自协和函数与多项式历时内点法

胡卫群 盛立人 杨明辉 蒋华松 编著

科学出版社

自协和函数与多项式时内点法

胡卫群　盛立人　杨明辉　蒋华松　编著

科学出版社
北京

内 容 简 介

本书旨在介绍最优化理论的内点法，重点介绍两位俄罗斯数学家 Nesterov 和 Nemirovski 所创造的采用自协和函数的内点法.正是这种新颖的方法使内点法几乎可应用于所有非线性规划问题.与 Nesterov 和 Nemirovski 的做法不同，我们在本书中采用 Renegar 方法，即参考内积的泛函方法来处理自协和性，这样做不但使叙述简明清晰，还有助于发掘内点法的理论深度，更能为微分几何学的研究提供一种新颖的方法.本书在讨论内点法的同时,也给出了这一理论在实用上 (见第 3 章和第 7 章)和理论上(见附录 B)的应用.

本书可作为计算机、数理、经济、工程等领域科研工作者的案头常备书，也可作为高校研究生及高年级本科生的参考书.

图书在版编目(CIP)数据

自协和函数与多项式历时内点法/胡卫群等编著. —北京：科学出版社，2012

ISBN 978-7-03-035457-0

Ⅰ.①自… Ⅱ.①胡… Ⅲ.①内点法-研究 Ⅳ.① O221

中国版本图书馆 CIP 数据核字(2012) 第 205311 号

责任编辑：顾 艳 胡 凯／责任校对：宋玲玲

责任印制：赵德静／封面设计：许 瑞

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

铭浩彩色印装有限公司印刷

科学出版社发行 各地新华书店经销

*

2012 年 12 月第 一 版 开本：B5 (720 × 1000)

2012 年 12 月第一次印刷 印张：16 3/4

字数：322 000

定价：59.00 元

前　　言

最优化理论中最基本的问题之一当推线性规划 (linear programming, LP) (Goemans, 1994; Goldfarb, Todd, 1989). 注意到 LP 在近二十年来以五个数量级的速度推进, 从这个意义上说, 我们断言 LP 在计算技巧与算法两方面已经产生了革命性的效应. 这也就意味着十几年前需要用一年时间才能解决的计算问题, 现在只需要几分钟.

长期以来, 人们对于 LP 的求解都选用 Dantzig 在 1947 年创造的单形法 (Dantzig, 1963, 1951). 一般认为它在求解诸如运筹、商业、经济与工程等复杂问题时十分可靠和有效. 单形法具备这种优越性的理论依据是计算复杂性理论, 该理论一直确信单形法是个多项式历时算法 (常称之为 P 问题), 即它的计算量有一个 (问题输入量的) 多项式的上界. 虽然后来被反例证明这个认定是不正确的, 但那样的反例在实用上极为稀少, 因而其影响微不足道. 事实上, 单形法在数学上的漏洞确实存在 (Gonzaga, 1989).

对于线性规划的 Dantzig 单形法, 历史上出现的第一次巅覆是 Khachiyan(1979) 的椭球法, 这件事甚至上了《纽约时报》的头条, 成为轰动一时的新闻.1979 年, 针对 LP 而不针对单形法, Khachiyan 利用处理非线性最优问题的椭球法, 证明了 LP 确实是一个多项式可解问题. 美中不足的是, 他的方法只有理论上的重大意义, 而对于一般问题的计算, 它的计算速度远比单形法要慢. 对 Dantzig 单形法构成真正巅覆的是 Karmarkar(1984) 给出的另一个针对 LP 的方法, 正是这个方法开了处理 LP 问题、甚至更复杂问题的先河, 这就是早先被称为投影法、后来被称为内点法 (interior point method, IPM) 的 Karmarkar 算法 (Anstreicher, Bosch, 1992). 首先, 此法是个多项式历时可解的运算; 其次, 它的计算幅度的上界远比 Khachiyan 的更精确; 最后, 它在实用上更高效! 内点法一词也从此得以正名. 对 Karmarkar 算法深入研究后还发现, 此法还可以用于二次规划、半定规划 (semidefinite programming, SDP) 等单形法无能为力的非线性问题.

可以说, 多项式历时内点法的近代理论是从 1984 年 N. Karmarkar 的讨论班论文开始的. 此后近十年间有近百位专家在这个领域内工作, 并涌现了近千份论文与预印本涉及该课题. 内点法的电子版文献库由 E. Kranich 编著, 那里已包含了

近 1500 个条目. 这些研究虽然大多集中于理论方面, 但实用方面的成果也相当可观, 这里只需举两个例子: 其一是线性规划中的内点法曾成功地帮助美国国防部在海湾战争中作出正确的物流计划; 另一个例子是, 该方法又成为非常有效又非常普及的软件包 OSL2 中的核心技术. 对此, 线性规划中单形法的创始人 Dantzig 曾说: "现在 (编著者注: 1990 年), 内点算法已经公开与各种单形法分庭抗礼." 这句话的意思是, 诞生近 50 年来被公认为强有力的单形法正面临着新生方法的严峻挑战.

说到 "对于 LP 的多项式历时算法", Khachiyan 证明对于一个具有 m 个约束、n 个变量的有理系数 LP 问题, 可以只用 $O(n^3(n+m)L)$ 个算术运算求解, 其中 L 是问题的输入长度, 即特定问题的数据的二进位总长度. 而 Karmarkar 的新方法的复杂度界限是 $O(m^{3/2}n^2L)$ 个运算. 假使我们在标准化要求下令 n 与 m 有同一阶量, 即 $m=O(n)$, 则前者是 $O(n^4L)$, 后者是 $O(n^{3.5}L)$, 所以说这是一种改进. 因此, Karmarkar 方法是个真正意义上的多项式可解算法.

另一方面, J. Renegar(1988) 发表了他对 LP 的循径 (path-following) 多项式历时内点算法, 其复杂性估计比 Karmarkar 方法更好, 即维数的立方, $O(n^3L)$. 这个结果相当于求解一个线性方程组. 这是迄今为止对 LP 所做的理论复杂性估计中最好的结果. 后来 C. C. Gonzaga 也得到了同一个结果.

据 Wright(1997) 的判断, 迄今为止已经创造出的有关 LP 的内点法总数已超过一百种. 全体内点法本质上大致可以分成四类: 循径法、仿射度量法、投射势函数法与仿射势函数法. 但真正使人们对内点法另眼相看的, 却是它关于半定规划理论的应用. SDP 的研究具有非常重要的理论与应用价值 (Boyd, Vandenberghe, 1996, 1995): 第一, 正半定 (或正定) 约束直接出现在大量的重要应用 (特别是工程问题) 中; 第二, 许多凸最优化问题, 例如 LP 与 (凸) 二次约束的二次规划, 都可以当作 SDP 问题处理, 因此 SDP 提供了一个统一方法来研究一大类凸最优化问题及其算法. 更重要的是, 不论是在理论上还是在实用上, SDP 均可以有效地求解.

从上述历史进程可以看出, 一切都是围绕着 LP 在进行, 我们至多是将有关方法扩充到与 LP 相近的线性约束二次规划等问题上. 若是从凸规划的角度来说, 至关重要的是, 应当如何在保持多项式历时的前提下, 拓宽内点法的应用范畴. 毫无疑问, 这绝对是一项挑战.

正当不少人踌躇满志于 LP 内点法的时候, 有两个独具慧眼的人却在那里思索. Nesterov 与 Nemirovski 发现 Karmarkar 等提出的内点障碍函数法只要略加修正便可用到更广泛的凸规划 [如半定规划、几何规划以及更一般的锥规划 (conic optimization)] 中去, 但要完成这个突破, 关键中的关键是要加细对经典 Newton 逼近法的估值. 完成这个突破的是一项堪称里程碑的工作: Nesterov 和 Nemirovski 于 1993 年出版了 Interior-Point Polynomial Algorithms in Convex Programming(《凸规划的多项式历时内点算法》). 在书中他们提出了一种被称为自协和性的数学性质, 并且证明, LP 的所有内点法原则上可以推广到所有凸最优问题中去. 具体说来, 他们在运用内点法中的 Newton 逼近时, 发现当障碍 (或罚) 函数具有这一类自协和性时, Newton 逼近将自然地产生多项式历时, 这束思想火花对最优化理论的计算复杂性研究无疑是一项重要突破. 现在的问题是, 并非每一个障碍函数都具有自协和性, 但内点法中所有常用的障碍函数 (如对数障碍、熵障碍等) 却都具有自协和性, 这就找到了迄今为止已有 LP 与 SDP 的内点法都是多项式历时的成因. Nesterov 和 Nemirovski 进而计算大量新的障碍函数, 它们都具有自协和性且二阶导数具有可计算性, 从而大大地扩充了内点法的应用范围, 使他们的理论得以进一步运用到锥规划、非凸规划、几何规划以及微分不等式等最优化理论中去. 这些理论都总结在上述那本名著里. 最后, 为了能用于实际问题, 障碍函数 (毋宁说是它的一阶与二阶导数) 必须具有可计算性. Nesterov 和 Nemirovski 又从理论上证明, 对每一凸集 (或一个锥), 这样的自协和障碍函数总是存在的 (Nemirovski, Todd, 2009; Nemirovski, 2004; Nesterov, Nemirovski, 1991, 1990a, 1990b, 1987, 1985).

现在我们可以说, Khachiyan、Karmarkar 与 Renegar 真正结束了人们对 LP 的遐想; 而 Nesterov 与 Nemirovski(1993) 在他们的名著中完成了对经典 Newton 逼近法加细估值的突破, 使凸规划理论的计算复杂性研究获得了巨大的飞跃, 堪称 “内点革命”.

上述名著能引起数学家们的兴趣, 还因为自协和性提供了微分几何内蕴性质研究的新路径, 但计算学家们却囿于那本名著的难读. 看来第一个读懂的人还是 Renegar, 他在 2001 年发表了着重于研究自协和泛函的分析讨论的一本小册子 (Renegar, 2001), 真正弥补了 Nesterov 与 Nemirovski 的原著许多不易看懂的想法和概念. 现在我们对自协和性本身的兴趣已经超过了内点法, 这是因为我们一直在期待有人能将前一时期非常流行的 LLL 算法 (由 A. K. Lenstra, H. W. Lenstra, L. Lovasz 三人提出) 与自协和性联系起来加以研究.

为使同行了解自协和性在多项式历时内点法中的作用, Nemirovski 曾在世界不少地方作了多次演讲 (Nemirovski, Todd, 2009; Nemirovski, 2004), 但那些演讲仍不足以真正体现他和 Nesterov 的基本思路. 实际上, 读了 Renegar 的书与 Nemirovski 的演讲稿, 你就可以顺利读懂 Nesterov 与 Nemirovski 的原著, 这样一来, 你就可以

彻底弄清楚多项式历时内点法.

本书将循着上述有关内点法的发展史, 陈述内点法与自协和函数论. 本书前三章旨在介绍内点法对 LP 及 SDP 的应用, 以及这些应用的应用 (熟悉内点法的读者可以直接跳过前两章). 本书的中心部分是第 4 章 $\sim$ 第 6 章, 我们将用自协和泛函研究内点法, 因此可将这一讨论视为内点法的最现代化陈述 (主要限于 Renegar 的研究). 第 7 章旨在讨论由 Nesterov 与 Nemirovski 首创的关于自协和函数的历时内点法在非线性凸规划中的许多应用. 为了更好地说明自协和函数的真正价值, 我们在本书附录 A 中介绍 Nesterov 和 Nemirovski 对于自协和函数与多项式历时内点法的原始理念以及将该方法用于凸规划的一些较新的结果. 另一方面, 作为一个尝试, 在本书附录 B 中, 我们通过 Grassmann 流形上的向量场和矩阵微分方程, 用内点法给出线性互补问题解的明显表达式 (Güler, 1994; Cottle et al., 1992; Kojima et al., 1991; Megiddo, 1986). 我们就这样界定本书的范围, 并期望它能为内点方法提供一份材料较完整和观念较新颖的资料.

本书的撰写曾得到多方面的帮助：南京林业大学理学院学术团队的主要成员石小平、陈菊珍、宋伟灵、李媛媛、黄一维等参与了部分章节的编写工作; 美国西北大学工业工程 (IE) 专业的几位博士也为我们提供了大量的、特别是一些在国内不易得到的文献资料, 在此一并致谢. 清华大学冯克勤教授一直关注着我们团队的工作, 也曾给予许多指导和鼓励, 在此谨致最衷心的谢意.

由于作者水平有限, 书中必定有不当和错误之处, 敬请读者批评指正.

作 者

2012 年 2 月

目　　录

第 1 章　概论——凸规划的一般问题与内点法

由于本书旨在讨论内点法 (interior point method, IPM) 对于凸规划 (convex programming, CP) 更近代的发展, 因而出于历史原因, 我们在本章中将首先概述最一般的凸规划的 IPM. 有两个 IPM 方法将被讨论: 对数障碍法与中心路径法, 它们对线性规划 (linear programming, LP) 与半定规划 (semidefinite programming, SDP) 的处理是相当成功的, 因为在一般情况下, 这两个方法能保证收敛性，且在理论实现上也不难, 但要扩充这两个方法时却并不像处理 LP 与 SDP 那么容易 (Boyd, Vandenberghe, 1996; Hertog, 1993; Huard, 1967; Frisch, 1956, 1955a, 1955b).

1.1　问　题　表

本章所有问题的讨论均在实数域$\mathbf{R}$上进行.

首先列出本书涉及的最重要和最基本的问题. 先引入下面一些熟知的记号: $\mathbf{R}^n$ 表示 n 维欧氏空间, $M^{m\times n}$ 表示 $m\times n$ 矩阵空间, $S^{n\times n}$ 表示 n 阶对称矩阵空间.

一般的最优化问题可以陈述如下: 给一个区域 D 及一个目标函数 $f:\ D\to\mathbf{R}$, 试找出

$$\begin{aligned} &\min && f(\boldsymbol{x}), \\ &\text{使满足} && \boldsymbol{x}\in D. \end{aligned} \tag{1.1}$$

如果 D 为有限维区域, 则最优化问题 (1.1) 称为**有限维最优化问题**, 简称**有限维问题**. 本书主要讨论有限维问题, 其中 $D\subset\mathbf{R}^n$. 有限维问题又可作如下分类:

(1) **无约束最优化问题**(unconstrained optimization problems): 此时 $D=\mathbf{R}^n$.

(2) **约束最优化问题**(constrained optimization problems): 此时

$$D=\{\boldsymbol{x}:g_i(\boldsymbol{x})\leqslant 0,\forall i=1,2,\cdots,m\},\quad g_i:\mathbf{R}^n\to\mathbf{R}. \tag{1.2}$$

(3) **判定问题**(decision problems): D 为一可行集, 对任一数 α,

$$D\subset\mathbf{R}^n,\quad f:D\to\mathbf{R},\quad f(\boldsymbol{x})<\alpha,\quad \forall\boldsymbol{x}\in D \tag{1.3}$$

是否成立?

(4) **可行性问题**(feasibility problems): 下列集合

$$\hat{D} = \{\boldsymbol{x} \in D : f(\boldsymbol{x}) \geqslant \alpha\} \tag{1.4}$$

是否为空集?

(5) **线性规划问题**: 设 $\boldsymbol{x}, \boldsymbol{c} \in \mathbf{R}^n, \boldsymbol{b} \in \mathbf{R}^m, \boldsymbol{A} \in \mathbf{R}^{m\times n}$, 求解

$$\begin{aligned} \min \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{使满足} \quad & \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}. \end{aligned} \tag{1.5}$$

(6) **二次规划问题**(quadratic programs, QP):

$$\begin{aligned} \min \quad & f(\boldsymbol{x}), \\ \text{使满足} \quad & \boldsymbol{A}\boldsymbol{x} \geqslant \boldsymbol{b}, \end{aligned} \tag{1.6}$$

其中 $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{D}\boldsymbol{x} + 2\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \boldsymbol{D} \in S^{n\times n}$.

(7) **半定规划问题**: 求解

$$\begin{aligned} \min \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}, \\ \text{使满足} \quad & \boldsymbol{F}_0 + \sum_{i=1}^{n} x_i \boldsymbol{F}_i \geqslant 0, \end{aligned} \tag{1.7}$$

其中 $\boldsymbol{F}_i \in S^{n\times n}, i = 0, 1, \cdots, n$.

注　下列问题又称为线性矩阵不等式问题:

集合 $\{\boldsymbol{x} : \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \in \mathbf{R}^n, \boldsymbol{F}_0 + \sum\limits_{i=1}^{n} x_i \boldsymbol{F}_i \geqslant 0\}$ 是否为空?

(8) **凸规划**(convex programs): 此时 f 为一般凸函数, D 为凸集合. 每一 LP 及 SDP 均为凸规划问题.

1.2　对数障碍法

本章先介绍两个方法: 对数障碍法与中心路径法, 可用于处理下面一般的最优化问题

$$\begin{aligned} \min \quad & f(\boldsymbol{x}), \\ \text{使满足} \quad & r_i(\boldsymbol{x}) \geqslant 0, i = 1, 2, \cdots, m, \end{aligned}$$

其中 $f : D(\subseteq \mathbf{R}^n) \to \mathbf{R}$ 为凸函数, $r_i : D \to \mathbf{R}$ 为凹函数, 并假设这 $m+1$ 个函数均二次可微.

以上问题是凸规划问题. 我们熟知, 在此假设下, 凸规划问题的任一局部极小均为全局极小.

给定凸规划问题, 定义对数障碍函数为

$$\phi(\boldsymbol{x}) := -\sum_{i=1}^{m} \ln r_i(\boldsymbol{x}),$$

它在可行域的内域 $\{\boldsymbol{x} \in D : r_i(\boldsymbol{x}) > 0, i = 1, 2, \cdots, m\}$ 中有定义, 但当 $\boldsymbol{x}$ 靠近边界点时, 对数障碍函数增至无穷大. 今后以 $\boldsymbol{x}^*$, f^* 分别表示极小点与极小值.

引理 1.1 对数障碍函数在可行域上为凸.

证明 只要证明每个 $-\ln r_i(\boldsymbol{x})$ 为凸即可, 但已设 r_i 为凹, 故对 $\forall \lambda \in [0,1]$ 有

$$r_i(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \geqslant \lambda r_i(\boldsymbol{x}) + (1-\lambda) r_i(\boldsymbol{y}).$$

因函数 $-\ln$ 为单调减且严格凸, 故对 $\forall \lambda \in [0,1]$, 有

$$-\ln r_i(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant -\ln[\lambda r_i(\boldsymbol{x}) + (1-\lambda) r_i(\boldsymbol{y})] \leqslant \lambda[-\ln r_i(\boldsymbol{x})] + (1-\lambda)[-\ln r_i(\boldsymbol{y})],$$

从而 $-\ln r_i(\boldsymbol{x})$ 为凸. □

所谓**对数障碍法**, 是通过求解一个无约束凸规划问题

$$\min \quad f(\boldsymbol{x}) + \frac{1}{\alpha}\phi(\boldsymbol{x}) \tag{1.8}$$

来求解约束的凸规划问题, 其中 α 任意给定. 对给定的 $\alpha > 0$, 目标函数 $f(\boldsymbol{x}) + \frac{1}{\alpha}\phi(\boldsymbol{x})$ 在可行域内部为凸且可微. 因此任一局部极小为全局极小, 而实际上, 极小可以用迭代方式求解. 项 $\frac{1}{\alpha}\phi(\boldsymbol{x})$ 作为一种障碍作用以 “排斥” 边界, 从而保证极小点恒在可行域内部. 这就是内点法的由来.

引理 1.2 设可行域 D 为有界, 则对 $\forall \alpha \in D$, 问题 (1.8) 的极小点 $\boldsymbol{x}_\alpha^*$ 存在, 且

$$\lim_{\alpha \to \infty} f(\boldsymbol{x}_\alpha^*) = f^*.$$

事实上对 $\forall \alpha > 0$ 有

$$0 \leqslant f(\boldsymbol{x}_\alpha^*) - f^* \leqslant \frac{m}{\alpha}, \tag{1.9}$$

其中 m 为凸规划问题中约束的个数.

证明 下面给一个有趣的证明, 我们利用这样的事实：对数障碍函数包含了 Lagrange 对偶关于规划问题下界 f^* 的信息. 记 $\boldsymbol{x}_\alpha^*$ 为极小点. 由可微性可知, 目标函数的梯度在 $\boldsymbol{x}_\alpha^*$ 处必为零, 即

$$\boldsymbol{\nabla}\left[f(\boldsymbol{x}) + \frac{1}{\alpha}\phi(\boldsymbol{x})\right]\Bigg|_{\boldsymbol{x}=\boldsymbol{x}_\alpha^*} = \boldsymbol{\nabla} f(\boldsymbol{x}_\alpha^*) - \sum_{i=1}^{m} \frac{1}{r_i(\boldsymbol{x}_\alpha^*)\alpha} \boldsymbol{\nabla} r_i(\boldsymbol{x}_\alpha^*) = \boldsymbol{0}. \tag{1.10}$$

但上述 $\boldsymbol{x}_\alpha^*$ 处的梯度也是另一个函数

$$f(\boldsymbol{x})-\sum_{i=1}^{m}\lambda_i r_i(\boldsymbol{x}),\qquad \text{其中 } \lambda_i:=\frac{1}{r_i(\boldsymbol{x})\alpha}$$

在 $\boldsymbol{x}_\alpha^*$ 处的梯度.

注意到 $\lambda_i(\boldsymbol{x}_\alpha^*)=1/[r_i(\boldsymbol{x}_\alpha^*)\alpha]$ 为正, 故上述函数为凸, 因此 $\boldsymbol{x}_\alpha^*$ 也是凸规划问题的极小点. 于是得出有关 f^* 的下界:

$$\begin{aligned} f^* &\geqslant \min_{\boldsymbol{x}\text{ 可行}}\left\{f(\boldsymbol{x})-\sum_{i=1}^{m}\lambda_i r_i(\boldsymbol{x})\right\}\quad [\lambda_i r_i(\boldsymbol{x})\geqslant 0]\\ &=f(\boldsymbol{x}_\alpha^*)-\sum_{i=1}^{m}\lambda_i r_i(\boldsymbol{x}_\alpha^*)\\ &=f(\boldsymbol{x}_\alpha^*)-\sum_{i=1}^{m}\frac{1}{\alpha}=f(\boldsymbol{x}_\alpha^*)-\frac{m}{\alpha}.\end{aligned}$$

由于 $f^*\leqslant f(\boldsymbol{x}_\alpha^*)$, 故引理得证. □

根据这些讨论, 我们给出算法如下:

算法 1.1　[SUMT (序列无约束极小技巧)]

Input 严格可行的 $\boldsymbol{x}^{(0)}$, 公差 $\varepsilon>0$, 初始障碍值 $\alpha^{(0)}$, 乘因子 $\beta>1$;

Output 一个可行值 $\boldsymbol{x}$, 使成立 $|f^*-f(\boldsymbol{x})|<\varepsilon$.

Assumptions 可行集为有界,

begin

$$\boldsymbol{x}=\boldsymbol{x}^{(0)},\quad \alpha:=\alpha^{(0)};$$

repeat

$$\boldsymbol{v}:=-\left[\boldsymbol{\nabla}^2 f(\boldsymbol{x})+\frac{1}{\alpha}\boldsymbol{\nabla}^2\phi(\boldsymbol{x})\right]^{-1}\left[\boldsymbol{\nabla} f(\boldsymbol{x})+\frac{1}{\alpha}\boldsymbol{\nabla}\phi(\boldsymbol{x})\right];\text{(牛顿方向)}$$

$$\delta^*:=\arg\min_{\delta}\left\{f(\boldsymbol{x}+\delta\boldsymbol{v})+\frac{1}{\alpha}\phi(\boldsymbol{x}+\delta\boldsymbol{v})\right\};\text{(线搜索)}$$

$$\boldsymbol{x}:=\boldsymbol{x}+\delta^*\boldsymbol{v};$$

until $||\boldsymbol{v}||$ 充分小;

return if

$$m/\alpha<\varepsilon;$$
$$\alpha:=\alpha\beta;$$

end

注　由于 SUMT 使用的是 Newton 方法, 故必须假设 Hessian 矩阵为正定. 但此假设条件很难满足, 原因是对一般函数, 除退化情形外, Hessian 矩阵常为非负定.

1.3 中心路径法

中心路径法为解凸规划问题 (1.8) 的另一种方法, 仍设 1.2 节的条件成立.

定义 1.1 不等式集 $r_i(\boldsymbol{x})>0, i=1,2,\cdots,m$ 的解析中心定义为

$$\boldsymbol{x}_{ac} := \arg\max_{\boldsymbol{x}\text{可行}} \prod_{i=1}^{m} r_i(\boldsymbol{x}).$$

严格说来, 解析中心并非可行域的性质, 因为一个可行域可以用不同的不等式组来描述, 而每一组均有一解析中心; 但解析中心肯定位于可行域内部.

凸规划的中心路径法的思想仍然是用无约束最优化问题来代替约束问题. 定义一个实函数 $\boldsymbol{x}_c^*(t)$ 如下:

$$\begin{aligned}\boldsymbol{x}_c^*(t) &:= f[<t; r_i>0(i=1,2,\cdots,m)]\text{的解析中心}\\ &= \arg\max_{\boldsymbol{x}}\left\{[t-f(\boldsymbol{x})]\prod_{i=1}^{m} r_i(\boldsymbol{x})\right\} \\ &= \arg\min_{\boldsymbol{x}}\left\{-\ln[t-f(\boldsymbol{x})]-\sum_{i=1}^{m}\ln[r_i(\boldsymbol{x})]\right\}.\end{aligned} \tag{1.11}$$

只有当不等式组 $f(\boldsymbol{x})<t, r_i(\boldsymbol{x})>0$ 至少在一点 $\boldsymbol{x}$ 满足时, $\boldsymbol{x}_c^*(t)$ 才有定义. 若找到这样一个点 $\boldsymbol{x}$, 则因式 (1.11) 为凸, 故 $\boldsymbol{x}_c^*(t)$ 不难找到.

上述概念可以用几何来表述.

引理 1.3 若 $t>f^*$ 且限制可行集 $\{\boldsymbol{x}: f(\boldsymbol{x})<t, r_i(\boldsymbol{x})\geqslant 0\}$ 有界, 则 $\boldsymbol{x}_c^*(t)$ 存在且有

$$0 \leqslant f(\boldsymbol{x}_c^*(t)) - f^* \leqslant m[t-f(\boldsymbol{x}_c^*)], \tag{1.12}$$

其中 m 为约束条件的个数. 特别地, 有 $\lim\limits_{t\to f^*} f(\boldsymbol{x}_c^*(t)) = f^*$.

证明 由可微性, 在 $\boldsymbol{x}=\boldsymbol{x}_c^*(t)$ 处有

$$\nabla\left\{-\ln[t-f(\boldsymbol{x})]-\sum_{i=1}^{m}\ln r_i(\boldsymbol{x})\right\} = \frac{1}{t-f(\boldsymbol{x})}\nabla f(\boldsymbol{x}) + \nabla\phi(\boldsymbol{x}) = \boldsymbol{0}.$$

由此推出式 (1.10) 对 $\alpha = 1/[t-f(\boldsymbol{x}_c^*(t))]$ 为零. 对此 α, 不等式 (1.9) 成为不等式 (1.12). □

同理, 我们给出一个算法如下:

算法 1.2 (中心方法)

Input 严格可行的 $\boldsymbol{x}^{(0)}$, 公差 $\varepsilon > 0$,

初始上界 $t > f(\boldsymbol{x}^{(0)})$且$0 < \theta < 1$.

Output 一个可行值 $\boldsymbol{x}$, 使有 $\mid f^* - f(\boldsymbol{x}) \mid < \varepsilon$.

Assumptions 可行域与 $\{\boldsymbol{x} : f(\boldsymbol{x}) < t\}$ 的交集为有界.

begin

$$\boldsymbol{x} = \boldsymbol{x}^{(0)};$$

repeat

$$\begin{aligned}\boldsymbol{v} :=& -\{\boldsymbol{\nabla}^2(-\ln[t - f(\boldsymbol{x})] + \ln\phi(\boldsymbol{x}))\}^{-1}\\ &\times\left\{\boldsymbol{\nabla}\left[-\ln(t - f(\boldsymbol{x}))\right] + \frac{1}{\alpha}\phi(\boldsymbol{x})\right\};(\text{牛顿方向})\end{aligned}$$

$$\delta^* := \arg\min_{\delta}\left\{-\ln[t - f(\boldsymbol{x} + \delta\boldsymbol{v})] + \frac{1}{\alpha}\phi(\boldsymbol{x} + \delta\boldsymbol{v})\right\};(\text{线搜索})$$

$$\boldsymbol{x} := \boldsymbol{x} + \delta^*\boldsymbol{v};$$

until $||\boldsymbol{v}||$ 充分小;

return if

$$m[t - f(\boldsymbol{x})] < \varepsilon;$$
$$t := (1 - \theta)f(\boldsymbol{x}) + \theta t;$$

end

引理 1.4　上述算法中的迭代收敛于 f^*.

证明　由引理 1.3, 我们有 $f^* \geqslant f(\boldsymbol{x}) - m[t - f(\boldsymbol{x})] = (m+1)f(\boldsymbol{x}) - mt$, 从而 $f(\boldsymbol{x}) \leqslant \dfrac{f^* + mt}{m+1}$. 令 t^+ 为更新值, $t^+ := (1-\theta)f(\boldsymbol{x}) + \theta t$, 则

$$t^+ - f^* = (1-\theta)f(\boldsymbol{x}) + \theta t - f^* \leqslant (1-\theta)\frac{f^* + mt}{m+1} + \theta t - f^* = \frac{m+\theta}{m+1}(t - f^*),$$

即误差 $t - f^*$ 在每一次迭代时至少减少一个因子 $(m+\theta)/(m+1)(<1)$, 因此它收敛于零. □

注　在实用上更有效的是中心路径法的另一个版本, 即定义

$$\begin{aligned}\boldsymbol{x}_c^*(t) :=& \arg\max_{\boldsymbol{x}}\left\{[t - f(\boldsymbol{x})]^m\prod_{i=1}^{m} r_i(\boldsymbol{x})\right\}\\ =& \arg\min_{\boldsymbol{x}}\left\{-m\ln[t - f(\boldsymbol{x})] - \sum_{i=1}^{m}\ln[r_i(\boldsymbol{x})]\right\}.\end{aligned}$$