)}, \quad \gamma^{(k)} = \nabla_x L(x^{(k+1)}, \lambda^{(k)}, \mu^{(k)}) - \nabla_x L(x^{(k)}, \lambda^{(k)}, \mu^{(k)}), \\ \theta_k &= \begin{cases} 1, & (\delta^{(k)})^T \gamma^{(k)} \geq 0.2 (\delta^{(k)})^T B_k \delta^{(k)}, \\ \frac{0.8 (\delta^{(k)})^T B_k \delta^{(k)}}{(\delta^{(k)})^T B_k \delta^{(k)} - (\delta^{(k)})^T \gamma^{(k)}}, & (\delta^{(k)})^T \gamma^{(k)} < 0.2 (\delta^{(k)})^T B_k \delta^{(k)}. \end{cases} \end{aligned}$$

还有一些其他修正公式,在此不一一列举.所以约束极值问题的 BFGS 公式为

$$B_{k+1} = B_k - \frac{B_k \delta^{(k)} (\delta^{(k)})^T B_k}{(\delta^{(k)})^T B_k \delta^{(k)}} + \frac{\eta^{(k)} (\eta^{(k)})^T}{(\delta^{(k)})^T \eta^{(k)}}. \quad (5.6.13)$$

在约束极值问题中, $x^{(k)}$ 通常不是可行解,特别是有非线性等式约束时.对于由 $x^{(k)}$ 产生的 $x^{(k+1)}$,有时不能要求 $f(x^{(k+1)}) < f(x^{(k)})$,而应要求 $x^{(k+1)}$ 更靠近可行域.这就需要有一个综合评价迭代解优劣的函数,即通常所称的罚函数.常用罚函数有

$$P_r(x) = f(x) + \sum_{i=1}^l r_i |g_i(x)| - \sum_{i=1}^m r_i \min\{0, g_i(x)\} - \sum_{i=1}^n r_{m+i} \min\{0, x_i - l_i\}$$

和

$$P_r(x) = f(x) + \sum_{i=1}^l r_i (g_i(x))^2 + \sum_{i=1}^m r_i (\min\{0, g_i(x)\})^2 + \sum_{i=1}^n r_{m+i} (\min\{0, x_i - l_i\})^2.$$

前者称为 L_1 罚函数,后者称为 L_2 罚函数,其中 $r_i > 0, i = 1, 2, \dots, m+n$.

用 $Q(B_k, x^{(k)})$ 表示将子问题(5.6.11)中的 $\nabla_x^2 L(x^{(k)}, \lambda^{(k)}, \mu^{(k)})$ 换为 B_k 得到的模型.原问题(5.6.10)的计算步骤简述如下,其中的罚函数 $P_r(x)$ 为 L_1 罚函数.

算法 5.6.1 模型(5.6.10)的拟牛顿法.

(i) 给定 $x^{(0)}$, 令 $\lambda^{(0)} = 0, \mu^{(0)} = 0, B_0 = I$, 置 $k = 0$.

(ii) 解 $Q(B_k, x^{(k)})$ 的 KT 条件, 设解为 $(s^{(k)}, \lambda^{(k+1)}, \mu^{(k+1)})$.

(iii) 如果 $\|s^{(k)}\| < \varepsilon$ ($\varepsilon = 10^{-4}$), 则 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 为所求的解, 停止.

否则, 如果

$$P_r(x^{(k)} + s^{(k)}) < P_r(x^{(k)}),$$

则令 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 并由(5.6.13)计算 B_{k+1} , 置 $k = k+1$ 转(ii); 否则计算

$$P_r(x^{(k)} + \alpha s^{(k)}), \quad \alpha = 0.1, 0.2, \dots, 0.9, 1.0,$$

以使 $P_r(x^{(k)} + \alpha s^{(k)})$ 达到最小的 $x^{(k)} + \alpha s^{(k)}$ 作为 $x^{(k+1)}$ 并由(5.6.13)计算 B_{k+1} , 置 $k = k+1$ 转(ii).

与许多文献介绍的同类算法相比 以上计算步骤较为简单 但也可解决不少问题.

例 5.6.1 用拟牛顿法解例 5.5.2.

解 令罚函数中的所有罚因子 $r_i = 1$,计算结果如表 5.37 所示.

表 5.37 拟牛顿法的计算结果

k	子问题的初始点 $x^{(k)}$ $\lambda^{(k)}$ 及计算结果 $s^{(k)}$ $\lambda^{(k+1)}$
0	$x^{(0)} = (1, 5, 5, 1)$, $\lambda^{(0)} = (0, 0)$, $s^{(0)} = (-0.3750004, -0.1250005, -1.125, 0.6250004)$, $\lambda^{(1)} = (-0.1510417, 0.4770834)$
1	$x^{(1)} = (0.6249996, 4.875, 3.875, 1.625)$, $s^{(1)} = (0.2584962, -0.1368132, -0.0573175, -0.1099963)$, $\lambda^{(2)} = (-0.1202335, 0.5555326)$
2	$x^{(2)} = (0.8834959, 4.738186, 3.817683, 1.515004)$, $s^{(2)} = (-0.1021793, -0.07062721, 3.123283 \times 10^{-5}, 0.2470938)$, $\lambda^{(3)} = (-0.1581881, 0.5820075)$
3	$x^{(3)} = (0.7915344, 4.674622, 3.817711, 1.737388)$, $s^{(3)} = (-0.03448588, -0.0352602, -0.01799512, 0.1293865)$, $\lambda^{(4)} = (-0.1783384, 0.5738308)$
4	$x^{(4)} = (0.7570485, 4.639362, 3.799716, 1.866775)$, $s^{(4)} = (-2.27946 \times 10^{-3}, 4.997253 \times 10^{-4}, -1.13194 \times 10^{-2}, 0.01750433)$, $\lambda^{(5)} = (-0.1792638, 0.5736848)$
5	$x^{(5)} = (0.7547691, 4.639862, 3.788397, 1.884279)$. $s^{(5)} = (-2.595782 \times 10^{-4}, -5.068779 \times 10^{-4}, 1.928806 \times 10^{-4}, 8.451939 \times 10^{-4})$. $\lambda^{(6)} = (-0.1798275, 0.5735294)$
6	$x^{(6)} = (0.7545095, 4.639355, 3.788589, 1.885125)$, $s^{(6)} = (-6.556511 \times 10^{-7}, 1.478195 \times 10^{-5}, -2.098083 \times 10^{-5}, 7.152557 \times 10^{-6})$, $\lambda^{(7)} = (-0.1797949, 0.5735409)$

其中 $x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$, $k = 0, 1, \dots, 6$. 除了 $\alpha_2 = 0.9$ 外 ,其他 $\alpha_k = 1$. 由于 $\|s^{(6)}\| < 10^{-4}$,所求的解为

$$x^* = (0.7545, 4.6394, 3.7886, 1.8851)^T,$$

对应的目标函数值 $f(x^*) = 16.84921$.

以上共进行了 42 次旋转运算,比 5.6 节的拉格朗日法效率高.但同样的算法用于别的问题,效果可能又差些.什么样的问题用什么样的算法求解最好,这个问题很值得探讨.

第六章 均值方差投资组合选择模型的解法

6.1 标准均值方差投资组合选择模型

6.1.1 问题的模型及其 KT 条件

1952 年美国学者马科维兹在 Journal of Finance 杂志上发表了一篇文章,题目是 Portfolio Selection. 这篇文章主张以收益率的方差作为风险的度量,并提出一个以极小化风险为目标的资产组合选择模型,称为均值方差模型. 这个模型及其所蕴涵的风险分散化思想是现代投资理论的基础.

设有 n 种资产可供选择,这 n 种资产的收益率为 $R_1, R_2, \dots, R_n$ (随机变量) 均值为 $r_1, r_2, \dots, r_n$, 协方差矩阵为 $H = (\sigma_{ij})_{n \times n}$, 其中 $\sigma_{ij} = \text{cov}(R_i, R_j)$. 如果每种资产占总资产的比例为 $x_1, x_2, \dots, x_n$, 则投资组合(或资产组合)的收益率 $R_p = R_1 x_1 + R_2 x_2 + \dots + R_n x_n$, 其均值

$$E(R_p) = E(R_1 x_1 + R_2 x_2 + \dots + R_n x_n) = r_1 x_1 + r_2 x_2 + \dots + r_n x_n,$$

方差风险

$$\begin{aligned} D(R_p) &= E((R_1 x_1 + R_2 x_2 + \dots + R_n x_n) - (r_1 x_1 + r_2 x_2 + \dots + r_n x_n))^2 \\ &= E((R_1 - r_1)x_1 + (R_2 - r_2)x_2 + \dots + (R_n - r_n)x_n)^2 \\ &= \sum_{i=1}^n \sum_{j=1}^n E((R_i - r_i)(R_j - r_j))x_i x_j \\ &= x^T H x, \end{aligned}$$

其中 $x = (x_1, x_2, \dots, x_n)^T$. 现在的问题是如何确定每种资产的比例,使资产组合的预期收益率达到一定水平并使风险最小,其数学模型如下:

$$\left. \begin{aligned} \min \quad & \frac{1}{2} x^T H x, \\ \text{s. t.} \quad & r_1 x_1 + r_2 x_2 + \dots + r_n x_n = r_p, \\ & x_1 + x_2 + \dots + x_n = 1, \\ & x_1, x_2, \dots, x_n \geq 0. \end{aligned} \right\} \quad (6.1.1)$$

马科维兹将此模型称为标准均值方差投资组合选择模型,只不过目标函数前没有 $\frac{1}{2}$. 其中的 r_p 是投资组合的预期收益率,为使问题有可行解,应该满足

$$\min\{r_1, r_2, \dots, r_n\} \leq r_p \leq \max\{r_1, r_2, \dots, r_n\}.$$

模型(6.1.1)的库恩-塔克条件(KT条件)为

$$\left. \begin{aligned} \sigma_{i1}x_1 + \dots + \sigma_{in}x_n - \lambda_1 - \lambda_2 r_i &\geq 0, i = 1, 2, \dots, n, \\ x_1 + x_2 + \dots + x_n &= 1, \\ r_1x_1 + r_2x_2 + \dots + r_nx_n &= r_p, \\ x_1, x_2, \dots, x_n &\geq 0, \\ (\sigma_{i1}x_1 + \dots + \sigma_{in}x_n - \lambda_1 - \lambda_2 r_i)x_i &= 0, i = 1, 2, \dots, n. \end{aligned} \right\} \quad (6.1.2)$$

不等式组(6.1.2)中的

$$\sigma_{i1}x_1 + \dots + \sigma_{in}x_n - \lambda_1 - \lambda_2 r_i \geq 0 \text{ 和 } x_i \geq 0$$

两个不等式称为互补不等式,它们的系数向量称为互补向量. 每个互补松弛条件等号的左边部分是两个互补不等式左边线性函数的乘积. 将 n 个互补松弛条件相加,整理后移项可得到 $x^T Gx = \lambda_1 + \lambda_2 r_p$,所以可用两种方式计算投资组合的风险.

6.1.2 模型的特点及算法

模型(6.1.1)是一个较为特殊的二次规划. 其目标函数不含一次项,仅有两个一般约束而且都是等式. 更为重要的是,协方差矩阵 H 正半定,是一个凸二次规划,所以可按第三章介绍的方法求解.

不等式组(6.1.2)有 $n+2$ 个变量,其中 λ_1 和 λ_2 是自由变量,需引入人工不等式 $\lambda_1 \geq 0$ 和 $\lambda_2 \geq 0$. 初始基本不等式组为

$$x_1 \geq 0, \dots, x_n \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0.$$

其第 i 式的系数向量是 $n+2$ 阶单位矩阵的第 i 行,记为 e_i . 初始基本解 $y^{(0)} = \begin{pmatrix} x^{(0)} \\ \lambda^{(0)} \end{pmatrix}$,其中 $x^{(0)} = (0, 0, \dots, 0)^T$, $\lambda^{(0)} = (0, 0)^T$. 其余线性(不)等式是非基(不)等式,它们的系数向量 $h_i = (\sigma_{i1}, \dots, \sigma_{in}, -1, -r_i)$, $e = (1, \dots, 1, 0, 0)^T = (r_1, \dots, r_n, 0, 0)^T$ 等是初始非基向量,偏差分别为 $h_i y^{(0)} = 0$, $ex^{(0)} - 1 = -1$, $x^{(0)} - r_p = -r_p$. 各非基向量关于基向量的组合系数及其偏差如表6.1所示.

表 6.1 初 始 表

	e_1	...	e_n	e_{n+1}	e_{n+2}	
h_1	σ_{11}	...	σ_{1n}	- 1	- r_1	0
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
h_n	σ_{n1}	...	σ_{nn}	- 1	- r_n	0
e	1	...	1	0	0	- 1
r	r_1	...	r_n	0	0	- r_p

由于 e 和 r 是等式的系数向量 , 首先需将它们换入基中. 也就是说要在非基向量 e 和 r 两行找两个非零元素进行两次旋转运算. 为了维持互补松弛条件 , 再以它们的对称元素为枢轴进行两次旋转运算 , 将人工不等式的系数向量 e_{n+1} 和 e_{n+2} 换为非基向量. 不妨设 $r_1 \neq r_2$, 作以下 4 次旋转运算 :

$$e \leftrightarrow e_1, \quad r \leftrightarrow e_2, \quad h_1 \leftrightarrow e_{n+1}, \quad h_2 \leftrightarrow e_{n+2}.$$

为了看出这 4 次旋转运算的结果 , 将表 6.1 的数据划分为 3×4 部分 , 如表 6.2 所示.

表 6.2 初始表的矩阵形式

	e_1	e_2	$e_3 \dots e_n$	e_{n+1}	e_{n+2}	
h_1	H_{11}		H_{12}	$-A_1^T$		0
h_2	H_{12}^T		H_{22}	$-A_2^T$		0
h_3						
$\vdots$						
h_n						
e	A_1		A_2	O		σ_a
r						

其中

$$H_{11} = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}, \quad H_{12} = \begin{pmatrix} \sigma_{13} & \dots & \sigma_{1n} \\ \sigma_{23} & \dots & \sigma_{2n} \end{pmatrix}, \quad H_{22} = \begin{pmatrix} \sigma_{33} & \dots & \sigma_{3n} \\ \vdots & & \vdots \\ \sigma_{n3} & \dots & \sigma_{nn} \end{pmatrix},$$
$$A_1 = \begin{pmatrix} 1 & 1 \\ r_1 & r_2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \dots & 1 \\ r_3 & \dots & r_n \end{pmatrix}, \quad \sigma_a = \begin{pmatrix} -1 \\ -r_p \end{pmatrix}.$$

上面提到的 4 次旋转运算相当于在表 6.2 中 , 先沿着 A_1 的主对角线进行两

次旋转运算,再沿着 $-A_1^T$ 的主对角线进行两次旋转运算.先后两次应用第一章的旋转运算公式(1.1.1) 结果见表 6.3.

表 6.3 预处理的结果

	e r	e ₃ ... e _n	h ₁ h ₂	
e _{n+1} e _{n+2}	G ₁₁	G ₁₂	- B ₁₁ ^T	σ' _I
h ₃ ⋮ h _n	G ₁₂ ^T	G ₂₂	- B ₁₂ ^T	σ' _{II}
e ₁ e ₂	B ₁	B ₂	O	σ' _{III}

其中

$$\begin{aligned} G_{11} &= A_1^{-T} H_{11} A_1^{-1}, \quad G_{12} = A_1^{-T} (H_{12} - H_{11} A_1^{-1} A_2), \\ G_{22} &= H_{22} - H_{12}^T A_1^{-1} A_2 - A_2^T A_1^{-T} (H_{12} - H_{11} A_1^{-1} A_2) \\ &= (-A_2^T A_1^{-T} \quad I) \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix} \begin{pmatrix} -A_1^{-1} A_2 \\ I \end{pmatrix}, \quad (6.1.3) \\ B_1 &= A_1^{-1}, \quad B_2 = -A_1^{-1} A_2, \\ \sigma'_I &= -A_1^{-T} H_{11} A_1^{-1} \sigma_a, \quad \sigma'_{II} = -H_{12}^T A_1^{-1} \sigma_a + A_2^T A_1^{-T} H_{11} A_1^{-1} \sigma_a, \\ \sigma'_{III} &= -A_1^{-1} \sigma_a. \end{aligned}$$

其中 A_1^{-T} 表示 A_1 的逆矩阵的转置, I 为单位矩阵.事实上

$$\sigma'_{III} = -A_1^{-1} \sigma_a = \left(\frac{r_2 - r_p}{r_2 - r_1}, \frac{r_p - r_1}{r_2 - r_1} \right)^T.$$

由于 $\min\{r_1, r_2, \dots, r_n\} \leq r_p \leq \max\{r_1, r_2, \dots, r_n\}$, 如果 r_1 和 r_2 分别是 $r_1, r_2, \dots, r_n$ 的最小值和最大值, 可使表 6.3 的两个非基约束向量的偏差不为负数. 由于协方差矩阵 $H = \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^T & H_{22} \end{pmatrix}$ 正半定, 由表 6.3 可见 G_{22} 正半定.

以上 4 次旋转运算称为预处理, 以后进行的旋转运算称为主迭代. 与一般二次规划的旋转算法一样, 预处理后, 基等式对应的 e 和 r 两列以及人工不等式对应的 e_{n+1} 和 e_{n+2} 两行都可以删掉. 在剩余的元素中, 以正对角元素为枢轴进行的

旋转运算称为主旋转,以关于主对角线对称的两个非零元素为枢轴进行的两次旋转运算称为双枢轴运算.双枢轴运算和主旋转一样,可以维持互补松弛条件.在入基向量确定的条件下,应该首先考虑主旋转,然后才是双枢轴运算,因为主旋转的数值稳定性非常好.

用 $r(H)$ 表示协方差矩阵 H 的秩,由 (6.1.3) 式可见, G_{22} 的秩不超过 $r(H)$. 所以在表 6.3 中至多可以沿着 G_{22} 的对角线进行 $r(H)$ 次旋转运算,将单位向量 $e_3, e_4, \dots, e_n$ 中的 $r(H)$ 个换为非基向量. $e_i (i = 1, 2, \dots, n)$ 是基向量时, $x_i \geq 0$ 是基不等式,基本解 $x_i = 0$. 只有当 e_i 是非基向量时 x_i 才可能不等于零. 所以每一投资组合最多包含 $r(H) + 2$ 种资产. 一般说来有以下结论.

结论 6.1.1 每一投资组合包含的资产数目不超过协方差矩阵的秩加 2.

下面假设在表 6.3 中沿 G_{22} 的对角线进行 $k - 2 (k > 3)$ 次主旋转,将 $e_3, \dots, e_k$ 也换为非基向量. 整理后的结果如表 6.4 所示.

表 6.4 一般表的矩阵形式

	$e \quad r$	$h_1 \quad \dots \quad h_k$	$e_{k+1} \quad \dots \quad e_n$	
e_{n+1}	L_{11}	$-E_1^T$	L_{12}	σ_λ
e_{n+2}				
e_1	E_1	D_{11}	E_2	σ_e
$\vdots$				
e_k				
h_{k+1}	L_{12}^T	$-E_2^T$	L_{22}	σ_h
$\vdots$				
h_n				

根据第一章的定理 1.1.2 和 1.1.4, L_{22} 和 D_{11} 正半定. 于是有以下结论.

结论 6.1.2 在标准均值方差模型的旋转算法表中,单位向量 $e_1, e_2, \dots, e_n$ 中的非基向量相应的主子矩阵以及拉格朗日向量 $h_1, h_2, \dots, h_n$ 中的非基向量相应的主子矩阵正半定.

下面除了假设 $e_1, e_2, \dots, e_n$ 的前 k 个是非基向量外,还假设它们相应的协方差矩阵正定,相应的收益率互不相等.

用 A 表示表 6.3 中 G_{22} 左上角的 $k - 2$ 阶主子矩阵,用 B 表示 $B_2 = -A_1^{-1}A_2$ 的前 $k - 2$ 列构成的矩阵,那么

$$B = - \begin{pmatrix} 1 & 1 \\ r_1 & r_2 \end{pmatrix}^{-1} \begin{pmatrix} 1 & \dots & 1 \\ r_3 & \dots & r_n \end{pmatrix} = - \frac{1}{r_2 - r_1} \begin{pmatrix} r_2 - r_3 & \dots & r_2 - r_k \\ r_3 - r_1 & \dots & r_k - r_1 \end{pmatrix}$$

如果 $k = 3$,可以算出 $BA^{-1}B^T$ (即非基向量 e_1 和 e_2 行)的两个主对角元素为

$$\frac{(r_2 - r_3)^2}{(r_2 - r_1)^2 A} \text{ 和 } \frac{(r_3 - r_1)^2}{(r_2 - r_1)^2 A},$$

它们均是正数,如果 $k \geq 4$, $BA^{-1}B^T$ 正定. 所以表 6.4 中 D_{11} 的主对角线元素均为正. 也就是说,如果 $e_1, e_2, \dots, e_k$ 中的任何一个入基,主旋转运算总是可以进行的. 为了看出非基向量 $h_i (i = k+1, k+2, \dots, n)$ 入基的情形,考虑表 6.4 与初始表之间的关系. 重新安排表 6.1 中的数据得表 6.5.

表 6.5 与表 6.4 对应的初始表

	e_{n+1} e_{n+2}	e_1 $\dots$ e_k	e_{k+1} $\dots$ e_n	
e	O	A_1	A_2	σ_a
r				
h_1	$-A_1^T$	H_{11}	H_{12}	0
$\vdots$				
h_k				
h_{k+1}	$-A_2^T$	H_{12}^T	H_{22}	0
$\vdots$				
h_n				

在这里

$$H_{11} = \begin{pmatrix} \sigma_{11} & \dots & \sigma_{1k} \\ \vdots & & \vdots \\ \sigma_{k1} & \dots & \sigma_{kk} \end{pmatrix}, \quad H_{12} = \begin{pmatrix} \sigma_{1,k+1} & \dots & \sigma_{1n} \\ \vdots & & \vdots \\ \sigma_{k,k+1} & \dots & \sigma_{kn} \end{pmatrix}, \quad H_{22} = \begin{pmatrix} \sigma_{k+1,k+1} & \dots & \sigma_{k+1,n} \\ \vdots & & \vdots \\ \sigma_{n,k+1} & \dots & \sigma_{nn} \end{pmatrix},$$

$$A_1 = \begin{pmatrix} 1 & \dots & 1 \\ r_1 & \dots & r_k \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & \dots & 1 \\ r_{k+1} & \dots & r_n \end{pmatrix}.$$

如果 H_{11} 正定且 $r_1, r_2, \dots, r_k$ 不全相等, 则

$$\begin{pmatrix} L_{11} & -E_1^T \\ E_1 & D_{11} \end{pmatrix} = \begin{pmatrix} O & A_1 \\ -A_1^T & H_{11} \end{pmatrix}^{-1}$$

$$= \begin{pmatrix} (A_1 H_{11}^{-1} A_1^T)^{-1} & -(A_1 H_{11}^{-1} A_1^T)^{-1} A_1 H_{11}^{-1} \\ H_{11}^{-1} A_1^T (A_1 H_{11}^{-1} A_1^T)^{-1} & H_{11}^{-1} - H_{11}^{-1} A_1^T (A_1 H_{11}^{-1} A_1^T)^{-1} A_1 H_{11}^{-1} \end{pmatrix}.$$

因此

$$L_{22} = H_{22} - \begin{pmatrix} -A_2^T & H_{12}^T \\ -A_1^T & H_{11} \end{pmatrix}^{-1} \begin{pmatrix} A_2 \\ H_{12} \end{pmatrix}$$

$$= (H_{22} - H_{12}^T H_{11}^{-1} H_{12}) + (A_2^T - H_{12}^T H_{11}^{-1} A_1^T)(A_1 H_{11}^{-1} A_1^T)^{-1}(A_2 - A_1 H_{11}^{-1} H_{12}).$$

可见 L_{22} 是两个正半定矩阵的和. 另一方面,

$$\begin{aligned}\sigma_h &= 0 - (-A_2^T, H_{12}^T) \begin{pmatrix} O & A_1 \\ -A_1^T & H_{11} \end{pmatrix}^{-1} \begin{pmatrix} \sigma_a \\ 0 \end{pmatrix} \\ &= (A_2^T - H_{12}^T H_{11}^{-1} A_1^T)(A_1 H_{11}^{-1} A_1^T)^{-1} \sigma_a,\end{aligned}$$

其中 $\sigma_a = \begin{pmatrix} -1 \\ -r_p \end{pmatrix}$. 如果 L_{22} 主对角线上第 i 个元素是 0, 那么 $A_2^T - H_{12}^T H_{11}^{-1} A_1^T$ 的第 i 行是零向量, 于是 σ_h 的第 i 个分量等于 0. 换句话说, 如果某个非基向量 h_i 的偏差不等于零, 特别当它是负数时, 则其相应的主元素一定是正数.

综上所述有以下结论.

结论 6.1.3 在旋转算法表中, 如果 $e_1, e_2, \dots, e_n$ 至少有三个是非基向量, 并且它们相应的协方差矩阵正定, 相应的收益率互不相等, 则仅需要主旋转运算.

从表 6.4 可见, 一般表的标量形式如表 6.6 所示.

表 6.6 一般表的标量形式

	e	r	h_i	e_j	
e_{n+1}	$w_{n+1, n+2}$	$w_{n+1, n+2}$	$-w_{i, n+1}$	$w_{j, n+1}$	σ_{n+1}
e_{n+2}	$w_{n+1, n+2}$	$w_{n+2, n+2}$	$-w_{i, n+2}$	$w_{j, n+2}$	σ_{n+2}
e_i	$w_{i, n+1}$	$w_{i, n+2}$	w_{ii}	w_{ij}	σ_i
h_j	$w_{j, n+1}$	$w_{j, n+2}$	$-w_{ij}$	w_{jj}	σ_j

此表相应的目标函数值为 $\lambda_1 + \lambda_2 r_p = \sigma_{n+1} + \sigma_{n+2} r_p$.

如果 e_i 入基 h_i 出基, e_{n+1} 和 e_{n+2} 的偏差分别为 $\frac{\sigma_{n+1} + \sigma_i w_{i, n+1}}{w_{ii}}$ 和

$\frac{\sigma_{n+2} + \sigma_i w_{i, n+2}}{w_{ii}}$. 所以新的目标函数值为

$$\left(\frac{\sigma_{n+1} + \sigma_i w_{i, n+1}}{w_{ii}} \right) + \left(\frac{\sigma_{n+2} + \sigma_i w_{i, n+2}}{w_{ii}} \right) r_p,$$

目标函数值的改变量为

$$\Delta V = \frac{(w_{i, n+1} + w_{i, n+2} r_p) \sigma_i}{w_{ii}}.$$

对于表 6.6, 用 a_j 表示除 e 和 r 外的基向量, 则

$$e_i = \sum_j w_{ij} a_j + w_{i, n+1} e + w_{i, n+2} r.$$

上式两边右乘当前基本解 $\bar{x}$ 得

$$\sigma_i = e_i \bar{x} = w_{i, n+1} + w_{i, n+2} r_p,$$

故 $\Delta V = \frac{\sigma_i^2}{w_{ii}}$. 同样可证, 如果在表 6.4 中以 h_j 入基 e_j 出基, 则目标函数的改变量

$$\Delta V = -\frac{\sigma_j^2}{w_{jj}}.$$

所以有以下结论.

结论 6.1.4 设在旋转算法表中, 非基向量 e_i (或 h_i) 行的主元素为 w_{ii} , 偏差为 σ_i . 如果以 e_i (或 h_i) 入基, 以 h_i (或 e_i) 出基, 则目标函数的改变量为 $\frac{\sigma_i^2}{w_{ii}}$

$$\left(\text{或} -\frac{\sigma_i^2}{w_{ii}} \right).$$

此结论说明, 如果一次主旋转运算使资产组合的资产减少, 则风险增加; 反之, 风险减小. 但这时的资产组合可能包含负资产, 相当于卖空.

6.1.3 旋转算法的计算步骤

下面介绍标准均值方差模型 KT 条件(6.1.2)的计算步骤. 在这种解法中每进行一次迭代, 表格中的所有元素都要更新, 因而将其称为全表格运算法. 有多种规则选择非基向量入基, 在此我们使用最小偏差规则, 即在所有偏差为负的非基向量中选择偏差最小者入基.

算法 6.1.1 (6.1.2)的全表格运算法.

步骤 1 初始步骤. 以 $x_1 \geq 0, \dots, x_n \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式组, 列出表 6.1 所示初始表.

步骤 2 预处理. 设 $r_1 = \min\{r_1, r_2, \dots, r_n\}$, $r_2 = \max\{r_1, r_2, \dots, r_n\}$, 进行以下 4 次旋转运算:

$$e \leftrightarrow e_1, \quad h_1 \leftrightarrow e_{n+1}, \quad r \leftrightarrow e_2, \quad h_2 \leftrightarrow e_{n+2}.$$

步骤 3 主要迭代(按最小偏差规则).

(i) 如果所有非基向量(不包括 e_{n+1} 和 e_{n+2})的偏差非负, 则当前基本解是不等式组(6.1.2)的解, 停止; 否则转(ii).

(ii) 以偏差最小的非基向量入基. 如果同行主对角线上的元素为正, 则以其为枢轴进行旋转运算; 转(i); 否则, 以同行最大的正元素及其对称元素为枢轴, 先后进行两次旋转运算; 转(i).

值得注意的是, 就某一个 r_p 的数值算得最优解后, 如果还要计算别的数值

相应的最优解,可使用参数化方法. 由于基等式 $r_1x_1 + r_2x_2 + \dots + r_nx_n = r_p$ 对应的基向量为 r , 当 r_p 变为 $r_p + \Delta$ 时, 将 r 所在列乘以 Δ 加到最后一列上去, 就得到 $r_1x_1 + r_2x_2 + \dots + r_nx_n = r_p + \Delta$ 相应的表.

下面用一个例子说明计算细节.

例 6.1.1 某投资者对三种风险资产感兴趣, 这三种资产收益率的均值为

0.06, 0.12, 0.09, 协方差矩阵 $H = \begin{pmatrix} 0.2 & 0.3 & -0.01 \\ 0.3 & 2.4 & 0.5 \\ -0.01 & 0.5 & 1.1 \end{pmatrix}$. 试为之构造风险

最小的资产组合, 使预期收益率 $r_p = 0.08, 0.10, 0.11$ 或 0.115 .

解 此问题的数学模型为:

$$\begin{aligned} \min f(x) &= \frac{1}{2} x^T H x, \\ \text{s. t. } 0.06x_1 + 0.12x_2 + 0.09x_3 &= r_p, \\ x_1 + x_2 + x_3 &= 1, \\ x_1, x_2, x_3 &\geq 0, \end{aligned}$$

其中 $x = (x_1, x_2, x_3)^T$, $r_p = 0.08, 0.10, 0.11$ 或 0.115 . $r_p = 0.08$ 时的 KT 条件为

$$\begin{aligned} 0.2x_1 + 0.3x_2 - 0.01x_3 - \lambda_1 - 0.06\lambda_2 &\geq 0, \\ 0.3x_1 + 2.4x_2 + 0.5x_3 - \lambda_1 - 0.12\lambda_2 &\geq 0, \\ -0.01x_1 + 0.5x_2 + 1.1x_3 - \lambda_1 - 0.09\lambda_2 &\geq 0, \\ x_1 + x_2 + x_3 &= 1, \\ 0.06x_1 + 0.12x_2 + 0.09x_3 &= 0.08, \\ x_1, x_2, x_3 &\geq 0, \\ (0.2x_1 + 0.3x_2 - 0.01x_3 - \lambda_1 - 0.06\lambda_2)x_1 &= 0, \\ (0.3x_1 + 2.4x_2 + 0.5x_3 - \lambda_1 - 0.12\lambda_2)x_2 &= 0, \\ (-0.01x_1 + 0.5x_2 + 1.1x_3 - \lambda_1 - 0.09\lambda_2)x_3 &= 0. \end{aligned}$$

令

$$\begin{aligned} h_1 &= (0.2, 0.3, -0.01, -1, -0.06), \\ h_2 &= (0.3, 2.4, 0.5, -1, -0.12), \\ h_3 &= (-0.01, 0.5, 1.1, -1, -0.09), \\ e &= (1, 1, 1, 0, 0), \\ r &= (0.06, 0.12, 0.09, 0, 0), \\ e_1 &= (1, 0, 0, 0, 0), \\ e_2 &= (0, 1, 0, 0, 0), \end{aligned}$$

表 6.9 第二次双枢轴运算结果

	e	r	e ₃	h ₁	h ₂	σ_i
e ₄	2	- 31.6667	0.33	- 2	1	- 0.5333
e ₅	- 31.6667	555.5557	- 9.8333	16.6667	- 16.6667	12.7778
h ₃	0.33	- 9.8333	1.41 *	0.5	0.5	- 0.4567
e ₁	2	- 16.6667	- 0.5	0	0	0.6667
e ₂	- 1	16.6667	- 0.5	0	0	0.3333

在表 6.9 中除了人工不等式的系数向量 e₄ 和 e₅ 行外, h₃ 的偏差是负数, 以该行的对角元素 1.41 为枢轴作旋转运算结果见表 6.10.

表 6.10 第一次主旋转结果

	e	r	h ₃	h ₁	h ₂	r _p = 0.08	r _p = 0.10	r _p = 0.11
e ₄	1.9228	- 29.3652	0.234	- 2.117	0.883	- 0.4265	- 1.0138	- 1.3074
e ₅	- 29.3652	486.978	- 6.974	20.1537	- 13.1797	9.593	19.3325	24.2023
e ₃	- 0.234	6.974	0.7092	- 0.3546	- 0.3546	0.3239	0.4634	0.5331
e ₁	2.117	- 20.1537	- 0.3546	0.1773 *	0.1773	0.5047	0.1017	- 0.0999
e ₂	- 0.883	13.1797	- 0.3546	0.1773	0.1773	0.1714	0.435	0.5668

由表 6.10 的 r_p = 0.08 列可见, 除了 e₄ 和 e₅ 行外, 所有非基向量偏差是正数, 最优资产组合为 (x₁, x₂, x₃) = (0.5047, 0.1714, 0.3239), 其方差风险为

$$\lambda_1 + \lambda_2 r_p = - 0.4265 + 9.593 \times 0.08 = 0.3410.$$

当 r_p = 0.10 时, 预期收益率的改变量 $\Delta = 0.10 - 0.08 = 0.02$. 根据参数公式(2.4.3) 将基向量 r 列的数据乘以 0.02 加到 r_p = 0.08 列得

$$\begin{aligned} & - 0.4265 - 29.3652 \times 0.02 = - 1.0138, \\ & 9.593 + 486.978 \times 0.02 = 19.3325, \\ & 0.3239 + 6.974 \times 0.02 = 0.4634, \\ & 0.5047 + (- 20.1537) \times 0.02 = 0.1017, \\ & 0.1714 + 13.1797 \times 0.02 = 0.435. \end{aligned}$$

以上数字见表 6.10 的 r_p = 0.10 栏. 由于 e₁, e₂, e₃ 的偏差仍然是正数, 最优资产组合为 (x₁, x₂, x₃) = (0.1017, 0.435, 0.4634), 其方差风险为

$$\lambda_1 + \lambda_2 r_p = - 1.0138 + 19.3325 \times 0.10 = 0.9195.$$

当 $r_p = 0.11$ 时,预期收益率的改变量 $\Delta = 0.11 - 0.08 = 0.03$. 将基向量 r 列的数据乘以 0.03 加到 $r_p = 0.08$ 列得

$$\begin{aligned} & -0.4265 - 29.3652 \times 0.03 = -1.3074, \\ & 9.593 + 486.978 \times 0.03 = 24.2023, \\ & 0.3239 + 6.974 \times 0.03 = 0.5331, \\ & 0.5047 + (-20.1537) \times 0.03 = -0.0999, \\ & 0.1714 + 13.1797 \times 0.03 = 0.5668. \end{aligned}$$

以上数字见表 6.10 的 $r_p = 0.11$ 栏. 由于 e_1 的偏差是负数, 以该行的对角元素 0.1773 为枢轴作旋转运算结果见表 6.11.

表 6.11 第二次主旋转结果

	e	r	h_3	e_1	h_2	$r_p = 0.11$	$r_p = 0.115$
e_4	27.2	-270	-4	-11.94	3	-2.5	-3.85
e_5	-270	2777.78	33.3333	113.6668	-33.3333	35.5556	49.4445
e_3	4	-33.3333	0	-2	0	0.3333	0.1667
h_1	-11.94	113.6668	2	5.64	-1	0.5633	1.1317
e_2	-3	33.3333	0	1	0	0.6667	0.8333

从表 6.11 的 $r_p = 0.11$ 列可见,除了 e_4 和 e_5 行外所有非基向量偏差为正,最优资产组合为 $(x_1, x_2, x_3) = (0, 0.6667, 0.3333)$, 其方差风险为

$$\lambda_1 + \lambda_2 r_p = -2.5 + 35.5556 \times 0.11 = 1.4111.$$

当 $r_p = 0.115$ 时,预期收益率的改变量 $\Delta = 0.115 - 0.11 = 0.005$. 将基向量 r 列的数据乘以 0.005 加到 $r_p = 0.11$ 列,结果见 $r_p = 0.115$ 列. 由于非基向量偏差的符号未变,最优资产组合为 $(x_1, x_2, x_3) = (0, 0.8333, 0.1667)$, 其方差风险为

$$\lambda_1 + \lambda_2 r_p = -3.85 + 49.4445 \times 0.115 = 1.8361.$$

以上各表都没有删掉基等式对应的 e 行和 r 列,也没有删掉其互补向量 e_4 行和 e_5 行. 这是为了利用 r 列修正不同预期收益率的偏差,利用 e_4 行和 e_5 行的偏差计算投资组合的方差风险. 保持基向量 e 列这一列是为了使表中数据看起来对称,否则可以删掉. 如果仅就 r_p 的一个数值计算投资组合,并利用 $x^T H x$ 计算方差风险,这两列和两行都可以删掉.

在模型(6.1.1)中,资产的预期收益率和协方差矩阵是给定的. 对于实际问题,原始数据通常为价格,如股票价格和外汇牌价. 如果计算它们的组合,需利用各资产连续若干时期的价格数据计算各期收益率,如第 i 种资产的 T 期收益率

$r_{i1}, r_{i2}, \dots, r_{iT}$ 则

$$r_i \approx \frac{1}{T} \sum_{t=1}^T r_{it}, \sigma_{ij} \approx \frac{1}{T-1} \sum_{t=1}^T (r_{it} - r_i)(r_{jt} - r_j), \quad i, j = 1, 2, \dots, n.$$

6.1.4 有效投资组合

设有 n 种资产或资产组合,第 k 种资产(组合)的预期收益率为 r_k ,标准差为 σ_k .理性的投资者不仅希望收益率高而且希望风险小.所以对于其中的两种资产 i 和 j ,如果

$$r_i < r_j, \quad \sigma_i \geq \sigma_j,$$

或者

$$r_i \leq r_j, \quad \sigma_i > \sigma_j,$$

资产 i 不是有效的.从 n 种资产中去掉无效的资产,剩下的资产是有效的.从收益率和风险两方面衡量,有效资产集合内的任何一种资产不优于另一种,其选择取决于投资者的收益/风险偏好.

在以标准差 σ 为横坐标、收益率 r 为纵坐标的直角坐标系中表示每一种资产,位于左上方的资产是有效的.在标准投资组合模型(6.1.1)中,让 r_p 在其最小值和最大值之间变化,可以算出无数投资组合,以它们的标准差为横坐标、收益率为纵坐标描点作图得到的曲线如图 6.1 所示.

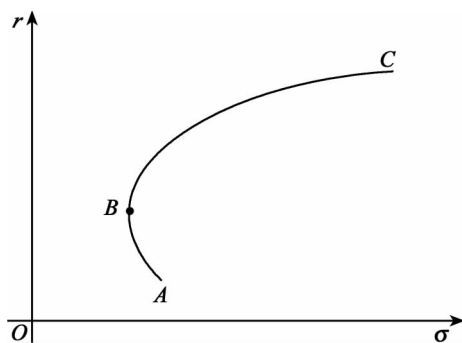


图 6.1 有效边界

由于 AB 弧段上每一点与其正上方的点相比,标准差相同,收益率小, AB 弧段上每一点不是有效的. BC 弧段上每一点是有效的.如果将 n 种资产中的每一种也在坐标系中用一点表示,所有点都在曲线 ABC 的右下方.一般来说,收益率最小和最大的两点为 A 和 C .有效投资组合的意义在于:其标准差(即收益的波

动性)比大多数资产小很多,而预期收益率仅比收益率最大的资产低一点.

如果将模型(6.1.1)的第一个约束改为不等式

$$r_1 x_1 + r_2 x_2 + \dots + r_n x_n \geq r_p,$$

当 r_p 在 A, B 两点的收益率之间变化时,结果都是 B 点,因为目标函数要求方差最小,所以计算结果为曲线 BC,即得到的每一个投资组合都是有效的.

6.2 资产有上界限制的均值方差模型

对于机构投资者,有的资产(如股票)占总资产的比例不允许超过一定限度,这时需使用资产有上界限制的模型.用 u_i 表示资产 x_i 的上界,它满足 $0 < u_i \leq 1$ ($i = 1, 2, \dots, n$)且 $u_1 + u_2 + \dots + u_n \geq 1$.模型的形式为

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} x^T H x, \\ \text{s. t. } r_1 x_1 + r_2 x_2 + \dots + r_n x_n &= r_p, \\ x_1 + x_2 + \dots + x_n &= 1, \\ u_i \geq x_i \geq 0, i &= 1, 2, \dots, n. \end{aligned} \right\} \quad (6.2.1)$$

此模型其他符号的含义与模型(6.1.1)相同.当然,我们可以把约束 $u_i \geq x_i$ ($i = 1, 2, \dots, n$)都看做一般约束,按一般凸二次规划的算法计算,但这样做计算量较大.由于 $x_i \geq 0$ 和 $-x_i \geq -u_i$ 两不等式系数向量的符号相反,可简化计算.

设 r_p 的取值使问题有可行解.模型(6.2.1)的 KT 条件为

$$\left. \begin{aligned} \sigma_{i1} x_1 + \dots + \sigma_{in} x_n - \lambda_1 - r_1 \lambda_2 - \mu_i + \bar{\mu}_i &= 0, \quad i = 1, 2, \dots, n, \\ \mu_i, \bar{\mu}_i &\geq 0, \quad i = 1, 2, \dots, n, \\ \mu_i x_i = 0, \quad \bar{\mu}_i (-x_i + u_i) &= 0, \quad i = 1, 2, \dots, n, \\ x_1 + x_2 + \dots + x_n &= 1, \\ r_1 x_1 + r_2 x_2 + \dots + r_n x_n &= r_p, \\ x_i \geq 0, \quad -x_i \geq -u_i, \quad i &= 1, 2, \dots, n. \end{aligned} \right\} \quad (6.2.2)$$

以上不等式组共有 $5n+2$ 个线性(不)等式和 $3n+2$ 个变量.如果仅消去 n 个 μ_i 及相应的非负不等式,还有 $4n+2$ 个线性(不)等式和 $2n+2$ 个变量,对应大小为 $(2n+2) \times (2n+3)$ 的旋转算法表,其大部分元素为零.下面说明如何消去所有 μ_i 和 $\bar{\mu}_i$,使问题的大小减少为 $(n+2) \times (n+3)$.

对于任何 $\bar{x} = (\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n)^T$ 和每个 $i \in \{1, 2, \dots, n\}$ $x_i \geq 0$ 和 $-x_i \geq -u_i$ 不可能都是紧约束,所以 μ_i 和 $\bar{\mu}_i$ 至少有一个为 0. 记

$$h_i(x, \lambda) = \sigma_{i1} x_1 + \dots + \sigma_{in} x_n - \lambda_1 - r_1 \lambda_2.$$

如果 $\dot{x}_i = 0$, 则 $\dot{\mu}_i = 0$ 因而 $\mu_i = h_i(x, \lambda) \geq 0$; 若 $\dot{x}_i = u_i$, 则 $\mu_i = 0$ 因而 $\dot{\mu}_i = -h_i(x, \lambda) \geq 0$. 若 $\dot{x}_i$ 既不等于 0 也不等于 u_i , 则 $\mu_i = \dot{\mu}_i = 0$, 于是 $h_i(x, \lambda) = 0$. 所以可在计算过程或者仅使用 $h_i(x, \lambda) \geq 0$ 或者仅使用 $-h_i(x, \lambda) \geq 0$. 不难验证, 如果求得不等式组

$$\begin{aligned} h_i(x, \lambda) &\geq 0, \quad x_i \geq 0, \quad h_i(x, \lambda)x_i = 0, \quad i \in I_1, \\ -h_i(x, \lambda) &\geq 0, \quad -x_i \geq -u_i, \quad -h_i(x, \lambda)(x_i - u_i) = 0, \quad i \in I_2, \\ x_1 + x_2 + \dots + x_n &= 1, \\ r_1 x_1 + r_2 x_2 + \dots + r_n x_n &= r_p \end{aligned} \quad (6.2.3)$$

的一个解, 并且它的每个分量还满足 $-x_i \geq -u_i (i \in I_1)$ 以及 $x_i \geq 0 (i \in I_2)$, 则此解是不等式组(6.2.2)的解. 其中 I_1 和 I_2 是 $\{1, 2, \dots, n\}$ 某个划分, 即 $I_1 \cup I_2 = \{1, 2, \dots, n\}$, $I_1 \cap I_2 = \emptyset$, I_1 和 I_2 可以有一个是空集. 不等式组(6.2.3)的特点是: 它仅包含 $x_i \geq 0$ 和 $-x_i \geq -u_i$ 两不等式中的一个. 其解法如下.

首先不考虑所有变量的上界, 按算法 6.1.1 的步骤 1 和步骤 2 解相应的 KT 条件, 即解不等式组(6.1.2). 为保持算法的完整性, 现将这两个步骤简述如下.

以 $x_1 \geq 0, \dots, x_n \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式组, 列出初始表如表 6.12 所示.

表 6.12 初 始 表

	e_1	...	e_n	e_{n+1}	e_{n+2}	
h_1	σ_{11}	...	σ_{1n}	-1	$-r_1$	0
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
h_n	σ_{n1}	...	σ_{nn}	-1	$-r_n$	0
e	1	...	1	0	0	-1
r	r_1	...	r_n	0	0	$-r_p$

其中 $h_i = (\sigma_{i1}, \dots, \sigma_{in}, -1, -r_i)$, $e = (1, \dots, 1, 0, 0)$, $r = (r_1, \dots, r_n, 0, 0)$, e_i 是 $n+2$ 阶单位矩阵的第 i 行, $i = 1, 2, \dots, n+2$. 然后进行以下 4 次交换(设 $r_1 \neq r_2$):

$$e \leftrightarrow e_1, \quad h_1 \leftrightarrow e_{n+1}, \quad r \leftrightarrow e_2, \quad h_2 \leftrightarrow e_{n+2}.$$

到目前为止, 我们还未考虑变量的上界限制, 下面分析在什么情况下需要考虑这些限制以及如何处理.

由于 $x_i \geq 0$ 的系数向量为 e_i , $-x_i \geq -u_i$ 的系数向量为 $-e_i$. e_i 和 $-e_i (i =$

$1, 2, \dots, n$) 仅有一个出现在同一张旋转算法表中. 如果 e_i 和 $-e_i$ 有一个是基向量, 另一个的偏差为 $u_i > 0$, 因而可不考虑. 现在设 e_i 是出现于当前表中的非基向量, 偏差为 σ_i , 则 $-e_i$ 的偏差为 $u_i - \sigma_i$. 若 $\sigma_i < 0$, $-e_i$ 的偏差 $u_i - \sigma_i > 0$, 因而暂时可不考虑 $-e_i$. 但是如果 $\sigma_i > 0$, $-e_i$ 的偏差 $u_i - \sigma_i$ 可能是负数, 因而是一个可能的入基向量. 如果 $-e_i$ 入基, 可先将表中的数据作如下更新: 将 e_i 换为 $-e_i$, 将 σ_i 换为 $u_i - \sigma_i$, 将第 i 行所有数据全部反号, 然后将第 i 列所有数据全部反号, 并将 h_i 换为 $-h_i$. 这相当于将非基不等式 $x_i \geq 0$ 换为 $-x_i \geq -u_i$, 将基不等式 $h_i(x, \lambda) \geq 0$ 换为 $-h_i(x, \lambda) \geq 0$. 由于这两个基不等式对应的等式均是 $h_i(x, \lambda) = 0$, 基本解不变, 所以各非基向量偏差不变. 同样, 若 $-e_i$ 出现于当前表中, 偏差为正, 但 e_i 的偏差为负并以 e_i 入基, 则可将 $-e_i$ 换为 e_i , 将 $-h_i$ 换为 h_i , 我们将这样的替换称为向量替换.

不难证明, 将一个正半定矩阵的第 i 行和第 i 列所有元素反号 (主对角线上第 i 个元素不变), 得到的矩阵仍然正半定, 所以旋转算法表的基本性质与变量没有上界的情况一样.

以下算法仍然按最小偏差规则选择入基向量, 这些向量不仅包括表中所列的非基向量, 还包括未列出的非基单位向量或负单位向量.

算法 6.2.1 不等式组 (6.2.3) 的计算步骤.

步骤 1 初始步骤. 以 $x_1 \geq 0, \dots, x_n \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式, 列出表 6.12 所示初始表.

步骤 2 预处理. 设 $r_1 = \min\{r_1, r_2, \dots, r_n\}$, $r_2 = \max\{r_1, r_2, \dots, r_n\}$, 进行以下 4 次旋转运算:

$$e \leftrightarrow e_1, h_1 \leftrightarrow e_{n+1}, r \leftrightarrow e_2, h_2 \leftrightarrow e_{n+2}.$$

步骤 3 主要迭代 (按最小偏差规则).

(i) 若所有非基向量偏差非负, 停止; 否则转 (ii).

(ii) 以偏差最小的非基向量入基. 若该向量未列于表中, 则进行一次向量替换. 如果同行主对角线上的元素为正, 以其为枢轴进行一次旋转运算; 转 (i); 否则以该行最大正元素及其对称元素为枢轴先后进行两次旋转运算; 转 (i).

下面说明资产组合的预期收益率 r_p 在什么范围内变化时, 模型 (4.2.1) 有可行解.

设 $r_1 \leq r_2 \leq \dots \leq r_n$, 将 1 依次分配给 $x_1, x_2, \dots, x_n$, 使它们尽可能达到上界, 如 $x_1 = u_1, x_2 = u_2, \dots, x_{k-1} = u_{k-1}, x_k = v$, 其中 $u_1 + u_2 + \dots + u_{k-1} + v = 1$, 则 r_p 的最小值为 $r_1 u_1 + r_2 u_2 + \dots + r_{k-1} u_{k-1} + r_k v$. 同样将 1 依次分配给 $x_n, x_{n-1}, \dots, x_1$ 可求出 r_p 的最大值.

下面用一个例子说明资产有上解限制的均值方差模型的解法.

例 6.2.1 在上节的例 6.1.1 中,如果要求每种资产的比例不超过 0.5,资产组合的预期收益率 $r_p = 0.08$,求一个最小风险资产组合.

解 先不考虑变量的上界约束,即按模型(6.1.1)求解.初始步骤及预处理与例 6.1.1 相同,结果见表 6.10,但不包括 $r_p = 0.10$ 和 $r_p = 0.11$ 这两列.由于仅要求计算一个最优解,该表的 e_4 和 e_5 这两行以及 e 和 r 这两列可删掉.

表 6.10 中,非基向量 e_3, e_1, e_2 的偏差都是正数,需计算 $-e_3, -e_1, -e_2$ 的偏差.由于 $u_i = 0.5 (i = 1, 2, 3)$ 这三个向量的偏差分别为

$$0.5 - 0.3239 = 0.1761, 0.5 - 0.5047 = -0.0047, 0.5 - 0.1714 = 0.3286.$$

非基向量 $-e_1$ 的偏差是负数, $-e_1$ 入基.但 $-e_1$ 未出现于表 6.10 中,所以将表中的 e_1 换为 $-e_1$,将该行的偏差 0.5047 换为 -0.0047,并将该行其余数字反号.相应地,将基向量 h_1 换为 $-h_1$ 并将该列所有数字反号.结果如表 6.13 所示,其中 $\bar{\sigma}_i$ 列的 0.1716, 0.5047, 0.3286 分别是 $-e_3, e_1, -e_2$ 的偏差.

表 6.13 向量替换的结果

	h_3	$-h_1$	h_2	σ_i	$\bar{\sigma}_i$
e_3	0.7092	0.3546	-0.3546	0.3239	0.1761
$-e_1$	0.3546	0.1773*	-0.1773	-0.0047	0.5047
e_2	-0.3546	-0.1773	0.1773	0.1714	0.3286

以 $-e_1$ 入基 $-h_1$ 出基的结果见表 6.14.

表 6.14 最后结果

	h_3	$-e_1$	h_2	σ_i	$\bar{\sigma}_i$
e_3	0	2	0	0.3333	0.1667
$-h_1$	-2	5.64	1	0.0267	—
e_2	0	-1	0	0.1667	0.3333

表中 e_3 和 e_2 的偏差是正数,可算得 $-e_3$ 和 $-e_2$ 的偏差分别为 0.1667 和 0.3333,也是正数.所以最优资产组合为 $x_1 = u_1 = 0.5, x_2 = 0.1667, x_3 = 0.3333$.其方差风险为

$$(0.5, 0.1667, 0.3333) \begin{pmatrix} 0.2 & 0.3 & -0.01 \\ 0.3 & 2.4 & 0.5 \\ -0.01 & 0.5 & 1.1 \end{pmatrix} \begin{pmatrix} 0.5 \\ 0.1667 \\ 0.3333 \end{pmatrix} = 0.3411.$$

6.3 具有交易成本的均值方差模型

交易成本可理解为购买某种资产(如股票)所支付的金额超过其现值的部分. 购买量较少时, 交易成本一般是购买金额的线性函数, 这时的均值方差资产组合选择模型仍然是二次规划. 如果购买量很大, 往往可以得到折价购买的优惠, 即单位成本将减少. 但是如果购买量太大, 引起供给短缺, 单位成本又会上升.

设有 n 种风险资产可供选择. 这 n 种资产收益率的均值为 $r_1, r_2, \dots, r_n$, 协方差矩阵为 $H = (\sigma_{ij})$, 其中 σ_{ij} 是第 i, j 两种资产收益率的协方差. 第 i 种资产的购买金额为 x_i , 其交易成本记为 $c_i(x_i)$. 现要求从这 n 种资产中选购若干种资产, 使投资收益率达到一定水平同时方差风险尽可能小. 此问题可用以下模型表示.

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} x^T H x, \\ \text{s. t. } g_1(x) &= x_1 + x_2 + \dots + x_n - 1 = 0, \\ g_2(x) &= \sum_{i=1}^n (r_i x_i - c_i(x_i)) - r_p = 0, \\ x_1, x_2, \dots, x_n &\geq 0, \end{aligned} \right\} \quad (6.3.1)$$

其中成本函数 $c_i(x_i)$ 可以是凸的、凹的或非凸非凹的函数, 可利用实际数据拟合. 图 6.2 显示的是凹成本函数.

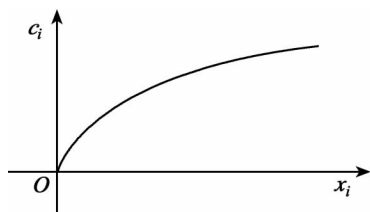


图 6.2 凹成本函数

根据笔者的实践经验, 序列二次规划法中的增广拉格朗日法适用于多种形式的成本函数. 这要求 $c_i(x_i)$ 在包含区间 $[0, 1]$ 的一个开区间内一阶和二阶可微. 给定 $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})^T$ 和 $\lambda^{(k)} = (\lambda_1^{(k)}, \lambda_2^{(k)})^T$, 子问题为

$$\begin{aligned}
& \min \frac{1}{2} s^T \nabla_x^2 P(x^{(k)}, \lambda^{(k)}, 0) s + \nabla f(x^{(k)})^T s, \\
& \text{s. t. } \nabla g_1(x^{(k)})^T s + g_1(x^{(k)}) = 0, \\
& \quad \nabla g_2(x^{(k)})^T s + g_2(x^{(k)}) = 0, \\
& \quad s_i \geq -x_i^{(k)}, \quad i = 1, 2, \dots, n.
\end{aligned} \tag{6.3.2}$$

以上模型的 KT 条件为

$$\left. \begin{aligned}
& \nabla_x^2 P(x^{(k)}, \lambda^{(k)}, 0) s - \lambda_1 \nabla g_1(x^{(k)}) - \lambda_2 \nabla g_2(x^{(k)}) + \nabla f(x^{(k)}) \geq 0, \\
& [\nabla_x^2 P(x^{(k)}, \lambda^{(k)}, 0) s - \lambda_1 \nabla g_1(x^{(k)}) - \lambda_2 \nabla g_2(x^{(k)}) + \nabla f(x^{(k)})]_i (s_i + x_i^{(k)}), \\
& \quad i = 1, 2, \dots, n, \\
& \nabla g_1(x^{(k)})^T s + g_1(x^{(k)}) = 0, \\
& \nabla g_2(x^{(k)})^T s + g_2(x^{(k)}) = 0, \\
& s_i \geq -x_i^{(k)}, \quad i = 1, 2, \dots, n.
\end{aligned} \right\} \tag{6.3.3}$$

其中 $[\cdot]_i$ 表示一个向量的第 i 个分量,

$$g_1(x^{(k)}) = \sum_{i=1}^n x_i^{(k)} - 1, \quad g_2(x^{(k)}) = \sum_{i=1}^n (r_i x_i^{(k)} - c_i(x_i^{(k)})) - r_p,$$

$$\nabla f(x^{(k)}) = H x^{(k)}, \quad \nabla g_1(x^{(k)}) = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \quad \nabla g_2(x^{(k)}) = \begin{pmatrix} r_1 - c'_1(x_1^{(k)}) \\ r_2 - c'_2(x_2^{(k)}) \\ \vdots \\ r_n - c'_n(x_n^{(k)}) \end{pmatrix}$$

$$\nabla_x^2 P(x^{(k)}, \lambda^{(k)}, 0) = \nabla^2 f(x^{(k)}) - \lambda_2^{(k)} \nabla^2 g_2(x^{(k)}) + r (\nabla^2 (g_1(x^{(k)}))^2 + \nabla^2 (g_2(x^{(k)}))^2),$$

其中 r 是个常数 (如 $r = 1$),

$$\begin{aligned}
\nabla^2 f(x^{(k)}) &= H, \quad \nabla^2 g_2(x^{(k)}) = \begin{pmatrix} -c''_1(x_1^{(k)}) & & & \\ & -c''_2(x_2^{(k)}) & & \\ & & \ddots & \\ & & & -c''_n(x_n^{(k)}) \end{pmatrix}, \\
\nabla^2 (g_1(x^{(k)}))^2 &= 2 \begin{pmatrix} 1 & 1 & \dots & 1 \\ 1 & 1 & \dots & 1 \\ \vdots & \vdots & & \vdots \\ 1 & 1 & \dots & 1 \end{pmatrix}_{n \times n},
\end{aligned}$$

其中 $c'_i(x_i)$ 和 $c''_i(x_i)$ 是 $c_i(x_i)$ 的一阶和二阶导数, $\nabla^2 (g_2(x^{(k)}))^2$ 是 n 阶方

阵,其第 i 行第 j 列的元素为

$$\frac{\partial^2 (g_2(x^{(k)}))^2}{\partial x_i \partial x_j} = 2 \left(\frac{\partial g_2(x^{(k)})}{\partial x_i} \frac{\partial g_2(x^{(k)})}{\partial x_j} + g_2(x^{(k)}) \frac{\partial^2 g_2(x^{(k)})}{\partial x_i \partial x_j} \right), \quad i, j = 1, 2, \dots, n.$$

初始解 $x^{(0)}$ 和 $\lambda^{(0)}$ 根据具体情况给定,例如可令 $x^{(0)} = 0$, $\lambda^{(0)} = 0$. 设模型 (6.3.3) 的解为 $s^{(k)}$ 和 $\lambda^{(k+1)}$, 下一个子问题的初始解为 $x^{(k+1)} = x^{(k)} + s^{(k)}$ 和 $\lambda^{(k+1)}$.

为了简单起见,仅介绍一个很小的算例.

例 6.3.1 5 种证券的平均收益率为 0.16, 0.11, 0.08, 0.13, -0.02, 样本协方差矩阵

$$H = \begin{pmatrix} 0.56 & 0.11 & 0.09 & 0.08 & 0.13 \\ 0.11 & 0.32 & 0.2 & 0.14 & 0.11 \\ 0.09 & 0.2 & 0.22 & 0.13 & 0.12 \\ 0.08 & 0.14 & 0.13 & 0.48 & 0.11 \\ 0.13 & 0.11 & 0.12 & 0.11 & 0.24 \end{pmatrix},$$

交易成本 $c_i(x_i) = 0.002x_i + 0.002(1 - e^{-x_i})$, $i = 1, 2, 3, 4, 5$. 要求构造一个证券组合使预期收益率为 10%, 风险尽可能小.

解 可以看出交易成本是个凹函数,下面用拉格朗日法计算.

(i) 令 $x^{(0)} = (0, 0, 0, 0, 0)^T$, $\lambda^{(0)} = (0, 0)^T$, $r \equiv 1$. $k = 0$ 时,子问题 (6.3.2) 的系数如下:

$$\nabla_x^2 P(x^{(0)}, \lambda^{(0)}, \rho) = \begin{pmatrix} 2.6083 & 2.1431 & 2.1137 & 2.1193 & 2.1225 \\ 2.1431 & 2.3421 & 2.2161 & 2.1667 & 2.1049 \\ 2.1137 & 2.2161 & 2.2312 & 2.1492 & 2.1164 \\ 2.1193 & 2.1667 & 2.1492 & 2.5114 & 2.1040 \\ 2.1225 & 2.1049 & 2.1164 & 2.1040 & 2.2408 \end{pmatrix},$$

$$\nabla f(x^{(0)}) = 0, \nabla g_1(x^{(0)}) = (1, 1, 1, 1, 1)^T, g_1(x^{(0)}) = -1,$$

$$\nabla g_2(x^{(0)}) = (0.156, 0.106, 0.076, 0.126, -0.024)^T, g_2(x^{(0)}) = -0.1,$$

$$l_b^{(0)} = (0, 0, 0, 0, 0)^T$$

其中 $l_b^{(k)}$ 表示第 k 个子问题各变量的下界构成的向量. 列出初始表后经过 7 次旋转运算, 结果为

$$s^{(0)} = (0.2198309, 0.116766, 0.4099581, 0.1883671, 0.06507802)^T,$$

$$\lambda^{(1)} = (2.136178, 0.5853165)^T,$$

$$x^{(1)} = x^{(0)} + s^{(0)}$$

$$= (0.2198309, 0.116766, 0.4099581, 0.1883671, 0.06507802)^T.$$

(ii) $k = 1$ 时, 子问题(6.3.2)的系数如下:

$$\nabla_x^2 P(x^{(1)} | \lambda^{(1)} | \rho) = \begin{pmatrix} 2.6080 & 2.1432 & 2.114 & 2.1195 & 2.1225 \\ 2.1432 & 2.3415 & 2.2163 & 2.1668 & 2.1049 \\ 2.114 & 2.2163 & 2.2310 & 2.1494 & 2.1163 \\ 2.1195 & 2.1668 & 2.1494 & 2.5110 & 2.1040 \\ 2.1225 & 2.1049 & 2.1163 & 2.1040 & 2.24 \end{pmatrix},$$

$$\nabla f(x^{(1)}) = \begin{pmatrix} 0.1964 \\ 0.1771 \\ 0.1656 \\ 0.1848 \\ 0.1270 \end{pmatrix},$$

$$\nabla g_1(x^{(1)}) = (1, 1, 1, 1, 1)^T, \quad g_1(x^{(1)}) = 0,$$

$$\nabla g_2(x^{(1)}) = (0.1564, 0.1062, 0.0767, 0.1263, -0.0239)^T, \quad g_2(x^{(1)}) \approx 0,$$

$$l_b^{(1)} = (-0.2198, -0.1168, -0.4010, -0.1884, -0.0651)^T.$$

经过 7 次旋转运算 结果为

$$s^{(1)} = (-6.204247 \times 10^{-4}, -7.857531 \times 10^{-4}, 4.282296 \times 10^{-4}, -4.783869 \times 10^{-4}, 1.456264 \times 10^{-3})^T,$$

$$\lambda^{(2)} = (0.1362734, 0.3822526)^T,$$

$$x^{(2)} = x^{(1)} + s^{(1)} = (0.2192104, 0.1159802, 0.4103863, 0.1878887, 6.653427 \times 10^{-2})^T.$$

(iii) $k = 2$ 时, 子问题(6.3.2)的系数如下:

$$\nabla_x^2 P(x^{(2)} | \lambda^{(2)} | \rho) = \begin{pmatrix} 2.6083 & 2.1432 & 2.114 & 2.1195 & 2.1225 \\ 2.1432 & 2.3419 & 2.2163 & 2.1668 & 2.1049 \\ 2.114 & 2.2163 & 2.2313 & 2.1494 & 2.1163 \\ 2.1195 & 2.1668 & 2.1494 & 2.5113 & 2.1040 \\ 2.1225 & 2.1049 & 2.1163 & 2.1040 & 2.2404 \end{pmatrix},$$

$$\nabla f(x^{(2)}) = \begin{pmatrix} 0.1961 \\ 0.1769 \\ 0.1656 \\ 0.1846 \\ 0.1271 \end{pmatrix},$$

$$\nabla g_1(x^{(2)}) = (1, 1, 1, 1, 1)^T, \quad g_1(x^{(2)}) = 0,$$

$$\nabla g_2(x^{(2)}) = (0.1564, 0.1062, 0.0767, 0.1263, -0.0239)^T, \quad g_2(x^{(2)}) = 0,$$

$$l_b^{(2)} = (-0.2192, -0.1160, -0.4104, -0.1879, -0.06653)^T.$$

经过 7 次旋转运算 结果为

$$s^{(2)} = (7.003546 \times 10^{-7}, 2.935529 \times 10^{-6}, -5.275011 \times 10^{-6}, 2.235174 \times 10^{-7}, 1.393259 \times 10^{-6})^T,$$

$$\lambda^{(3)} = (0.1362741, 0.3827342)^T,$$

$$x^{(3)} = x^{(1)} + s^{(1)}$$

$$= (0.219211, 0.1159831, 0.410381, 0.1878892, 6.653566 \times 10^{-2})^T.$$

由于 $\|s^{(2)}\| < 10^{-4}$, 计算停止. 五种证券的比例为

$$x_1 = 21.92\%, x_2 = 11.60\%, x_3 = 41.04\%, x_4 = 18.79\%, x_5 = 6.65\%.$$

其方差风险为 0.1746 标准差 0.4179 预期收益率 10%. 以上共进行了 21 次旋转运算.

6.4 自融资均值方差模型

自融资投资组合是净投资为零的投资组合. 设 n 种资产收益率的均值为 $r_1, r_2, \dots, r_n$, 协方差矩阵为 $H = (\sigma_{ij})_{n \times n}$, 其中 σ_{ij} 是第 i 种和第 j 种资产收益率的协方差. 设现有资产的总价值为 M_0 , 第 i 种资产的比例为 $d_i (\geq 0)$, $i = 1, 2, \dots, n$. 用 x_i 表示第 i 种资产比例的改变量, 自融资均值方差投资组合问题用以下模型表示.

$$\left. \begin{aligned} \min \quad & \frac{1}{2} x^T H x, \\ \text{s. t.} \quad & x_1 + x_2 + \dots + x_n = 0, \\ & r_1 x_1 + r_2 x_2 + \dots + r_n x_n \geq r_p, \\ & u_i \geq x_i \geq l_i, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (6.4.1)$$

其中 $x = (x_1, x_2, \dots, x_n)^T$, r_p 是投资组合预期收益率的下限, 各变量的下界和上界由投资者的意愿决定. 例如, 假设目前第一种资产的比例为 20%, 希望以后比例不超过 50%, 并且不准备卖空, 则 $l_1 = -0.2$, $u_1 = 0.3$; 若希望卖空, 但卖空金额不超过总资产的 10%, 则 $l_1 = -0.3$.

模型 (6.4.1) 是个凸二次规划且各变量有上界, 其解法与变量有上界的均值方差模型 (6.2.1) 类似. 首先不考虑所有变量的上界, 其 KT 条件为

$$\left. \begin{aligned} \sigma_{i1} x_1 + \dots + \sigma_{in} x_n - \lambda_1 - \lambda_2 r_i &\geq 0, \quad i = 1, 2, \dots, n, \\ x_1 + x_2 + \dots + x_n &= 0, \\ r_1 x_1 + r_2 x_2 + \dots + r_n x_n &\geq r_p, \\ \lambda_2 &\geq 0, \quad x_i \geq l_i, \quad i = 1, 2, \dots, n, \\ (\sigma_{i1} x_1 + \dots + \sigma_{in} x_n - \lambda_1 - \lambda_2 r_i)(x_i - l_i) &= 0, \quad i = 1, 2, \dots, n, \\ (r_1 x_1 + r_2 x_2 + \dots + r_n x_n - r_p) \lambda_2 &= 0. \end{aligned} \right\} (6.4.2)$$

引入人工不等式 $\lambda_1 \geq 0$, 记

$$h_i = (\sigma_{i1}, \dots, \sigma_{in}, -1, -r_i), e = (1, \dots, 1, 0, 0), r = (r_1, \dots, r_n, 0, 0),$$

并用 e_i 表示 $n+2$ 阶单位矩阵的第 i 行. 算法的初始基本不等式组为

$$x_1 \geq l_1, \dots, x_n \geq l_n, \lambda_1 \geq 0, \lambda_2 \geq 0.$$

初始基本解为 $y^{(0)} = \begin{pmatrix} x^{(0)} \\ \lambda^{(0)} \end{pmatrix}$, 其中 $x^{(0)} = (l_1, l_2, \dots, l_n)^T$, $\lambda^{(0)} = (0, 0)^T$. 初

始非基向量 h_i, e 和 r 的偏差分别为

$$\sigma_i = h_i y^{(0)} = \sum_{j=1}^n \sigma_{ij} l_j, \quad i = 1, 2, \dots, n; \quad \sigma_{n+1} = ex^{(0)} = \sum_{j=1}^n l_j,$$

$$\sigma_{n+2} = rx^{(0)} - r_p = \sum_{j=1}^n r_j l_j - r_p.$$

各非基向量关于基向量的组合系数及其偏差如表 6.15 所示.

表 6.15

初 始 表

	e_1	...	e_n	e_{n+1}	e_{n+2}	
h_1	σ_{11}	...	σ_{1n}	-1	$-r_1$	σ_1
$\vdots$	$\vdots$		$\vdots$	$\vdots$	$\vdots$	$\vdots$
h_n	σ_{n1}	...	σ_{nn}	-1	$-r_n$	σ_n
e	1	...	1	0	0	σ_{n+1}
r	r_1	...	r_n	0	0	σ_{n+2}

模型(6.4.1)仅有一个等式约束,其系数向量为 e ,所以首先将 e 换入基中. 出基向量可以是 $e_1, e_2, \dots, e_n$ 中的任何一个,不妨以 e_1 出基,为此进行 2 次交换: $e \leftrightarrow e_1, h_1 \leftrightarrow e_{n+1}$. 由于 e_1 是等式约束的系数向量, e_1 列和 e_{n+1} 行可以删掉.

以上是旋转算法的前两个步骤,即初始步骤和预处理. 至此我们还没有考虑约束条件 $-x_i \geq -u_i$, 其系数向量为 $-e_i, i = 1, 2, \dots, n$. 计算要点如下:

(i) 从正负非基单位向量中找出可能的入基向量.

设 e_i 是出现于当前表的非基向量,偏差为 σ_i . 如果 $\sigma_i < 0$, e_i 是一个可能的

入基向量. 如果 $\sigma_i > 0$, 但 $u_i - l_i - \sigma_i < 0$, $-e_i$ 是一个可能的入基向量. 同样, 若 $-e_i$ 是出现于当前表的非基向量, 偏差为 σ_i . 当 $\sigma_i < 0$ 时, $-e_i$ 是一个可能的入基向量. 当 $\sigma_i > 0$, 但 $u_i - l_i - \sigma_i < 0$, e_i 是一个可能的入基向量.

(ii) 向量替换.

设 e_i 是出现于当前表的非基向量, 偏差为 σ_i , 但入基向量为 $-e_i$. 将表中的 e_i 换为 $-e_i$, 将其偏差 σ_i 换为 $u_i - l_i - \sigma_i$, 将该行其余数字全部反号. 再将表中的 h_i 换为 $-h_i$, 并将该列所有数字反号. 同样, 若 $-e_i$ 是出现于当前表的非基向量, 偏差为 σ_i , 但入基向量为 e_i . 将表中的 $-e_i$ 换为 e_i , 将其偏差 σ_i 换为 $u_i - l_i - \sigma_i$, 将该行其余数字全部反号. 再将表中的 $-h_i$ 换为 h_i , 并将该列所有数字反号.

算法 6.4.1 自融资均值方差模型(6.4.1)的计算步骤.

步骤 1 初始步骤. 以 $x_1 \geq l_1, \dots, x_n \geq l_n, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不等式, 列表 6.15 所示初始表.

步骤 2 预处理. 任选一个 $i \in \{1, 2, \dots, n\}$, 进行 2 次旋转运算 $x \leftrightarrow e_i, h_i \leftrightarrow e_{n+1}$, 然后删掉入基向量 e 所在列和出基向量 e_{n+1} 所在行.

步骤 3 主要迭代(按最小偏差规则).

(i) 若所有非基向量偏差非负, 停止. 否则转(ii).

(ii) 以偏差最小的非基向量入基. 若该向量未列于表中, 则进行一次向量替换. 如果该行没有正元素, 原问题无可行解, 停止. 如果该行的对角元素为正, 以其为枢轴进行一次旋转运算. 转(i); 否则以该行最大正元素及其对称元素为枢轴先后进行两次旋转运算. 转(i).

下面用一个例子说明自融资均值方差投资组合模型的解法.

例 6.4.1 某投资者持有 A, B, C 三种资产分别为 20%, 30%, 50%, 另外他还对资产 D 感兴趣. 这 4 种资产收益率的均值分别为 3.5%, 8%, 6.25% 和 9%, 协方差矩阵为

$$H = \begin{pmatrix} 0.195 & 0.225 & 0.255 & 0.315 \\ 0.225 & 0.725 & 0.49 & 0.755 \\ 0.255 & 0.49 & 0.42125 & 0.5875 \\ 0.315 & 0.755 & 0.5875 & 0.86 \end{pmatrix}.$$

此投资者希望构造一个净投资额为零的投资组合, 使预期收益率至少为 3.2%. 他还希望仅对资产 A 卖空, 但卖空的比例不超过现有总资产的 10%. 试为之构造一个方差风险最小的投资组合.

解 此问题的模型为

$$\min f(x) = \frac{1}{2} x^T H x ,$$

$$\text{s. t. } x_1 + x_2 + x_3 + x_4 = 0 ,$$

$$0.035x_1 + 0.08x_2 + 0.0625x_3 + 0.09x_4 \geq 0.032 ,$$

$$0.8 \geq x_1 \geq -0.3 , \quad 0.7 \geq x_2 \geq -0.3 , \quad 0.5 \geq x_3 \geq -0.5 , \quad 1 \geq x_4 \geq 0 .$$

其中 $x = (x_1, x_2, x_3, x_4)^T$.

先不考虑变量的上界约束 其 KT 条件为

$$0.195x_1 + 0.225x_2 + 0.255x_3 + 0.315x_4 - \lambda_1 - 0.035\lambda_2 \geq 0 ,$$

$$0.225x_1 + 0.725x_2 + 0.49x_3 + 0.755x_4 - \lambda_1 - 0.08\lambda_2 \geq 0 ,$$

$$0.255x_1 + 0.49x_2 + 0.42125x_3 + 0.5875x_4 - \lambda_1 - 0.0625\lambda_2 \geq 0 ,$$

$$0.315x_1 + 0.755x_2 + 0.5875x_3 + 0.86x_4 - \lambda_1 - 0.09\lambda_2 \geq 0 ,$$

$$x_1 + x_2 + x_3 + x_4 = 0 ,$$

$$0.035x_1 + 0.08x_2 + 0.0625x_3 + 0.09x_4 \geq 0.032 ,$$

$$x_1 \geq -0.3 , \quad x_2 \geq -0.3 , \quad x_3 \geq -0.5 , \quad x_4 \geq 0 , \quad \lambda_2 \geq 0 ,$$

$$(0.195x_1 + 0.225x_2 + 0.255x_3 + 0.315x_4 - \lambda_1 - 0.035\lambda_2)(x_1 + 0.3) = 0 ,$$

$$(0.225x_1 + 0.725x_2 + 0.49x_3 + 0.755x_4 - \lambda_1 - 0.08\lambda_2)(x_2 + 0.3) = 0 ,$$

$$(0.255x_1 + 0.49x_2 + 0.42125x_3 + 0.5875x_4 - \lambda_1 - 0.0625\lambda_2)(x_3 + 0.5) = 0 ,$$

$$(0.315x_1 + 0.755x_2 + 0.5875x_3 + 0.86x_4 - \lambda_1 - 0.09\lambda_2)x_4 = 0 ,$$

$$(0.035x_1 + 0.08x_2 + 0.0625x_3 + 0.09x_4 - 0.032) \lambda_2 = 0 .$$

令

$$h_1 = (0.195, 0.225, 0.255, 0.315, -1, -0.035) ,$$

$$h_2 = (0.225, 0.725, 0.49, 0.755, -1, -0.08) ,$$

$$h_3 = (0.255, 0.49, 0.42125, 0.5875, -1, -0.0625) ,$$

$$h_4 = (0.315, 0.755, 0.5875, 0.86, -1, -0.09) ,$$

$$e = (1, 1, 1, 1, 0, 0) ,$$

$$r = (0.035, 0.08, 0.0625, 0.09, 0, 0) ,$$

$$e_1 = (1, 0, 0, 0, 0, 0) ,$$

$$e_2 = (0, 1, 0, 0, 0, 0) ,$$

$$e_3 = (0, 0, 1, 0, 0, 0) ,$$

$$e_4 = (0, 0, 0, 1, 0, 0) .$$

$$e_6 = (0, 0, 0, 0, 0, 1) .$$

引入人工不等式 $\lambda_1 \geq 0$ 并令

$$e_5 = (0, 0, 0, 0, 1, 0) .$$

以 $x_1 \geq -0.3, x_2 \geq -0.3, x_3 \geq -0.5, x_4 \geq 0, \lambda_1 \geq 0, \lambda_2 \geq 0$ 为初始基本不

同样 在表 6.18 中以 r 入基以 e_6 出基 结果见表 6.19.

表 6.19 第二次主旋转结果

	h_1	e_2	e_3	h_4	r	σ_i	$\dot{\sigma}_i$
e_1	0	- 0.1818	- 0.5	0	- 18.1818	0.0227	1.0773
h_2	0.1818	0.0836	0	0.8182	1.1322	0.0112	—
h_3	0.5	0	0	0.5	0	0	—
e_4	0	- 0.8182	- 0.5	0	18.1818	1.0773	- 0.0773
e_6	18.1818	1.1322	0	- 18.1818	140.4958	4.1562	—

由表 6.19 可见 ,只有 $- e_4$ 的偏差 $- 0.0773 < 0$,应以 $- e_4$ 入基. 但 $- e_4$ 未列在表中 ,因此将 e_4 换为 $- e_4$,将 e_4 的偏差 1.0773 换为 $- 0.0773$,将该行其他数字反号 ,再将 h_4 换为 $- h_4$,并将该列所有数字反号 ,结果见表 6.20.

表 6.20 向量替换的结果

	h_1	e_2	e_3	$- h_4$	r	σ_i	$\dot{\sigma}_i$
e_1	0	- 0.1818	- 0.5	0	- 18.1818	0.0227	1.0773
h_2	0.1818	0.0836	0	- 0.8182	1.1322	0.0112	—
h_3	0.5	0	0	- 0.5	0	0	—
$- e_4$	0	0.8182	0.5	0	- 18.1818	- 0.0773	1.0773
e_6	18.1818	1.1322	0	18.1818	140.4958	4.1562	—

在表 6.20 中 , $- e_4$ 行主对角线上的元素为 0 ,但 0.8182 是该行最大正元素. 以该元素及其对称元素为枢轴向后进行两次旋转运算 结果见表 6.21.

表 6.21 双枢轴运算结果

	h_1	$- e_4$	e_3	h_2	r	σ_i	$\dot{\sigma}_i$
e_1	0	- 0.2222	- 0.3889	0	- 22.2222	0.0056	1.0944
$- h_4$	0.2222	0.1249	- 0.0624	- 1.2222	3.6543	0.0233	—
h_3	0.3889	- 0.0624	0.0312 *	0.6111	- 1.8272	- 0.116	—
e_2	0	1.2222	- 0.6111	0	22.2222	0.0944	0.9056
e_6	22.2222	3.6543	- 1.8272	- 22.2222	232.0988	4.6864	—

在表 6.21 中只有 h_3 的偏差是负数,以 h_3 入基以 e_3 出基,结果见表 6.22.

表 6.22 第三次主旋转结果

	h_1	$-e_4$	h_3	h_2	r	σ_i	$\dot{\sigma}_i$
e_1	4.8443 *	- 1	- 12.4568	7.6125	- 44.9827	- 0.1394	1.2394
$-h_4$	1	0	- 2	0	0	0	—
e_3	- 12.4568	2	32.0317	- 19.5749	58.527	0.3729	0.6271
e_2	7.6125	0	- 19.5749	11.9624	- 13.5443	- 0.1334	1.1334
e_6	44.9827	0	- 58.527	13.5443	125.1606	4.0051	—

在表 6.22 中以 e_1 入基以 h_1 出基,结果见表 6.23.

表 6.23 第四次主旋转结果

	e_1	$-e_4$	h_3	h_2	r	σ_i	$\dot{\sigma}_i$
h_1	0.2064	0.2064	2.5714	- 1.5714	9.2857	0.0288	—
$-h_4$	0.2064	0.2064	0.5714	- 1.5714	9.2857	0.0288	—
e_3	- 2.5714	- 0.5714	0	0	- 57.1429	0.0143	0.9857
e_2	1.5714	1.5714	0	0	57.1429	0.0857	0.9143
e_6	9.2857	9.2857	57.1429	- 57.1429	542.8571	5.3	—

至此,所有非基向量偏差非负,问题的解为

$$\begin{aligned}x_1 &= l_1 = -0.3, \quad x_4 = u_4 = 1, \quad x_3 = 0.0143 + l_3 = 0.0143 - 0.5 = -0.4857, \\x_2 &= 0.0857 + l_2 = 0.0857 - 0.3 = -0.2143.\end{aligned}$$

所以自融资投资组合为(- 30% , - 21.43% , - 48.57% , 100%).

假设初始资产总值为 100 万元,由于初始投资组合为(20% , 30% , 50% , 0) ,应将资产 A 全部卖掉,还卖空 10 万元.资产 B 卖掉 21.43 万元,资产 C 卖掉 48.57 万元,所得款项用于购买资产 D.可以算出自融资投资组合的预期收益率 $r_p = 3.2\%$,标准差 $\sigma_p = 36.36\%$.

6.5 V 形交易成本的均值方差模型

随着市场形势的变化,投资组合需要不断调整.就收益和风险而言,调整投资组合的动机不外乎是:或者对当前投资组合的预期收益(率)不满意,希望提

高收益水平;或者认为风险太大,希望降低风险.下面我们考虑前一问题,即如何增加或减少一些资产,但不卖空,使新的投资组合有更高的收益且风险较小.无论是买进或是卖出资产都要支付交易费用,而且交易金额越大费用越高,所以成本函数呈现 V 字形.图 6.3 显示的是交易费用率为常数的 V 形交易成本.

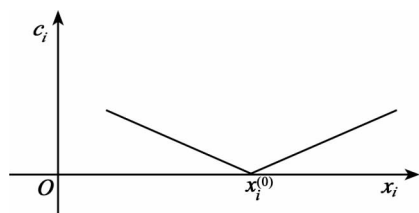


图 6.3 V 形交易成本

设某人持有 n 种证券, 市场价值分别为 $x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}$. 这 n 种证券收益率的均值为 $r_1, r_2, \dots, r_n$, 协方差矩阵为 $H = (\sigma_{ij})_{n \times n}$, 其中 σ_{ij} 是第 i, j 两种资产收益率的协方差. 设调整后各资产的金额为 $x_1, x_2, \dots, x_n$, 交易成本 $c_i(x_i) = \kappa |x_i - x_i^{(0)}|$, $i = 1, 2, \dots, n$. 其中 κ 为交易费用率, 如 $\kappa = 0.4\%$. 由于存在交易成本, 必然有 $\sum_{i=1}^n x_i < \sum_{i=1}^n x_i^{(0)}$, 但希望将来的资本盈余更多些. 问题的模型为

$$\left. \begin{aligned} \min f(x) &= \frac{1}{2} \frac{\sum_{i=1}^n \sum_{j=1}^n \sigma_{ij} x_i x_j}{\left(\sum_{i=1}^n x_i\right)^2}, \\ \text{s. t. } g_1(x) &= \sum_{i=1}^n (x_i + \kappa |x_i - x_i^{(0)}|) - \sum_{i=1}^n x_i^{(0)} = 0, \\ g_2(x) &= \sum_{i=1}^n (r_i x_i - \kappa |x_i - x_i^{(0)}|) - \beta \sum_{i=1}^n r_i x_i^{(0)} \geq 0, \\ x_i &\geq 0, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (6.5.1)$$

其中 $2f(x)$ 是投资组合的方差风险, $\sum_{i=1}^n r_i x_i^{(0)}$ 是当前投资组合的预期资本盈余, β 是一个大于 1 的数. 模型(6.5.1) 可用序列二次规划法中的牛顿法求解. 利用微分学知识可得

$$\frac{\partial f(x)}{\partial x_i} = \frac{\sum_{k=1}^n \sigma_{ik} x_k}{d^2} - \frac{\sum_{u=1}^n \sum_{v=1}^n \sigma_{uv} x_u x_v}{d^3}, \quad i = 1, 2, \dots, n,$$

$$\frac{\partial^2 f(\mathbf{x})}{\partial x_i \partial x_j} = \frac{\sigma_{ij}}{d^2} - \frac{2 \sum_{k=1}^n \sigma_{ik} x_k}{d^3} - \frac{2 \sum_{k=1}^n \sigma_{jk} x_k}{d^3} + \frac{3 \sum_{u=1}^n \sum_{v=1}^n \sigma_{uv} x_u x_v}{d^4}, \quad i, j = 1, 2, \dots, n.$$

其中 $d = \sum_{i=1}^n x_i$.

成本函数 $c_i(x_i) = \kappa |x_i - x_i^{(0)}|$ 在 $x_i^{(0)}$ 处不可微,是个“尖点”,但我们可以想象用一个光滑的曲线,如一个半径很小的圆弧或抛物线代替尖点,并使之与两边的直线段相切.不妨设成本函数在 $x_i^{(0)}$ 处的斜率为 0,所以

$$\frac{\partial g_1(x)}{\partial x_i} = \begin{cases} 1 + \kappa, & x_i > x_i^{(0)}, \\ 0, & x_i = x_i^{(0)}, \\ 1 - \kappa, & x_i < x_i^{(0)}; \end{cases} \quad \frac{\partial g_2(x)}{\partial x_i} = \begin{cases} r_i - \kappa, & x_i > x_i^{(0)}, \\ 0, & x_i = x_i^{(0)}, \\ r_i + \kappa, & x_i < x_i^{(0)}; \end{cases} \quad i = 1, 2, \dots, n.$$

对于给定的 $\mathbf{x}^{(k)} \in \mathbb{R}^n$, 子问题的模型为

$$\left. \begin{aligned} \min \quad & \frac{1}{2} \mathbf{s}^T \nabla^2 f(\mathbf{x}^{(k)}) \mathbf{s} + \nabla f(\mathbf{x}^{(k)})^T \mathbf{s}, \\ \text{s. t.} \quad & \nabla g_1(\mathbf{x}^{(k)})^T \mathbf{s} + g_1(\mathbf{x}^{(k)}) = 0, \\ & \nabla g_2(\mathbf{x}^{(k)})^T \mathbf{s} + g_2(\mathbf{x}^{(k)}) \geq 0, \\ & s_i \geq -x_i^{(k)}, \quad i = 1, 2, \dots, n. \end{aligned} \right\} \quad (6.5.2)$$

最好以初始投资组合作为 $\mathbf{x}^{(0)}$.

例 6.5.1 5 种证券的预期收益率和协方差矩阵见例 6.3.1. 当前投资组合中每种证券各占 10 万元,交易费用率为 0.4%. 求一个新的投资组合,使扣除交易成本后的预期资本盈余至少增加 5% (或 10%).

解 可以算得当前投资组合的预期资本盈余为 4.6 万元,方差风险为 0.1704. 以当前投资组合作为牛顿法的初始点,即 $\mathbf{x}^{(0)} = (10, 10, 10, 10, 10)^T$, $\beta = 1.05$ 时子问题的系数如下:

$$\nabla^2 f(\mathbf{x}^{(0)}) = \begin{pmatrix} 0.0001 & 0 & 0 & 0.0001 & 0 \\ 0 & 0.0001 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.0001 & 0 & 0 & 0.0001 & 0 \\ 0 & 0 & 0 & 0 & 0.0001 \end{pmatrix}, \quad \nabla f(\mathbf{x}^{(0)}) = \begin{pmatrix} 0.0005 \\ 0.0001 \\ -0.0004 \\ 0.0004 \\ -0.0006 \end{pmatrix},$$

以上的 0 在计算机上实际显示的是 0.0000 或 -0.0000. 另外

$$\nabla g_1(\mathbf{x}^{(0)}) = (0.996, 0.996, 0.996, 0.996, 0.996)^T, \quad g_1(\mathbf{x}^{(0)}) = 0,$$

$$\nabla g_2(\mathbf{x}^{(0)}) = (0.164, 0.114, 0.084, 0.134, -0.016)^T, \quad g_2(\mathbf{x}^{(0)}) = -0.23,$$

$$l_b = (-10, -10, -10, -10, -10)^T.$$

其中 l_b 是子问题模型(6.5.2)中各变量的下界构成的向量. 列出初始表后, 经过 7 次旋转运算得

$$s^{(0)} = (6.334019 \times 10^{-2}, -4.653862, 10.07246, -1.224302, -4.257633)^T,$$

$$x^{(1)} = x^{(0)} + s^{(0)} = (10.06334, 5.346138, 20.07246, 8.775698, 5.742367)^T.$$

同样求解其他子问题, 结果如表 6.24 所示, 其中 $x^{(k+1)} = x^{(k)} + s^{(k)}.$

表 6.24 子问题的初始解及计算结果

子问题	初始解 $x^{(k)}$ 及计算结果 $s^{(k)}$ 和 $x^{(k+1)}$
0	$x^{(0)} = (10, 10, 10, 10, 10),$ $s^{(0)} = (6.334019 \times 10^{-2}, -4.653862, 10.07246, -1.224302, -4.257633)$
1	$x^{(1)} = (10.06334, 5.346138, 20.07246, 8.775698, 5.742367),$ $s^{(1)} = (4.610062 \times 10^{-2}, 1.22952, -1.601856, 0.3972416, -0.1399269)$
2	$x^{(2)} = (10.10944, 6.575658, 18.4706, 9.172939, 5.60244),$ $s^{(2)} = (3.709793 \times 10^{-4}, -2.476692 \times 10^{-3}, 3.713608 \times 10^{-3}, -5.550385 \times 10^{-4}, -1.08242 \times 10^{-3})$
3	$x^{(3)} = (10.10981, 6.573182, 18.47432, 9.172384, 5.601357),$ $s^{(3)} = (1.907349 \times 10^{-6}, -9.059906 \times 10^{-6}, 7.629395 \times 10^{-6}, 9.536743 \times 10^{-7}, 2.384186 \times 10^{-6})$
4	$x^{(4)} = (10.10981, 6.573173, 18.47432, 9.172385, 5.60136)$

以上共经过 28 次旋转运算. 由于 $\|s^{(3)}\| < 10^{-4}$, 计算停止. 求得的投资组合为

$$(10.1098, 6.5732, 18.4743, 9.1724, 5.6014),$$

单位为万元, 其和为 49.9311 万元. 预期资本盈余 4.8989 万元, 扣除交易成本 0.06895 万元后为 4.8300 万元, 比当前投资组合的预期资本盈余 4.6 万元多 5%. 方差风险为 0.1707, 略高于当前值 0.1704.

如果令 $\beta = 1.1$, 得到的投资组合为 $(10.6807, 7.0731, 18.4653, 9.6229, 4.0845)$, 其和为 49.9265 万元. 预期资本盈余 5.1335 万元, 扣除交易成本 0.07346 万元后为 5.0600 万元, 比当前投资组合的预期资本盈余多 10%. 方差风险为 0.1741, 高于当前值 0.1704.

参 考 文 献

- [1] 曹志浩, 张玉得, 李瑞遐. 矩阵计算和方程求根. 北京: 人民教育出版社, 1979
- [2] 王德人. 非线性方程组解法与最优化方法. 北京: 人民教育出版社, 1980
- [3] 席少霖, 赵凤治. 最优化计算方法. 上海: 上海科学技术出版社, 1983
- [4] 赵瑞安, 吴方. 非线性最优化理论和方法. 杭州: 浙江科学技术出版社, 1992
- [5] 袁亚湘, 孙文瑜. 最优化理论与方法. 北京: 科学出版社, 2003
- [6] 余俊, 周济. 优化方法程序库 OPB-1——原理及使用说明. 北京: 机械工业出版社, 1989
- [7] 菲赫金哥耳茨著. 微积分学教程(第一卷第二分册). 叶彦谦等译. 北京: 人民教育出版社, 1964
- [8] 谢帮杰. 线性代数. 北京: 人民教育出版社, 1978
- [9] [日] 须田信英等著. 自动控制中的矩阵理论. 曹长修译. 北京: 科学出版社, 1979
- [10] 余湄, 董洪斌, 汪寿阳. 摩擦市场下的投资组合与无套利分析. 北京: 科学出版社, 2005
- [11] 张忠桢. 马科维兹资产组合选择模型的一种快速算法. 中国学术期刊文摘, 2001, 7 (5): 656-659
- [12] 陈铁英, 张忠桢. 自融资均值方差投资组合模型的旋转算法. 系统工程理论与实践, 2004, 24 (6): 98-103
- [13] 张忠桢. 凸规划——投资组合与网络优化的旋转算法. 武汉: 武汉大学出版社, 2004
- [14] Bazarra M S, Shetty C M. Nonlinear Programming: Theory and Algorithms. New York: John Wiley & Sons, 1979
- [15] Best M J. Equivalence of some quadratic programming algorithms. Mathe-

- mathematical Programming , 1984 , 30 (9) : 71-87
- [16] Chang Yow-Yieh , Cottle R W. Least-index resolution of degeneracy in quadratic programming. Mathematical Programming , 1980 , 18 : 127-137
- [17] Chvatal V. Linear Programming. New York : W. H. Freeman and Company , 1983
- [18] Dai Yuhong. Convergence properties of the BFGS algorithm. SIAM Journal on Optimization , 2002 , 13 (3) : 693-701
- [19] Feinstein C D , Thapa M N. Notes : a reformulation of mean-absolute deviation portfolio optimization model. Management Science , 1993 , 39 (12) : 1552-1553.
- [20] Fletcher R. Practical Method of Optimization : Constrained Optimization. New York : John Wiley & Sons , 1981
- [21] Fletcher R , Leyffer S. Nonlinear programming without a penalty function , 2002 , 91 (1) : 239-269
- [22] Francis J C. Investments : Analysis and Management. New York : McGraw-Hill , 1991
- [23] Hansen P , Jaumard B , Ruiz M , Xiong J. Global minimization of indefinite quadratic functions subject to box constraints. Naval Research Logistics , 1993 , 40 : 373-392
- [24] Konno H , Yamazaki H. Mean-absolute deviation optimization model and its applications to Tokyo stock market. Management Science , 1991 , 37 (5) : 519-531
- [25] Konno H , Wijayanayake A. Portfolio optimization under concave transaction costs and minimal transaction unit constraints. Mathematical Programming , 2001 , Ser. B 89 : 233-250.
- [26] Markowitz H M , Todd G P. Mean-Variance Analysis in Portfolio Choice and Capital Markets. Pennsylvania : Frank J. Fabozzi Associates , 2000
- [27] Markowitz H M , Todd G P , Xu G , Yamane Y. Fast computation of mean-variance efficient sets using historical covariance. Journal of Financial Engineering , 1992 , 1 (2) : 117-132
- [28] Mascarenhas W F. The BFGS method with exact line searches fails for non-convex objective functions. Mathematical Programming , 2004 , 99 : 49-61
- [29] Murty K G , Kabadi S N. Some NP-complete problems in quadratic and

- linear programming. Mathematical Programming , 1987 , 37 (11) : 117-129
- [30] Pang Jong-Shi. An equivalence between algorithms for quadratic programming. Mathematical Programming , 1981 , 20 : 152-165
- [31] Perold A F. Large-scale portfolio optimization. Management Science , 1984 , 30 (10) : 1143-1160
- [32] Powell M J D. On the quadratic programming algorithm of Goldfarb and Idnani. Mathematical Programming , 1985 , study 25 (10) : 46-61
- [33] Ulbrich S. On the superlinear local convergence of a filter-SQP method. Mathematical Programming , Ser. B , 2004 , 100 : 217-245
- [34] Vandenbussche D , Nemhauser G L. A branch-and-cut algorithm for non-convex quadratic programs with box constraints. Mathematical Programming , 2005 , 102 : 559-575
- [35] Vavasis S A. Approximation algorithm for indefinite quadratic programming. Mathematical Programming , Ser. A , 1992 , 57 : 279-311