

普通高等学校信息与计算科学专业系列丛书

教育科学“十五”国家规划课题研究成果

最 优 化 方 法

孙文瑜 徐成贤 朱德通

高等教育出版社

内容提要

本书既系统地介绍最优化方法的基本理论和有效算法,又反映了目前该学科的发展动态。主要内容包括:基本概念,线性规划,线性搜索和信赖域方法,无约束最优化,线性与非线性最小二乘问题,二次规划,约束最优化等。全书深入浅出,理论、计算与实际应用相结合,尽可能避免较深的数学推导。每章后都有一个小结,并附有习题,以利于教学。

本书可作为信息与计算科学、数学与应用数学、统计、运筹学、管理科学与工程、计算机、经济与金融,以及有关理工科专业的本科生作为教材或教学参考书。具有微积分和高等代数基础的科技人员可自学本书。

图书在版编目(CIP)数据

最优化方法/ 孙文瑜, 徐成贤, 朱德通 .—北京:
高等教育出版社, 2004. 7
ISBN 7 - 04 - 014375 - 5

最 孙 ... 徐 ... 朱 ... 最优
化-高等学校-教材 .O224

中国版本图书馆 CIP 数据核字(2004)第 054929 号

出版发行	高等教育出版社	购书热线	010 - 64054588
社 址	北京市西城区德外大街 4 号	免费咨询	800 - 810 - 0598
邮政编码	100011	网 址	http:// www hep .edu .cn
总 机	010 - 82028899		http:// www hep .com .cn

经 销 新华书店北京发行所
印 刷

开 本	787 × 960 1/ 16	版 次	年 月 第 1 版
印 张	14	印 次	年 月 第 次印刷
字 数	260 000	定 价	17. 90 元

本书如有缺页、倒页、脱页等质量问题,请到所购图书销售部门联系调换。
版权所有 侵权必究

策划编辑	王 瑜
责任编辑	王 瑜
封面设计	王凌波
责任绘图	朱 静
版式设计	马静如
责任校对	尤 静
责任印制	

信息与计算科学专业系列教材编委会

顾 问	李大潜	刘应明		
主 任	徐宗本			
副主任	王国俊	马富明	胡德焜	
委 员	(以姓氏笔画为序)			
	韦志辉	叶中行	白峰杉	羊丹平
	孙文瑜	吕 涛	阮晓青	陈发来
	沈世镒	陈 刚	张志让	吴 微
	柳重堪	凌永祥	徐 刚	徐树方
	黄象鼎	雍炯敏		
秘 书	李水根	王 瑜		

总 序

根据教育部 1998 年颁布的普通高等院校专业目录,“信息与计算科学”专业被列为数学类下的一个新专业(它覆盖原有的计算数学及其应用软件、信息科学与运筹控制等专业)。这一新专业的设置很好地适应了新世纪以信息技术为核心的全球经济发展格局下的数学人才培养与专业发展的需要。然而,作为一个新专业,对其专业内涵、专业规范、教学内容与课程体系等有一个自然的认识与探索过程。教育部数学与统计学教学指导委员会数学类专业教学指导分委员会(下称教指委)经过过去两年艰苦细致的工作,对这些问题现在已有了比较明确的指导意见,发表了《关于信息与计算科学专业办学现状与专业建设相关问题的调查报告》及《信息与计算科学专业教学规范》(讨论稿)(见《大学数学》第 19 卷 1 期(2003))。为此,全国高等学校教学研究中心在承担全国教育科学“十五”国家级规划课题——“21 世纪中国高等教育人才培养体系的创新与实践”研究工作的基础上,根据教指委所颁布的新的教学规范,组织国内各高校的专家教授,进行其子项目课题“21 世纪中国高等学校信息与计算科学专业教学内容与课程体系的创新与实践”的研究与探索。为推动本专业的教材建设,该项目课题小组与高等教育出版社联合成立了“信息与计算科学专业系列教材编委会”,邀请有多年教学和科研经验的教师编写系列教材,由高等教育出版社独家出版,并冠以教育科学“十五”国家规划课题研究成果。

按照新的《信息与计算科学专业教学规范》(讨论稿),信息与计算科学专业是以信息技术和计算技术的数学基础为研究对象的理科类专业。其目标是培养学生具有良好的数学基础和数学思维能力,掌握信息与计算科学基础理论、方法与技能,受到科学研究的训练,能解决信息技术和科学与工程计算中的实际问题的高级专门人才。毕业生能在科技、教育、信息产业、经济与金融等部门从事研究、教学、应用开发和管理工作的,能继续攻读研究生学位。根据这一专业目标定位和落实“强基础、宽口径、重实际、有侧重、创特色”的办学指导思想,我们认为,本专业在数学基础、计算机基础、专业基础方面应该得到加强,而各学校在这三个基础方面可大体一致,但专业课(含选修课)允许各校自主选择、体现各自特点。考虑到已有大量比较成熟的数学基础与计算机基础课程教材,本次教材编写主要侧重于专业基础课与专业课(含选修课)方面。

信息与计算科学,就其范畴与研究内容而言,是数学、计算机科学和信息工

程等学科的交叉,已远远超出数学学科的范畴。但作为数学学科下的一个理科学专业,信息与计算科学专业则主要研究信息技术的核心基础与运用现代计算工具高效求解科学与工程问题的数学理论与方法(或更简明地说,研究定向于信息技术与计算技术的数学基础),这一专业定位明显地与计算机科学与信息工程专业构成区别。基于这一定位,信息与计算科学专业可包括信息科学与科学计算(计算数学)两个大的方向。科学计算方向在我国已有长期的办学经验,通常被划分为偏微分方程数值解、最优化理论与方法、数值逼近与数值代数、计算基础等学科子方向。然而,对于信息科学,它到底应该怎样划分学科子方向?应该怎样设置专业与专业基础课?所有这些都仍是正在探索的问题。

任何技术都可以认为是延伸与扩展人的某种功能的方式与方法,所以信息技术可以认为是扩展人的信息器官功能的技术。人的信息器官主要包括感觉器官、传导器官(传导神经网络)、思维器官和效应器官四大类型,其功能则主要是信息获取、信息传输、信息处理和信息应用(控制),因而感测技术、通信技术、智能技术与控制技术通常被认为是最基本的信息技术(常称之为信息技术的四基元),其他信息技术可认为是这四种基本技术的高阶逻辑综合或分解衍生。所以可以把信息科学理解为是“有关信息获取、信息传输、信息处理与信息控制基础的科学”。从这个意义上,我们认为:信息处理(包括图像处理、信号分析等)、信息编码与信息安全、计算智能(人工智能、模式识别等)、自动控制等可构成信息科学的主要学科子方向。这一认识也是教指委设置信息与计算科学专业信息科学方向课程的基本依据。

本系列教材正是基于以上认识,为落实新的《信息与计算科学专业教学规范》(讨论稿)而组织编写的。我们相信,该系列教材的出版对缓解本专业教材的紧缺局面,对推动信息与计算科学专业的快速与健康发展会大有裨益。

从长远的角度看,为适应不同类型院校和不同层次要求的课程需求,本系列教材编委会还将不断组织教材的修订和编写新的教材,从而使本专业的教学用书做到逐步充实、完善和多样化。我们诚恳希望采用本系列教材的教师、同学们及广大读者对书中存在的问题及时指正并提出修改意见和建议。

信息与计算科学专业系列教材编委会

2003年8月31日

前 言

最优化是一门应用十分广泛的学科,它研究在有限种或无限种可行方案中挑选最优方案,构造寻求最优解的计算方法。由于最优化在科学、工程、国防、交通、管理、经济、金融、计算机等领域的广泛应用,现在许多高校理科、工科、管理科学、经济与金融等学科都把最优化开设为一门必修或选修课程。

为了给我国高等院校理工科和管理类本科生提供一本现代、实用、合适的“最优化方法”教材,我们根据教育部信息与计算科学教材编写组的要求,结合我们自己在最优化领域的科研和教学体会,把这个领域最基本、最重要、最实用、最有效的现代最优化方法的内容介绍给学生,使本科生对最优化方法有一个基本的和较全面的了解,为进一步从事最优化方法的理论、算法、软件与应用打下一个较好的基础。

本书深入浅出,通俗易懂。我们尝试努力讲清每种方法的动机、算法、性质、收敛性及数值例子,尽可能避免较深较难的数学推导,力争把所需要的数学知识降到最低。对于个别难度较大的章节,为了全书的连贯性,我们仍然给出,但用*号表示,仅供教师和感兴趣的学生参考。例如§4.4.3拟牛顿法的收敛性,§5.1.3解线性不等式约束的线性最小二乘问题等。对于这些章节,教师可根据具体情况予以跳过。对于§1.2凸集与凸函数中后面部分的理论内容教师可以选讲。

本书为本科生一学期用教材,材料覆盖最优化领域的核心内容。第一章介绍最优化问题的基本概念;第二章介绍在实际上有最广泛应用的线性规划;第三章讨论线性搜索与信赖域方法,它们提供了最优化方法的总体收敛策略;第四章研究无约束最优化方法,这一章是全书的一个中心,这不仅因为无约束最优化本身的重要性,而且众多约束最优化问题也都是转化为无约束问题来处理;第五章讨论一类特殊的最优化问题——线性与非线性最小二乘问题。这些问题在实际中广泛存在。我们向大家提供解这类问题的基本的现代方法,相信这会对大家今后的工作是相当有益的。第六和第七两章讨论约束最优化,其中我们将一类特殊的和重要的约束优化类型——二次规划单独构成一章,对约束最优化的理论与算法我们进行了简单的讨论。

由于这本书是作为本科生教材编写的,我们不希望本书内容太多、太全、太深,但也必须使学生了解这门学科的全貌,了解与掌握必要的方法、理论与算法

软件。应该说,本书包含的内容是最优化的核心部分,特别是实际中用得较多的内容,我们力求多讲一些。这本书可供一学期 48 课时(每周 3 课时 \times 16 周)~72 课时(每周 4 课时 \times 18 周)的各类学校采用。教师可根据本校具体情况适当增删内容。本书适用面较广,既适用于重点院校,也适用于一般院校;既适用于信息与计算科学、数学与应用数学、统计与运筹等数学类专业,也适用于计算机、理学类、工程类、管理类、金融经济等学科。我们建议这门课在三年级讲授,使学生预先具有高等数学和线性代数的基本知识。同时,在讲授这门课时,最好要求学生上机编程序做练习,因为用最优化方法解题都需要在计算机上完成。

本书每章后面都有一个小结,概述了这一章的主要内容、背景,及一些进展。同时每章给出了一些习题供教师选择使用。全书后面附有一个附录和参考文献。附录中给出了求解无约束最优化和约束最优化的一些标准试验函数,这些试验函数供数值试验用。书末列出的参考文献供教师和感兴趣的学生作进一步深入研究探讨时参考。

在本书写作和录入的过程中,作者的研究生们提供了不少建议,并协助做了许多工作,谨致谢意。

南京航空航天大学理学院院长倪勤教授认真审阅了全书手稿,并提出了宝贵的修改意见。作者谨向他致以衷心的感谢。

教育部信息与计算科学教材编委会、高等教育出版社的领导及王瑜编辑对本书的写作给予了热情指导和关心,谨向他们致谢。

作者还要感谢国家自然科学基金委员会多年来的资助,感谢南京师范大学、西安交通大学和上海师范大学对作者工作的一贯支持。

尽管本书作者多年来一直从事最优化方法的研究和教学,但限于水平和时间,本书难免有不妥和错误之处,欢迎读者批评指正,以便再版时修改更正。

孙文瑜 徐成贤 朱德通

2004 年 2 月 15 日

目 录

第 I 章	基本概念	1
§ 1.1	最优化问题简介	1
§ 1.2	凸集和凸函数	5
§ 1.3	最优性条件	16
§ 1.4	最优化方法概述	27
小结	31
习题	32
第 2 章	线性规划	34
§ 2.1	基本性质	34
§ 2.2	单纯形方法	45
§ 2.3	线性规划的对偶与对偶单纯形法	59
§ 2.4	线性规划的内点算法	71
小结	79
习题	81
第 3 章	线性搜索与信赖域方法	85
§ 3.1	线性搜索	85
§ 3.2	0.618 法和 Fibonacci 法	86
3.2.1	0.618 法	87
3.2.2	Fibonacci 法	88
3.2.3	二分法	90
§ 3.3	逐次插值逼近法	90
§ 3.4	精确线性搜索方法的收敛性	94
§ 3.5	不精确线性搜索方法	96
3.5.1	Goldstein 准则	97
3.5.2	Wolfe 准则	98
3.5.3	Armijo 准则	99
3.5.4	不精确线性搜索的收敛性	99

§ 3.6	信赖域方法思想和算法框架	102
§ 3.7	信赖域方法的收敛性	103
§ 3.8	解信赖域子问题	107
	小结	110
	习题	111
第4章	无约束最优化方法	112
§ 4.1	最速下降法	112
§ 4.2	牛顿法	115
§ 4.3	共轭梯度法	119
4.3.1	共轭方向法	119
4.3.2	共轭梯度法	121
4.3.3	对于非二次函数的共轭梯度法	125
§ 4.4	拟牛顿法	128
4.4.1	拟牛顿条件	128
4.4.2	DFP 校正和 BFGS 校正	130
4.4.3	拟牛顿法的收敛性*	135
	小结	139
	习题	139
第5章	线性与非线性最小二乘问题	141
§ 5.1	线性最小二乘问题的解法	142
5.1.1	解线性最小二乘问题	142
5.1.2	解线性等式约束的线性最小二乘问题	145
5.1.3	解线性不等式约束的线性最小二乘问题*	149
§ 5.2	非线性最小二乘的 Gauss Newton 法	151
§ 5.3	信赖域方法	155
§ 5.4	对 Gauss Newton 矩阵的拟牛顿修正	162
	小结	168
	习题	169
第6章	二次规划	172
§ 6.1	二次规划	172
§ 6.2	等式约束二次规划问题	174
§ 6.3	凸二次规划的有效集方法	180

小结	184
习题	184
 第7章 约束最优化的理论与方法	186
§ 7.1 约束最优化问题与最优性条件	186
§ 7.2 二次罚函数方法	191
§ 7.3 内点障碍罚函数法	193
§ 7.4 序列二次规划方法	196
小结	202
习题	202
 附录 试验函数	203
§ 1 无约束最优化问题的试验函数	203
§ 2 约束最优化问题的试验函数	204
 参考文献	207

第 I 章

基 本 概 念

§ 1.1 最优化问题简介

最优化技术在国民经济的许多领域如国防、工农业生产、交通运输、金融、贸易、管理、科学研究中有广泛的应用.所谓最优化就是在众多可行的方案或方法中找到最好的方案或方法.例如在确定投资项目时希望选择期望收益最大或风险最小的项目;空间飞船的设计要求在有限的空间放置尽可能多的设备;一个建筑在满足设计要求的条件下建筑费用应尽可能少;两地之间的输送管道或运输路线在满足要求的条件下应尽可能短.为应用最优化技术确定最优的方案,需要针对具体的实际问题建立相应的最优化模型,再根据模型的具体形式和特性选择适当的最优化方法求解.本书旨在介绍最优化的基本理论和常用的最优化方法.

最优化问题的数学模型一般形式为

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i = 1, 2, \dots, m, \\ & c_i(x) \leq 0, \quad i = m+1, \dots, p, \end{aligned} \tag{1.1.1}$$

其中 $x = (x_1, x_2, \dots, x_n)^T \in \mathbf{R}^n$, $f: \mathbf{R}^n \rightarrow \mathbf{R}^1$, $c_i: \mathbf{R}^n \rightarrow \mathbf{R}^1$ ($i = 1, 2, \dots, p$) 为连续函数,通常还要求连续可微. x 称为决策变量, $f(x)$ 为目标函数, $c_i(x)$, $i = 1, 2, \dots, p$ 为约束函数, $c_i(x) = 0$, $i = 1, 2, \dots, m$ 为等式约束, $c_i(x) \leq 0$, $i = m+1, \dots, p$ 为不等式约束. \min 和 s.t. 分别是英语单词 minimize(极小化)和 subject to(受约束)的缩写.

根据实际问题的不同要求,最优化模型有不同的形式,但经过适当的变换都可以转换成上述一般的形式.例如,对于求目标函数 $f(x)$ 极大的问题

$$\max \quad f(x)$$

可转换成求 $-f(x)$ 极小的问题

$$\min \quad (-f(x))$$

其中 $(-f)(x) = -f(x)$.又如对于形如

$$c_i(x) \leq 0$$

的不等式约束,可同样转换成上述形式的不等式约束

$$h_i(x) \leq 0,$$

其中 $h_i(x) = -c_i(x)$.还有像

$$a(x) \leq b(x) + c$$

的不等式约束,可通过令 $h(x) = b(x) - a(x) + c$ 转换成

$$h(x) \leq 0$$

的不等式约束形式 通过后面进一步深入的学习,我们会看到转换在最优化理论和方法的研究和应用中有重要作用 .

问题(1.1.1)是最优化问题的一般数学表现形式 .只要在问题中存在任何约束条件,就称为约束最优化问题 .只有等式约束时,

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i = 1, 2, \dots, m \end{aligned}$$

称为等式约束最优化问题 .只有不等式约束时,

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & c_i(x) \leq 0, \quad i = 1, 2, \dots, m \end{aligned}$$

称为不等式约束最优化问题 .如果既有等式约束,又有不等式约束,则称为混合约束问题 .如果问题中无任何约束条件,则称为无约束最优化问题 .无约束最优化问题的数学模型为

$$\min \quad f(x), \quad x \in \mathbf{R}^n, \quad (1.1.2)$$

一般简记为

$$\min \quad f(x).$$

无约束最优化问题是最优化的基础,一则很多实际的最优化问题本身就是无约束最优化问题,二则许多约束最优化方法都是通过变换把约束最优化问题转换成无约束最优化问题后,用适当的无约束优化方法求解 .

根据模型(1.1.1)中函数的具体性质和复杂程度,最优化问题又有许多不同的类型 .根据决策变量的取值是离散的还是连续的分为离散最优化和连续最优化,离散最优化通常又称组合最优化,如整数规划、资源配置、邮路问题、生产安排等问题都是离散最优化问题的典型例子 .离散最优化问题的求解较之连续最优化问题的求解难度更大,本书只介绍连续最优化的理论与方法 .根据连续最优化模型中函数的光滑与否又分为光滑最优化与非光滑最优化 .如果模型(1.1.1)中的所有函数都连续可微,则称为光滑最优化;只要有一个函数非光滑,则相应的优化问题就是非光滑优化问题 .本书只研究光滑最优化问题的求解方法,即在本书的大多数章节我们都假定模型中的函数连续可微,有时二阶或更高阶连续可微 .对于连续光滑的最优化问题,如果所有函数都是变量 $x = (x_1, x_2, \dots, x_n)^T$ 的线性函数,则称(1.1.1)为线性规划问题 .线性规划问题的一般形式为

$$\begin{aligned} \min \quad & c_1 x_1 + c_2 x_2 + \dots + c_n x_n, \\ \text{s.t.} \quad & a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n = b_1, \\ & \dots\dots\dots \\ & a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n = b_m, \\ & a_{m+1,1} x_1 + \dots + a_{m+1,n} x_n = b_{m+1}, \\ & \dots\dots\dots \\ & a_{p1} x_1 + a_{p2} x_2 + \dots + a_{pn} x_n = b_p. \end{aligned}$$

有时对变量 x 还有非负的约束

$$x_i \geq 0, \quad i = 1, 2, \dots, n,$$

或有界的约束

$$l_i \leq x_i \leq u_i, \quad i = 1, 2, \dots, n,$$

但这些约束都可以包括在上述模型的不等式约束中.上述线性规划模型的矩阵向量表示为

$$\begin{aligned} \min \quad & c^T x, \\ \text{s.t.} \quad & A_1 x = b^1, \\ & A_2 x \leq b^2, \end{aligned}$$

其中 $c = (c_1, c_2, \dots, c_n)^T$, $b^1 = (b_1, b_2, \dots, b_m)^T$, $b^2 = (b_{m+1}, \dots, b_p)^T$,

$$A_1 = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \dots & \dots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{m+1,1} & a_{m+1,2} & \dots & a_{m+1,n} \\ \dots & \dots & & \dots \\ a_{p1} & a_{p2} & \dots & a_{pn} \end{pmatrix}.$$

当模型(1.1.1)中的目标函数 $f(x)$ 是变量 x 的二次函数,而所有约束都是 x 的线性函数时称为二次规划问题.二次规划问题的一般形式为

$$\begin{aligned} \min \quad & q(x) = \frac{1}{2} x^T G x + c^T x + d, \\ \text{s.t.} \quad & A_1 x = b^1, \\ & A_2 x \leq b^2, \end{aligned}$$

其中 A_1, A_2 的表示同线性规划模型类似, $c = (c_1, c_2, \dots, c_n)^T$, d 为纯量, G 为 $n \times n$ 阶对称矩阵

$$G = \begin{pmatrix} G_{11} & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} & \dots & G_{2n} \\ \dots & \dots & & \dots \\ G_{n1} & G_{n2} & \dots & G_{nn} \end{pmatrix},$$

满足 $G_{ij} = G_{ji}$, " $i = j$ ".只要模型(1.1.1)的函数中有一个关于 x 是非线性的,就称为非线性最优化问题.非线性最优化问题是最一般的最优化问题,而线性规划和二次规划问题却是相当重要的特殊的最优化问题,这同样因为在实际中形成

的许多最优化问题是线性规划问题或二次规划问题,而且在用迭代法求非线性最优化问题的最优解时我们常常用线性规划或二次规划来局部近似原非线性最优化问题,并通过求所得近似问题的最优解来对已有最优解的估计进行改进.

如果点 $x \in \mathbf{R}^n$ 满足最优化模型 (1.1.1) 中的所有约束条件,就称其为可行点(feasible point),所有可行点的全体称为可行域(feasible region),在本书中我们用 F 表示可行域,即

$$F = \{x / c_i(x) = 0, i = 1, 2, \dots, m, \quad c_i(x) \leq 0, i = m + 1, \dots, p\}.$$

对于一个可行点 \bar{x} , 考虑不等式约束 $c_i(x) \leq 0$, 如果有 $c_i(\bar{x}) = 0$, 就称约束 $c_i(x) \leq 0$ 在点 \bar{x} 是有效约束或起作用约束(active constraint), 并称可行点 \bar{x} 位于约束 $c_i(x) \leq 0$ 的边界. 如果有 $c_i(\bar{x}) < 0$, 就称不等式约束 $c_i(x) \leq 0$ 在点 \bar{x} 是无效约束或不起作用约束(inactive constraint), 并称 \bar{x} 是约束 $c_i(x) \leq 0$ 的内点. 在任意可行点, 所有的等式约束都被认为是有效约束. 在一个可行点, 所有有效约束的全体被称为该可行点的有效集, 并记为

$$A = \{i / i = 1, 2, \dots, m, m + 1, \dots, p, c_i(\bar{x}) = 0\}.$$

对于一可行点 \bar{x} , 如果没有一个不等式约束是有效的, 就称 \bar{x} 是可行域的内点. 不是内点的可行点就是可行域的边界点. 显然, 在边界点至少有一个不等式约束是有效约束. 当存在等式约束时, 任何可行点都要满足等式约束, 因此不可能是等式约束的内点.

图 1.1 给出了由下述约束条件给出的可行域 F :

$$\begin{aligned} c_1(x) &= 2x_1 + 3x_2 + x_3 - 6 = 0, \\ x_1 &\leq 0, \\ x_2 &\leq 0, \\ x_3 &\leq 0. \end{aligned}$$

对于可行点 x^1 , 约束 $x_1 \leq 0$ 和 $x_3 \leq 0$ 是有效约束, 而 $x_2 \leq 0$ 是无效约束. 对于可行点 x^2 , 则刚好相反, 约束 $x_2 \leq 0$ 是有效约束, 而 $x_1 \leq 0$ 和 $x_3 \leq 0$ 是无效约束. 而对于可行点 x^3 , 三个不等式约束都是无效约束. 图中可行域的边界由粗线表示.

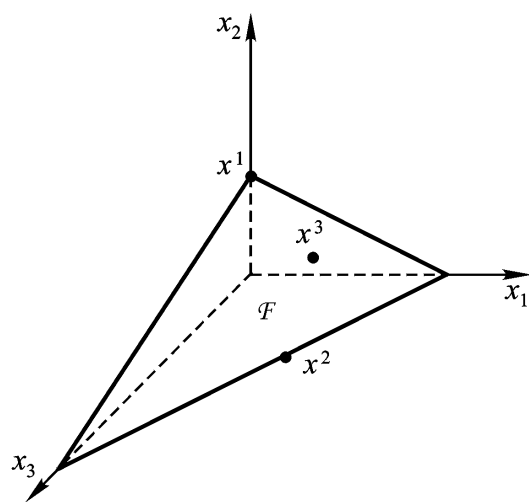


图 1.1 可行域与可行域的边界

一个可行点 $x^* \in F$ 称为问题 (1.1.1) 的(全局或总体)最优解, 如果有

$$f(x^*) \leq f(x), \quad \forall x \in F;$$

如果上述不等式对所有不同于 x^* 的可行点 x 严格成立, 即

$$f(x^*) < f(x), \quad \forall x \in F, x \neq x^*,$$

则称 x^* 为严格(全局或总体)最优解. 对于可行点 x^* , 如果存在一个邻域

$$N(x^*) = \{x \mid \|x - x^*\| \leq \delta\}$$

使得成立

$$f(x^*) \leq f(x), \quad \forall x \in N(x^*) \cap F,$$

则称 x^* 为优化问题 (1.1.1) 的局部最优解, 其中 $\delta > 0$ 是一个小的正数, 范数 $\|\cdot\|$ 可以是任意向量范数, 但一般常用 l_2 范数

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

如果上述不等式对所有 $x \in N(x^*) \cap F$ 严格成立, 则称 x^* 为严格局部小点.

图 1.2 就单变量函数的情形给出了全局最优解和局部最优解的一个例子. 在这个图中, 点 x_1 是所示函数的严格全局极小解, x_2 和 x_3 则是局部极小解, 其中 x_3 是严格局部极小解, 而 x_2

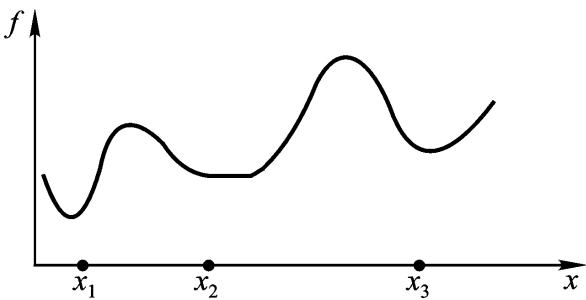


图 1.2 全局极小点与局部极小点

是非严格局部极小解. 需要注意的是并非所有的连续可微函数都有极小解, 单变量函数 $f(x) = x^3$ 就是一个最简单而又容易理解的例子. 即使问题有最优解, 最优解未必唯一, 也未必是全局最优解. 本书介绍的方法一般只能用于确定最优化问题的局部最优解, 有关确定全局最优解的最优化方法属于最优化研究的另一个领域——全局最优化. 然而, 如果最优化问题的目标函数是凸的, 而可行域是凸集, 则问题的任何最优解 (不一定唯一) 必是全局最优解, 这样的最优化问题又称凸规划.

§ 1.2 凸集和凸函数

正如我们在上一节所指出的目标函数为凸函数, 可行集为凸集的凸规划问题可确保任意最优解就是问题的全局最优解. 因此, 凸集与凸函数在最优化理论与算法的研究与学习中起着重要作用. 本节我们介绍凸集与凸函数的某些基本概念和基本性质.

定义 1.2.1 集合 $D \subset \mathbb{R}^n$ 称为凸的, 如果对于任意 $x, y \in D$ 有

$$x + (1 - \lambda)y \in D, \quad \forall \lambda \in [0, 1].$$

换句话说, 如果 $x, y \in D$, 则连接 x 与 y 的直线段上的所有点都在 D 内.

图 1.3 给出了平面上凸集与非凸集的例子.

有关凸集有下面这样一些有用的性质: 两个凸集的交, 和以及差仍然是凸集. 即如果 $D_1, D_2 \subset \mathbf{R}^n$ 是凸集, 则 $D_1 \cap D_2 = \{x \mid x \in D_1, x \in D_2\}$, $D_1 + D_2 = \{x + y \mid x \in D_1, y \in D_2\}$ 和 $D_1 - D_2 = \{x - y \mid x \in D_1, y \in D_2\}$ 都是凸集; 对于任意非零实数 λ , 集合 $D_\lambda = \{x \mid x \in D_1\}$ 也是凸集. 根据凸集的定义, 我们还可以得到凸集的任意有限个点的凸组合仍属于凸集的结论, 即我们有下述定理.

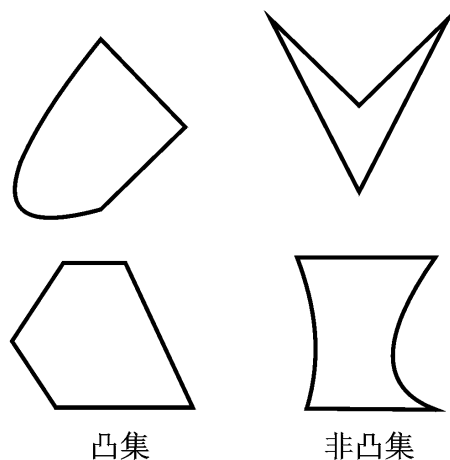


图 1.3 凸集与非凸集

定理 1.2.2 设 $D \subset \mathbf{R}^n$ 是凸集, 则任意 m 个点 $x^{(i)} \in D (i = 1, 2, \dots, m)$ 的凸组合仍属于 D , 即有

$$\sum_{i=1}^m \lambda_i x^{(i)} \in D, \quad (1.2.1)$$

其中 $\lambda_i \geq 0, i = 1, 2, \dots, m, \sum_{i=1}^m \lambda_i = 1$.

证明 定理的结论可以用归纳法证明. 根据凸集的定义, 定理的结论在 $m = 2$ 时显然成立. 设结论在 $m = k$ 时成立, 我们证明结论在 $m = k + 1$ 时也成立. 设 $x^{(i)} \in D, i = 1, 2, \dots, k + 1$, 考虑任意一组实数 $\lambda_i \geq 0, i = 1, 2, \dots, k + 1, \sum_{i=1}^{k+1} \lambda_i = 1$, 则

$$\begin{aligned} \sum_{i=1}^{k+1} \lambda_i x^{(i)} &= \sum_{i=1}^k \lambda_i x^{(i)} + \lambda_{k+1} x^{(k+1)} \\ &= \left(1 - \lambda_{k+1}\right) \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^{(i)} + \lambda_{k+1} x^{(k+1)}. \end{aligned}$$

由于 $1 - \lambda_{k+1} = \sum_{i=1}^k \lambda_i$, 我们有

$$\sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} = 1.$$

由归纳假设有

$$\bar{x} = \sum_{i=1}^k \frac{\lambda_i}{1 - \lambda_{k+1}} x^{(i)} \in D.$$

再由凸集的定义得

$$\sum_{i=1}^{k+1} \lambda_i x^{(i)} = \left(1 - \lambda_{k+1}\right) \bar{x} + \lambda_{k+1} x^{(k+1)} \in D.$$

这就完成了定理的证明.

凸集除了在刻画可行域的特性方面起重要作用外, 下面的凸集分离定理在分析最优解的最优性条件时起重要作用.

定义 1.2.3 设 $D_1, D_2 \subset \mathbf{R}^n$ 为两非空凸集, 若存在非零向量 $a \in \mathbf{R}^n$ 和实数 α , 使得

$$\begin{aligned} D_1 \cap H^+ &= \{x \in \mathbf{R}^n \mid a^T x \geq \alpha\}, \\ D_2 \cap H^- &= \{x \in \mathbf{R}^n \mid a^T x \leq \alpha\}, \end{aligned}$$

则称超平面

$$H = \{x \in \mathbf{R}^n \mid a^T x = \alpha\} \quad (1.2.2)$$

分离了集合 D_1 和 D_2 . 如果更有

$$\begin{aligned} D_1 \cap H_0^+ &= \{x \in \mathbf{R}^n \mid a^T x > \alpha\}, \\ D_2 \cap H_0^- &= \{x \in \mathbf{R}^n \mid a^T x < \alpha\}, \end{aligned}$$

则称超平面 H 严格分离 D_1 和 D_2 , 其中 H_0^+, H_0^- 分别表示集合 H^+, H^- 的内部.

定理 1.2.4 设 $D \subset \mathbf{R}^n$ 是非空闭凸集, $y \in \mathbf{R}^n$ 但 $y \notin D$, 则

(1) 存在唯一的点 $\bar{x} \in D$, 使得集合 D 到点 y 的距离最小, 即

$$\|\bar{x} - y\| = \inf\{\|x - y\| \mid x \in D\}.$$

(2) $\bar{x} \in D$ 是点 y 到集合 D 的最短距离点的充分必要条件为

$$(x - \bar{x})^T (\bar{x} - y) = 0, \quad \forall x \in D.$$

证明 (1) 取充分大的 $\mu > 0$ 使得

$$D_\mu = D \cap (y + \mu S) \neq \emptyset,$$

其中 \emptyset 表示空集, S 为空间 \mathbf{R}^n 中的单位球

$$S = \{x \in \mathbf{R}^n \mid \|x\| = 1\},$$

则集合 D_μ 是非空有界闭凸集, 因此连续函数

$$f(x) = \|x - y\|$$

在集合 D_μ 上必取得最小值. 设这个最小值在点 $\bar{x} \in D_\mu$ 取得, 则 \bar{x} 就是集合 D 到 y 距离最短的点.

再证点 \bar{x} 的唯一性, 设有 $\tilde{x} \in D$, $\tilde{x} \neq \bar{x}$ 使得

$$\|y - \tilde{x}\| = \|y - \bar{x}\| = r,$$

则由范数的三角不等式有

$$\left\| y - \frac{\bar{x} + \tilde{x}}{2} \right\| = \left\| \frac{1}{2}(y - \bar{x}) + \frac{1}{2}(y - \tilde{x}) \right\| \leq \frac{1}{2}\|y - \bar{x}\| + \frac{1}{2}\|y - \tilde{x}\| = r.$$

由于 D 是凸集, 故有 $(\bar{x} + \tilde{x})/2 \in D$. 又因 r 是集合 D 到 y 的最短距离, 上式应有等号成立. 由此, 存在实数 λ 使成立

$$y - \bar{x} = \lambda(y - \tilde{x}).$$

又由 $\|y - \bar{x}\| = \|y - \tilde{x}\| = r$, 应有 $|\lambda| = 1$. 若 $\lambda = -1$, 则由 $y - \bar{x} = -y + \tilde{x}$ 得 $y = (\bar{x} + \tilde{x})/2 \in D$, 与 $y \notin D$ 相矛盾, 因此必有 $\lambda = 1$, 由此得 $\tilde{x} = \bar{x}$. 这就证明了 \bar{x} 的唯一性.

(2) 先证充分性. 设定理中所述的条件成立, 则对任意 $x \in D$, $x \neq \bar{x}$ 有

$$\begin{aligned} \|x - y\|^2 &= \|x - \bar{x} + \bar{x} - y\|^2 \\ &= \|x - \bar{x}\|^2 + \|\bar{x} - y\|^2 + 2(x - \bar{x})^T(\bar{x} - y). \end{aligned}$$

由 $\|x - \bar{x}\|^2 > 0$ 以及 $(x - \bar{x})^T(\bar{x} - y) = 0$, 得

$$\|x - y\|^2 > \|\bar{x} - y\|^2, \quad \forall x \in D.$$

这就证明了 \bar{x} 是 D 到 y 距离最短的点.

再证必要性. 设 \bar{x} 是 D 到点 y 距离最短的点, 即

$$\|\bar{x} - y\|^2 \leq \|x - y\|^2, \quad \forall x \in D. \quad (1.2.3)$$

由于 D 是凸集及 $\bar{x} \in D$, 对于任意给定的 $x \in D$ 以及任意的 $\lambda \in (0, 1)$ 有

$$\bar{x} + \lambda(x - \bar{x}) = \lambda x + (1 - \lambda)\bar{x} \in D.$$

用此代替式 (1.2.3) 中的 x 得

$$\begin{aligned} \|\bar{x} - y\|^2 &\leq \|y - \bar{x} - \lambda(x - \bar{x})\|^2 \\ &= \|y - \bar{x}\|^2 + \lambda^2 \|x - \bar{x}\|^2 - 2\lambda(y - \bar{x})^T(x - \bar{x}). \end{aligned}$$

因此对任意 $\lambda \in (0, 1)$ 得

$$\lambda^2 \|x - \bar{x}\|^2 - 2\lambda(y - \bar{x})^T(x - \bar{x}) \leq 0.$$

将此式改写为

$$\lambda \|x - \bar{x}\|^2 + 2(\bar{x} - y)^T(x - \bar{x}) \leq 0.$$

在上式两端同除以 λ , 再令 $\lambda \rightarrow 0$ 便得

$$(x - \bar{x})^T(\bar{x} - y) \leq 0, \quad \forall x \in D.$$

这就完成了定理的证明.

利用上述定理我们可以得到下述点与凸集的分离定理.

定理 1.2.5 设 $D \subset \mathbf{R}^n$ 为非空闭凸集, $y \in \mathbf{R}^n, y \notin D$, 则存在非零向量 $a \in \mathbf{R}^n$ 和实数 α , 使得

$$a^T x \leq \alpha < a^T y, \quad \forall x \in D,$$

即存在超平面 $H = \{x \in \mathbf{R}^n \mid a^T x = \alpha\}$ 严格分离点 y 与凸集 D .

证明 由 D 是非空闭凸集, $y \notin D$ 以及定理 1.2.4 知存在唯一的点 $\bar{x} \in D$ 使集合 D 到点 y 的距离最短, 且有

$$(x - \bar{x})^T(\bar{x} - y) = 0, \quad \forall x \in D$$

成立. 将此式改写为

$$x^T(y - \bar{x}) = \bar{x}^T(y - \bar{x}) \quad \forall x \in D.$$

由此得

$$\begin{aligned} \|y - \bar{x}\|^2 &= (y - \bar{x})^T(y - \bar{x}) = y^T(y - \bar{x}) - \bar{x}^T(y - \bar{x}) \\ &= y^T(y - \bar{x}) - x^T(y - \bar{x}), \quad \forall x \in D. \end{aligned}$$

令 $a = y - \bar{x}$, 则有 $a \neq 0$ 且

$$0 < a^T a = a^T y - a^T x, \quad \forall x \in D.$$

即有 $a^T x < a^T y$, $\forall x \in D$. 令 $\alpha = \sup\{a^T x \mid x \in D\}$ 即得

$$a^T x < a^T y, \quad \forall x \in D,$$

这就完成了定理的证明.

由定理(1.2.5)可以推得下述在最优化理论研究中起重要作用的 Farkas 引理.

引理 1.2.6 设 $A \in \mathbf{R}^{m \times n}$ 为 $m \times n$ 阶矩阵, $b \in \mathbf{R}^m$, 则下述两组方程中有且仅有一组有解

$$Ax = 0, \quad b^T x > 0, \quad (1.2.4)$$

$$A^T y = b, \quad y \geq 0, \quad (1.2.5)$$

其中 $x \in \mathbf{R}^n, y \in \mathbf{R}^m$.

证明 设方程组(1.2.5)有解, 若有 $x \in \mathbf{R}^n$ 使得 $Ax = 0$, 则由 $y \geq 0$ 得 $b^T x = y^T Ax = 0$. 这就证明了在(1.2.5)有解时方程组(1.2.4)无解.

反之, 设方程组(1.2.5)无解. 记 $D = \{z \mid z = A^T y, y \geq 0\}$, 则 $D \subset \mathbf{R}^n$ 为非空闭凸集, 且有 $b \notin D$. 由定理 1.2.5 存在非零向量 $a \in \mathbf{R}^n$ 和实数 α 使得

$$a^T x < \alpha, \quad \forall x \in D.$$

由于 $0 \in D$, 从上式知 $\alpha \geq 0$, 因而 $a^T b > 0$. 同样从上式还可得到

$$a^T x = a^T A^T y = y^T Aa, \quad \forall y \geq 0.$$

由于 $y \geq 0$ 可以任意大, 从上式知必有 $Aa = 0$, 由此及 $a^T b > 0$ 知向量 a 满足方程组(1.2.4), 即 a 是方程组(1.2.4)的解, 这就完成了引理的证明.

在明白了点与凸集的分离关系之后, 我们再学习两个凸集的分离关系. 首先给出凸集支撑超平面的概念.

定义 1.2.7 设 $D \subset \mathbf{R}^n$ 为非空集合, 且点 $\bar{x} \in \mathbf{R}^n$ 属于集合 D 的边界, 即 $\bar{x} \in D$. 若存在非零向量 $a \in \mathbf{R}^n$, 使成立

$$D \cap H_{\bar{x}}^+ = \{x \in \mathbf{R}^n \mid a^T (x - \bar{x}) \geq 0\}$$

或者

$$D \cap H_{\bar{x}}^- = \{x \in \mathbf{R}^n \mid a^T (x - \bar{x}) \leq 0\},$$

则称超平面

$$H_{\bar{x}} = \{x \in \mathbf{R}^n \mid a^T (x - \bar{x}) = 0\}$$

是集合 D 在其边界点 \bar{x} 的支撑超平面.

下面的定理说明非空凸集在其任一个边界点处都存在支撑超平面.

定理 1.2.8 设 $D \subset \mathbf{R}^n$ 是非空凸集, $\bar{x} \in D$, 则存在非零向量 $a \in \mathbf{R}^n$, 使成立

$$a^T x \leq a^T \bar{x}, \quad \forall x \in \bar{D},$$

即凸集 $D \subset \mathbf{R}^n$ 在其边界点处存在支撑超平面, 其中 \bar{D} 表示集合 D 的闭包.

证明 因为 $\bar{x} \in D$, 存在收敛序列 $y^{(k)} \in D, y^{(k)} \rightarrow \bar{x}$. 由点与凸集分离定理 1.2.5, 对于每一 $y^{(k)}$, 存在非零向量 $a^{(k)}$, 有

$$a^{(k)} = 1, \quad a^{(k)T} x < a^{(k)T} y^{(k)}, \quad \forall x \in \bar{D}.$$

由序列 $\{a^{(k)}\}$ 的有界性, 知存在收敛子序列 $\{a^{(k)}\}$, 有 $a^{(k)} \rightarrow a, a \in \mathbf{R}^n$ 且 $a = 1$. 对于这一收敛子序列有

$$a^{(k)T} x < a^{(k)T} y^{(k)}, \quad \forall x \in \bar{D}.$$

在上式中令 $k \rightarrow \infty$, 并利用 $y^{(k)} \rightarrow \bar{x}$ 的事实即得定理的结论.

组合定理 1.2.5 和定理 1.2.8, 我们可以得到下列推论.

推论 1.2.9 设 $D \subset \mathbf{R}^n$ 是非空凸集, $\bar{x} \notin D$, 则存在非零向量 $a \in \mathbf{R}^n$ 使成立

$$a^T x \leq a^T \bar{x}, \quad \forall x \in \bar{D}.$$

由这一推论, 我们可以得到下列两凸集分离定理.

定理 1.2.10 设 D_1, D_2 是 \mathbf{R}^n 的两个非空凸集, 且 $D_1 \cap D_2 = \emptyset$, 则存在超平面分离 D_1 和 D_2 , 即存在非零向量 $a \in \mathbf{R}^n$ 使得

$$a^T x \leq a^T y, \quad \forall x \in \bar{D}_1, y \in \bar{D}_2.$$

证明 令 $D = D_1 - D_2$, 由凸集的性质知 D 为凸集, 且零向量不属于 D , 这是因为如果有 $0 \in D$, 则有 $D_1 \cap D_2 \neq \emptyset$, 这与定理的条件相矛盾. 由 $0 \notin D$, 从推论 1.2.9 知存在非零向量 $a \in \mathbf{R}^n$ 使得

$$a^T z \leq a^T 0 = 0, \quad \forall z \in D.$$

由 D 的定义, 任意 $z \in D$ 可表示为 $z = x - y, x \in D_1, y \in D_2$, 代入上式即得定理的结论. 定理证毕.

根据上述两凸集分离定理可以得到在非线性最优化理论中具有重要作用的 Gordan 择一定理.

定理 1.2.11 设 $A \in \mathbf{R}^{m \times n}$ 为 $m \times n$ 阶矩阵, 则或者存在 $x \in \mathbf{R}^n$ 使

$$Ax < 0, \tag{1.2.6}$$

或者存在 $y \in \mathbf{R}^m$ 使

$$A^T y = 0, \quad y \geq 0, \quad y \neq 0; \tag{1.2.7}$$

且两者不能同时成立.

证明 设若存在 $x \in \mathbf{R}^n, y \in \mathbf{R}^m$ 使 (1.2.6) 和 (1.2.7) 同时成立, 则有

$$y^T Ax < 0, \quad (y^T Ax)^T = x^T A^T y = 0.$$

这是一个矛盾, 这表明 (1.2.6) 和 (1.2.7) 不能同时成立.

再证明两者中必有一个成立. 如果有 $x \in \mathbf{R}^n$ 使 (1.2.6) 成立, 则结论已得证. 现设不存在 $x \in \mathbf{R}^n$ 使 (1.2.6) 成立, 即对任意的 $x \in \mathbf{R}^n$ 都有 $Ax \geq 0$. 令

$$Z_1 = \{z \in \mathbf{R}^m \mid z < 0\}, \quad Z_2 = \{z \in \mathbf{R}^m \mid z = Ax \geq 0, x \in \mathbf{R}^n\}.$$

显然 Z_1, Z_2 是非空凸集, 且 $Z_1 \cap Z_2 = \emptyset$. 由两个凸集的分离定理 1.2.10 知存在非零向量 $a \in \mathbf{R}^m$ 使

$$a^T z \leq a^T (Ax), \quad \forall z \in \bar{Z}_1, x \in \mathbf{R}^n.$$

由于 $z \in Z_1$ 可以任意小, 因此有 $a \geq 0$. 在上式中取 $z = 0$ 又得

$$a^T A x = 0, \quad \forall x \in \mathbf{R}^n.$$

对此取 $x = -A^T a$, 则有 $-A^T a = 0$, 即得 $A^T a = 0$. 这就证明了向量 $a \in \mathbf{R}^m$ 满足方程组 (1.2.7). 定理证毕.

在明白了凸集的基本概念和有关的性质之后, 我们转而学习凸函数.

定义 1.2.12 设函数 $f(x)$ 在凸集 D 上有定义, 如果对任意 $x, y \in D$ 和任意 $\lambda \in [0, 1]$ 有

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y),$$

则称 $f(x)$ 是凸集 D 上的凸函数. 如果上述不等式对 x, y 与任意 $\lambda \in (0, 1)$ 严格成立, 则称 f 是凸集 D 上的严格凸函数.

凸函数的定义表明, 如果 $f(x)$ 是凸集 D 上的凸函数, 则对于凸集 D 上的任意两点 x, y , 连结点 $(x, f(x))$ 与点 $(y, f(y))$ 之间的直线段位于函数图形 (曲线或曲面) 的上方, 图 1.4 给出了凸函数的几何直观表示.

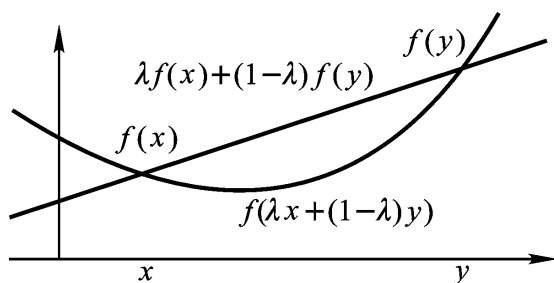


图 1.4 凸函数

同凸函数相对应的是凹函数, 一个函数 $f(x)$ 称为是凸集 D 上的 (严格) 凹函数, 如果 $-f(x)$ 是凸集 D 上的 (严格) 凸函数.

下面的定理表明凸规划 (可行域是凸集, 目标函数是凸函数的最优化) 问题的局部最优解必是全局最优解.

定理 1.2.13 设 x^* 是凸规划问题的一个局部最优解, 则 x^* 也是全局最优解. 如果目标函数是严格凸的, 则 x^* 是唯一全局最优解.

证明 定理的结论可用反证法加以证明. 设 x^* 是凸规划问题的一个局部最优解, 但不是全局最优解, 则存在另一可行点, 设为 y 满足 $f(y) < f(x^*)$. 由可行集的凸性, 对于任意 $0 < \lambda < 1$, 点 $x^* + (1-\lambda)y$ 都是可行点. 又根据目标函数的凸性有

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

这表明在 x^* 的任意小的邻域内都存在函数值小于 $f(x^*)$ 的可行点, 这与 x^* 是局部最优解相矛盾. 因此函数值小于 $f(x^*)$ 的可行点不存在, x^* 必是一个全局最优解.

当目标函数严格凸时可用类似的方法证明 x^* 是全局最优解, 留作练习.

最优化问题的可行域大都如同模型 (1.1.1) 是由约束条件确定的. 什么样的约束确定的可行域是凸的呢? 下面的定理对此给出了答案.

定理 1.2.14 考虑非空可行域

$$F = \{ x / c_i(x) = 0, \quad i = 1, 2, \dots, m \}.$$

如果每一个约束函数 $c_i(x)$ 是凹函数, 则可行域 F 是凸集.

证明 设 x, y 为两不同可行点, 即有 $c_i(x) = 0, c_i(y) = 0, i = 1, 2, \dots, m$. 由函数 $c_i(x), i = 1, 2, \dots, m$ 的凹性, 知 $-c_i(x), i = 1, 2, \dots, m$ 是凸函数. 再由凸函数的定义, 对 $\lambda \in (0, 1)$, 有

$$-c_i(\lambda x + (1 - \lambda)y) = -\lambda c_i(x) - (1 - \lambda)c_i(y) = 0, \quad i = 1, 2, \dots, m,$$

即有

$$c_i(\lambda x + (1 - \lambda)y) = 0, \quad \lambda \in (0, 1), \quad i = 1, 2, \dots, m.$$

这就证明了

$$\lambda x + (1 - \lambda)y \in F,$$

即 F 是凸集.

这一定理表明当最优化问题(1.1.1)中没有等式约束, 而不等式约束函数都是凹函数时, 则可行域是凸集. 由于线性约束 $a^T x + b = 0$ 可以表示成 $c(x) = 0$ 和 $-c(x) = 0$, 而 $c(x)$ 既是凸函数又是凹函数, 故构成凸规划可行域的约束中可以包含线性等式约束. 对于可行域是凸集的最优化问题, 如果目标函数是凸函数, 则所论的最优化问题是一个凸规划问题. 那么哪些函数是凸函数呢?

最常见也是最简单的凸函数是变量 x 的线性函数, 例如线性规划中的目标函数

$$c^T x = c_1 x_1 + c_2 x_2 + \dots + c_n x_n.$$

需要指出的是线性函数既是凸函数也是凹函数. 另一类常见的凸函数是二次函数

$$\begin{aligned} q(x) &= \frac{1}{2} x^T G x + b^T x + c \\ &= \frac{1}{2} \sum_{i,j=1}^n x_i G_{ij} x_j + \sum_{i=1}^n b_i x_i + c, \end{aligned}$$

其中

$$G = \begin{pmatrix} G_{11} & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} & \dots & G_{2n} \\ \dots & \dots & & \dots \\ G_{n1} & G_{n2} & \dots & G_{nn} \end{pmatrix}$$

是 $n \times n$ 阶对称矩阵, 即有 $G_{ji} = G_{ij}, i, j = 1, 2, \dots, n$. 当对称矩阵 G 正半定时, $q(x)$ 是凸函数; 当 G 是正定时, $q(x)$ 是严格凸函数; 当 G 是负半定时, $q(x)$ 是凹函数; 当 G 是不定矩阵时, $q(x)$ 既不是凸的也不是凹的. 下面的几个定理分别用函数值, 一阶导数和二阶导数刻画了一般的非线性函数是凸函数的基本特性, 他们也被称为凸函数的判别定理.

定理 1.2.15 函数 $f(x)$ 是 \mathbf{R}^n 上凸函数的充分必要条件是对任意 $x, y \in \mathbf{R}^n$, 单变量函数 $\varphi(t) = f(x + t(y-x))$ 是关于 t 的凸函数.

证明 先证必要性. 设 t_1, t_2 是 $[0, 1]$ 的任意两个值, 由 $\varphi(t)$ 的定义和 $f(x)$ 的凸性, 对任意 $\lambda \in [0, 1]$, 有

$$\begin{aligned} \varphi(\lambda t_1 + (1-\lambda)t_2) &= f(x + (\lambda t_1 + (1-\lambda)t_2)(y-x)) \\ &= f(x + \lambda t_1(y-x) + (1-\lambda)t_2(y-x)) \\ &= f(\lambda(x + t_1(y-x)) + (1-\lambda)(x + t_2(y-x))) \\ &= \lambda f(x + t_1(y-x)) + (1-\lambda)f(x + t_2(y-x)) \\ &= \lambda \varphi(t_1) + (1-\lambda)\varphi(t_2). \end{aligned}$$

由凸函数的定义知 $\varphi(t)$ 是关于 t 的凸函数.

再证充分性. 设定理的条件成立, 即对任意 $x, y \in \mathbf{R}^n$, 单变量函数 $\varphi(t) = f(x + t(y-x))$ 是关于 t 的凸函数. 任取 $x, y \in \mathbf{R}^n$, 设 $\bar{z} = x + t_1(y-x)$, $\hat{z} = x + t_2(y-x)$ 为方向 $x + y$ 上的两个点, 则对于任意 $\lambda \in [0, 1]$, 有

$$\begin{aligned} f(\lambda \bar{z} + (1-\lambda)\hat{z}) &= f(\lambda(x + t_1(y-x)) + (1-\lambda)(x + t_2(y-x))) \\ &= f(x + (\lambda t_1 + (1-\lambda)t_2)(y-x)) \\ &= \varphi(\lambda t_1 + (1-\lambda)t_2) \\ &= \lambda \varphi(t_1) + (1-\lambda)\varphi(t_2) \\ &= \lambda f(x + t_1(y-x)) + (1-\lambda)f(x + t_2(y-x)) \\ &= \lambda f(\bar{z}) + (1-\lambda)f(\hat{z}). \end{aligned}$$

由 $x, y \in \mathbf{R}^n$ 的任意性即知 $f(x)$ 是凸函数.

定理 1.2.16 设 $f(x)$ 是定义在凸集 D 上的一阶连续可微函数, 则

(1) $f(x)$ 是 D 上凸函数的充分必要条件是

$$f(y) - f(x) \leq f'(x)^T (y - x), \quad \forall x, y \in D. \quad (1.2.8)$$

(2) $f(x)$ 是 D 上严格凸函数的充分必要条件是

$$f(y) - f(x) > f'(x)^T (y - x), \quad \forall x, y \in D, x \neq y. \quad (1.2.9)$$

证明 (1) 先证必要性. 设 $f(x)$ 是凸集 D 上的凸函数, 则对任意 $x, y \in D$ 和任意 $\lambda \in (0, 1)$ 有

$$f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda)f(x),$$

由此得

$$\frac{f(x + \lambda(y-x)) - f(x)}{\lambda} \leq f(y) - f(x). \quad (1.2.10)$$

由泰勒展开式有

$$f(x + \lambda(y-x)) = f(x) + \lambda f'(x)^T (y-x) + o(\lambda \|y-x\|),$$

代入(1.2.10)式得

$$f'(x)^T (y-x) + \frac{o(\lambda \|y-x\|)}{\lambda} \leq f(y) - f(x).$$

两边关于 λ 取极限即得

$$f(y) - f(x) = f'(x)^T (y - x).$$

再证充分性. 设函数 $f(x)$ 满足条件 (1.2.8). 对任意 $x, y \in D$ 取 $\bar{x} = x + (1 - \lambda)y, \lambda \in (0, 1)$, 由 D 是凸集知 $\bar{x} \in D$. 根据条件 (1.2.8) 我们分别有

$$\begin{aligned} f(\bar{x}) + f'(\bar{x})^T (x - \bar{x}) &= f(x), \quad \forall x \in D, \\ f(\bar{x}) + f'(\bar{x})^T (y - \bar{x}) &= f(y), \quad \forall y \in D. \end{aligned}$$

用 λ 乘以第一式, 用 $1 - \lambda$ 乘以第二式后两式再相加得

$$\lambda f(\bar{x}) + (1 - \lambda) f(\bar{x}) = \lambda f(x) + (1 - \lambda) f(y),$$

由于 $\bar{x} = x + (1 - \lambda)y$, 由上式即得对任意 $x, y \in D$ 以及任意 $\lambda \in (0, 1)$ 有

$$f(x + (1 - \lambda)y) = \lambda f(x) + (1 - \lambda) f(y).$$

由凸函数的定义知 $f(x)$ 是凸集 D 上的凸函数. 这就完成了第一部分的证明.

定理的第二部分结论可用类似的方法证明, 留作练习.

定理 1.2.17 设 $f(x)$ 是非空开凸集 $D \subset \mathbf{R}^n$ 上的二阶连续可微函数. 则

(1) $f(x)$ 是 D 上凸函数的充分必要条件是 $f(x)$ 的 Hesse 矩阵 (二阶导数矩阵) $\nabla^2 f(x)$ 在 D 上正半定.

(2) 如果 $f(x)$ 的 Hesse 矩阵 $\nabla^2 f(x)$ 在 D 上正定, 则 $f(x)$ 是 D 上的严格凸函数; 反之, 如果 $f(x)$ 是 D 上的严格凸函数, 则 $\nabla^2 f(x)$ 在 D 上正半定.

证明 (1) 先证必要性. 任取 $\bar{x} \in D$, 由 D 是开凸集, 对于任意 $s \in \mathbf{R}^n$ 存在 $\lambda > 0$ 使对任意 $\lambda \in [0, \lambda_0]$ 有 $\bar{x} + \lambda s \in D$. 由 $f(x)$ 是 D 上的凸函数, 根据定理 1.2.16 有

$$f(\bar{x} + \lambda s) \leq f(\bar{x}) + \lambda f'(\bar{x})^T s, \quad \forall s \in \mathbf{R}^n, \quad \lambda \in [0, \lambda_0]. \quad (1.2.11)$$

由于 $f(x)$ 在 D 上二阶连续可微, 将 $f(\bar{x} + \lambda s)$ 在 \bar{x} 处二阶泰勒展开, 得

$$f(\bar{x} + \lambda s) = f(\bar{x}) + \lambda f'(\bar{x})^T s + \frac{1}{2} \lambda^2 s^T \nabla^2 f(\bar{x}) s + o(\lambda^2 s^2). \quad (1.2.12)$$

将 (1.2.12) 代入 (1.2.11) 得

$$\frac{1}{2} \lambda^2 s^T \nabla^2 f(\bar{x}) s + o(\lambda^2 s^2) \leq 0, \quad \forall s \in \mathbf{R}^n.$$

上式两边除以 λ^2 , 并令 $\lambda \rightarrow 0$, 取极限得

$$s^T \nabla^2 f(\bar{x}) s \leq 0, \quad \forall s \in \mathbf{R}^n, \quad \bar{x} \in D,$$

这就证明了 $\nabla^2 f(x)$ 在 D 上是正半定的.

再证充分性. 设 $\nabla^2 f(x)$ 在 D 上正半定. 任取 $x \in D$, 将 $f(x)$ 在点 $\bar{x} \in D$ 展开, 有

$$f(x) = f(\bar{x}) + f'(\bar{x})^T (x - \bar{x}) + \frac{1}{2} (x - \bar{x})^T \nabla^2 f(\xi) (x - \bar{x}), \quad (1.2.13)$$

其中 $\xi = \bar{x} + \lambda(x - \bar{x}) = x + (1 - \lambda)\bar{x}, 0 < \lambda < 1$. 由集合 D 的凸性知 $\xi \in D$, 因而 $\nabla^2 f(\xi)$ 正半定, 故有 $(x - \bar{x})^T \nabla^2 f(\xi) (x - \bar{x}) \geq 0$, 代入 (1.2.13) 即得

$$f(x) - f(\bar{x}) \leq f'(\bar{x})^T (x - \bar{x})$$

对任意 $x, \bar{x} \in D$ 成立. 根据定理 1.2.16 的结论(1)即知 $f(x)$ 是 D 上的凸函数.

(2) 设 $f(x)$ 的 Hesse 矩阵 $H^2 f(x)$ 在 D 上正定, 任取两不同点 $x, y \in D$, 将 $f(y)$ 在点 x 处展开, 有

$$f(y) = f(x) + f'(x)^T (y - x) + \frac{1}{2} (y - x)^T H^2 f(\xi) (y - x), \quad (1.2.14)$$

其中 $\xi = x + \theta(y - x) \in D, \theta \in (0, 1)$. 由 $H^2 f(x)$ 在 D 的正定性以及 $x \neq \bar{x}$ 有 $(x - \bar{x})^T H^2 f(\xi) (x - \bar{x}) > 0$, 代入(1.2.14)即得

$$f(y) - f(x) > f'(x)^T (y - x)$$

对任意不同的 $x, y \in D$ 成立, 根据定理 1.2.16 的结论(2)知 $f(x)$ 是 D 上的严格凸函数.

当 $f(x)$ 是 D 上的严格凸函数时, $f(x)$ 也是 D 上的凸函数, 由本定理的第一部分立得 $H^2 f(x)$ 在 D 上正半定的结论.

在这个定理中, 并没有严格凸函数的 Hesse 矩阵在 D 上正定的结论, 这可以用一个简单的单变量函数的例子 $f(x) = x^4$ 来说明这样的结论并不成立. 事实上, 这一函数显然是严格凸的, 但在点 $x = 0$ 处我们有 $H^2 f(x) = 12x^2 = 0$, 即 $H^2 f(x)$ 在 $x = 0$ 处不是正定的, 只是正半定.

根据定理 1.2.17 就可以明白为什么前述二次函数 $q(x)$ 在 $n \times n$ 阶对称矩阵 G 正定时是严格凸的, 在 G 正半定时是凸的, 在 G 负定时是严格凹的, 在 G 负半定时是凹的.

和凸函数关系密切的是水平集. 下面几个定理讨论与凸函数对应的水平集的性质.

定理 1.2.18 设 $D \subset \mathbf{R}^n$ 是非空凸集, f 是定义在 D 上的凸函数, α 是一个实数, 则水平集 $L = \{x \in D / f(x) \leq \alpha\}$ 是一个凸集.

证明 设 $x_1, x_2 \in L$, 于是 $x_1 \in D, x_2 \in D, f(x_1) \leq \alpha, f(x_2) \leq \alpha$. 今设 $\theta \in (0, 1), x = \theta x_1 + (1 - \theta)x_2$. 由 D 的凸性知 $x \in D$. 又由于 f 是凸函数, 故

$$f(x) \leq \theta f(x_1) + (1 - \theta)f(x_2) \leq \theta\alpha + (1 - \theta)\alpha = \alpha.$$

因此, $x \in L$, 从而 L 是凸集.

进一步, 若 f 是 D 上的连续凸函数, 则水平集 L 是闭凸集.

定理 1.2.19 设 f 是非空凸集 $D \subset \mathbf{R}^n$ 上的连续凸函数, 则水平集 $L = \{x \in D / f(x) \leq \alpha\}$ 是闭凸集.

证明 我们只要证明 L 是闭的即可. 设 x^* 是 L 的极限点, 则存在序列 $\{x_k\} \subset L, \{x_k\} \rightarrow x^*$. 于是, 由 f 的连续性,

$$f(x^*) = \lim_{k \rightarrow \infty} f(x_k),$$

从而 $x^* \in L$. 这表明 L 是闭集.

定理 1.2.20 设 f 在非空凸集 $D \subset \mathbf{R}^n$ 上二次连续可微, 且存在常数 $m > 0$, 使得

$$u^T \nabla^2 f(x) u \geq m \|u\|^2, \quad \forall x \in L, \quad u \in \mathbf{R}^n, \quad (1.2.15)$$

则水平集 $L(x_0) = \{x \in D \mid f(x) = f(x_0)\}$ 是有界闭凸集.

证明 由上面两个定理知 $L(x_0)$ 是闭凸集. 现在证明 $L(x_0)$ 的有界性.

因为水平集 $L(x_0)$ 是凸的, 由 (1.2.15) 有, $\forall x, y \in L(x_0)$,

$$m \|y - x\|^2 \leq (y - x)^T \nabla^2 f(x + (y - x))(y - x), \quad (0, 1).$$

由于 f 在 D 上二次连续可微, 故

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} (y - x)^T \nabla^2 f(\xi) (y - x)$$

$$= f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} m \|y - x\|^2,$$

其中 $\xi = x + \theta(y - x) \in L(x_0)$, m 与 x, y 无关. 因此, 对任意 $y \in L(x_0)$, $y \neq x_0$,

$$\begin{aligned} f(y) - f(x_0) &= \nabla f(x_0)^T (y - x_0) + \frac{1}{2} m \|y - x_0\|^2 \\ &\geq -\|\nabla f(x_0)\| \|y - x_0\| + \frac{1}{2} m \|y - x_0\|^2. \end{aligned} \quad (1.2.16)$$

由于 $f(y) = f(x_0)$, 故

$$\|y - x_0\| \leq \frac{2}{m} \|\nabla f(x_0)\|.$$

这表明水平集 $L(x_0) = \{x \in D \mid f(x) = f(x_0)\}$ 有界.

§ 1.3 最优性条件

最优性条件是指最优化问题 (1.1.1) 的最优解 (局部的或全局的) 所必须满足的条件, 常见并常用的有一阶必要条件和二阶必要条件. 至于充分必要条件只对一些特殊的最优化问题存在. 最优性条件不仅对于最优化理论的研究具有重要意义, 而且对最优化算法的设计和终止条件的确定起重要作用. 我们在本节首先介绍最优性的一阶必要条件, 其次再学习有关的二阶条件.

在给出方便使用的最优性条件的数学表达式之前, 我们先从几何直观上了解最优性条件的几何意义. 为此先引入两个同最优解相关的基本概念——可行方向和下降方向.

定义 1.3.1 设 $f(x)$ 为定义在空间 \mathbf{R}^n 上的连续函数, 点 $\bar{x} \in \mathbf{R}^n$, 若对于方向 $s \in \mathbf{R}^n$ 存在数 $\alpha > 0$ 使成立

$$f(\bar{x} + \alpha s) < f(\bar{x}), \quad \forall \alpha \in (0, \alpha_0),$$

则称 s 为 $f(x)$ 在 \bar{x} 处的一个下降方向. 在点 \bar{x} 处的所有下降方向的全体记为 $D(\bar{x})$.

下面的定理给出了在 $f(x)$ 连续可微时下降方向同函数 $f(x)$ 的梯度 $\nabla f(x)$

之间的关系.

定理 1.3.2 设函数 $f(x)$ 在点 \bar{x} 处连续可微, 如存在非零向量 $s \in \mathbf{R}^n$ 使成立

$$f(\bar{x})^T s < 0,$$

则 s 是 $f(x)$ 在点 \bar{x} 处的一个下降方向.

证明 对于充分小的 $\alpha > 0$, 将 $f(\bar{x} + \alpha s)$ 在点 \bar{x} 处展开有

$$f(\bar{x} + \alpha s) = f(\bar{x}) + \alpha f(\bar{x})^T s + o(\alpha \|s\|).$$

由 $\alpha > 0$ 以及 $f(\bar{x})^T s < 0$ 知, 存在 $\delta > 0$,

使对任意 $\alpha \in (0, \delta)$ 有

$$f(\bar{x} + \alpha s) < f(\bar{x}).$$

结合这两式有

$$f(\bar{x} + \alpha s) < f(\bar{x}), \quad \forall \alpha \in (0, \delta),$$

这就证明了 s 是 $f(x)$ 在点 \bar{x} 处的下降方向.

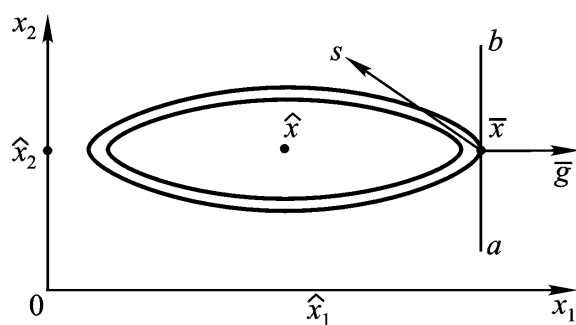


图 1.5 下降方向

图 1.5 给出了两个变量的函数

$$f(x) = f(x_1, x_2) = \frac{(x_1 - \hat{x}_1)^2}{a_1} + \frac{(x_2 - \hat{x}_2)^2}{a_2}$$

在点 \bar{x} 处的下降方向, 图中的椭圆曲线是函数的等高线 (也叫同位线). 点 $\hat{x} = (\hat{x}_1, \hat{x}_2)$ 是 $f(x)$ 的极小值点. 在点 \bar{x} , 向量 \bar{g} 是 $f(x)$ 在该点的梯度, 从点 \bar{x} 出发, 指向直线 ab 左侧的任何方向都是 $f(x)$ 在点 \bar{x} 的下降方向, 例如图中的方向 s .

设 $\bar{x} \in F$ 是一个可行点, 直观上一个非零向量 s 是最优化问题 (1.1.1) 在点 \bar{x} 的一个可行方向, 如果存在 $\alpha > 0$ 使得

$$\bar{x} + \alpha s \in F, \quad \forall \alpha \in (0, \delta).$$

然而可行方向的这种表达方式对于我们确定最优解的最优性条件无任何帮助. 为此给出可行方向的下述严格的定义.

定义 1.3.3 设 $\bar{x} \in F$ 为一可行点, $s \in \mathbf{R}^n$, 若存在非零方向序列 $s^{(k)}$, $k = 1, 2, \dots$ 和正数序列 α_k , $k = 1, 2, \dots$ 使成立

$$\bar{x} + \alpha_k s^{(k)} \in F, \quad \forall k,$$

且有 $\alpha_k > 0$, $s^{(k)} \rightarrow s$, 则称 s 是在 \bar{x} 处的一个可行方向. 在点 \bar{x} 的所有可行方向的全体记为 $F(\bar{x})$.

有了可行方向和下降方向的概念, 我们就很容易直观的理解最优化问题最优解所满足的最优性条件. 显然在一个最优解处, 不可能有任何一个既是可行的又是下降的方向. 因为根据可行方向和下降方向的定义, 如果存在任何可行的下降方向, 我们就能沿着这个方向找到使函数值更小的可行点, 这与最优解的定义相矛盾. 这就是由下述定理给出的一个最优解所必须满足的条件.

定理 1.3.4 考虑最优化问题(1.1.1), 设 x^* 是问题的一个局部最优解, 函数 $f(x)$ 连续可微, 则成立有

$$F(x^*) \cap D(x^*) = \emptyset.$$

证明 对于任意的可行方向 $s \in F(x^*)$, 我们证明 $s \notin D(x^*)$. 对可行方向 $s \in F(x^*)$, 存在可行点序列 $x^{(k)} = x^* + \alpha_k s^{(k)}$ 收敛于 x^* , 其中 $s^{(k)} \neq 0$, $\alpha_k > 0$, 且 $s^{(k)} \rightarrow s$, $\alpha_k \rightarrow 0$. 由泰勒展开式有

$$f(x^{(k)}) = f(x^* + \alpha_k s^{(k)}) = f(x^*) + \alpha_k \nabla f(x^*)^T s^{(k)} + o(\alpha_k \|s^{(k)}\|),$$

由于 x^* 是局部最优解, 对充分大的 k 有 $f(x^{(k)}) \geq f(x^*)$. 由此得

$$\alpha_k \nabla f(x^*)^T s^{(k)} + o(\alpha_k \|s^{(k)}\|) \geq 0.$$

在上式两端除以 α_k , 再令 $k \rightarrow \infty$ 取极限得

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

这就证明了 $s \notin D(x^*)$.

从图 1.6 我们可以进一步明白上述定理的意义, 图中的可行域 F 是由下述三个约束条件确定的

$$\begin{aligned} (x_1 - 2)^2 + (x_2 - 1)^2 &\leq 1, \\ (x_1 - 3)^2 + (x_2 - 1)^2 &\leq 1.1, \\ x_2 &\geq 0, \end{aligned}$$

$f(\bar{x})$ 和 $f(x^*)$ 给出了函数 $f(x)$ 的两

条等高线, 其中 \bar{x} 是可行点, x^* 是可行而且最优的点. 在可行点 \bar{x} , 由方向 $s^{(1)}$ 和方向 $s^{(2)}$ 所确定的伞形区域内的所有方向都是问题在点 \bar{x} 处的可行下降方向. 因此, 只要沿其中的任何一个方向移动, 都可以找到函数值小于 $f(\bar{x})$ 的可行点, 所以点 \bar{x} 不是问题的最优解. 而在可行点 x^* , 不存在任何可行的下降方向. 沿任何方向的移动, 或者可找到可行点但不能使函数值下降, 如方向 $s^{(3)}$; 或者可使函数值下降, 但却是不可行的, 如方向 $s^{(4)}$; 因此可行点 x^* 是问题的一个最优解.

最优化问题的可行域一般是由于等式或不等式表示的约束条件确定的, 然而由定义 1.3.3 所给出的可行方向同具体的约束条件无任何直接的联系, 而由定理 1.3.4 给出的最优性的条件对确定最优解没有任何直接的帮助. 为此, 有必要给出由确定可行域的约束条件表示的可行方向.

定义 1.3.5 设 $\bar{x} \in F$ 为一可行点, 可行域 F 由问题(1.1.1)中的约束条件确定. 设向量 $s \in \mathbb{R}^n$ 非零, 且有

$$s^T \nabla g_i(\bar{x}) = 0, \quad i = 1, 2, \dots, m,$$

$$s^T \nabla g_i(\bar{x}) < 0, \quad i \in I(\bar{x}),$$

则称 s 是可行域 F 在可行点 \bar{x} 的约束线性化后的可行方向, 其中 $I(\bar{x})$ 表示在可

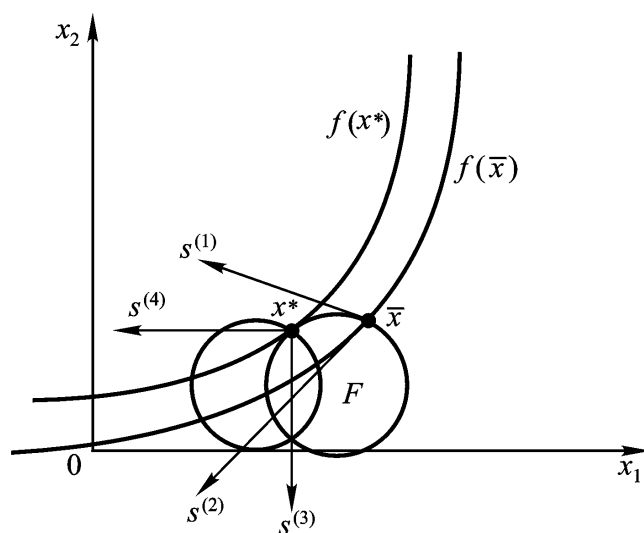


图 1.6 最优解与非最优解

点 \bar{x} 处的有效不等式约束集合. 记这样的可行方向的全体为 $L(\bar{x})$.

对于这样定义的可行方向, 如果在最优解 x^* 处有

$$F(x^*) = L(x^*)$$

成立, 那么定理 1.3.4 就可以表示成

$$L(x^*) \cap D(x^*) = \emptyset.$$

由于集合 $L(x^*)$ 是用问题的约束函数表示, 而集合 $D(x^*)$ 用问题的目标函数表示, 我们就可用问题中的函数来表达最优解的最优性条件. 下面的定理给出了在某一可行点 \bar{x} 处 $F(\bar{x}) = L(\bar{x})$ 成立的条件.

定理 1.3.6 设所有约束函数 $c_i(x)$, $i = 1, 2, \dots, p$ 在可行点 \bar{x} 处连续可微, 则有

$$(1) \quad F(\bar{x}) \subseteq L(\bar{x});$$

(2) 如果或者所有 $c_i(x)$, $i = 1, 2, \dots, m$, $i \in I(\bar{x})$ 是 x 的线性函数, 或者 $c_i(\bar{x})$, $i = 1, 2, \dots, m$, $i \in I(\bar{x})$ 线性无关, 则 $F(\bar{x}) = L(\bar{x})$ 成立.

证明 (1) 任取一个非零的可行方向 $s \in F(\bar{x})$, 则存在可行点序列 $x^{(k)} = \bar{x} + \lambda_k s^{(k)}$, 其中 $\lambda_k > 0$, $\lambda_k \rightarrow 0$ 和 $s^{(k)} \rightarrow s$. 由 $x^{(k)}$ 的可行性以及有效集的定义有

$$\begin{aligned} 0 &= c_i(x^{(k)}) = c_i(\bar{x} + \lambda_k s^{(k)}) \\ &= c_i(\bar{x}) + \lambda_k c_i(\bar{x})^T s^{(k)} + o(\lambda_k s^{(k)}) \\ &= \lambda_k c_i(\bar{x})^T s^{(k)} + o(\lambda_k s^{(k)}), \quad i \in I(\bar{x}), \\ 0 &= c_i(x^{(k)}) = c_i(\bar{x} + \lambda_k s^{(k)}) \\ &= c_i(\bar{x}) + \lambda_k c_i(\bar{x})^T s^{(k)} + o(\lambda_k s^{(k)}) \\ &= \lambda_k c_i(\bar{x})^T s^{(k)} + o(\lambda_k s^{(k)}), \quad i \in I(\bar{x}). \end{aligned}$$

在上述两式的两端都同除以 λ_k 后再对 $k \rightarrow \infty$ 取极限, 由函数的连续可微性得

$$\begin{aligned} c_i(\bar{x})^T s &= 0, \quad i \in I(\bar{x}), \\ c_i(\bar{x})^T s &\leq 0, \quad i \in K(\bar{x}), \end{aligned}$$

这就证明了 $s \in L(\bar{x})$. 由 $s \in F(\bar{x})$ 的任意性证明了 $F(\bar{x}) \subseteq L(\bar{x})$.

(2) 根据(1)的结论, 对于这两种情形我们只须证明 $L(\bar{x}) \subseteq F(\bar{x})$ 成立即得部分的结论. 首先考虑 $c_i(x)$, $i = 1, 2, \dots, m$, $i \in I(\bar{x})$ 是线性函数的情形, 这时数 $c_i(x)$ 可表示成 $c_i^T x + b_i$ 的形式, 其中 $c_i \in \mathbf{R}^n$ 即为函数的梯度向量, b_i 为纯量. 任取 $s \in L(\bar{x})$. 令 $x^{(k)} = \bar{x} + \lambda_k s^{(k)}$, 其中 $s^{(k)} = s$, $\lambda_k \rightarrow 0$. 将 $x^{(k)}$ 代入约束函数由可行性得

$$\begin{aligned} c_i(x^{(k)}) &= c_i^T (\bar{x} + \lambda_k s^{(k)}) + b_i \\ &= c_i^T \bar{x} + b_i + \lambda_k c_i^T s^{(k)} \\ &= \lambda_k c_i^T s^{(k)} = 0, \quad i \in I(\bar{x}), \\ c_i(x^{(k)}) &= c_i^T (\bar{x} + \lambda_k s^{(k)}) + b_i \\ &= c_i^T \bar{x} + b_i + \lambda_k c_i^T s^{(k)} \end{aligned}$$

$$= {}_k C_i^T s^{(k)} = 0, \quad i \in I(\bar{x}),$$

这就证明了 $s \in F(\bar{x})$, 再由 $s \in L(\bar{x})$ 的任意性得 $L(\bar{x}) = F(\bar{x})$.

再证 $c_i(\bar{x}), i=1, 2, \dots, m, i \in I(\bar{x})$ 线性无关的情形. 为简化符号, 不失一般性, 设 $I(\bar{x}) = \{m+1, \dots, k\}, (k \leq n)$. 由 $c_i(\bar{x}), i=1, 2, \dots, k$ 的线性无关性, 存在向量 a_{k+1}, \dots, a_n 使得 n 个向量 $c_1(\bar{x}), \dots, c_k(\bar{x}), a_{k+1}, \dots, a_n$ 线性无关. 任取 $s \in L(\bar{x})$, 考察下列含有参数 μ 的 n 个方程的方程组

$$r(x, \mu) = 0, \quad (1.3.1)$$

其中

$$\begin{aligned} r_i(x, \mu) &= c_i(x) - \mu s^T c_i(\bar{x}), \quad i=1, 2, \dots, k, \\ r_i(x, \mu) &= (x - \bar{x})^T a_i - \mu s^T a_i, \quad i=k+1, \dots, n. \end{aligned}$$

由于 $c_i(\bar{x}) = 0, i=1, 2, \dots, k$, 显然, 当 $\mu=0$ 时 $x=\bar{x}$ 是方程组的一个解. 又由于该方程组关于 x 的雅可比矩阵

$${}_x r(x, \mu) = [c_1(x), \dots, c_k(x), a_{k+1}, \dots, a_n]^T$$

在点 \bar{x} 非奇异, 由隐函数存在定理存在 $\mu=0$ 的某个邻域 μ 和 \bar{x} 的某个邻域 \bar{x} , 使得对任意 $\mu \in \mu$, 方程组 (1.3.1) 的解存在且唯一, 记为 $x(\mu) \in \bar{x}$. 作为 μ 的函数, $x(\mu)$ 在邻域 μ 内连续可微, 且有 $x(0) = \bar{x}$. 由复合函数的求导法则可以由方程组

$$\frac{dr_i}{d\mu} = \frac{r_i}{\mu} + \sum_{j=1}^n \frac{r_i}{x_j} \frac{dx_j}{d\mu} = 0, \quad i=1, 2, \dots, n$$

确定 $x(\mu)$ 关于 μ 的导数 $\frac{dx}{d\mu}$. 在 $\mu=0$ 时上述方程组可表示为

$$[c_1(\bar{x}), \dots, c_k(\bar{x}), a_{k+1}, \dots, a_n]^T \frac{dx}{d\mu} \Big|_{\mu=0} - s = 0.$$

由于此方程组的系数矩阵非奇异, 方程组只有零解, 即得

$$\frac{dx}{d\mu} \Big|_{\mu=0} = s.$$

根据 $x(\mu)$ 的定义, 在邻域 μ 内我们有

$$r(x(\mu), \mu) = 0,$$

对于其中的 $i=1, 2, \dots, k$ 有

$$r_i(x(\mu), \mu) = c_i(x(\mu)) - \mu s^T c_i(\bar{x}) = 0.$$

由 $s \in L(\bar{x})$ 的定义有

$$\begin{aligned} c_i(x(\mu)) &= \mu s^T c_i(\bar{x}) = 0, \quad i=1, 2, \dots, m, \\ c_i(x(\mu)) &= \mu s^T c_i(\bar{x}) = 0 \quad (\mu > 0), \quad i=m+1, \dots, k. \end{aligned}$$

而对于非有效的不等式约束, 由于 $c_i(\bar{x}) > 0, i=k+1, \dots, p$, 对充分小的 $\mu > 0$ 有

$$c_i(x(\mu)) = c_i(\bar{x}) + c_i(\bar{x})^T (x(\mu) - \bar{x}) + o(\|x(\mu) - \bar{x}\|) > 0.$$

这就证明了 $x(\mu)$ 是一个可行点序列. 再取 μ_k 为一单调递减趋于零的正序列, 利

用 $x(\mu)$ 是 μ 连续可微函数的性质有

$$x^{(k)} = x(\mu_k) = \bar{x} + \mu_k s + o(\mu_k s),$$

即 $x^{(k)}$ 是具有可行方向 s 的可行点序列, 因而 $s \in F(\bar{x})$. 再由 $s \in L(\bar{x})$ 的任意性 $L(\bar{x}) \subset F(\bar{x})$, 这就完成了定理的证明.

定理 1.3.6 的(2)中关于约束函数的条件称为约束规范, 有许多不同的约束规范条件和表现形式, 但最常见也是最方便使用的还是由上述定理的(2)所给出的约束规范条件. 有了上面的准备, 我们现在就可以给出最优解的一阶必要条件, 又称 Kuhn Tucker 条件.

定理 1.3.7 设 $x^* \in F$ 是最优化问题 (1.1.1) 的一个局部最优解, $f(x)$, $c_i(x)$, $i=1, 2, \dots, p$ 在 x^* 的一个邻域内连续可微. 如果对所有的等式约束和在 x^* 的有效约束, 或者都是 x 的线性函数, 或者他们在点 x^* 的梯度向量线性无关, 则存在向量 $\lambda^* = (\lambda_1^*, \lambda_2^*, \dots, \lambda_p^*)^T$ 使成立

$$\begin{aligned} f(x^*) - \sum_{i=1}^p \lambda_i^* c_i(x^*) &= 0, \\ \lambda_i^* &= 0, \quad i = m+1, \dots, p, \\ \lambda_i^* c_i(x^*) &= 0, \quad i = m+1, \dots, p. \end{aligned} \quad (1.3.2)$$

证明 根据定理的条件, 由定理 1.3.6 在点 x^* 成立有 $F(x^*) = L(x^*)$. 再据定理 1.3.4 有 $L(x^*) \cap D(x^*) = \emptyset$. 根据集合 $D(x^*)$ 与 $L(x^*)$ 的定义, 这以表示成下述不等式组无解,

$$\begin{aligned} f(x^*)^T s &< 0, \\ c_i(x^*)^T s &= 0, \quad i = 1, 2, \dots, m, \\ c_i(x^*)^T s &\leq 0, \quad i \in I(x^*). \end{aligned}$$

将上述不等式组改写成

$$\begin{aligned} -f(x^*)^T s &> 0, \\ -c_i(x^*)^T s &\leq 0, \quad i = 1, 2, \dots, m, \\ c_i(x^*)^T s &\leq 0, \quad i = 1, 2, \dots, m, \\ -c_i(x^*)^T s &\leq 0, \quad i \in I(x^*), \end{aligned}$$

或矩阵向量的表示

$$As \leq 0, \quad -f(x^*)^T s > 0, \quad (1.3.3)$$

其中 $A = [-c_i(x^*)^T (i=1, 2, \dots, m), -c_i(x^*)^T (i=1, 2, \dots, m), -c_i(x^*)^T (i \in I(x^*))]$. 由不等式组 (1.3.3) 无解, 根据引理 1.2.5 存在非负向量 $y = (y^{(1)}, y^{(2)}, y^{(3)})^T$ 使得

$$A^T y = -f(x^*),$$

其中 $y^{(1)} \in \mathbf{R}^m, y^{(2)} \in \mathbf{R}^m, y^{(3)} \in \mathbf{R}^{|\mathcal{I}(x^*)|}$. 将上式展开得

$$\begin{aligned} -f(x^*) &= \sum_{i=1}^m y_i^{(1)} [-c_i(x^*)] + \sum_{i=1}^m y_i^{(2)} c_i(x^*) + \sum_{i \in \mathcal{I}(x^*)} y_i^{(3)} [-c_i(x^*)] \\ &= -\sum_{i=1}^m (y_i^{(1)} - y_i^{(2)}) c_i(x^*) - \sum_{i \in \mathcal{I}(x^*)} y_i^{(3)} c_i(x^*). \end{aligned}$$

令 $\bar{y}_i = y_i^{(1)} - y_i^{(2)}, i = 1, 2, \dots, m, \bar{y}_i = y_i^{(3)}, i \in \mathcal{I}(x^*)$, 并对非有效约束指标 i 置 $\bar{y}_i = 0$, 则有

$$f(x^*) = \sum_{i=1}^p \bar{y}_i c_i(x^*),$$

对于不等式约束还有 $\bar{y}_i = 0, i = m+1, \dots, p$. 又因对有效的不等式约束有 $c_i(x^*) = 0, \bar{y}_i = 0$, 对非有效的不等式约束有 $\bar{y}_i = 0$, 所以有 $\bar{y}_i c_i(x^*) = 0, i = m+1, \dots, p$. 这就完成了定理的证明.

在大部分最优化研究的文献中, 称最优解 x^* 所满足的一阶必要条件 (1.3.2) 为 Kuhn Tucker 条件或 KKT 条件, 满足 Kuhn Tucker 条件的点为 Kuhn Tucker 点或 KKT 点; 称式 (1.3.2) 中的第三个等式为互补松弛 (Complementarity) 条件, 如果对于任意 $i = m+1, \dots, p, c_i(x^*)$ 和 \bar{y}_i 中有且仅有一个取零值, 则称严格互补松弛条件成立.

由于无约束最优化问题中无任何约束条件, 由定理 1.3.7 立即可以得到无约束最优化问题最优解的一阶必要条件是

$$f'(x^*) = 0, \quad (1.3.4)$$

即在无约束最优化问题的最优解处, 任何方向都不是目标函数的下降方向. 习惯上把满足条件 (1.3.4) 的点称为平稳点或驻点, 这是因为无约束问题的最优点一定满足条件 (1.3.4), 但满足 (1.3.4) 的点不一定是无约束问题的局部最优解, 单变量函数 $f(x) = x^3$ 就提供了这样的一个例子. 在点 $x^* = 0$, 有 $f'(x^*) = 0$, 但 $x^* = 0$ 却不是其最优解. 这种情况同样适用于约束最优化问题, 即约束最优化问题的最优解在约束规范条件满足时必定是 Kuhn Tucker 点, 但满足 Kuhn Tucker 条件的可行点未必是最优解. 大部分最优化方法都是确定满足 Kuhn Tucker 条件的 KKT 点. 下面的定理表明对于凸规划问题 Kuhn Tucker 条件却是最优解的充分条件.

定理 1.3.8 设凸规划问题的目标函数与约束函数都连续可微, 如果在可行点 x^* 处约束规范条件和 Kuhn Tucker 条件成立, 则 x^* 是问题的全局最优解.

证明 设 x^* 是 Kuhn Tucker 点, \bar{y} 是相应的向量. 根据凸规划对函数的要求, 我们知, $f(x)$ 是凸函数, $c_i(x), i = 1, \dots, m$ (如果 $m = 0$) 是线性函数, $c_i(x), i = m+1, \dots, p$ (如果 $p > m$) 是凹函数. 因此对任意 $x \in F$ 有

$$\begin{aligned}
 f(x) &= f(x^*) + f'(x^*)^T (x - x^*), \\
 c_i(x) &= c_i(x^*) + c_i'(x^*)^T (x - x^*), \quad i = 1, 2, \dots, m, \\
 -c_i(x) &= -c_i(x^*) - c_i'(x^*)^T (x - x^*), \quad i = m+1, \dots, p.
 \end{aligned}$$

由于 $c_i^* = 0, i = m+1, \dots, p$, 对任意 $x \in F$ 有 $\sum_{i=1}^p c_i^*(x) = 0$. 因此对任意 $x \in F$ 得

$$\begin{aligned}
 f(x) &= f(x) - \sum_{i=1}^p c_i^*(x) \\
 &= f(x^*) + f'(x^*)^T (x - x^*) - \sum_{i=1}^p c_i^* [c_i(x^*) + c_i'(x^*)^T (x - x^*)] \\
 &= f(x^*) - \sum_{i=1}^p c_i^*(x^*) + \sum_{i=1}^p c_i^*(x^*)^T (x - x^*) \\
 &= f(x^*).
 \end{aligned}$$

这就证明了 x^* 是凸规划问题的全局最优解.

Kuhn Tucker 条件中的向量 λ^* 称为最优 Lagrange 乘子. 事实上, 引入问题 (1.1.1) 的 Lagrange 函数

$$L(x, \lambda) = f(x) - \sum_{i=1}^p \lambda_i c_i(x),$$

那么我们可以看到条件 (1.3.2) 中的第一个方程可以写成

$$\nabla_x L(x^*, \lambda^*) = 0,$$

即 Lagrange 函数 $L(x, \lambda)$ 关于 x 的一阶偏导数在 (x^*, λ^*) 处取零值. 如果不考虑在点 x^* 的无效约束, 则在点 x^* 的可行性条件 $c_i(x^*) = 0, i = 1, 2, \dots, m, i \in I(x^*)$, 又可表示成

$$L(x^*, \lambda^*) = 0.$$

即 Lagrange 函数关于 λ 的一阶偏导数在 (x^*, λ^*) 处也取零值, 因而 (x^*, λ^*) 是 Lagrange 函数的一个平稳点.

下面我们讨论最优解的二阶最优性条件, 为简化讨论, 假定在点 x^* 处严格互补松弛条件成立, 并定义在点 x^* 处的有效约束的零可行方向集.

定义 1.3.9 设在可行点 x^* 处严格互补松弛条件成立, 如果存在非零向量序列 $s^{(k)}$ 和正数序列 $\alpha_k > 0$ 使有

$$x^{(k)} = x^* + \alpha_k s^{(k)} \in F, \quad \alpha_k \rightarrow 0,$$

$$c_i(x^{(k)}) = 0, \quad i = 1, 2, \dots, m, i \in I(x^*),$$

且有 $s^{(k)} \rightarrow s, \alpha_k \rightarrow 0$, 则称 s 为可行域在点 x^* 处的零可行方向, 所有这些方向的集

合称为零可行方向集,记为 $Z(x^*)$ 称满足

$$s^T c_i(x^*) = 0, \quad i = 1, 2, \dots, m, \quad i \in I(x^*)$$

的非零向量 s 为约束线性化后的零可行方向,所有这些方向的集合称为约束线性化后的零可行方向集,记为 $L(x^*)$.

显然,集合 $Z(x^*)$ 是集合 $F(x^*)$ 的子集,而集合 $L(x^*)$ 是集合 $L(x^*)$ 的子集.同样可以证明,在定理 1.3.6 的约束规范条件下成立

$$Z(x^*) = L(x^*).$$

留作练习.

下面的定理给出了最优解的二阶必要条件.

定理 1.3.10 考虑最优化问题 (1.1.1), 设 $x^* \in F$ 是其最优解, 且函数 $f(x)$ 与 $c_i(x)$, $i = 1, 2, \dots, p$ 二阶连续可微. 又设定理 1.3.6 的约束规范条件在 x^* 成立, 从而存在 Lagrange 乘子向量 λ^* 使 Kuhn Tucker 条件成立. 设严格补松弛条件成立, 则有

$$s^T \frac{\partial^2}{\partial x x} L(x^*, \lambda^*) s \geq 0, \quad \forall s \in Z(x^*),$$

其中 $\frac{\partial^2}{\partial x x} L(x^*, \lambda^*)$ 是 Lagrange 函数在 (x^*, λ^*) 处关于 x 的二阶偏导数矩阵,

$$\frac{\partial^2}{\partial x x} L(x, \lambda) = \frac{\partial^2}{\partial x x} f(x) - \sum_{i=1}^p \lambda_i \frac{\partial^2}{\partial x x} c_i(x).$$

证明 任取非零可行方向 $s \in Z(x^*)$, 则存在可行点序列 $x^{(k)} = x^* + s^{(k)}$ 使得 $c_i(x^{(k)}) = 0$, $i = 1, 2, \dots, m$, $i \in I(x^*)$, 其中 $s^{(k)} \rightarrow s$. 由于对于无效不行约束有 $\lambda_i^* = 0$, 因而对这样的序列有

$$\sum_{i=1}^p \lambda_i^* c_i(x^{(k)}) = 0.$$

由各函数的二阶连续可微性并利用 Kuhn Tucker 条件有

$$\begin{aligned} f(x^{(k)}) - f(x^*) &= \sum_{i=1}^p \lambda_i^* c_i(x^{(k)}) \\ &= L(x^* + s^{(k)}, \lambda^*) \\ &= L(x^*, \lambda^*) + \frac{1}{k} \frac{\partial}{\partial x} L(x^*, \lambda^*)^T s^{(k)} + \\ &\quad \frac{1}{2} \frac{\partial^2}{\partial x x} L(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2) \\ &= f(x^*) + \frac{1}{2} \frac{\partial^2}{\partial x x} L(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2). \end{aligned}$$

由于 x^* 是局部最优解, 对充分大的 k 有 $f(x^{(k)}) - f(x^*) \geq 0$ 成立, 由此得

$$\frac{1}{2} \frac{\partial^2}{\partial x x} L(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2) \geq 0.$$

在上式两边同除以 $\|s^{(k)}\|^2$ 后再令 $k \rightarrow \infty$ 取极限得

$$s^T \frac{\partial^2}{\partial x x} L(x^*, \lambda^*) s \geq 0.$$

由 s 在集合 $Z(x^*)$ 中的任意性证明了定理的结论.

由于定理 1.3.6 的约束规范条件成立时有 $Z(x^*) = L(x^*)$ 成立, 上述定理二阶必要条件可以用下述更直接的方式给出.

定理 1.3.11 设 x^* 是问题 (1.1.1) 的最优解且函数 $f(x)$ 与 $g_i(x)$, $i = 1, 2, \dots, p$ 二阶连续可微. 又设定理 1.3.6 的约束规范条件在点 x^* 成立, 从而存在 Lagrange 乘子向量 λ^* 使 Kuhn Tucker 条件成立. 如果严格互补松弛条件在 x^* 成立, 则

$$s^T \frac{1}{2} L(x^*, \lambda^*) s = 0$$

对一切满足

$$s^T \nabla g_i(x^*) = 0, \quad i = 1, 2, \dots, m, \quad i \in I(x^*)$$

的方向 s 均成立.

下面的定理则给出了最优解的二阶充分条件.

定理 1.3.12 考虑最优化问题 (1.1.1), 函数 $f(x)$ 与 $g_i(x)$, $i = 1, 2, \dots, p$ 均二阶连续可微. 设对于可行点 x^* 存在 Lagrange 乘子向量 λ^* 使 Kuhn Tucker 条件成立. 若成立有

$$s^T \frac{1}{2} L(x^*, \lambda^*) s > 0, \quad \forall s \in Z(x^*),$$

则 x^* 是问题 (1.1.1) 的严格局部最优解.

证明 定理的证明采用反证法. 设 x^* 不是问题的严格局部最优解, 则存在收敛于 x^* 的可行点序列 $x^{(k)}, x^{(k)} \rightarrow x^*, k = 1, 2, \dots$, 使成立

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

不失一般性, 设

$$s^{(k)} = \frac{x^{(k)} - x^*}{\|x^{(k)} - x^*\|} \quad s,$$

并令 $\alpha_k = \|x^{(k)} - x^*\|$, 则 $x^{(k)} = x^* + \alpha_k s^{(k)}$ 且 $\alpha_k \rightarrow 0$, 因此 $s \in F(x^*) = L(x^*)$.

由函数 $f(x)$ 的连续可微性有

$$f(x^*) - f(x^{(k)}) = f(x^*) - f(x^* + \alpha_k s^{(k)}) = -\alpha_k \nabla f(x^*)^T s^{(k)} + o(\alpha_k s^{(k)}),$$

由此得

$$\alpha_k \nabla f(x^*)^T s^{(k)} + o(\alpha_k s^{(k)}) = 0.$$

在上式两端除以 α_k 后再令 $k \rightarrow \infty$ 取极限得

$$\nabla f(x^*)^T s = 0. \quad (1.3.5)$$

由于 $s \in L(x^*)$ 而 $Z(x^*) = L(x^*)$, 有两种可能性: 即 $s \in Z(x^*)$ 和 $s \notin Z(x^*)$. 首先证明不可能有 $s \notin Z(x^*)$. 为此设 $s \notin Z(x^*)$ 即存在有指标 $i_0 \in I(x^*)$ 使得 $s^T \nabla g_{i_0}(x^*) > 0$, 这时由 Kuhn Tucker 条件有

$$s^T \nabla f(x^*) = -\sum_{i=1}^p \lambda_i^* s^T \nabla g_i(x^*) > 0,$$

这与式 (1.3.5) 相矛盾, 这一矛盾表明 $s \in Z(x^*)$ 不成立, 因而有 $s \notin Z(x^*)$.

由 $x^{(k)}$ 的可行性和 $\lambda_i^* = 0, i = m+1, \dots, p$, 有

$$\begin{aligned} f(x^{(k)}) &= f(x^*) - \sum_{i=1}^p \lambda_i^* c_i(x^{(k)}) = L(x^{(k)}, \lambda^*) \\ &= L(x^*, \lambda^*) + \frac{1}{2} s^{(k)T} L_{xx}(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2) \\ &= f(x^*) + \frac{1}{2} s^{(k)T} L_{xx}(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2), \end{aligned}$$

注意到 $f(x^{(k)}) \geq f(x^*)$, 由上式得

$$\frac{1}{2} s^{(k)T} L_{xx}(x^*, \lambda^*) s^{(k)} + o(\|s^{(k)}\|^2) \geq 0.$$

两边除以 $\|s^{(k)}\|^2$, 并令 $k \rightarrow \infty$ 取极限得

$$s^T L_{xx}(x^*, \lambda^*) s \geq 0,$$

这与定理的条件相矛盾, 由此证明了 x^* 是问题的严格局部最优解.

当定理 1.3.6 的约束规范条件在 x^* 处成立时, 有 $Z(x^*) = L(x^*)$ 成立, 这上述二阶充分条件可用下述更直接的方式给出.

定理 1.3.13 设最优化问题 (1.1.1) 的函数 $f(x)$ 与 $c_i(x), i = 1, 2, \dots, p$ 均二阶连续可微, 在可行点 x^* 处定理 1.3.6 的约束规范条件成立. 若存在 Lagrange 乘子向量 λ^* 使 Kuhn-Tucker 条件和严格松弛互补条件成立, 且对所有满足

$$s^T c_i(x^*) = 0, i = 1, 2, \dots, m, \quad i \in I(x^*)$$

的非零向量 s 有

$$s^T L_{xx}(x^*, \lambda^*) s > 0,$$

则 x^* 是问题 (1.1.1) 的一个严格局部最优解.

由于无约束最优化问题无任何约束, 由上述几个最优解的二阶条件 (必要的和充分的), 直接可以得到无约束最优化问题最优解的下列二阶必要条件和二阶充分条件.

定理 1.3.14 (二阶必要条件) 设 x^* 是无约束最优化问题的一个最优解, $f(x)$ 在 x^* 的一个邻域内二阶连续可微, 则有 $\nabla f(x^*) = 0$, 且 $f(x)$ 在 x^* 的二阶 Hesse 阵正半定, 即成立

$$s^T \nabla^2 f(x^*) s \geq 0, \quad \forall s \in \mathbf{R}^n.$$

定理 1.3.15 (二阶充分条件) 设 $f(x)$ 在 x^* 的一个邻域内二阶连续可微, 且有 $\nabla f(x^*) = 0$, $\nabla^2 f(x^*)$ 正定, 即成立

$$s^T \nabla^2 f(x^*) s > 0, \quad \forall s \in \mathbf{R}^n, s \neq 0,$$

则 x^* 是 $f(x)$ 的无约束优化问题的一个严格局部最优解.

最后需要提醒的一点是, 上述有关严格互补松弛条件的假定完全可以放宽. 在严格互补松弛条件不成立时, 我们需要对不等式约束的非零拉格朗日乘子定

义一个指标集 $I(x^*) = \{i \mid \bar{c}_i > 0, i = m+1, \dots, p\}$. 然后对指标集 $i = 1, 2, \dots, m, i \notin I(x^*)$ 定义相应的零可行方向集合 $Z(x^*)$ 和 $L(x^*)$, 可以得到同上述一个定理同样的二阶最优性条件.

§ 1.4 最优化方法概述

现实生活中大量的最优化问题, 除极个别例子外, 即使一些看似简单的问题, 一般都不可能给出问题的解的显式表达式. 以下述简单的无约束最优化问题为例

$$\min f(x) = e^{2x} + x^2,$$

根据最优性的一阶必要条件, 最优解必定是方程

$$f'(x) = 2e^{2x} + 2x = 0$$

的解. 由 $f(x)$ 的连续性, 当 $x \rightarrow -\infty$ 有 $f(x) \rightarrow \infty$ 和 $x \rightarrow \infty$ 有 $f(x) \rightarrow \infty$, 上述方程的解存在, 但我们却无法得出解的任何解析表达式. 因此求最优化问题的解, 一般用迭代的方法, 其基本思想为, 给定最优解的一个初始估计, 记为 $x^{(0)}$, 方法产生一个逐步改善的有限或无限的迭代序列 $\{x^{(k)}\}$, 在 $\{x^{(k)}\}$ 是有限点列时, 它的最后一个点是 Kuhn Tucker 点; 在 $\{x^{(k)}\}$ 是无限点列时, 其任意一个聚点是 Kuhn Tucker 点, 并在对最优解的估计满足指定的精度要求时停止迭代. 根据最优性的一阶必要条件, 最优解一定是 Kuhn Tucker 点, 因此理论上由迭代法所确定的解一般是 Kuhn Tucker 点, 再由方法的其他一些特性, 如下降性可以确保所得的 Kuhn Tucker 点是所论问题的最优解或最优解的近似.

随着最优化理论和最优化技术的发展, 以及最优化越来越广泛的应用, 已经有大量的最优化方法适用于各种不同类型的最优化问题的求解. 本书作为最优化理论和方法的一本教科书, 将重点介绍一些最基本, 也是最普遍使用的最优化方法. 尽管最优化问题的类型和求解的方法千差万别, 但所有的方法都具有下述基本的迭代格式:

算法 1.4.1 (最优化方法的基本迭代格式)

1. 给定最优解的一个初始估计 $x^{(0)}$, 置 $k=0$;
2. 如果 $x^{(k)}$ 满足对最优解估计的终止条件, 停止迭代;
3. 确定一个改善 $x^{(k)}$ 的修正量 $\Delta x^{(k)}$;
4. 得到最优解的一个更好的估计 $x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$, 置 $k=k+1$ 后转步 2.

在上述迭代法的基本格式中涉及初始点的选取; 迭代点好坏的判定; 迭代的终止条件; 以及最重要也是最关键的修正量 $\Delta x^{(k)}$ 的确定. 下面对这些分别加以简单的讨论.

初始点的选取依赖于方法的收敛性能. 一个算法称为收敛的, 如果算法产生的序列 $\{x^{(k)}\}$ 满足

$$\lim_k x^{(k)} - x^* = 0,$$

其中 x^* 是问题的 Kuhn Tucker 点. 一个算法如果对于任意给定的初始点都能够收敛, 就说这个方法全局收敛或整体收敛. 有些算法只有当初始点接近或充分接近最优解时才有收敛性, 称这样的算法为局部收敛的方法. 因此对于全局收敛的算法, 初始点的选取可以没有任何的限制, 而对于局部收敛的算法, 则要求初始点应尽可能接近最优解. 然而由于最优解是未知的, 选取一个好的初始点也是一个困难的问题. 对于大量实际的最优化问题一般可以从以前的实践经验确定合适的最优解的初始估计.

在最优化方法中, 一般要选用一个评价函数 (Merit Function) 来评价一个迭代点的好坏. 对于无约束最优化问题, 由于没有约束条件, 通常就用目标函数 $f(x)$ 作为评价函数. 以无约束极小化问题

$$\min f(x)$$

为例, 如果有 $f(x^{(k+1)}) < f(x^{(k)})$, 就说点 $x^{(k+1)}$ 要好于点 $x^{(k)}$, 即要求产生的迭代序列 $\{x^{(k)}\}$ 使评价函数值单调下降. 对于约束最优化问题, 则要复杂一些. 如果迭代点都是可行点, 当然可以直接用目标函数作评价函数, 这适用于迭代点都是可行点的方法. 但是对于那些迭代点不是可行点的方法, 判定一个点的好坏既要考虑目标函数值的大小, 还要考虑这个点的可行程度 (离可行域的距离). 因此这类方法采用的评价函数中一般既包含目标函数又包含约束函数. 对此有许多不同类型的评价函数, 有关这些评价函数我们将在第七章学习约束最优化方法时加以介绍.

迭代的终止条件在不同的最优化方法中也是不同的. 理论上, 根据最优性的一阶必要条件, 以及算法的设计思想, Kuhn Tucker 条件是最合适的迭代终止条件. 以无约束最优化问题为例, 可以用

$$f(x^{(k)}) \quad (1.4.1)$$

来终止迭代, 其中 $\epsilon > 0$ 是给定的精度要求. 可是这样的终止准则依赖于函数在极小点邻域内的性质. 图 1.7 就单变量函数的情形给出了一个例子. 图中点 x_1^* 和 x_2^* 是函数 $f(x)$ 的两个无约束局部最优解. 点 \bar{x}_1 在 x_1^* 的邻域内, 点 \bar{x}_2 在 x_2^* 的邻域内, 从图上可以看出, 尽管点 \bar{x}_1 离最优解 x_1^* 还有一段距离, 但由于函数在这个邻域内比较

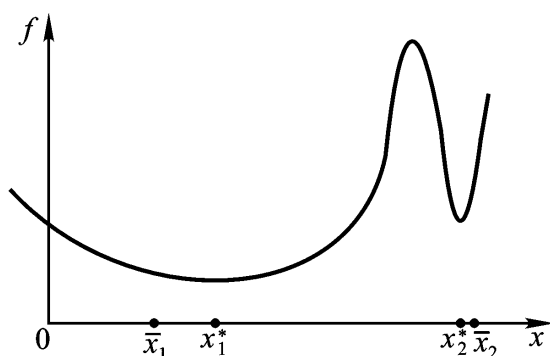


图 1.7 极小点邻域内的梯度

平坦,在点 \bar{x}_1 的梯度值已经很小;而对于点 \bar{x}_2 ,尽管它已十分接近最优解 x_2^* ,但由于函数在这个邻域内呈现比较陡的特性,在该点的梯度值还相对较大.正因为如此,准则(1.4.1)一般适用于收敛速度比较慢的算法.

收敛速度是迭代方法的又一重要性质.对于一个不可能在有限步内找到最优解的最优化方法,我们不仅要求它收敛,还要求它有较快的收敛速度,这是因为一个收敛很慢的方法往往需要很长的时间才能得到满足精度要求的最优解的近似,因而不是一个有效的方法.设向量序列 $\{x^{(k)}\} \in \mathbf{R}^n$ 收敛于 x^* ,定义误差序列

$$e_k = x^{(k)} - x^*.$$

如果存在常数 C 和 r 使成立

$$\lim_k \frac{e_{k+1}}{e_k} = C,$$

就说序列 $\{x^{(k)}\}$ 以 C 为因子 r 阶收敛于 x^* .最常见的为 $r=1$ 与 $r=2$ 的情形,当 $r=1, 0 < C < 1$ 时称为线性收敛,这时的误差序列具有如下的性能:

$$e_{k+1} = C e_k,$$

依最简单的情形,在上式中取等号,设初始误差为 1,如果 $C=0.5$,则误差序列为

$$1, 0.5, 0.25, 0.125, 0.0625, \dots,$$

如果 $C=0.1$,则误差序列为

$$1, 0.1, 0.01, 0.001, \dots,$$

可以看出 C 越小,收敛越快.如果从同一个初始点开始,收敛快的算法可以用较少的迭代次数达到预定的精度,而收敛慢的算法则需要较多的迭代次数才能得到相同精度要求的点.

当 $r=1, C=0$ 时,称序列 $\{x^{(k)}\}$ 超线性收敛于 x^* ,超线性收敛是一种比线性收敛快的收敛.从后面的学习我们可以看到,多数的最优化方法具有超线性收敛的特性,在上述收敛率的定义中,所有 $r>1$ 的收敛都属于超线性收敛.事实上,由

$$\lim_k \frac{e_{k+1}}{e_k} = C,$$

有

$$\lim_k \frac{e_{k+1}}{e_k} = \lim_k \frac{e_{k+1}}{e_k} e_k^{r-1} = C \times \lim_k e_k^{r-1} = 0.$$

称 $r=2$ 时的收敛为二次收敛,这时误差序列的性能可以用下述不等式表示

$$e_{k+1} = C e_k^2.$$

二次收敛是一种更快的收敛,还是考察上式取等式时的简单情形,设初始误差为 1, $C=1$,则误差序列为

$$1, 0.01, 0.0001, 0.00000001, \dots,$$

可以看到,二次收敛的方法每迭代一次近似解的精度就增加一倍.

一个理想的算法终止准则为

$$x^{(k)} - x^*,$$

然而由于 x^* 是未知的,这样的准则并不具有任何实用价值.但是由

$$\begin{aligned} \frac{x^{(k+1)} - x^*}{x^{(k)} - x^*} &= \frac{x^{(k+1)} - x^{(k)} + x^{(k)} - x^*}{x^{(k)} - x^*} \\ &= \left| \frac{x^{(k+1)} - x^{(k)} - x^{(k)} + x^*}{x^{(k)} - x^*} \right| \\ &= \left| \frac{x^{(k+1)} - x^{(k)}}{x^{(k)} - x^*} - 1 \right| \quad 0, \end{aligned}$$

在序列 $\{x^{(k)}\}$ 超线性收敛于 x^* 时,我们可以得到

$$\lim_k \frac{x^{(k+1)} - x^{(k)}}{x^{(k)} - x^*} = 1.$$

上式表明,对于一个超线性收敛的算法, $x^{(k+1)} - x^{(k)}$ 是 $x^{(k)} - x^*$ 的一个估计.因此对于具有超线性收敛速度的方法,

$$x^{(k+1)} - x^{(k)},$$

是一个比较合适的终止准则.同样也可以用评价函数值序列来确定终止准则,还是以无约束优化问题为例,由于 x^* 未知,

$$f(x^{(k)}) - f(x^*),$$

不可能直接用作收敛准则,但当 $f(x)$ 二次连续可微时可以推得

$$|f(x^{(k+1)}) - f(x^{(k)})| = O(x^{(k+1)} - x^{(k)})^2,$$

因此对于快速收敛的算法,

$$|f(x^{(k+1)}) - f(x^{(k)})|$$

也是一个相当有效的收敛准则.在有些方法中为确保所得的是最优解的满意的估计,往往采用两个或几个收敛准则同时使用的方法.究竟选用什么样的收敛准则取决于方法的性能以及用户具体的爱好.

如果最优解的一个近似不能满足要求的精度,方法需要计算一个修正量 $s^{(k)}$ 来得到最优解的一个更好的近似 $x^{(k+1)} = x^{(k)} + s^{(k)}$. $s^{(k)}$ 的计算是最优化方法最关键和最主要的计算工作.对于大多数的最优化方法来说, $s^{(k)}$ 是通过求解一个相对简单易于求解的最优化问题(通常称为子问题)确定的.由于最优化要求确定使目标函数值取最优(最小)的可行点,修正量 $s^{(k)}$ 的确定要求在新点 $x^{(k+1)}$ 处的评价函数值小于在点 $x^{(k)}$ 处的评价函数值,称这样每次迭代都使评价函数值下降的算法为单调下降算法,本书介绍的都是这类单调下降算法.在 20 世纪 90 年代,出现了所谓的非单调算法,这类算法不要求迭代点每次都下降,代之以要求每隔几步使评价函数值下降,对这类方法本书不作介绍.

小 结

本章作为全书的开篇,主要介绍了同最优化方法和技术有关的基本概念和基本的理论,全章共分 4 节.

第 1.1 节在于帮助读者认识 and 了解最优化问题,给出了最优化问题一般形式的数学模型,以作为后述各章节学习的基础;指出了这种一般形式的模型同各种具体问题模型之间的关系和相互转换;给出了几类主要的最优化问题的标准形式,如线性规划、二次规划、无约束最优化、等式约束最优化,和不等式约束最优化问题等;介绍了最优化问题的一些基本定义,如可行点、可行域、起作用约束、局部最优解、整体最优解等,以及它们之间的关系.

第 1.2 节介绍了最优化的基本理论之一——凸集和凸函数.凸集和凸函数在理解和以后进一步学习、应用乃至研究最优化时起着重要的作用.该节首先通过引入凸集的定义,分析了凸集同最优化直接相关的性质和特性,其重点为在最优化理论中起重要作用的凸集分离定理.由于任何凸函数的任何局部最优解必是它的整体最优解,对凸函数的判定和凸函数性质的了解尤为重要.给出了一个函数是凸函数的一阶充分必要条件,二阶充分和必要条件,介绍了目标函数为凸函数,可行域为凸集的凸规划问题,证明了凸规划问题的任何最优解必为全局最优解,给出了可行域是凸集的条件和要求.

第 1.3 节讲述最优化的基础——最优性条件.直观上一个可行点是一个最优化问题的最优解,如果在该点处不存在任何既可行又下降的方向.但要给出一个可实际用于判定一个可行点是最优解的条件并非那么简单,需要有一定的理论支持.第 1.3 节对此作了简要的介绍.引入了可行方向,下降方向的定义,再根据凸集分离定理给出了最优化问题最优解的一阶必要条件,又称 KKT 条件;然后在假定所有函数二阶连续可微的条件下给出了一般最优化问题最优解的二阶必要条件和二阶充分条件.这些条件在有关的最优化算法设计和研究中起着重要的作用,是算法设计和研究的基础,也是理解和学好具体最优化算法所必需的.

作为一本最优化方法的教科书,本章的第 1.4 节用于介绍一般最优化方法的基本特征和要求,以后各章节的各式各样的最优化方法大部分都具有这些特征.对于现实生活中的大量最优化问题,不可能直接给出问题的具有解析表达式的解,确定问题的最优解一般采用迭代法,即从一个给定的初始点开始,方法逐步产生一个越来越接近最优解的点的序列,并在一定的条件得到满足时取相应的迭代点作为所求最优解的一个近似.我们在这一节给出了一般方法的迭代格

式,评价一个点好坏的准则和方法、终止迭代的准则、衡量一个方法性能的收敛性(包括局部收敛性和全局收敛性)和收敛速度.



1. 考虑由约束

$$x_1^2 + x_2^2 = 1, \quad 1 - x_2 + x_1 = 0, \quad x_1 = 0$$

确定的可行域 F . 判定点 $x^{(1)} = -\frac{1}{2}, \frac{1}{2}^T$, $x^{(2)} = (-1, 1)^T$, $x^{(3)} = (-1, 0)^T$, $x^{(4)} = 0, -\frac{1}{2}^T$ 和 $x^{(5)} = -\frac{1}{2}, -\frac{1}{2}^T$ 是否是可行点? 如果是可行点是内点还是边界点? 是个约束的边界点?

2. 考虑下述约束最优化问题

$$\begin{aligned} \min \quad & x_1, \\ \text{s.t.} \quad & x_1^2 + (x_2 - 2)^2 = 3, \\ & x_1^2 = 1, \end{aligned}$$

画出问题的可行域和目标函数的等位线,并由此确定问题的所有局部最优解和全局最优解.

3. 设 $F = \{x \mid g_i(x) = 0, i = 1, 2, \dots, m\}$ 其中所有 $g_i(x)$, $i = 1, 2, \dots, m$ 都连续. 证明点 $\bar{x} \in F$ 是内点当且仅当对某一 $\delta > 0$ 有

$$\{x \mid \|x - \bar{x}\| < \delta\} \subset F.$$

4. (1) 证明有限个凸集的交集仍然是凸集.

(2) 设 $D_1 = \{x \mid x_1 + x_2 = 1, x_1 = 0\}$, $D_2 = \{x \mid x_1 - x_2 = 0, x_1 = 0\}$.

令 $D = D_1 \cup D_2$. 证明 D_1, D_2 均为凸集,但 D 却不是凸的,由此得出凸集的并集未必是凸集.

5. 设 $f_i(x)$, $i = 1, 2, \dots, m$ 是定义在凸集 $D \subset \mathbf{R}^n$ 上的凸函数,证明函数

$$g(x) = \sum_{i=1}^m \lambda_i f_i(x)$$

也是 D 上的凸函数,其中 $\sum_{i=1}^m \lambda_i = 1$, $\lambda_i \geq 0$, $i = 1, 2, \dots, m$,即凸函数的凸组合还是凸函数.

6. 设 $f(x)$ 为定义在凸集 $D \subset \mathbf{R}^n$ 上的凸函数, k 为一个给定的实数,称集合

$$T = \{x \mid f(x) = k\}$$

为函数 $f(x)$ 关于实数 k 的水平集. 证明对任意实数 k , 集合 T 是凸集.

7. 设多变量函数 $f: \mathbf{R}^n \rightarrow \mathbf{R}^1$ 和单变量函数 $g: \mathbf{R}^1 \rightarrow \mathbf{R}^1$ 都是凸函数,证明复合函数 $h(x) = (g \circ f)(x)$, $x \in \mathbf{R}^n$ 是凸函数.

8. 设 $c_i(x)$, $i = 1, 2, \dots, m$ 为 \mathbf{R}^n 上的凹函数, $f(x)$ 是 \mathbf{R}^n 上的凸函数,证明函数

$$P(x) = f(x) - \mu \sum_{i=1}^m \log c_i(x)$$

在集合 $D = \{x \mid c_i(x) > 0\}$ 上是凸函数.

9. 设 x^* 是凸规划问题的一个解. 证明如果目标函数严格凸,则 x^* 是唯一全局最优解.

10. 设 $f(x)$ 定义在集合 $D \subset \mathbf{R}^n$ 且连续可微. 证明 $f(x)$ 是 D 上严格凸函数的充分必要条件为

$$f(y) - f(x) > f'(x)^T (y - x), \quad \forall x, y \in D, x \neq y.$$

11. 设在点 \bar{x} 有 $f'(\bar{x}) \neq 0$, 证明在所有长度为 $\|s\| = 1$ 的方向 s 中, 方向 $s = -f'(\bar{x})/\|f'(\bar{x})\|$ 是 $f(x)$ 的最速下降方向.

12. 求出函数

$$f(x) = 2x_1^3 - 3x_1^2 - 6x_1x_2(x_1 - x_2 - 1)$$

的所有稳定点, 其中哪一个点是极小值点? 哪一个点是极大值点? 有没有既不是极大又不是极小的点?

13. 确定线性函数 $f(x) = 2x_1 - x_2 + 3x_3$ 的所有下降方向. 请问这样的下降方向是否同所在点的位置有关?

14. 考虑约束最优化问题

$$\begin{aligned} \min \quad & -x_1 + x_2, \\ \text{s.t.} \quad & x_1 + 2x_2 = 2, \\ & x_1 \geq 0, \quad x_2 \geq 0, \end{aligned}$$

(1) 确定在点 $x^{(1)} = (0, 0)^T$, $x^{(2)} = (0, 1)^T$, $x^{(3)} = (1, \frac{1}{2})^T$ 和 $x^{(4)} = (2, 0)^T$ 处的可行方向.

(2) 在这些点是否存在可行的下降方向, 并由此从中确定最优解.

15. 设

$$q(x) = \frac{1}{2} x^T B x + b^T x + c,$$

$x \in \mathbf{R}^n$, $B \in \mathbf{R}^{n \times n}$ 为对称矩阵, $b \in \mathbf{R}^n$, $c \in \mathbf{R}$. 证明 $q(x)$ 有唯一极小点的充分必要条件是矩阵 B 正定, 并给出这个极小解.

16. 在定理 1.3.6 约束规范条件下证明 $Z(x^*) = L(x^*)$.

17. 考虑下述三个数列:

$$u_k = \frac{1}{c^{2^{-k}}}, \quad (c > 0),$$

$$v_k = \frac{1}{k^k},$$

$$w_k = d^{p^k}, \quad (0 < a < 1, p > 1),$$

证明它们分别具有线性, 超线性和阶数为 p 的收敛率.

18. 设 $a_0 = b$, 考虑数列

$$a_{k+1} = \frac{1}{2} \left(a_k + \frac{b}{a_k} \right),$$

证明数列 $\{a_k\}$ 的极限为 $a^* = \sqrt{b}$, 且收敛率是二次的.

第 2 章

线 性 规 划

当最优化问题中的目标函数与约束函数都是变量 $x \in \mathbf{R}^n$ 的线性函数时称为线性规划. 工程与管理科学中大量的问题都是变量数目为成百上千, 乃至上万或数十万的线性规划问题. 学习和研究线性规划的求解方法, 不仅可以用于求解大量的实际线性规划问题, 而且可以用于非线性最优化问题的求解, 这是因为当用迭代法求解一个非线性最优化问题时, 如果我们在迭代点对问题中的有关函数取局部线性近似, 所得的问题就是一个线性规划问题. 我们在本章将学习求解线性规划问题的常用方法, 包括单纯形法、对偶单纯形法, 以及用于求解大规模线性规划问题的内点算法. 在具体学习这些方法之前, 我们先了解线性规划问题的有关概念和基本性质.

§ 2.1 基本性质

在具体学习线性规划的有关概念和基本性质之前, 我们先从几何直观了解线性规划的可行域和最优解的性能. 考察下列含有两个变量的线性规划问题.

例 2.1.1

$$\begin{aligned} \min \quad & 3x_1 - 2x_2, \\ \text{s.t.} \quad & -x_1 + x_2 \leq 3, \\ & -2x_1 + x_2 \leq 2, \\ & 4x_1 + x_2 \leq 16, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

图 2.1 给出了这一问题的可行域 F , 它是由线段 AB, BC, CD, DE, EA 围成凸多边形(凸集), A, B, C, D, E 是这个凸集的五個顶点. 写有 $f = 6, f = 0$ 与 $f = -5$ 的三条平行斜线是目标函数 $f(x) = 3x_1 - 2x_2$ 不同取值的同位线, 可以看出随着同位线的向左移动, 目标函数值逐渐减小, $f = -5$ 的同位线同可行域交于可行域的顶点 A . 如果把 $f = -5$ 的同位线向左作任何一点点的移动, 尽目标函数值会有所减小, 但同可行域不再有任何交点. 也就是说不存在任何使标函数值小于 -5 的可行点, 因此可行域的顶点 A 是上述线性规划问题的最

解, 最优目标函数值为 -5 . 由图 2.1 我们还可以看出在顶点 A , 约束

$$-x_1 + x_2 = 3, \quad -2x_1 + x_2 = 2$$

是有效或起作用的约束, 而其余三个约束在点 A 是不起作用的约束. 据此, 最优解 A 可以由方程组

$$-x_1 + x_2 = 3, \quad -2x_1 + x_2 = 2$$

的解直接确定, 即有 $x_1^* = 1, x_2^* = 4$.

从上述这个直观的例子, 我们似乎可以得出如下的结论:

- (1) 线性规划问题的可行域如果非空, 则是一个凸集-凸多面体;
- (2) 如果线性规划问题有最优解, 那么最优解可在可行域的顶点中确定;
- (3) 如果可行域有界, 可行域只有有限个顶点, 问题的最优解必存在, 并在这有限个顶点中确定;
- (4) 最优解可由最优顶点处的有效约束形成的方程组的解确定.

关于可行域的凸性可以直接从第一章的定理 1.2.14 得出, 这是因为线性规划问题的任何形式的约束, 都可以转换成 $a_i(x) \leq 0$ 的形式, 而对于线性函数 $a_i(x)$, 它们既是凸函数又是凹函数.

线性规划问题的可行域, 在两个变量的情形, 是由约束所表示的一些直线围成的凸多边形, 在三个变量的情形是由约束所表示的一些平面所围成的凸多面体 (见图 2.2), 而在维数大于 3 时, 是由线性约束所表示的一些超平面所围成的超凸多面体. 根据约束条件的不同, 凸多面体可能有界, 也可能无界, 图 2.2 的可行域由约束条件

$$x_1 + x_2 + x_3 = 1, \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0$$

确定, 可行域有界. 而图 2.3 的可行域由约束

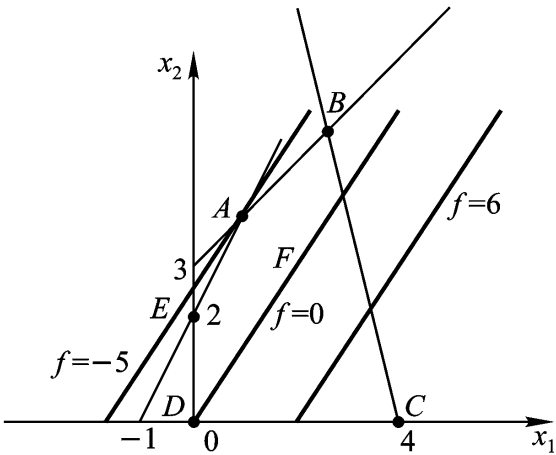


图 2.1 图形解

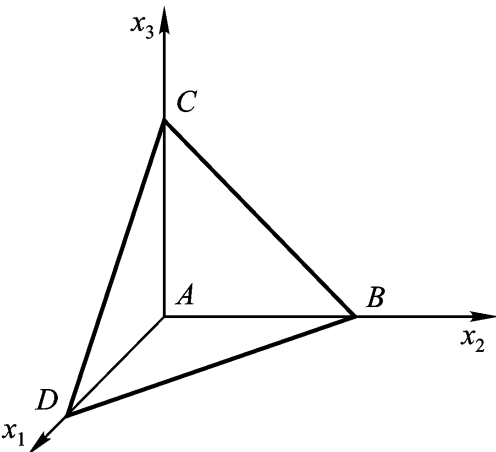


图 2.2 有界可行域

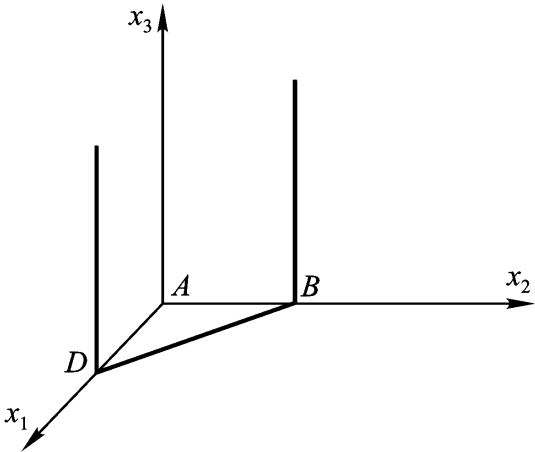


图 2.3 无界可行域

$$x_1 + x_2 = 1, \quad x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 = 0$$

确定,是无界的可行域.

可以把可行域分为内部与边界,边界又分顶点(如图 2.1 的 A, B, C, D, E , 图 2.2 中的 A, B, C, D), 边(如图 2.1 中的边 AB, BC, CD, DE, EA , 图 2.2 中的 AB, AC, AD, BC, BD, CD), 以及界面(如图 2.2 的面 ABC, ABD, ADC, BDC).

对于二维空间,可行域的边也就是广义意义下的面.

对于线性规划问题可以很容易地证明其最优解如果存在,必定在可行域的边界上.这是因为如果最优解在可行域的内部,则由于在可行域的内部所有的约束都不起作用,可以把它看作一个无约束最优化问题.根据无约束最优解的一阶

必要条件,目标函数 $f(x) = \sum_{i=1}^n c_i x_i$ 的一阶导数为零,即有

$$f'(x) = (c_1, c_2, \dots, c_n)^T = 0.$$

这是明显不可能的(除非所有 $c_j = 0, j = 1, 2, \dots, n$, 这又没有意义).至于最优解必定可在可行域的顶点中确定我们将在后面作进一步的分析.

对于变量多于 3 的线性规划问题,已不可能用图形画出其可行域,再根据目标函数的同位线确定问题的最优顶点来得出最优解.下一节学习的单纯形方法是在问题的基本可行解中确定最优解.事实上,下面进一步的分析将表明在基本可行解和可行域的顶点之间有等价关系,也就是说基本可行解就是可行域顶点的代数表示.

再回到图 2.1,假设我们要确定问题的最优解并已知顶点 C , 根据最优解必在可行域顶点的结论,我们只需在可行域的顶点中逐个确定最优解.因此我们需要确定一个比顶点 C 更优的顶点,从图上可以看出从点 C 有两种移动的可能到达另一个顶点,一是沿边 CB 移动到 B 点,另一个是沿边 CD 移动到 D 点.从图中可以看出沿这两个边移动都可以使目标函数值下降,从中选择一个使目标函数下降快的边,设为 CD , 由此可以得到顶点 D . 由于顶点 D 不是最优的,在顶点 D 再找一个使目标函数下降快的边 DE , 沿此边移动得更好的顶点 E , 由于 E 还不是最优解,可以找到使目标函数下降的边 EA , 沿此边移动便可确定问题的最优顶点.我们在下一节学习的单纯形方法就是根据这一思路逐步确定问题的最优解,即在非最优解的顶点,确定一条使目标函数下降比较快的边,沿此边移动产生一个更优的顶点,重复这一过程或者得到问题的一个最优解,或者确定问题无最优解.

单纯形方法要求线性规划具有所谓的标准形式,本节接下来的讨论将针对标准形的线性规划问题展开,因此我们先给出线性规划问题的标准形式,并指出如何把其他形式的线性规划转换成标准形.

标准形式的线性规划问题只有线性等式约束和对变量的非负约束,即

$$\begin{aligned}
 \min \quad & f(x) = c_1 x_1 + c_2 x_2 + \dots + c_n x_n, \\
 \text{s.t.} \quad & a_{i1} x_1 + a_{i2} x_2 + \dots + a_{in} x_n = b_i, \quad i = 1, 2, \dots, m, \\
 & x_j \geq 0, \quad j = 1, 2, \dots, n,
 \end{aligned} \tag{2.1.1}$$

或矩阵向量的表达方式

$$\begin{aligned}
 \min \quad & f(x) = c^T x, \\
 \text{s.t.} \quad & Ax = b, \\
 & x \geq 0,
 \end{aligned} \tag{2.1.2}$$

其中 $c = (c_1, c_2, \dots, c_n)^T$ 称为价值系数向量, $c_j, j = 1, 2, \dots, n$ 称为价值系数,

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

是约束的 $m \times n$ 阶系数矩阵, $b = (b_1, b_2, \dots, b_m)^T \in \mathbf{R}^m$ 为约束的右端向量并要求 $b \geq 0$. 对于系数矩阵 A , 我们假定 $m < n$ 且行满秩, 这是因为如果 A 不是行满秩, 则表示约束中有线性相关的方程, 通过删除相关的约束并不影响问题的可行域而使剩下的约束形成的系数矩阵满秩. 至于 m , 如果 $m > n$, 约束方程组或者不相容, 可行域为空集; 或者删除相关约束得 $m = n$. 如果 $m = n$ 且满秩, 则约束方程组有唯一解, 如果这个唯一的解还满足变量的非负条件, 则它就是最优解, 否则就没有最优解. 因此只有在 $m < n$ 时, 才需要在无限多个满足约束方程的可行解中确定使目标函数值取最优的解.

任何其他形式的线性规划问题都可以转换成这种标准形式. 求函数 $f(x)$ 极大的优化问题可以转换成求 $-f(x)$ 极小的优化问题. 例如, 如果原问题求函数 $f(x) = 4x_1 - 5x_2 + 2x_3$ 的极大, 将目标函数改成 $(x) = -f(x) = -4x_1 + 5x_2 - 2x_3$ 后求极小, 在求得最优解后, 再将最优目标函数值乘以 -1 即得原问题要求的最优函数值. 如果目标函数中有常数项, 删去常数项并不对最优解的确定产生任何影响, 只需在最优解确定后, 需对目标函数的最优值加上删去的常数项.

如果某个约束的右端项为负数, 用 -1 同时乘约束的两端即可得非负的右端项, 如可将 $2x_1 - x_2 = -1$ 改成 $-2x_1 + x_2 = 1$.

对于不等式约束, 如果右端项是负的, 先用 -1 乘不等式的两端并改变不等号的方向将其改成右端项非负的不等式, 如将 $3x_1 + 2x_2 - 4x_3 \leq -12$ 改成 $-3x_1 - 2x_2 + 4x_3 \geq 12$, 将 $x_1 + x_2 \geq -1$ 改成 $-x_1 - x_2 \leq 1$. 对于右端项非负的不等式约束则通过引入非负的松弛变量使其成为等式约束. 如对约束 $-x_1 - x_2 \leq 1$ 引入非负变量, 设为 $x_3 \geq 0$ 使成为 $-x_1 - x_2 + x_3 = 1$, 而对于约束 $-3x_1 - 2x_2 + 4x_3 \geq 12$, 则需引入称为乘余变量的变量, 设为 $x_4 \geq 0$ 使成为 $-3x_1 - 2x_2 + 4x_3 - x_4 = 12$.

如果对某个变量, 设为 x_i , 无非负的限制, 则通过设两个非负的变量 $x_i = 0$, $x_i = 0$, 并令 $x_i = x_i - x_i$ 来代替原问题中的变量 x_i . 设最后所得的最优解为 \bar{x}_i , \bar{x}_i , 则原问题的最优解为 $x_i^* = \bar{x}_i - \bar{x}_i$, 如果 $\bar{x}_i = \bar{x}_i = 0$, 则有 $x_i^* = 0$, 如果有 $\bar{x}_i > \bar{x}_i = 0$, 则有 $x_i^* < 0$. 如果对某个变量不仅要求其非负还要求其大于某个正的下界, 例如要求 $x_1 \geq 3$, 则令 $x_1 = x_1 - 3$, 代入原问题即可将 $x_1 \geq 3$ 替换为 $x_1 \geq 0$.

现在我们就标准形的线性规划问题讨论其可行域的顶点与基本可行解之间的关系. 顶点是对于凸集而言的, 凸集 D 的一个点称为顶点, 如果它不可能表示成 D 的任何两个不同于该点的点的凸组合, 即不可能表示成

$$x = y + (1 - \lambda)z, \quad (0, 1),$$

其中 $y, z \in D, y \neq x, z \neq x$. 对于线性规划问题的可行域, 其顶点, 设为 x , 不可能表示成其任意两个不同于 x 的可行点的凸组合. 图 2.4 给出了凸集的顶点, 其中第二个凸集的弧 CD 上的所有边界点都是它的顶点, 但我们在本章关注的线性规划问题的可行域及其顶点类似于第一个图的情形.

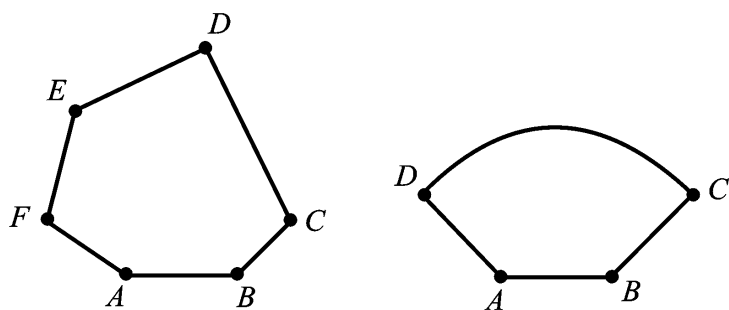


图 2.4 凸集的顶点

考虑标准形线性规划问题(2.1.1)中的等式约束

$$Ax = b. \quad (2.1.3)$$

由于 $m \times n$ 阶矩阵 A 满秩, 可以确定矩阵 A 的一个分块表示

$$A = [B \ N],$$

其中 B 为 $m \times m$ 阶的非奇异矩阵, N 为 A 的剩余的 $(n - m)$ 列形成的矩阵, 将向量 x 作相应的分块

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}$$

其中 x_B 为相应于矩阵 B 的 x 的分量组成的 m 阶向量, x_N 为相应于矩阵 N 的 x 的分量组成的 $(n - m)$ 维向量. 于是方程组(2.1.3)可以改写成

$$Ax = Bx_B + Nx_N = b,$$

由于 B 非奇异, 我们有

$$x_B = B^{-1}(b - Nx_N) = B^{-1}b - B^{-1}Nx_N.$$

由此可以得到

$$x = \begin{matrix} x_B \\ x_N \end{matrix} = \begin{matrix} B^{-1} b - B^{-1} N x_N \\ x_N \end{matrix}. \quad (2.1.4)$$

对于任意 $x_N \in \mathbf{R}^{(n-m)}$, 上述 x 都是方程组 (2.1.3) 的解. 特别, 取 $x_N = 0$ 所得的解

$$x = \begin{matrix} B^{-1} b \\ 0 \end{matrix}$$

称为基本解, x_B 称为基本变量, x_N 称为非基本变量, 矩阵 B 为基矩阵, N 为非基矩阵 (在下面不引起混淆的情况下, 我们同样用 B, N 分别表示相应的基变量和非基变量指标的集合). 基本解未必是线性规划问题 (2.1.1) 的可行解, 如果一个基本解还是可行的, 即有

$$x = \begin{matrix} B^{-1} b \\ 0 \end{matrix} \geq 0$$

则称其为基本可行解.

由于 B 是从矩阵 A 的 n 个列中选取 m 列所形成的, 因此对于矩阵 B 至多有

$$\frac{n}{m} = \frac{n!}{m! (n-m)!}$$

个选择, 也就是说基本解的个数是有限的, 这里 $n!$ 为 n 的阶乘, 即 $n! =$

$n(n-1)(n-2)\dots 3 \cdot 2 \cdot 1$. $\frac{n}{m}$ 仅仅是基本解的一个上界, 因为并非 A 的每一个分块表示都使得 B 非奇异. 作为基本解的一个子集, 基本可行解 (即顶点) 的个数要少于基本解的个数.

下面考察由例 2.1.1 的约束构成的可行域, 并验证其基本可行解同可行域顶点的对应关系. 首先通过引入三个松弛变量, 设为 x_3, x_4, x_5 将约束转换为标准形

$$\begin{aligned} -x_1 + x_2 + x_3 &= 3, \\ -2x_1 + x_2 + x_4 &= 2, \\ 4x_1 + x_2 + x_5 &= 16, \\ x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0. \end{aligned}$$

由于 $n=5, m=3$, 对于 B 共有 10 种可能的选择. 下面给出了由此得出的 10 个基本解

$$x^{(1)} = (0, 0, 3, 4, 16)^T$$

$$x^{(2)} = (0, 2, 1, 0, 14)^T$$

$$x^{(3)} = (4, 0, 7, 12, 0)^T$$

$$x^{(4)} = \left(\frac{13}{5}, \frac{28}{5}, 0, \frac{18}{5}, 0\right)^T$$

$$x^{(5)} = (1, 4, 0, 0, 8)^T$$

$$x^{(6)} = \left(\frac{14}{6}, \frac{40}{6}, -\frac{8}{6}, 0, 0\right)^T$$

$$x^{(7)} = (-3, 0, 0, -2, 28)^T$$

$$x^{(8)} = (-1, 0, 2, 0, 20)^T$$

$$x^{(9)} = (0, 3, 0, -1, 13)^T$$

$$x^{(10)} = (0, 16, -13, -12)^T.$$

在这 10 个基本解中, 由于后面 5 个有负的分量, 因而不是基本可行解. 前 5 个基本解由于没有负分量, 因而是基本可行解. 而这 5 个基本可行解刚好对应可行域的 5 个顶点: $x^{(1)}$ 对应顶点 D , $x^{(2)}$ 对应顶点 E , $x^{(3)}$ 对应顶点 C , $x^{(4)}$ 对应顶点 B , $x^{(5)}$ 对应顶点 A .

下述定理从理论上给出了基本可行解同可行域顶点之间的一一对应关系.

定理 2.1.2 可行点 $x = \{x/Ax=b, x \geq 0\}$ 是 F 的顶点的充分必要条件 x 是一个基本可行解.

证明 先证充分性. 设 x 是一基本可行解, 即存在 A 的一个分块表示 $A = [B \ N]$, B 非奇异, 使得

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix} = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}.$$

为方便, 我们假定 B 由 A 的前 m 列构成, N 由 A 的后 $(n-m)$ 列构成. 假设 x 不是可行域的顶点, 则存在两个不同于 x 的可行点 y 和 z , 以及某一实数 $(0, 1)$, 使得

$$x = \alpha y + (1 - \alpha) z.$$

将 y 和 z 同样表示成

$$y = \begin{pmatrix} y_B \\ y_N \end{pmatrix}, \quad z = \begin{pmatrix} z_B \\ z_N \end{pmatrix}.$$

由 y, z 的可行性有 $y_N \geq 0, z_N \geq 0$. 由 $0 = x_N = \alpha y_N + (1 - \alpha) z_N, 0 < \alpha < 1$, 必有 $y_N = 0, z_N = 0$. 又由于 x, y, z 是可行解, 我们有

$$Bx_B = By_B = Bz_B = b.$$

由于 B 非奇异, 上述等式表明有 $x_B = y_B = z_B$. 由此得出 $x = y = z$ 的结论, 这与 y, z 不同于 x 相矛盾, 这就证明了 x 是 F 的顶点.

再证必要性 设 x 为一顶点, 我们证明它是一个基本可行解. 由 x 的可行性, 即 $Ax=b, x \geq 0$, 可以把 x 的分量分成正的和零两部分, 即

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}, \quad x_B > 0, \quad x_N = 0.$$

其中 x_B 的分量数不一定为 m , 设为 p . 为方便, 同样假定 x_B 由 x 的前 p 个分量

组成. 对矩阵 A 作同样的分块表示得

$$A = [B \ N],$$

其中 B 为 $m \times p$ 阶矩阵. 首先证明 B 的列线性无关.

假定 B 的列线性相关. 则存在一个非零向量, 设为 $w \in \mathbf{R}^p$ 使成立 $Bw = 0$. 从 $Bx_B = b$ 得

$$B(x_B \pm w) = Bx_B \pm Bw = b$$

对 \mathbf{R} 的所有取值都成立. 由于 $x_B > 0$, 存在充分小的 $\epsilon > 0$ 使得

$$x_B + \epsilon w > 0, \quad x_B - \epsilon w > 0,$$

令

$$y = \begin{pmatrix} x_B + \epsilon w \\ 0 \end{pmatrix}, \quad z = \begin{pmatrix} x_B - \epsilon w \\ 0 \end{pmatrix}.$$

则 y 和 z 都是可行的且有 $y \neq x, z \neq x$. 由于 $x = (y + z)/2$, 这与 x 是可行域顶点的假定相矛盾, 这就证明了 B 的列线性无关.

再证 x 为一基本可行解. 由于矩阵 B 的列线性无关, 则必有 $p = m$. 如果 $p < m$, 则 B 非奇异, 由 $x_N = 0$ 得

$$Ax = Bx_B + Nx_N = Bx_B = b, \quad x_B = B^{-1}b > 0.$$

因此 x 是基本可行解. 如果 $p < m$, 则由于 A 满秩, 可从 N 中选取 $m - p$ 个列加入到矩阵 B 中形成一个新的 $m \times m$ 阶非奇异矩阵 \tilde{B} , 并把 N 中剩余的列组成新的 N 矩阵, 对解向量 x 作同样的重新组合得 x_B 和 x_N . 显然, 对此有 $x_B = \tilde{B}^{-1}b > 0, x_N = 0$, 即 x 是一个基本可行解. 这就完成了必要性的证明.

从定理的证明过程可以看到, 对于一个基本可行解并非它的所有基本变量都可能取正值. 如果某个基本可行解出现零值的基本变量, 则称其为退化的基本可行解(顶点). 一个有退化顶点的线性规划称为退化的线性规划. 我们通常对线性规划作非退化的假定, 这是因为退化现象是由于约束条件中存在可删除的约束而出现的.

为证明线性规划的最优解必定在可行域的某个顶点取得, 我们需要先给出下面的线性规划的代表定理.

定理 2.1.3 设 $m \times n$ 阶矩阵 A 满秩, 可行域 $F = \{x \mid Ax = b, x \geq 0\}$ 有界, 则任何可行点 $x \in F$ 都可以表示成 F 的顶点的凸组合, 即设 $V = \{v_1, v_2, \dots, v_k\}$ 为 F 的顶点集合, 则对任意 $x \in F$ 都有

$$x = \sum_{i=1}^k \theta_i v_i,$$

其中 $\theta_i \geq 0, \sum_{i=1}^k \theta_i = 1, i = 1, 2, \dots, k$.

在给出定理的严格证明之前,我们先通过图 2.5 来说明证明的基本思路.考察图 2.5 中的点 x ,它显然不是一个顶点(如果它是顶点,则已得定理的结论)对这样一个点,我们可以确定一个方向,如图中的方向 w ,沿这个方向我们可以确定两个可行点,如图中的点 y_1 和 y_2 ,点 x 可以表示成这两个点的凸组合. y_1 和 y_2 是可行域的边界点,同点 x 相比,这两个点的向量表示中,至少比 x 要多一个零分量.如果 y_1 和 y_2 已是顶点,则已得结论,否则,继续这个过程,分别把 y_1 和 y_2 表示成另外两组点的凸组合,这两组点同 y_1 和 y_2 相比又要分别比 y_1 和 y_2 至少要多一个零分量,如果这两组点已经是可行域的顶点,则已得结论,(如同图 2.5,点 y_1 可表示成顶点 v_1 和 v_4 的凸组合, y_2 则是 v_3 和 v_6 的凸组合,因而 x 可表示成 v_1, v_3, v_4 和 v_6 的凸组合,再令关于 v_2 和 v_5 的组合系数为零,即可把 x 表示成所有顶点的凸组合.)如果这两组点中还有不是顶点的点,对这些点重复这一过程,由于每进行一次组合,新点的零分量至少增加一个,而分量数 n 是有限的,因而经有限步组合后,所有的点必定都是可行域的顶点,而 x 可表示成所有这些顶点的凸组合,再把那些没有在这些点中的顶点的组合系数置为零,即得 x 可表示成可行域所有顶点的凸组合.

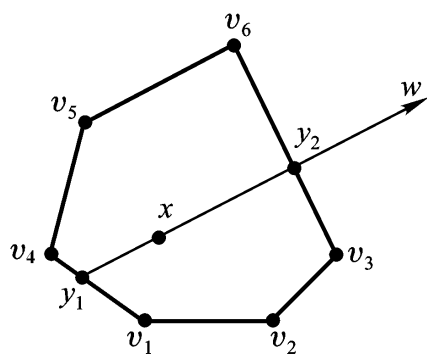


图 2.5 顶点的凸组合

定理 2.1.3 的证明 设 $x \in F$ 为任一可行点,如果 $x \in V$,即 x 为一顶点,设

$$x = v_j, 1 \leq j \leq k, \text{ 则取 } x_j = 1, x_i = 0, i \neq j, \text{ 于是我们有 } x = \sum_{i=1}^k x_i v_i.$$

如果 x 不是一个顶点,那么它也不是一个基本可行解,根据定理 2.1.1,矩阵 A 的相应于 x 的正分量的列必线性相关.于是我们可以找到一个非零的可行方向,设为 w ,满足

$$\begin{aligned} Aw &= 0, & w &\geq 0, \\ w_i &= 0, & \text{如果 } x_i &= 0. \end{aligned}$$

对于这样一个可行方向,必有符号不同的非零分量.这是因为如果它的所有非零分量同号,设为正号,则我们有

$$A(x + w) = Ax + Aw = b, \quad x + w \geq 0, \quad w \geq 0,$$

这表明可行域无界,同可行域有界的假定相矛盾.

沿方向 w 分别确定两个点 $y_1 = x + \theta_1 w$ 和 $y_2 = x - \theta_2 w$, 其中

$$\begin{aligned} \theta_1 &= \min \left\{ \frac{x_i}{w_i} \mid w_i < 0 \right\}, \\ \theta_2 &= \min \left\{ \frac{x_i}{w_i} \mid w_i > 0 \right\}, \end{aligned}$$

显然这两个点是可行点,且每个点的零分量个数至少比 x 的零分量个数要多 1

个,同时 x 可以表示成这两个点的凸组合.如果 y_1 和 y_2 已是可行域的顶点,则用同样的方法置其他顶点的组合系数为零即得定理的结论.否则将 y_1 和 y_2 中不是可行域顶点的每一个点表示成另外两个至少还多一个零分量的可行点的凸组合.对过程中产生的不是顶点的点重复这一过程,使得最终所有的表示点都是可行域的顶点.再置没包含在内的顶点的组合系数为零,即得定理的结论.

在上述定理的证明中要用到下面一个简单的结论:如果 x 是 y 和 z 的一个凸组合,而 y 和 z 分别是 y_1, y_2, \dots, y_p 和 z_1, z_2, \dots, z_q 的凸组合,则 x 也是 $y_1, y_2, \dots, y_p, z_1, z_2, \dots, z_q$ 的一个凸组合.此结论的证明留给读者作为练习.

为了给出无界可行域的表示定理,需要引入所谓的无界方向的概念.设 $D \subset \mathbf{R}^n$ 是一个凸集,方向 d 称为集合 D 的无界方向,如果

$$x + \lambda d \in D$$

对所有的 $x \in D$ 和所有的 $\lambda \geq 0$ 成立.也就是说,从 D 中的任意一个点沿方向 d 的正方向上的任何点都是 D 的点.如果凸集 D 有不只一个线性无关的无界方向,那么这些方向的任何凸组合也是 D 的一个无界方向.图 2.6 中的两个方向 d_1 和 d_2 即为可行域 F 的两个无界方向.

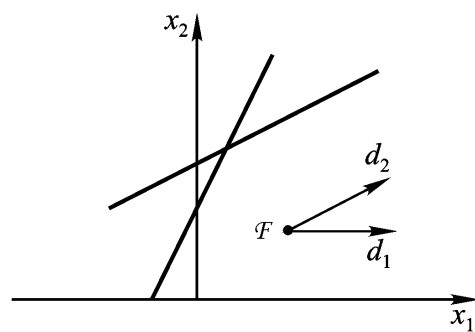


图 2.6 无界方向

定理 2.1.4 设 $m \times n$ 阶矩阵 A 满秩,可行域 $F = \{x \mid Ax = b, x \geq 0\}$ 无界,则任何可行点 x 都可以表示成

$$x = \sum_{i=1}^k \lambda_i v_i + d,$$

其中

$$\sum_{i=1}^k \lambda_i = 1, \quad \lambda_i \geq 0, \quad i = 1, 2, \dots, k,$$

$V = \{v_1, v_2, \dots, v_k\}$ 是可行域 F 的所有顶点的集合, d 或为零向量或为 D 的一个无界方向.

证明 证明类似于定理 2.1.3 的证明.设 x 为 F 的任意一点,如果 $x \in V$,则结论已成立.如 $x \notin V$,则同样可选取一个可行方向 w 满足

$$\begin{aligned} Aw &= 0, & w &\geq 0, \\ w_i &= 0, & \text{如果 } x_i &= 0. \end{aligned}$$

这时存在三种可能性:

(1) w 的非零分量有不同的符号,则如同定理 2.1.3 那样可以找到两个可行点 y_1 和 y_2 ,点 x 是 y_1 与 y_2 的凸组合,且 y_1 与 y_2 的零分量数至少比 x 的零分量数少 1.如果 y_1 和 y_2 都是顶点,则结论已得,否则需对其中非顶点的可行点

作进一步的组合表示.

(2) $w \neq 0$, 则由于 $A(x + w) = b$, $x + w \geq 0$ 对所有 $\lambda \geq 0$ 成立, w 是 F 的一无界方向. 这时沿 w 的反方向可以确定一个不同于 x 的可行点, 设为 y_1 , 它的分量数比 x 的零分量数至少要少一个, 且 x 可以表示成

$$x = y_1 + w$$

的形式. 如果 $y_1 \in V$, 则取 $d = -w$, 并置其他顶点的组合系数为零, 即得定理的结论; 如果 y_1 不是 F 的顶点, 继续对 y_1 作进一步的组合表示.

(3) $w = 0$, 则 $-w$ 是 F 的一个无界方向, 同(2)一样, 沿 w 可确定一个可行 y_1 , 其或为顶点或需要作进一步的组合表示.

重复对过程中产生的非顶点的可行点作组合表示, 或者可以将其表示为两个具有更多零分量的可行点的凸组合, 或者将其表示成一个具有更多零分量的可行点和一个无界方向的某个正倍数之和.

重复这一过程, 最终或者把 x 表示成 F 的若干个顶点的凸组合, 这时取 $d = 0$, 并置其他顶点的组合系数为零即得定理的结论; 或者把 x 表示成 F 的若干个顶点的凸组合和若干个 F 的无界方向的凸组合之和. 由于无界方向的凸组合然是一个无界方向, 取 d 为这一凸组合方向, 再置其他顶点的组合系数为零即得本定理的结论.

根据上述线性规划可行域的表示定理, 可以得出线性规划问题如果有最优解, 则必可在其可行域的顶点中确定的结论.

首先证明对于可行域无界的线性规划问题, 如果有有限的最优解, 则必有 $c^T d \geq 0$ 成立, 其中 d 为可行域的任一无界方向.

引理 2.1.5 考虑标准形线性规划问题(2.1.2), 设可行域无界, 但有有限的最优解, 则对于 F 的任意无界方向 d 均有

$$c^T d \geq 0.$$

证明 设 x^* 为标准形线性规划问题的最优解, 则根据无界方向的定义有 $x^* + \lambda d \in F$ 对所有 $\lambda \geq 0$ 成立. 由点 x^* 的最优性, 我们有

$$c^T (x^* + \lambda d) = c^T x^* + \lambda c^T d \leq c^T x^*.$$

由 $\lambda > 0$ 得出 $c^T d \leq 0$ 的结论, 证毕.

定理 2.1.6 标准形线性规划问题(2.1.2)如果有有限的最优解, 则必可在其可行域的某个顶点取得.

证明 设可行域 F 的顶点集合为 $V = \{v_1, v_2, \dots, v_k\}$. 又设 v_j 是这些顶点中目标函数取值最小的顶点, 即有

$$c^T v_j = \min \{c^T v_i \mid v_i \in V\}.$$

我们证明顶点 v_j 就是问题的一个最优解. 考察任意可行点 $x \in F$, 根据可行点的

示定理,有

$$x = \sum_{i=1}^k \alpha_i v_i + d,$$

其中 d 或为零, 或为 F 的一个无界方向. 于是根据引理 2.1.5, 顶点 v_j 在顶点集

中的最优性, 以及 $\alpha_i = 1, \alpha_i = 0, i = 1, 2, \dots, k$ 得

$$c^T x = \sum_{i=1}^k \alpha_i c^T v_i + c^T d = \sum_{i=1}^k \alpha_i c^T v_j = c^T v_j.$$

由于 x 在可行域 F 中的任意性, 这就证明了顶点 v_j 就是问题的最优解.

§ 2.2 单纯形方法

求线性规划问题最优解的单纯形方法是由 G. B. Dantzig 在 1947 年提出的, 由于这一方法的有效性, 几十年来一直在几乎所有的领域得到广泛的应用. 近年来, 对于大规模的线性规划问题, 尽管它受到了内点算法的挑战, 但单纯形方法还是受到广大用户的青睐.

单纯形方法同其他的数值求解方法一样是一种迭代法, 它根据线性规划问题的特点在问题可行域的顶点中逐步确定问题的最优解. 在每一个是基本可行解的迭代点, 如果它不是最优的, 单纯形法从与该顶点相联结的边中确定一个使目标函数值下降的边, 沿该边移动可以确定一个与该顶点相邻, 且目标函数又优于该顶点的新顶点(基本可行解). 由于可行域的顶点数是有限的, 如果每一次的移动都能使目标函数值下降, 则经过有限次的移动方法必终止于问题的一个最优顶点.

考察标准形线性规划问题

$$\begin{aligned} \min \quad & f(x) = c^T x, \\ \text{s.t.} \quad & Ax = b, x \geq 0. \end{aligned} \quad (2.2.1)$$

对其中有关符号的意义及假定同本章第 2.1 节. 设 $x^{(k)} \in F$ 为它的一个基本可行解, 单纯形方法首先检验它的最优性. 如果它不是最优的, 确定与该顶点相连的一条使目标函数下降的边; 接下来确定沿这个边移动多远可以到达另一个更优的相邻顶点, 也就是得出一个新的基本可行解.

最优性检验

设相应于基本可行解 $x^{(k)}$ 的基矩阵为 B_k , 非基矩阵为 N_k , 相应的基变量为

$x_B^{(k)}$, 非基变量为 $x_N^{(k)}$, 则问题的任意可行解可表示为

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}, \quad x_B = B_k^{-1} b - B_k^{-1} N_k x_N = x_B^{(k)} - B_k^{-1} N_k x_N \geq 0, \quad x_N \geq 0,$$

这里 $x_B^{(k)} = B_k^{-1} b$. 将 x 代入目标函数得

$$\begin{aligned} f(x) &= c^T x = c_B^{(k)T} x_B + c_N^{(k)T} x_N \\ &= c_B^{(k)T} x_B^{(k)} + c_N^{(k)T} - c_B^{(k)T} B_k^{-1} N_k x_N \\ &= y_k^T b + c_N^{(k)T} - y_k^T N_k x_N, \end{aligned} \quad (2.2.2)$$

其中 $y_k = B_k^{-T} c_B^{(k)}$ 称为单纯形乘子. 基本可行解 $x^{(k)}$ 的取值以及相应的目标函数值为

$$x^{(k)} = \begin{pmatrix} B_k^{-1} b \\ 0 \end{pmatrix}, \quad f_k = f(x^{(k)}) = c_B^{(k)T} B_k^{-1} b = y_k^T b.$$

为检验点 $x^{(k)}$ 是否是问题的最优解, 根据可行性的要求我们观察当非基变量由零向正方向变动时目标函数值的变化. 由式(2.2.2), 令

$$\hat{c}_N^{(k)} = c_N^{(k)} - N_k^T y_k,$$

称 $\hat{c}_j^{(k)}, j \in N_k$ (这里 N_k 表示所有非基变量的指标集合) 为相应于非基变量 $x_j, j \in N_k$ 的简约价值系数. 从式(2.2.2)可以看出, 如果 $\hat{c}_j^{(k)} > 0$, 增大非基变量 x_j 将使目标函数值增大; 如果 $\hat{c}_j^{(k)} = 0$, 目标函数值不受 x_j 变化的影响; 如果 $\hat{c}_j^{(k)} < 0$, 增大 x_j 将可使目标函数值下降. 因此如果对所有 $j \in N_k$ 有 $\hat{c}_j^{(k)} \geq 0$, 则由增大任何非基变量所引起的从 $x^{(k)}$ 向可行域内的移动只能导致目标函数值的增加, 因此 $x^{(k)}$ 就是问题的最优解. 相反, 如果有某一简约价值系数 $\hat{c}_j^{(k)} < 0, j \in N_k$, 则通过增大相应的非基变量 x_j 的值可使目标函数值得到改善, 因而 $x^{(k)}$ 不是问题的最优解.

确定新的基本可行解

设 $x^{(k)}$ 不是最优的基本可行解, 则需要确定一个使目标函数值更优的基本可行解, 也就是确定一组新的基矩阵和非基矩阵. 在单纯形方法中新的基矩阵是通过把原非基矩阵的一列 (称相应的变量为入基变量) 和原基矩阵的一列 (相应的变量为出基变量) 交换得到的.

设 $\hat{c}_p^{(k)} < 0, p \in N_k$, 增大 x_p 可使目标函数值下降. 可取 $x_p, p \in N_k$ 为入基变量, 即把矩阵 A 中相应于变量 x_p 的列, 记为 A_p 从 N_k 移入 B_k . 由于 x_p 将由非基变量成为基变量, 其取值将由零成为正值. 为此需要确定 x_p 取值的大小.

由

$$x_B = x_B^{(k)} - B_k^{-1} N_k x_N = B_k^{-1} b - B_k^{-1} N_k x_N,$$

当 x_p 由零增为正值而 x_N 的其余分量保持取零值时, 上式可改写为

$$x_B = B_k^{-1} b - B_k^{-1} A_p x_p = b^{(k)} - \hat{A}_p^{(k)} x_p,$$

其中 $b^{(k)} = B_k^{-1} b$, $\hat{A}_p^{(k)} = B_k^{-1} A_p$. 为保持可行性, x_p 的取值应确保 $x_B \geq 0$, 即有

$$x_j = b_j^{(k)} - x_p \hat{a}_{jp}^{(k)} \geq 0, \quad j \in B_k,$$

其中 $\hat{a}_{jp}^{(k)}$ 是 m 维向量 $\hat{A}_p^{(k)}$ 的第 j 个分量 (这里 B 是所有基变量指标的集合).

如果 $\hat{a}_{jp}^{(k)} < 0$, x_j 将随 x_p 的增大而增大; 如果 $\hat{a}_{jp}^{(k)} = 0$, $x_j = b_j^{(k)}$ 在 x_p 增大时保持不变; 如果 $\hat{a}_{jp}^{(k)} > 0$, x_j 将随 x_p 的增大而减小. 因此, 在 $\hat{a}_{jp}^{(k)} > 0$ 时, 为保持可

行性, x_p 的取值不能大于 $\frac{b_j^{(k)}}{\hat{a}_{jp}^{(k)}}$. 如果 $\hat{A}_p^{(k)}$ 有几个分量为正值, x_p 的取值为

$$x_p = \frac{b_q^{(k)}}{\hat{a}_{qp}^{(k)}} = \min \frac{b_j^{(k)}}{\hat{a}_{jp}^{(k)}}, \hat{a}_{jp}^{(k)} > 0, j \in B_k, \quad (2.2.3)$$

就可保持迭代点的可行性, 其中 q 是使上式达到极小的指标. 可以看出这时 x_q 从原来取正值的基变量成为取零值的非基变量, 也就是说 x_q 是要选的出基变量.

记相应于变量 x_q 的矩阵 A 中的列为 A_q , 将 B_k 中的列 A_q 与 N_k 中的列 A_p 交换即得到新的基矩阵和非基矩阵, 分别记为 B_{k+1} 和 N_{k+1} . 由此可以确定新的基本可行解 $x^{(k+1)}$ 以及相应的目标函数值,

$$x_p^{(k+1)} = x_k = \min \frac{b_j^{(k)}}{\hat{a}_{jp}^{(k)}} \mid \hat{a}_{jp}^{(k)} > 0, j \in B_k$$

$$x_q^{(k+1)} = 0, x_j^{(k+1)} = 0, j \in B_k, j \neq p, \quad (2.2.4)$$

$$x_j^{(k+1)} = x_j^{(k)} - x_p^{(k+1)} \hat{a}_{jp}^{(k)}, j \in B_k, j \neq q,$$

$$f_{k+1} = y_k^T b + c_p x_p^{(k+1)}. \quad (2.2.5)$$

对点 $x^{(k+1)}$ 重复上述过程直至确定问题的一个最优解或者确定问题无有界的最优解. 这是因为如果对所有 $j \in B_k$ 都有 $\hat{a}_{jp}^{(k)} \leq 0$, 则对任意 $x_p > 0$, 都有 $x_B = b^{(k)} - x_p \hat{A}_p^{(k)} > 0$, 且目标函数随着 x_p 的不断增大而无限地减小, 因此问题无有界的最优解.

根据上述分析, 求解标准形线性规划问题的单纯形方法步骤如下.

算法 2.2.1 (单纯形方法)

步 1. 给定一个初始基本可行解, 记迭代次数 $k=1$;

步 2. 计算单纯形乘子 $y_k = B_k^{-T} c_B^{(k)}$ 和简约价值系数向量 $\hat{c}_N^{(k)} = c_N^{(k)} - N_k^T y_k$;

步 3. 最优性检验: 计算 $\hat{c}_p^{(k)} = \min \hat{c}_j^{(k)} / j \in N_k$, 如果 $\hat{c}_p^{(k)} \geq 0$, 则 $x^{(k)}$ 为最优解, 停止迭代; 否则选 x_p 为入基变量;

步 4. 确定出基变量: 计算 $\hat{A}_p^{(k)} = B_k^{-1} A_p$. 如果对所有 $j \in B_k$ 有 $\hat{a}_{jp}^{(k)} \geq 0$, 则问题无有界的最优解, 停止迭代; 否则确定出基变量指标

$$\frac{b_q^{(k)}}{\hat{a}_{qp}^{(k)}} = \min \frac{b_j^{(k)}}{\hat{a}_{jp}^{(k)}}, \hat{a}_{jp}^{(k)} > 0, j \in B_k;$$

步 5. 交换 B_k 的列 A_q 与 N_k 的列 A_p 得新的基矩阵 B_{k+1} 和非基矩阵 N_{k+1} , 计算新的基本可行解 $x^{(k+1)}$, 置 $k := k + 1$ 后转步 2.

在上述算法中, 当存在不止一个简约价值系数 $\hat{c}_j^{(k)} < 0$ 时, 选取最负的 $\hat{c}_j^{(k)}$ 的指标为 p , 并以 x_p 作为入基变量.

为帮助理解上述单纯形法, 我们对下述例题用经典的单纯形表的形式求它的最优解. 考虑问题

$$\begin{aligned} \min \quad & -2x_1 - 3x_2, \\ \text{s.t.} \quad & -x_1 + x_2 = 3, \\ & -2x_1 + x_2 = 2, \\ & 4x_1 + x_2 = 16, \\ & x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

这一问题同本章第 2.1 节的例 2.1.1 有相同的可行域, 但目标函数不同, 其最优解在可行域的顶点 B 取得. 为用单纯形方法求其最优解, 首先引入三个松弛变量 x_3, x_4, x_5 将其转换成标准形的线性规划问题

$$\begin{aligned} \min \quad & -2x_1 - 3x_2, \\ \text{s.t.} \quad & -x_1 + x_2 + x_3 = 3, \\ & -2x_1 + x_2 + x_4 = 2, \\ & 4x_1 + x_2 + x_5 = 16, \\ & x_i \geq 0, i = 1, 2, \dots, 5. \end{aligned}$$

在用单纯形表的方式求线性规划问题的最优解时, 对于给定的初始基本可行解, 我们将问题用下述表的形式给出, 其中 B 和 N 的位置分别用于给出基矩阵和非基矩阵, C_B^T 和 C_N^T 分别为目标函数的价值向量, b 为约束右端项, $-f$ 用于给出目标函数在当前的负值, x_B 和 x_N 用于指明哪些是基变量, 哪些是非基变量.

基变量	x_B	x_N	右端项
负函数值	C_B^T	C_N^T	$(-f)$
x_B	B	N	b

对这样一个单纯形表, 我们首先对行 $[B \ N \ b]$ 应用高斯消去法将 B 转化成单位矩阵, 从而得到

$$[I \ B^{-1}N \ B^{-1}b].$$

这时 $B^{-1}b$ 即为当前基本可行解的基变量 x_B 的取值.再用向量 $-c_B$ 同矩阵 $[I \ B^{-1}N \ B^{-1}b]$ 的每一列分别取内积后把所得的值加至行 $[c_B^T \ c_N^T \ 0]$ 的相应元素,得到下表

基变量	x_B	x_N	右端项
负函数值	0	$c_N^T - c_B^T B^{-1}N$	$-c_B^T B^{-1}b$
x_B	I	$B^{-1}N$	$B^{-1}b$

此表中的右端项即为相应的基本可行解基变量的取值,非基变量则取零值.检验 $c_N^T - c_B^T B^{-1}N$ 中各分量的正负即可确定该基本可行解是否是最优解.如有负分量,取相应于最负分量的变量作为入基变量,并检查同该变量对应的 $B^{-1}N$ 相应列 \hat{A}_p 的各分量.如果所有分量都非正,则问题无有界的最优解,迭代可以终止.否则通过计算列 \hat{A}_p 的正分量同 $B^{-1}b$ 的相应分量的比值确定出基变量,这也就确定了进一步进行高斯消去法的主元 $a_{qp}^{(k)}$.

为在新的单纯形表上确定新的基本可行解,我们不必要对新的基矩阵重复前面的高斯消去过程.由于新的基矩阵同原基矩阵只相差一列,我们只需在当前的单纯形表上以 $a_{qp}^{(k)}$ 为主元,将新加入列的其他元素消去使成为单位向量即可得新的基本可行解.

下面给出对上述问题求解过程中的各单纯形表.对于上述引入松弛变量后已标准化的线性规划问题, $x_1 = 0, x_2 = 0, x_3 = 3, x_4 = 2, x_5 = 16$ 是其一个初始可行解(相应于顶点 D).

第一次迭代($k=1$):初始单纯形表为

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	-2	-3	0	0	0	0
x_3	-1	1	1	0	0	3
x_4	-2	1	0	1	0	2
x_5	4	1	0	0	1	16

由于矩阵 B 已是单位矩阵,不需要进行高斯消去而已得基本可行解.又因为 $c_B^T = (0,0,0)$,对 $-f$ 行进行运算所得仍为原来的行.因此可以直接在这表上进行最优性检验.由于 $a_1 = -2, a_2 = -3$ 都是负值,该基本可行解不是最优解.又由 $a_2 = -3 < a_1 = -2$,取 x_2 为入基变量.

再确定出基变量.计算表中右端项的列与 x_2 所在列各正分量的比值有 $3/1$, $2/1$ 和 $16/1$,通过比较大小取 x_4 为出基变量.据此得新的单纯形表:

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	-2	-3	0	0	0	0
x_3	-1	1	1	0	0	3
x_4	-2	<u>1</u>	0	1	0	2
x_5	4	1	0	0	1	16

其中有下列划线的数字 1 是选定的主元 a_{qp} .

第二次迭代($k=2$):以选定的主元消去 x_2 所在列的其余两个非零元素使其成为单位向量,得下表:

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	-2	-3	0	0	0	0
x_3	1	0	1	-1	0	1
x_2	-2	1	0	1	0	2
x_5	6	0	0	-1	1	14

新的基本可行解的基变量取值为 $x_2 = 2, x_3 = 1, x_5 = 14$,这相应于可行域的顶点 E .以向量 $c_B^{(2)} = (0, -3, 0)^T$ 对 $-f$ 行进行计算得下表(以向量 $c_B^{(2)}$ 先同第一列 $(1, -2, 6)^T$ 取内积后将值加至 -2,再以 $c_B^{(2)}$ 同第二列 $(0, 1, 0)^T$ 取内积将值加至 -3,继续这一过程直至最后一列.在这一过程中,对基变量所在列的计算可省略,因为相应于基变量的简约价值系数为零,因此可直接设其值为零即可.)

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	-8	0	0	3	0	6
x_3	1	0	1	-1	0	1
x_2	-2	1	0	1	0	2
x_5	6	0	0	-1	1	14

由于 $\hat{c}_1^{(2)} = -8 < 0$,该基本可行解还不是最优解.取 x_1 为入基变量.计算表中右端项的列与变量 x_1 所在列各正分量的比值 $1/1$ 和 $14/6$,比较大小,应取 x_3 为出基变量.据此得单纯行表如下,其中有下列划线的数字 1 是选定的新主元 a_{qp} .

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	- 8	0	0	3	0	6
x_3	1	0	1	- 1	0	1
x_2	- 2	1	0	1	0	2
x_5	6	0	0	- 1	1	14

第三次迭代 ($k = 3$): 以选定的主元进行高斯消去得单纯形表:

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	- 8	0	0	3	0	6
x_1	1	0	1	- 1	0	1
x_2	0	1	2	- 1	0	4
x_5	0	0	- 6	5	1	8

新的基本可行解为 $x_1 = 1$, $x_2 = 4$ 和 $x_5 = 8$, 这相应于可行域的顶点 A . 用 $\hat{C}_B^{(3)} = (-8, 0, 0)^T$ 对 $-f$ 行进行运算得下表 (用 $\hat{C}_B^{(3)}$ 同矩阵各列分别取内积后将内积值加至 $-f$ 所在行的相应元素):

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	0	0	8	- 5	0	14
x_1	1	0	1	- 1	0	1
x_2	0	1	2	- 1	0	4
x_5	0	0	- 6	5	1	8

由于 $\hat{C}_4^{(3)} = -5 < 0$, 这个基本可行解还不是最优解, 取 x_4 为入基变量. 计算表中右端项的列与 x_4 所在列各正分量的比值 $5/8$, 于是取 x_5 为出基变量 (主元为 5).

第四次迭代 ($k = 4$): 以选定的主元进行高斯消去得新的单纯形表:

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	0	0	8	- 5	0	14
x_1	1	0	$-\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{13}{5}$
x_2	0	1	$\frac{4}{5}$	0	$\frac{1}{5}$	$\frac{28}{5}$
x_4	0	0	$-\frac{6}{5}$	1	$\frac{1}{5}$	$\frac{8}{5}$

以 $\hat{C}_B^{(4)} = (0, 0, -5)^T$ 对 $-f$ 行的各非基变量的元素作同前面一样的运算得下表:

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	0	0	14	0	1	22
x_1	1	0	$-\frac{1}{5}$	0	$\frac{1}{5}$	$\frac{13}{5}$
x_2	0	1	$\frac{4}{5}$	0	$\frac{1}{5}$	$\frac{28}{5}$
x_4	0	0	$-\frac{6}{5}$	1	$\frac{1}{5}$	$\frac{8}{5}$

至此,由于对所有的 $j \in N_4$ 有 $\hat{c}_j^{(4)} \leq 0$, 已得问题的最优解 $x_1 = \frac{13}{5}, x_2 = \frac{28}{5}, x_3 = 0, x_4 = \frac{8}{5}, x_5 = 0$. 这刚好相应于可行域的顶点 B 最优函数值为 22 的负值, 即为 - 22 .

需要指出的是这里用单纯形表的形式求线性规划问题的解只是为了帮助对单纯形方法的理解, 它不是一个有效的求解过程. 这是因为首先在单纯形表的每一次迭代中, 我们都要计算矩阵 $B^{-1}N$, 而在实际上只有其中的列 $B^{-1}A_p$ 在进一步的计算中起作用, 对矩阵 $B^{-1}N$ 的其他列的计算完全是无用的. 对于中大规模的线性规划问题, 由这样无用的计算所需的计算工作量是十分巨大的. 因此, 一般的计算程序都采用所谓修正单纯形的方法, 也就是前面叙述的单纯形方法. 在这个方法中每次迭代需要计算基矩阵 B_k 的逆矩阵 B_k^{-1} , 以及三个向量, 即

$b_k = B_k^{-1}b, y_k = B_k^{-T}c_B^{(k)}$ 和 $\hat{A}_p^{(k)} = B_k^{-1}A_p$. 至于其他的量, 如

$$\hat{c}_j^{(k)} = c_j^{(k)} - c_B^{(k)T} B_k^{-1}A_j = c_j^{(k)} - y_k^T A_j, \quad j \in N_k$$
$$f_k = c_B^{(k)T} B_k^{-1}b = c_B^{(k)T} b_k,$$

以及计算 b_k 与 \hat{A}_p 相应正分量的比值都可由这些量求得 .

其次, 直接计算矩阵 B_k 的逆, 再用逆矩阵进行相关的计算也不是一个好的选择. 对于大量的实际问题而言, 约束系数矩阵都是具有大量零元素的稀疏矩阵, 因而每次迭代的基矩阵也是稀疏的. 对稀疏矩阵直接求逆所得的逆矩阵通常不再具有原有的稀疏结构, 而是具有大量非零元数的稠密的矩阵, 这不仅破坏了矩阵的稀疏结构, 也大大增加了运算的工作量. 同时直接计算矩阵的逆在数值计算上也可能是不稳定的 .

为提高算法的有效性和稳定性, 对矩阵 B_k 的逆矩阵的计算可以采用逆矩阵的乘积表示. 设第 k 次迭代的基矩阵为 B_k , 相应的基本可行解不是问题的最优解, 并已确定入基变量为 x_p , 出基变量为 x_q , 由此得新的基矩阵 B_{k+1} , 即用 N_k 中的列 A_p 代替 B_k 中的列 A_q 得 B_{k+1} . 由于 $\hat{A}_p^{(k)} = B_k^{-1}A_p$, 我们有

$$B_{k+1} = B_k M_k,$$

其中

$$M_k = \begin{pmatrix} 1 & & & & \hat{a}_{1p}^{(k)} \\ & 1 & & & \hat{a}_{2p}^{(k)} \\ & & \ddots & & \\ & & & \ddots & \hat{a}_{qp}^{(k)} \\ & & & & \ddots \\ & & & & \hat{a}_{mp}^{(k)} & \ddots \\ & & & & & 1 \end{pmatrix}.$$

于是矩阵 B_{k+1} 的逆可表示为

$$B_{k+1}^{-1} = P_k B_k^{-1}, \quad (2.2.6)$$

其中 P_k 是矩阵 M_k 的逆矩阵. 可以直接验证矩阵 P_k 具有形式

$$P_k = \begin{pmatrix} 1 & & & & \hat{a}_{1p}^{(k)} \\ & 1 & & & \hat{a}_{2p}^{(k)} \\ & & \ddots & & \\ & & & \ddots & \hat{a}_{qp}^{(k)} \\ & & & & \ddots \\ & & & & \hat{a}_{mp}^{(k)} & \ddots \\ & & & & & 1 \end{pmatrix},$$

$$\hat{a}_{jp}^{(k)} = -\frac{\hat{a}_{jp}^{(k)}}{\hat{a}_{qp}^{(k)}}, j \neq q, \quad \hat{a}_{qp}^{(k)} = \frac{1}{\hat{a}_{qp}^{(k)}}.$$

称 M_k 为第 k 次迭代的变换矩阵.

设初始基本可行解的基矩阵为单位矩阵, 则由 (2.2.6) 得第二次迭代的基矩阵的逆矩阵为 $B_2^{-1} = P_1 B_1^{-1} = P_1$, 第三次迭代的基矩阵的逆矩阵为 $B_3^{-1} = P_2 B_2^{-1} = P_2 P_1$. 依此递推得第 $k+1$ 次迭代的基矩阵的逆矩阵为

$$B_{k+1}^{-1} = P_k P_{k-1} \dots P_2 P_1.$$

在具体计算过程中, 我们并不需要存储每一个矩阵 P_k , 而只需存储矩阵 P_k 的列向量 $\hat{a}_p^{(k)} = \begin{pmatrix} \hat{a}_{1p}^{(k)} \\ \hat{a}_{2p}^{(k)} \\ \vdots \\ \hat{a}_{mp}^{(k)} \end{pmatrix}^T$ 以及相应的列指标. 至于三个向量 $b_k = B_k^{-1} b$,

$y_k = B_k^{-T} c_B^{(k)}$ 和 $\hat{A}_p^{(k)} = B_k^{-1} A_p$ 的计算按下述方式逐步进行. 对于向量 $A_p^{(k)}$, 由

$$\hat{A}_p^{(k)} = B_k^{-1} A_p = P_{k-1} P_{k-2} \dots P_2 P_1 A_p,$$

记 $\hat{A}_p^{(1)} = A_p$, 则上述表达式可表示成

$$\hat{A}_p^{(i)} = P_{i-1} \hat{A}_p^{(i-1)}, i = 2, 3, \dots, k,$$

依次计算 $\hat{A}_p^{(2)}, \hat{A}_p^{(3)}, \dots$, 直至 $\hat{A}_p^{(k)}$. 对于向量 b_k 可以作同样的处理, 只不过更为有效; 这是因为由 $b_k = B_k^{-1} b$ 以及 $B_k^{-1} = P_{k-1} B_{k-1}^{-1}$, 我们得

$k+1$ 后转步 2 .

单纯形方法的迭代始于一个基本可行解,对于由约束形式 $Ax \leq b$ 转化成标准形的线性规划问题,取所有的松弛变量为基变量即得一个初始基矩阵为单位矩阵的基本可行解.但是对于一般的等式约束 $a_i^T x = b$ 和形如 $a_i^T x \leq b$ 的不等式约束,在通常情况下一般难于确定一个初始基本可行解.考察下列线性规划问题

$$\begin{aligned} \min \quad & -x_1, \\ \text{s.t.} \quad & 2x_1 + 3x_2 = 7, \\ & 2x_1 - 3x_2 \leq 6, \\ & 4x_1 + x_2 \leq 4, \\ & x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

对第二与第三个约束引入松弛变量 x_3 和乘余变量 x_4 将其化成标准形

$$\begin{aligned} \min \quad & -x_1, \\ \text{s.t.} \quad & 2x_1 + 3x_2 = 7, \\ & 2x_1 - 3x_2 + x_3 = 6, \\ & 4x_1 + x_2 - x_4 = 4, \\ & x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

对此,除 x_3 明确可以作为一个基变量外,我们难于确定其他两个基变量.这还是一个十分简单的问题,对于复杂的线性规划问题确定初始基本可行解就更难了.对这类难于确定初始基本可行解的线性规划问题通常的做法是引入所谓的人工变量,形成明确具有初始基本可行解的线性规划,再用两阶段方法或大 M 方法求解.

依上述问题为例,对第一和第三个约束引入人工变量,记为 a_1 和 a_2 ,则有

$$\begin{aligned} \min \quad & -x_1, \\ \text{s.t.} \quad & 2x_1 + 3x_2 + a_1 = 7, \\ & 2x_1 - 3x_2 + x_3 = 6, \\ & 4x_1 + x_2 - x_4 + a_2 = 4, \\ & x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, a_1 \geq 0, a_2 \geq 0, \end{aligned}$$

这时 $a_1 = 7, x_3 = 6, a_2 = 4$, 其他变量取零即构成问题的一个基本可行解,取其为初始基本可行解即可开始单纯形迭代.

所谓两阶段方法就是依上面得到的基本可行解作为初始基本可行解,用单纯形方法先解下述人工的辅助线性规划问题

$$\begin{aligned}
& \min \quad a_1 + a_2, \\
& \text{s.t.} \quad 2x_1 + 3x_2 + a_1 = 7, \\
& \quad \quad 2x_1 - 3x_2 + x_3 = 6, \\
& \quad \quad 4x_1 + x_2 - x_4 + a_2 = 4, \\
& \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, a_1 \geq 0, a_2 \geq 0,
\end{aligned}$$

这个问题的目标函数不是原问题的目标函数,而是两个人工变量的和.如果原问题有可行解,则上述人工问题的最优目标函数值为零,如果人工问题的最优目标函数值大于零,则原问题无可行解.显然,在迭代开始时,所有的人工变量都包括在初始基本可行解中.在迭代进行过程中,人工变量逐步成为出基变量,而原问题的变量中有一些逐步选为入基变量成为基变量.对于一个人工变量,一旦它被选为出基变量成为非基变量(取零值),就可以将其从上述辅助问题中删除.只要原问题有可行解,所有的人工变量都将逐步被删除,最后得到的基本可行解就可以用作求原问题最优解的初始基本可行解,继续对原问题进行单纯形迭代以确定其或者无有界的最优解,或得到一个最优解.把上述求含有人工变量辅助问题以确定初始基本可行解的过程作为第一阶段,而把由确定的初始基本可行解开始求原问题解的过程作为第二阶段.

所谓大 M 方法实质上是通过对人工变量施加惩罚,把两阶段方法的两个阶段合并成一个阶段的方法.还是用上面的例子为例,大 M 方法以 $a_1 (= 7)$, $x_3 (= 6)$, $a_2 (= 4)$ 为基变量,其他变量为非基变量作为初始基本可行解,用单纯形方法求下述线性规划问题的最优解

$$\begin{aligned}
& \min \quad -x_1 + M(a_1 + a_2), \\
& \text{s.t.} \quad 2x_1 + 3x_2 + a_1 = 7, \\
& \quad \quad 2x_1 - 3x_2 + x_3 = 6, \\
& \quad \quad 4x_1 + x_2 - x_4 + a_2 = 4, \\
& \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, a_1 \geq 0, a_2 \geq 0,
\end{aligned}$$

其中人工变量的和同 M 相乘后作为惩罚项被加在目标函数上, M 称为惩罚因子,取充分大的一个正值.如果原问题有可行解,由于 M 的取值很大,只要有人工变量留在基变量中,相应的目标函数值就为大大的大于目标函数在任何可行点的目标函数值,因此单纯形迭代会逐步把人工变量逼出基变量而成为非基变量.同样一旦一个人工变量成为出基变量,就可以从问题中删除.当所有的人工变量被删除后所得到的基本可行解即为原问题的一个基本可行解,同时惩罚项也被从目标函数中移走.继续进行迭代以确定问题的最优解或判定其无有界的最优解.如果直至迭代终止时基变量中还有人工变量则表明原问题无可行解.

在这一节的最后我们再讲两个理论问题:最优解的唯一性与迭代的有限终止性.设 x^* 是一个线性规划问题的最优解,其中基变量为 $x_B^* \geq 0$,非基变量为

$x_N^* = 0$, 相应的基矩阵与非基矩阵为 B 与 N , 目标函数的价值向量为 c_B 和 c_N , 则所有可行解可表示为

$$x = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} + \begin{pmatrix} -B^{-1}N \\ I \end{pmatrix} x_N,$$

目标函数可表示为

$$f(x) = c^T x = c_B^T B^{-1}b + (c_N^T - c_B^T B^{-1}N) x_N = f(x^*) + x_N^T \hat{c}_N^*,$$

其中 $\hat{c}_N^* = c_N^T - N^T B^{-T} c_B$. 根据可行性有 $x_N \geq 0$. 根据最优性向量 \hat{c}_N^* 的所有分量均非负. 如果 \hat{c}_N^* 的所有分量均为正, 则 x_N 的任何分量由零向正值的微小变化都会使目标函数值增大, 这表明问题不可能有其他取值同为 $f(x^*)$ 的可行点, 因而 x^* 是问题的一个唯一局部最优解. 又因为线性规划是凸规划, 其局部最优解也就是它的全局最优解. 如果向量 \hat{c}_N^* 有零分量, 则 x^* 不再是唯一的最优解. 设 $\hat{c}_j^* = 0$, 则增大相应的非基变量 x_j 并保持其他的非基变量取零值直至另一个基本可行解(可行域的另一顶点)都将保持点的可行性, 而相应的函数值为 $f(x) = f(x^*) + x_j \hat{c}_j^* = f(x^*)$ 也是最优的. 因而最优解不唯一. 不难证明如果一个线性规划有多于一个的最优顶点, 则所有这些顶点的任何凸组合也是其最优解. 从这里我们也可以看出如果线性规划问题的最优解不唯一, 则有无穷多个最优解.

对于非退化的线性规划问题, 单纯形方法可以确保经有限次迭代终止于问题的一个最优解. 所谓非退化的线性规划是指问题的每个基本可行解都是非退化的, 也就是没有取零值的基变量. 事实上只要迭代过程中的所有顶点都是非退化的就可以得出有限终止的结论.

定理 2.2.3 设单纯形法迭代过程中的每一个点都是非退化的, 则方法必定经有限次迭代或者确定问题的一个最优解, 或者判定问题无有界的最优解.

证明 对于一个迭代点, 设为 $x^{(k)}$, 如有 $\hat{c}_N^{(k)} \geq 0$, 则 $x^{(k)}$ 即为最优解. 否则有 $\hat{c}_p^{(k)} < 0$, 并选 x_p 为入基变量. 如果 $\hat{A}_p^{(k)}$ 的所有分量非正, 则可以断定问题无有界的最优解; 否则可确定变量 x_p 的取值为

$$x_p = \theta_k = \min \left\{ \frac{b_j^{(k)}}{\hat{a}_{jp}^{(k)}} \mid \hat{a}_{jp}^{(k)} > 0 \right\}.$$

由于点 $x^{(k)}$ 是非退化的, 因而对所有 $j = 1, \dots, k$ 有 $b_j^{(k)} > 0$. 由此得 $\theta_k > 0$ 而相应的新处的目标函数值为

$$f_{k+1} = f_k + \theta_k \hat{c}_p^{(k)},$$

由于 $\hat{c}_p^{(k)} < 0$, $\theta_k > 0$, 我们得

$$f_{k+1} < f_k,$$

也就是目标函数值严格单调下降. 正因为目标函数值严格单调下降, 故保证了迭代过程中的基矩阵不可能出现循环. 由于线性规划只有有限个基本可行解, 故只

要迭代过程中的每一个点都是非退化的,单纯形迭代必经有限次迭代后终止 .

对非退化的线性规划问题,单纯形方法确保每次迭代都能使目标函数值有一定的下降,但是对于退化的线性规划问题,在退化点就未必能使函数值有所下降 .下面的例子就可说明这一点 .

$$\begin{aligned} \min \quad & -5x_1 - 3x_2, \\ \text{s.t.} \quad & x_1 + 2x_2 = 3, \\ & 2x_1 + x_2 = 3, \\ & x_1 = 1.5, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

对这样一个问题,我们引入三个松弛变量 $x_3 \geq 0, x_4 \geq 0, x_5 \geq 0$ 后可得初始单纯形表 .

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	-5	-3	0	0	0	0
x_3	1	2	1	0	0	3
x_4	2	1	0	1	0	3
x_5	1	0	0	0	1	1.5

选 x_1 为入基变量, x_5 为出基变量得第二个单纯形表

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	0	-3	0	0	5	7.5
x_3	0	2	1	0	-1	1.5
x_4	0	1	0	1	-2	0
x_1	1	0	0	0	1	1.5

从这个表我们看到基变量 x_4 取零,因而是一个退化的基本可行解 选 x_2 为入基变量, x_4 为出基变量继续迭代得第三个单纯形表

基变量	x_1	x_2	x_3	x_4	x_5	右端项
$-f$	0	0	0	3	-1	7.5
x_3	0	0	1	-2	3	1.5
x_2	0	1	0	1	-2	0
x_1	1	0	0	0	1	1.5

经过一次迭代,尽管 x_2 由非基变量成了基变量,但其取值仍为零,相应的目标函数值并没有任何下降.在这样一种情况下,我们不能从理论上保证单纯形方法必能有限次迭代能取得最优解或确定其无解.由于函数值不下降而保持不变,有可能出现迭代在几个退化的基本可行解之间循环重复,从而导致迭代不能终止(见练习题[6]).

已经有不少技术用于确保单纯形方法对退化线性规划问题的终止性,其中常用的有扰动方法和 Bland 方法.在 Bland 方法中,要求根据变量的自然顺序,而不是数值的大小来确定入基变量和出基变量.具体来说,设问题中变量的自然顺序为 $x_1, x_2, x_3, \dots, x_n$,在单纯形法的每一次迭代,如果 $\hat{c}_N^{(k)}$ 有不正一个负分量,应把负的简约价值系数中指标最小的变量选为入基变量,而不是把相应于最负分量的变量选为入基变量;在确定出基变量时,如果有多于一个指标使 $\frac{b^{(k)}}{a_{jp}^{(k)}}$ 取得最小值,也取其中最先出现的变量作为出基变量. Bland(1977)已经证明采用上述措施确定入基和出基变量的单纯形方法必具有有限终止性.

§ 2.3 线性规划的对偶与对偶单纯形法

对每一个线性规划问题,可以构造另一个与之相应且密切相关的线性规划问题,如果称前者为原始问题,后者就称为对偶问题.线性规划的原始问题和对偶问题无论从数学的角度还是从经济的角度都有十分密切的关系.线性规划的对偶在求解算法的设计中也起重要作用.例如,近十年来得到广泛关注和应用的内点算法同时利用了线性规划的原始和对偶信息.

让我们从下述形式的线性规划问题开始,引入其对偶问题,再得出其他形式的线性规划问题的对偶问题.考虑经典的线性规划问题

$$\begin{aligned} \min \quad & f(x) = c^T x, \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned}$$

经典线性规划问题不同于标准形线性规划问题,这里的约束都是 型的不等式约束,对变量同样有非负的限制.它的对偶问题具有对称形式

$$\begin{aligned} \max \quad & z(y) = b^T y, \\ \text{s.t.} \quad & A^T y \leq c, \\ & y \geq 0. \end{aligned}$$

首先原始问题是对目标函数求极小,对偶问题是对目标函数求极大;其次对偶问题中变量 $y \in \mathbf{R}^m$,其分量的个数等于原问题的约束数;对偶问题的目标函数 $b^T y$,其价值系数向量 b 刚好是原问题约束的右端项;对偶问题的约束 $A^T y \leq c$,其系数矩阵是原约束系数矩阵的转置,右端项是原问题目标函数的价值系数向量 c ;对于对偶问题中的变量 y 同样有非负的限制.可以看出经典线性规划问题同它的对偶问题之间有所谓的对称关系,这体现在

$$\min \iff \max, \quad Ax \leq b \iff A^T y \leq c, \quad c \iff b.$$

从最优化问题的最优性条件我们知道,如果上述经典线性规划问题有最优解,则在最优解,设为 x^* 处,存在最优的拉格朗日乘子向量 $y^* \in \mathbf{R}^m$ 和 $\lambda^* \in \mathbf{R}^n$ 满足

$$c = A^T y^* + \lambda^*, \quad y^* \geq 0, \quad \lambda^* \geq 0.$$

考虑到 $\lambda^* \geq 0$,上式又可以写成

$$c \geq A^T y^*, \quad y^* \geq 0,$$

即在原问题最优解处,相应于约束 $Ax \leq b$ 的最优拉格朗日乘子是对偶问题的一个可行解.

如同其他形式的线性规划问题可以转换成标准形式,我们同样可以把其他形式的线性规划问题转换成经典形式后再得出其对偶问题.例如对于 的约束,通过两边乘 -1 转换成 的约束;对于等式约束,可转换成两个不等式约束;例如约束 $2x_1 + 3x_2 = 6$ 可转换成

$$2x_1 + 3x_2 \leq 6, \quad -2x_1 - 3x_2 \leq -6$$

两个约束.至于对变量 x 非负限制的转换,我们已在前面作过介绍,这里不再重复.

下面给出确定一般线性规划问题对偶问题的基本规则.考虑一般形式的线性规划问题

$$\begin{aligned} \min \quad & f(x) = c^T x, \\ \text{s.t.} \quad & A_1 x \leq b_1, \\ & A_2 x \leq b_2, \\ & A_3 x = b_3, \\ & x \geq 0, \end{aligned}$$

首先将其转换成经典形式

$$\begin{aligned} \min \quad & f(x) = c^T x, \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

其中

$$A = \begin{bmatrix} A_1 & & & \\ & -A_2 & & \\ & & A_3 & \\ & & & -A_3 \end{bmatrix}, \quad b = \begin{bmatrix} b \\ -b \\ b \\ -b \end{bmatrix}.$$

则其对偶问题为

$$\begin{aligned} \max \quad & z(\bar{y}) = b^T \bar{y}_1 - b^T \bar{y}_2 + b^T \bar{y}_3 - b^T \bar{y}_4, \\ \text{s.t.} \quad & A^T \bar{y} = A_1^T \bar{y}_1 - A_2^T \bar{y}_2 + A_3^T \bar{y}_3 - A_3^T \bar{y}_4 = c, \\ & \bar{y} \geq 0, \end{aligned}$$

其中 $\bar{y}^T = (\bar{y}_1^T, \bar{y}_2^T, \bar{y}_3^T, \bar{y}_4^T)$. 令 $y_1 = \bar{y}_1, y_2 = -\bar{y}_2, y_3 = \bar{y}_3 - \bar{y}_4$, 得对偶问题为

$$\begin{aligned} \max \quad & z(y) = b^T y_1 + b^T y_2 + b^T y_3, \\ \text{s.t.} \quad & A_1^T y_1 + A_2^T y_2 + A_3^T y_3 = c, \\ & y_1 \geq 0, y_2 \geq 0. \end{aligned}$$

注意相应于原问题 的约束, 在对偶问题中的相应变量 y_1 有非负()的限制; 对原问题 型约束, 对偶问题中的相应变量 y_2 有同样方向的约束 ; 而对应于原问题的等式 (=) 约束, 对偶问题中的变量则无限制 . 由

原问题约束	对偶问题变量
=	无限制

此我们就原线性规划问题中的约束与其对偶问题的变量间的对偶关系得到上面的表 .

再考虑原始线性规划问题中对变量的不同限制,

$$\begin{aligned} \min \quad & f(x) = c^T x, \\ \text{s.t.} \quad & Ax = b, \\ & x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

这里 $A = [A_1 \ A_2 \ A_3]$, $c^T = (c_1^T, c_2^T, c_3^T)$, $x^T = (x_1^T, x_2^T, x_3^T)$, 对变量 x_3 无限制 . 同样将其转换成经典形式

$$\begin{aligned} \min \quad & c_1^T \bar{x}_1 - c_2^T \bar{x}_2 + c_3^T \bar{x}_3 - c_3^T \bar{x}_4, \\ \text{s.t.} \quad & A_1 \bar{x}_1 - A_2 \bar{x}_2 + A_3 \bar{x}_3 - A_3 \bar{x}_4 = b, \\ & \bar{x}_1 \geq 0, \bar{x}_2 \geq 0, \bar{x}_3 \geq 0, \bar{x}_4 \geq 0, \end{aligned}$$

其中 $\bar{x}_1 = x_1, \bar{x}_2 = -x_2, x_3 = \bar{x}_3 - \bar{x}_4$.这一问题的对偶为

$$\begin{aligned} \max \quad & z(y) = b^T y, \\ \text{s.t.} \quad & \bar{A}^T y \leq \bar{c}, \\ & y \geq 0, \end{aligned}$$

其中

$$\bar{A} = [A_1 \quad -A_2 \quad A_3 \quad -A_3], \bar{c} = (c^T, -c^T, c^T, -c^T).$$

即有

$$\begin{aligned} \max \quad & z(y) = b^T y, \\ \text{s.t.} \quad & A_1^T y \leq c, \\ & -A_2^T y \leq -c, \\ & A_3^T y \leq c, \\ & -A_3^T y \leq -c, \\ & y \geq 0, \end{aligned}$$

得对偶问题

$$\begin{aligned} \max \quad & z(y) = b^T y, \\ \text{s.t.} \quad & A_1^T y \leq c, \\ & A_2^T y \geq c, \\ & A_3^T y = c, \\ & y \geq 0. \end{aligned}$$

由此我们就原线性规划问题的变量与其对偶问题的约束间的对偶关系得到右面的表 .

利用上述规则就可以很容易地得到任何形式线性规划问题的对偶问题 .例如考虑线性规划问题

原问题变量	对偶问题约束
无限制	=

$$\begin{aligned} \min \quad & f(x) = x_1 + 2x_2 + 3x_3, \\ \text{s.t.} \quad & 3x_1 - 2x_2 + x_3 = 6, \\ & 4x_1 - 2x_2 + 3x_3 \leq 5, \\ & x_1 + 3x_2 + 5x_3 \geq 9, \\ & x_1 \geq 0, x_2 \geq 0, x_3 \text{ 无限制}. \end{aligned}$$

根据上述规则,这一问题的对偶为

$$\begin{aligned}
 \max \quad & z(y) = 6y_1 + 5y_2 + 9y_3, \\
 \text{s.t.} \quad & 3y_1 + 4y_2 + y_3 = 1, \\
 & -2y_1 - 2y_2 + 3y_3 = 2, \\
 & y_1 + 3y_2 + 5y_3 = 3, \\
 & y_1 \text{ 无限制}, y_2 \geq 0, y_3 \geq 0.
 \end{aligned}$$

还需要指出的一点是,对任何线性规划问题,其对偶的对偶还是原问题.为简单起见,只考虑经典的线性规划问题,已知其对偶形式,将其对偶转换成经典形式得

$$\begin{aligned}
 \min \quad & z(y) = -b^T y, \\
 \text{s.t.} \quad & -A^T y = -c, \\
 & y \geq 0,
 \end{aligned}$$

它的对偶是

$$\begin{aligned}
 \max \quad & f(x) = -c^T x, \\
 \text{s.t.} \quad & -Ax = -b, \\
 & x \geq 0,
 \end{aligned}$$

经整理即得原线性规划问题.据此,在线性规划的原始和对偶两个问题之间,不必过于强调哪一个是原始问题,哪一个是对偶问题,其中的任意一个都可以是原始问题,而另一个就是它的对偶问题.

线性规划的弱对偶定理和强对偶定理描述了线性规划的原始问题的最优值和对偶问题的最优值之间的关系.由于任何形式的线性规划问题都可以转换成标准形式,下面我们就标准型的线性规划问题分析这两个定理,有关的结论适用于其他形式的线性规划.

考虑标准型的线性规划问题

$$\begin{aligned}
 \min \quad & f(x) = c^T x, \\
 \text{s.t.} \quad & Ax = b, \\
 & x \geq 0.
 \end{aligned}$$

根据上述对偶规则,可得其对偶问题为

$$\begin{aligned}
 \max \quad & z(y) = b^T y, \\
 \text{s.t.} \quad & A^T y = c,
 \end{aligned}$$

由于原始问题中的约束是等式约束,对偶问题中没有对对偶变量的非负的限制.

定理 2.3.1 (弱对偶定理) 设 x 和 y 分别是原始问题和对偶问题的可行解,则有

$$f(x) = c^T x = b^T y = z(y).$$

证明 根据 y 的对偶可行性有 $A^T y = c$.再由 $Ax = b, x \geq 0$ 得

$$f(x) = c^T x \quad y^T A x = y^T b = z(y).$$

弱对偶定理表明,一个极小化线性规划问题在其某一可行点处的函数值,给出了另一个作为其对偶的极大化线性规划问题的最优值的一个上界;或者说,一个极大化线性规划问题在其某个可行点处的函数值,给出了另一个作为其对偶的极小化线性规划问题的最优值的下界.称在两个可行解的函数值之间的差 $c^T x - b^T y$ 为对偶间隙.

从这一定理,我们还可以得出下列两个有用的结论:

(1) 如果两个互为对偶的线性规划问题中有一个无有界的最优解,则另一个问题必定不可行.这是因为如果另一个问题有可行解,则其相应的目标函数值为另一个问题的最优函数值提供了一个界,这同它的无界性假设相矛盾.

(2) 如果两个问题各有一个可行解,且相应的目标函数值相同,则这两个可行解各是相应问题的最优解.设 x^* 和 y^* 各是上述原问题和对偶问题的可行解,且 $c^T x^* = b^T y^*$,则根据弱对偶定理对任意原问题的可行解 x 有 $c^T x - b^T y^* = c^T x - c^T x^*$,因而 x^* 是原问题的最优解;同样可以得出 y^* 是对偶问题的最优解.弱对偶定理还表明,我们可以不通过求解过程来确定一个可行点是否是最优的.

定理 2.3.2 (强对偶定理) 如果互为对偶的两个线性规划问题中一个有最优解,则另一个也有最优解,且两者的最优目标函数值相等.

证明 考虑标准型线性规划问题,设 x^* 是其一个最优解.根据标准型线性规划问题最优解的理论,将 x^* 分成基本解和非基本解两部分

$$x^* = \begin{pmatrix} x_B^* \\ x_N^* \end{pmatrix}.$$

对系数矩阵 A 和目标函数的价值向量 c 作相应的分块表示

$$A = [B \ N], \quad c = \begin{pmatrix} c_B \\ c_N \end{pmatrix},$$

则 $x_B^* = B^{-1} b$. 由于 x^* 是最优解,根据本章第 2.1 节的分析有

$$c_N^T - c_B^T B^{-1} N = 0.$$

由此得 $c_N^T = N^T B^{-T} c_B^T$. 令 $y^* = B^{-T} c_B^T$, 则有

$$A^T y^* = \begin{pmatrix} B^T \\ N^T \end{pmatrix} B^{-T} c_B^T = \begin{pmatrix} c_B \\ N^T B^{-T} c_B \end{pmatrix} = \begin{pmatrix} c_B \\ c_N \end{pmatrix} = c.$$

这表明 y^* 是对偶可行的. 又由于

$$b^T y^* = b^T B^{-T} c_B^T = c_B^T B^{-1} b = c_B^T x_B^* = c^T x^*.$$

根据弱对偶定理知 y^* 是对偶问题的最优解. 这就完成了定理的证明.

强对偶定理表明,如果一个线性规划有最优解,则在其最优解处,对偶间隙为零.从定理证明的过程可以看到,如果 x^* 是原始问题的最优解,则通过令

$y^* = B^{-T} c_B$, 即得对偶问题的最优解. 我们知道, 在单纯形迭代的每一步都要计算单纯形乘子向量 $y = B^{-T} c_B$. 由于在非最优解处, $c_N - N^T B^{-T} c_B = 0$ 不满足, 因而

$$A^T y = \begin{matrix} B^T \\ N^T \end{matrix} B^{-T} c_B = \begin{matrix} c_B \\ N^T B^{-T} c_B \end{matrix} \begin{matrix} c_B \\ c_N \end{matrix} = c$$

并不成立, 即 $y = B^{-T} c_B$ 不是对偶可行的. 我们称这样的基本可行解为对偶不可行的基本可行解. 而对于最优解它不但是基本可行解, 而且也是对偶可行的. 因此对单纯形方法可以作这样的解释: 单纯形方法的每次迭代确定一个对偶不可行的基本可行解, 逐步减少其对偶不可行性, 一旦得到一个对偶可行的基本可行解即已得最优解, 迭代终止. 在本节后面要介绍的对偶单纯形方法则刚好相反, 每次迭代产生一个对偶可行的基本解(对原问题不可行), 逐步增加其对原问题的可行性, 一旦在某步迭代得到一个对偶可行的基本可行解, 即是最优解, 迭代终止.

再考虑对偶可行性的含义, 对于标准型的线性规划问题, 由最优性的一阶条件, 在其最优解 x^* 处存在拉格朗日乘子向量 y^* 和 $\bar{c} = 0$ 使成立

$$c - A^T y^* = \bar{c}, \quad x^{*T} \bar{c} = 0.$$

因此, 对偶可行性就是确保相应于约束条件 $x \geq 0$ 的拉格朗日乘子 \bar{c} 非负. 对于单纯形方法来说一旦在一个基本可行解处, 其拉格朗日乘子非负(对偶可行), 则这个基本可行解就是最优解. 同样, 对于对偶单纯形方法来说, 它的每一次迭代都确保相应的拉格朗日乘子非负(对偶可行). 一旦在某步迭代, 基本解又是可行解, 则最优解的条件得到满足, 该点是问题的最优解.

至于条件 $x^{*T} \bar{c} = x^{*T} (c - A^T y^*) = 0$, 就是在最优性条件一节提过的互补松弛条件. 显而易见这个条件在单纯形迭代的每一步都得到满足. 这是因为根据乘子 $y = B^{-T} c_B$ 的定义, 我们有

$$c - A^T y = \begin{matrix} c_B \\ c_N \end{matrix} - \begin{matrix} B^T \\ N^T \end{matrix} B^{-T} c_B = \begin{matrix} 0 \\ c_N - N^T B^{-T} c_B \end{matrix}.$$

而相应的基本可行解为 $x^T = (x_B^T, x_N^T)$, 其中 $x_B = B^{-1} b$, $x_N = 0$, 因此有

$$x^T (c - A^T y) = x_B^T 0 + x_N^T (c_N - N^T B^{-T} c_B) = 0.$$

注意, 在非最优的迭代点, 或者有 $x_B = 0$, 但 $c_N - N^T B^{-T} c_B = 0$ 不成立(单纯形法), 或者有 $c_N - N^T B^{-T} c_B = 0$, 但 $x_B = 0$ 不成立(对偶单纯形法).

对偶单纯形法

在正式介绍对偶单纯形方法之前, 我们再通过一个简单的例子说明对偶问题的经济意义. 设某个企业用同样的三种原料生产两种产品, 已知销售一个单位的第 i 种产品可获利 c_i , 生产一个单位的第 i 种产品所用三种原料的量分别为

a_{ji} , $i = 1, 2$, $j = 1, 2, 3$, 企业可提供的三种原料的总量分别是 b_j , $j = 1, 2, 3$, 则使企业获利最大的生产计划安排须由下述线性规划确定

$$\begin{aligned} \max \quad & c_1 x_1 + c_2 x_2, \\ \text{s.t.} \quad & a_{11} x_1 + a_{12} x_2 \leq b_1, \\ & a_{21} x_1 + a_{22} x_2 \leq b_2, \\ & a_{31} x_1 + a_{32} x_2 \leq b_3, \\ & x_1 \geq 0, x_2 \geq 0, \end{aligned}$$

这里 x_1 和 x_2 分别表示两种产品的计划生产量, $c_1 x_1 + c_2 x_2$ 是企业从销售所生产的产品获取的利润, 应取极大. 显然, 如果企业增加原料的供应, 就可增加产品的生产 (假定有足够的生产能力) 从而增加赢利. 但是, 不适当的增加也会导致可能的损失, 因而有一个如何合理增加不同原料的问题. 上述线性规划只提供了在各种原料给定的情况下如何组织生产的信息, 并没有提供如何增加不同原料供应的任何信息. 为此我们再看这一问题的对偶

$$\begin{aligned} \min \quad & b_1 y_1 + b_2 y_2 + b_3 y_3, \\ \text{s.t.} \quad & a_{11} y_1 + a_{21} y_2 + a_{31} y_3 \leq c_1, \\ & a_{12} y_1 + a_{22} y_2 + a_{32} y_3 \leq c_2, \\ & y_1 \geq 0, y_2 \geq 0. \end{aligned}$$

如果把 y_i , $i = 1, 2, 3$ 看作三种原料的投料比列, 则对偶问题可理解为确定最优的投料比例使损失最小, 而相应于原问题最优解的对偶变量 y 的取值, 确定了最优生产计划下的不同原料的投料比列, 从而也相应的确定了如何合理地增加原料供应. 因此通常把对偶问题中的变量 y 又称为影子价格 (Shadow Price).

根据前面的叙述, 对偶单纯形方法通过产生一个保持对偶可行性, 并使原始可行性得到逐步改善的基本解系列, 以最终得到一个最优的基本可行解. 因此对偶单纯形法从一个对偶可行的基本解开始, 先检验其最优性; 如果不是最优的, 则从取负值的基变量中确定一个出基变量, 以增加新迭代点的原始可行性; 选好出基变量后, 根据新迭代点应确保对偶可行的原则确定一个入基变量; 选定好出基和入基变量后, 就可以对基矩阵修正得到新的对偶可行的基本解, 重复这个过程, 直至或者确定问题无可行解, 或者得到一个最优解. (注意, 同单纯形法不同, 对偶单纯形法确定的是或者问题的最优解, 或者问题无可行解.)

设 $x^{(k)}$ 是标准型线性规划问题的一个对偶可行的基本解, 相应的基矩阵和非基矩阵分别为 B_k 和 N_k , 则

$$x_B^{(k)} = B_k^{-1} b = \bar{b}^{(k)}, x_N^{(k)} = 0.$$

记相应的价值系数向量为 $c_B^{(k)}$ 和 $c_N^{(k)}$, 由对偶可行性有

$$c_N^{(k)} - N_k^T B_k^{-T} c_B^{(k)} \leq 0.$$

对这样的解 $x^{(k)}$, 如果有

$$b^{(k)} \leq 0,$$

则已有原始可行性,因而 $x^{(k)}$ 即为要求的一个最优基本可行解.如果 $b^{(k)}$ 有负的分量,则 $x^{(k)}$ 原始不可行,取其中的一个负分量作为出基变量,将其值从负增加至零而成为非基变量,设取 $x_q^{(k)} = b_q^{(k)} < 0$ 为出基变量.

为确定入基变量,考察在点 $x^{(k)}$ 的约束方程

$$x_B^{(k)} + B_k^{-1} N_k x_N^{(k)} = B_k^{-1} b = b^{(k)}.$$

则相应于基变量 $x_q^{(k)}$ 的方程为

$$x_q^{(k)} + \sum_{j \in N_k} \hat{a}_{qj}^{(k)} x_j = b_q^{(k)},$$

其中 $\hat{a}_{qj}^{(k)}$ 为矩阵 $B_k^{-1} N_k$ 中相应于基变量 $x_q^{(k)}$ 所在行的元素.通过选择一个入基变量,设为 x_p ,使其取正值并使 x_q 取零值(成为非基变量).由于 $b_q^{(k)} < 0$,如果有某一元素 $\hat{a}_{qp}^{(k)} < 0$, $p \in N_k$,则保持其他非基变量取零值,并令

$$x_p = \frac{b_q^{(k)}}{\hat{a}_{qp}^{(k)}} > 0,$$

就可使 x_q 取零值而成为非基变量,且方程 $Ax = b$ 成立.如果有不止一个元素 $\hat{a}_{qj}^{(k)} < 0$, $j \in N_k$,对入基变量 x_p 的选取则依赖于确保对偶可行性.设 x_p 为入基变量,则元素 $\hat{a}_{qp}^{(k)}$ 就成为对新的基矩阵进行高斯消去的主元,对价值系数 $\hat{c}_j^{(k)}$ 进行同单纯形算法一样的运算后得

$$\hat{c}_j^{(k+1)} = \hat{c}_j^{(k)} - \hat{c}_p^{(k)} \frac{\hat{a}_{qj}^{(k)}}{\hat{a}_{qp}^{(k)}}, \quad j \in N_{k+1}.$$

为保持对偶可行性,要求 $\hat{c}_j^{(k+1)} \geq 0$, $j \in N_{k+1}$,这只有当

$$\left| \frac{\hat{c}_j^{(k)}}{\hat{a}_{qj}^{(k)}} \right| \leq \left| \frac{\hat{c}_p^{(k)}}{\hat{a}_{qp}^{(k)}} \right|$$

对所有 $\hat{a}_{qj}^{(k)} < 0$, $j \in N_k$ 成立时才满足.因此入基变量 x_p 的选择依据为

$$\left| \frac{\hat{c}_p^{(k)}}{\hat{a}_{qp}^{(k)}} \right| = \min \left| \frac{\hat{c}_j^{(k)}}{\hat{a}_{qj}^{(k)}} \right| \mid \hat{a}_{qj}^{(k)} < 0, \quad j \in N_k.$$

出基变量和入基变量一经选定,以元素 $\hat{a}_{qp}^{(k)}$ 为主元进行高斯消去并对价值系数向量进行修正即得新的对偶可行的基本解 $x^{(k+1)}$.

我们先依下述例子用单纯形表的方式来说明对偶单纯形方法,然后再给出对偶单纯形方法的迭代步骤.考虑线性规划问题

$$\begin{aligned} \min \quad & 6.8x_1 + 3x_2, \\ \text{s.t.} \quad & 2.6x_1 + x_2 \leq 800, \\ & 3.8x_1 + 3x_2 \leq 1000, \\ & 1.6x_1 + x_2 \leq 100, \\ & 6x_1 + 10x_2 \leq 6000, \\ & x_1 \geq 0, x_2 \geq 0. \end{aligned}$$

对这样一个问题引入剩余变量 $x_i \geq 0, i = 3, 4, 5, 6$, 将其转化成标准型

$$\begin{aligned} \min \quad & 6.8x_1 + 3x_2, \\ \text{s.t.} \quad & 2.6x_1 + x_2 - x_3 = 800, \\ & 3.8x_1 + 3x_2 - x_4 = 1\,000, \\ & 1.6x_1 + x_2 - x_5 = 100, \\ & 6x_1 + 10x_2 - x_6 = 6\,000, \\ & x_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6. \end{aligned}$$

对此, 显然难于确定问题的一个初始基本可行解. 如用单纯形方法求解则需再引入人工变量后用两阶段方法或大 M 方法进行求解. 由于在问题的目标函数中所有的价值系数分量都非负, 可以想象如果引入松弛变量 $x_i \geq 0, i = 3, 4, 5, 6$ 将其转换成标准型式后, 取松弛变量为初始的基变量, 虽然原始可行性不能得到满足, 但对偶可行性却得到满足, 因而可以用对偶单纯形方法求解. 为此先把约束转换成 \leq 型, 再引入松弛变量后得

$$\begin{aligned} \min \quad & 6.8x_1 + 3x_2, \\ \text{s.t.} \quad & -2.6x_1 - x_2 + x_3 = -800, \\ & -3.8x_1 - 3x_2 + x_4 = -1\,000, \\ & -1.6x_1 - x_2 + x_5 = -100, \\ & -6x_1 - 10x_2 + x_6 = -6\,000, \\ & x_i \geq 0, \quad i = 1, 2, 3, 4, 5, 6. \end{aligned}$$

取 x_3, x_4, x_5, x_6 为初始基变量即得初始对偶可行的基本解, 相应的单纯形表如下:

基变量	x_1	x_2	x_3	x_4	x_5	x_6	右端项
$-f$	6.8	3	0	0	0	0	0
x_3	-2.6	-1	1	0	0	0	-800
x_4	-3.8	-3	0	1	0	0	-1\,000
x_5	-1.6	-1	0	0	1	0	-100
x_6	-6	-10	0	0	0	1	-6\,000

右端项全负, 不是最优解. 我们取最负的分量, 即 x_6 为出基变量. 比较单纯形表 x_6 所在行中相应于非基变量的负元素和 $-f$ 行相应分量的比值有 $\left| \frac{6.8}{-6} \right| > \left| \frac{3}{-10} \right|$, 因此取 x_2 为入基变量. 以 -10 为主元进行高斯消去后, 再把 $-f$ 行的各元素减去向量 $(0, 0, 0, 3)^T$ 与该元素下矩阵所在列向量的内积得新的对偶单纯形表

基变量	x_1	x_2	x_3	x_4	x_5	x_6	右端项
$-f$	5	0	0	0	0	0.3	-1 800
x_3	-2	0	1	0	0	-0.1	-200
x_4	-2	0	0	1	0	-0.3	800
x_5	-1	0	0	0	1	-0.1	500
x_2	0.6	1	0	0	0	-0.1	600

右端项有一个负分量,相应的基本解不可行,因而不是最优解.取 x_3 为出基变量,比较 x_3 所在行内相应于非基变量的负元素与 $-f$ 行相应元素的比值有 $\left| \frac{5}{-2} \right| < \left| \frac{0.3}{-0.1} \right|$,取 x_1 为入基变量.依 -2 为主元进行高斯消去,再把 $-f$ 行的各元素减去向量 $(5,0,0,0)^T$ 与该元素下向量的内积得下述单纯形表

基变量	x_1	x_2	x_3	x_4	x_5	x_6	右端项
$-f$	0	0	2.5	0	0	0.05	-2 300
x_1	1	0	-0.5	0	0	0.05	100
x_4	0	0	-1	1	0	-0.02	1 000
x_5	0	0	-0.5	0	1	-0.05	600
x_2	0	1	0.3	0	0	-0.13	540

至此,右端项的所有分量都已非负,当前的迭代点已是一个对偶可行的基本可行解,因而也是最优解,即最优解为 $x^* = (100, 540)^T$,相应的最优目标函数值为 $f^* = 2\,300$.

在进一步理解了对偶单纯形法以及它的迭代过程后,我们再给出对偶单纯形方法可在计算机上实现的迭代步骤.如同采用单纯形表的过程一样,在用对偶单纯形表求解时,要计算整个矩阵 $B_k^{-1} N_k$,这是没有必要的,我们只需要计算矩阵 $B_k^{-1} N_k$ 的相应于出基变量所在行的元素 $\hat{a}_{qj}^{(k)}, j = 1, \dots, n$ 即可.这可以通过先计算阵 B_k 的逆 B_k^{-1} ,记 B_k^{-1} 的相应于出基变量所在的行为 $\hat{B}_q^{(k)}$,矩阵 N_k 中的列向为 $A_j, j = 1, \dots, n$,则有

$$\hat{a}_{qj}^{(k)} = (\hat{B}_q^{(k)})^T A_j, \quad j = 1, \dots, n.$$

另一方面,如果对所有 $j = 1, \dots, n$ 有 $\hat{a}_{qj}^{(k)} \geq 0$,则由

$$x_q^{(k)} + \sum_{j=1}^n \hat{a}_{qj}^{(k)} x_j = b_q^{(k)} < 0,$$

可以看出不可能存在任何 $x_j > 0$ 使 $\sum_{j=1}^n x_j \hat{a}_{qj}^{(k)} = b_q^{(k)}$ 成立,从而使 $x_q = 0$,即原问题不存在可行解,由弱对偶定理知对偶问题无有界的最优解.据此,可得对偶单纯形法的迭代步骤如下:

最优解的基有没有发生变化.这是因为,如果最优基没有发生变化,尽管最优解由于数据的变化而发生变化,但由于最优基没变,从原来的最优基可以很快由 $B^{-1}b$ 根据新的数据得到新的最优解.

其次,如果最优基发生变化,如何由原来的最优解确定新的最优解.

根据前面的单纯形法和对偶单纯形法的学习,我们知道,对于标准型的线性规划问题,确定一个基本解是否是最优的在于其可行性及对偶可行性两个方面,即要求

$$B^{-1}b \geq 0, c_N - N^T B^{-T} c_B \leq 0.$$

两个条件中只要有一个得不到满足,就不是最优的.因此为确定最优基是否发生变化,只需对变化的量从上述关系中确定一个范围,在这个范围内上述条件仍然得到满足.那么变化只要限于这个范围内,最优基就不会发生变化.如果变化超出了这个范围,则最优基发生变化,需要重新确定新的最优解.求解可以从原最优解开始,或者用单纯形法或者用对偶单纯形法确定.如果变化后的数据使对偶可行性得不到满足就用单纯形法,如果原始可行性得不到满足,就用对偶单纯形法.

§ 2.4 线性规划的内点算法

单纯形方法自其产生之日起,因为其能有效地求解各类线性规划问题而得到越来越广泛的应用.随着线性规划问题的规模越来越大,对单纯形方法性能的了解变得十分必要.大量的统计分析和理论研究表明,用单纯形方法求解线性规划问题的平均迭代次数,是问题约束个数 m 的一个不大的倍数.我们知道,每次单纯形迭代需要计算 $\hat{c}_j^{(k)} (j = 1, \dots, n)$ 以检验最优性,其计算需要 $O(mn)$ 次算术运

算;如果不是最优解,则需要计算比值 $\frac{b_j}{\hat{a}_{jp}^{(k)}}, \hat{a}_{jp}^{(k)} > 0 (j = 1, \dots, n)$ 以确定出基变量,这同

需要 $O(mn)$ 次算术运算;由矩阵 B_k 的逆 B_k^{-1} 计算矩阵 B_{k+1} 的逆 B_{k+1}^{-1} 则需要 $O(m^3)$ 次算术运算;而计算新的基本解 $b_{k+1} = B_{k+1}^{-1}b = P_k b_k$ 需要 $2m$ 次算术运算.因此,每次单纯形迭代需要 $O(m^3 + mn)$ 阶的运算工作量,因而算法的平均运算工作量为 $O(m^4 + m^2 n)$ 阶.这一数字是 m 和 n 的一个多项式.一个算法如果其求得问题的解所用的运算工作量是问题的参数 m 和 n 的多项式,则称这一算法的复杂性为多项式时间算法;如果所需的运算工作量的阶数是以 m 或 n 为幂的指数 2^m 或 2^n ,则称为复杂性是指数时间的算法.下面的表比较了 m 不同时 m^4 和 2^m 的值.

m	m^4	2^m
5	625	32
10	1×10^4	1 024
50	6.25×10^6	1.13×10^{15}
100	1×10^8	1.26×10^{30}
1 000	1×10^{12}	1×10^{301}

从这个表可以看出,随着 m 的不断增大,指数 2^m 的增长是十分惊人的.我们以 $m=100$ 为例,假设有一个计算机每秒可完成 $m^4 = 100^4 = 1 \times 10^8$ 阶的运算,则完成 $2^{100} = 1.26 \times 10^{30}$ 运算需要 4×10^{14} 年,也就是说多项式时间算法一至两秒可完成的问题,用指数时间算法可能需要 4×10^{14} 年.由此可见,指数时间算法不是可取的算法.

既然,单纯形算法的平均运算工作量是多项式时间的,那么对于一个具体的问题其算法的复杂性是否也是多项式的呢?对此问题的回答是否定的.Klee 和 Minty 在 1972 年给出了下面这样一个例子

$$\begin{aligned} \max \quad & f(x) = \sum_{i=1}^m 10^{m-i} x_i, \\ \text{s.t.} \quad & 2 \sum_{i=1}^{j-1} 10^{j-i} x_i + x_j \leq 100^{j-1}, \quad j = 1, 2, \dots, m, \\ & x_i \geq 0, \quad i = 1, 2, \dots, m. \end{aligned}$$

这样一个线性规划问题的可行域有 2^m 个顶点,当引入 m 个松弛变量使其化成标准型后用单纯形法求解,方法从初始顶点 $x=0$ 开始,如果每次迭代以取最负的价值系数的非基变量为入基变量,则迭代将经过问题的每一个顶点(或 $2^m - 1$ 次迭代)后方达到最优顶点.图 2.7 给出了问题在 $m=3$ 时的可行域,以及单纯形迭代的迭代点的进展过程.

尽管这一问题是为人为设计的,也是最坏的一种可能,但它却表明,单纯形方法作为线性规划的一个求解算法,其算法的复杂性是指数时间的.这一结论并没有对单纯形方法的应用带来任何负面影响,方法一如既往地得到广泛的应用,但对研究工作者却提出了两个问题:

(1)对线性规划问题是否存在时间复杂性是多

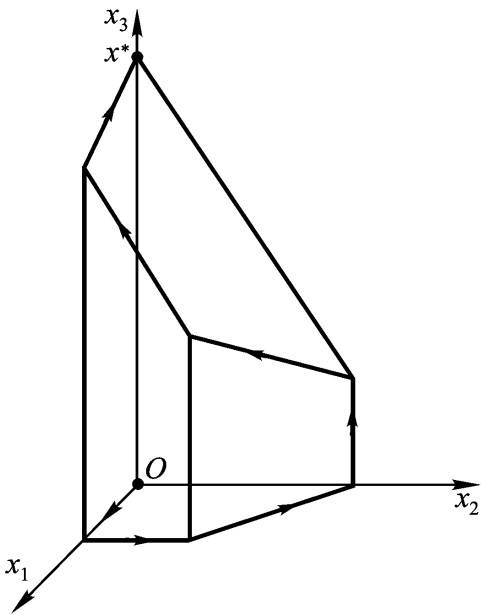


图 2.7 $n=3$ 时的 Klee Minty 问题

项式的算法;(2)如果存在多项式时间算法,如何设计这样的算法.

对第一个问题,前苏联数学家 Khachiyan 在 1979 年作了正面回答.他当时提出了一个椭球算法求不等式问题的解,并证明了算法的时间复杂性是多项式的.利用对偶理论,线性规划问题可以转换成不等式问题,这就明确回答了对线性规划存在多项时间算法.但是计算的实际表明,椭球算法的效果要比单纯形方法差得多,并不是一个有实际应用价值的算法.

对第二个问题的回答始于在美国贝尔实验室工作的印度数学家 Karmarkar 在 1984 年的杰出工作.他对线性规划的求解提出了一个具有多项式时间复杂性的内点算法.算法的提出在当时的学术界引起了轰动,并由此引发了对内点算法研究的热潮.现今,内点算法已经成为求解大规模线性规划问题的有效算法之一.

我们已经知道,求线性规划问题解的单纯形方法在问题的基本可行解中确定最优解,在几何上,每次迭代它是沿着可行域的边界从一个顶点向另一个更好的顶点移动来实现的.内点算法则完全不同,它是从可行域的一个严格内点开始,产生一个使目标函数值逐步改善的严格内点序列,并最终收敛于问题的最优解(见图 2.8).

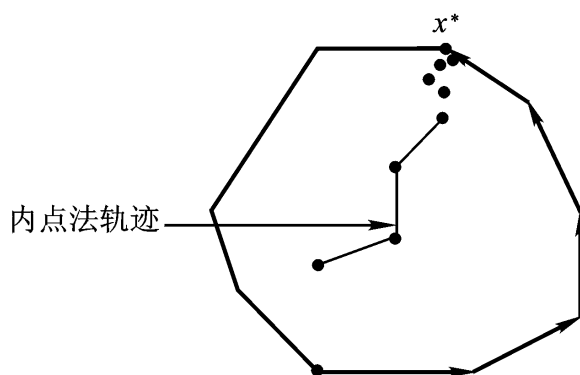


图 2.8 单纯形法与内点法轨迹

经过近 20 年的研究与发展,如今已有大量求解线性规划问题的内点算法,它们可以分成三类:路径跟踪算法,仿射调比算法,和原始对偶内点算法.我们在这里只介绍已得到广泛应用,又易于理解的原始对偶内点算法.

考虑标准型线性规划问题

$$\begin{aligned} \min \quad & c^T x, \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0, \end{aligned}$$

其相应的对偶问题可表示成

$$\begin{aligned} \min \quad & b^T y, \\ \text{s.t.} \quad & z = c - A^T y, \\ & z \geq 0, \end{aligned}$$

其中 $z \geq 0$ 是引入的松弛变量. 在这里我们假定 A 行满秩; 原始问题和对偶问题可行域的内部 $F_p = \{x / Ax = b, x > 0\}$, $F_d = \{y, z / z = c - A^T y, z > 0\}$ 都非空. 根

原始可行性与对偶可行性, 以及强对偶定理, 上述问题有最优解, 而且最优解 x^* 和 (y^*, z^*) 应满足下述方程组

$$\begin{aligned} Ax &= b, \\ z &= c - A^T y, \\ c^T x &= b^T y, \\ x &\geq 0, z \geq 0. \end{aligned}$$

由于 $c^T x - b^T y = c^T x - y^T Ax = (c - A^T y)^T x = x^T z$, 上述方程组可以重写成

$$\begin{aligned} Ax &= b, \\ z &= c - A^T y, \\ x^T z &= 0, \\ x &\geq 0, z \geq 0. \end{aligned} \tag{2.4.1}$$

这是一个非线性的方程组, 所谓原始对偶内点算法就是用大家熟知的牛顿法求上述方程组的解, 并要求所有的迭代点满足 $x > 0, z > 0$, 称这样的点为内点. 如记

$$x_i z_i = \mu, \quad (\mu > 0), \quad i = 1, 2, \dots, n,$$

则 x_i 和 z_i 两者都为正, 且在 $\mu \rightarrow 0$ 时, 有 $x_i z_i \rightarrow 0, i = 1, 2, \dots, n$, 即 $x^T z \rightarrow 0$. 因此, 对逐步减小并收敛于零的正数序列 μ , 求方程组

$$\begin{aligned} Ax &= b, \\ z &= c - A^T y, \\ x_i z_i &= \mu, \quad i = 1, 2, \dots, n, \\ x &\geq 0, z \geq 0 \end{aligned} \tag{2.4.2}$$

的解, 就可保持解的严格可行性, 并在 $\mu \rightarrow 0$ 时最终确定原问题的最优解 x^* 和相应的对偶问题的最优解 (y^*, z^*) .

在上述方程组中第一和第二个方程是线性的, 而第三个方程是非线性的, 因此需要用迭代法求解. 设 (x, y, z) 是方程组 (2.4.2) 解的一个近似, 迭代法要求确定修正量 dx, dy, dz 以得到方程组 (2.4.2) 解的一个更好的近似

$$x^+ = x + dx, \quad y^+ = y + dy, \quad z^+ = z + dz.$$

要求新的近似 x^+, y^+, z^+ 满足

$$\begin{aligned}
 A(x + dx) &= b, \\
 z + dz &= c - A^T(y + dy), \\
 (x_i + dx_i)(z_i + dz_i) &= \mu^+, \quad i = 1, 2, \dots, n, \\
 x + dx &> 0, \quad z + dz > 0,
 \end{aligned}$$

其中要求 $\mu^+ < \mu$. 考虑到 (x, y, z) 的可行性, 展开第三个方程并略去关于 dx_i 和 dz_i 的高次项可得上述关于 dx, dy 和 dz 的方程组的线性近似

$$\begin{aligned}
 Adx &= 0, \\
 dz + A^T dy &= 0, \\
 x_i dz_i + z_i dx_i &= \mu^+ - x_i z_i, \\
 x + dx &> 0, \quad z + dz > 0,
 \end{aligned} \tag{2.4.3}$$

这就是确定修正量 dx, dy, dz 的线性方程组. 引入对角矩阵

$$X = \text{diag}(x_1, x_2, \dots, x_n), \quad Z = \text{diag}(z_1, z_2, \dots, z_n)$$

和向量 $e = (1, 1, \dots, 1)^T$, 上述方程组可表示成

$$\begin{pmatrix} Z & O & X & dx \\ A & O & O & dy \\ O & A^T & I & dz \end{pmatrix} = \begin{pmatrix} (\mu^+ I - XZ)e \\ 0 \\ 0 \end{pmatrix}. \tag{2.4.4}$$

解上述方程组, 设所得解为 (dx, dy, dz) . 如果

$$\mu^+ = \mu, \quad 0 < \theta < 1,$$

则对于接近 1 的 θ 值, 如 $\theta = 1 - \theta/n, 0 < \theta < 1$, 可确保 $x + dx > 0, z + dz > 0$, 且由此迭代格式形成的算法的时间复杂性是多项式的. 但是这样确定 θ 值使得 μ 的减少很慢, 导致迭代次数的增加, 方法收敛慢. 因此 θ 的取值一般应能使 μ 较

快下降收敛于零. 由 $x_i z_i = \mu, i = 1, 2, \dots, n$, 有 $x^T z = \sum_{i=1}^n \mu = n\mu$. 随着迭代的进行, $x^T z > 0$ 的值逐渐收敛于零, 因而 μ 的另一种取法为

$$\mu = \frac{x^T z}{n}.$$

当 μ 的值以较快的速度下降时, $x + dx > 0, z + dz > 0$ 不再得到保证, 为确保迭代点的严格可行性, 取

$$x^+ = x + \theta dx, \quad y^+ = y + \theta dy, \quad z^+ = z + \theta dz,$$

其中步长 θ 的选取应在确保 $x^+ > 0, z^+ > 0$ 的前提下尽可能的大.

方程组 (2.4.3) 有 $2n + m$ 个方程, 含 $2n + m$ 个未知量, 为确定它的解, 由第三个方程得

$$dz = -A^T dy,$$

将其代入第一个方程有

$$Zdx - XA^T dy = (\mu^+ I - XZ)e,$$

用 AZ^{-1} 左乘这个方程的两边, 并利用第二个方程得

$$-AZ^{-1}XA^T dy = AZ^{-1}(\mu^+ I - XZ)e.$$

记 $Z^{-1}X = W$, 则由于 X, Z 是对角矩阵, W 也是对角元素全为正的对角矩阵. 由于矩阵 A 行满秩, 矩阵 AWA^T 非奇异, 因而由上述方程解得

$$dy = -(AWA^T)^{-1}AZ^{-1}(\mu^+ I - XZ)e,$$

据此得

$$dz = A^T dy = -A^T(AWA^T)^{-1}AZ^{-1}(\mu^+ I - XZ)e,$$

$$dx = Z^{-1}[(\mu^+ I - XZ)e - Xdz].$$

在 dx, dy, dz 确定后, 步长 $\theta > 0$ 的选取由

$$\theta = \max$$

确定, 其中 θ_{\max} 为使 $x + \theta dx \geq 0, z + \theta dz \geq 0$ 所容许的最大步长, 即

$$\theta_{\max} = \min\{\theta_P, \theta_D\},$$

$$\theta_P = \min_i \left\{ -\frac{x_i}{dx_i} \mid dx_i < 0 \right\}, \quad \theta_D = \min_i \left\{ -\frac{z_i}{dz_i} \mid dz_i < 0 \right\}.$$

根据步长的选取应在确保 $x^+ \geq 0, z^+ \geq 0$ 的前提下尽可能大的原则, θ 值应小于 1, 并尽可能接近 1, 如 0.99, 或 0.999 乃至更接近 1.

至此, 我们可以给出求线性规划问题最优解的原始对偶内点算法的迭代步骤如下:

算法 2.4.1 (原始对偶内点算法)

步 1. 给定参数值 $\mu_1 > 0$, 初始严格可行内点 $x^{(1)} > 0, y^{(1)}, z^{(1)} > 0$, 初始 $\mu_1 > 0$, 置 $k = 1$;

步 2. 如果 $(x^{(k)})^T z^{(k)} \leq \epsilon$, 停止迭代, $(x^{(k)}, y^{(k)}, z^{(k)})$ 即为满足要求的近似最优解;

步 3. 解方程组 (2.4.3) 得解 $dx^{(k)}, dy^{(k)}, dz^{(k)}$;

步 4. 计算 θ_P, θ_D , 确定最大步长 θ_{\max} , 计算 $x^{(k+1)}, y^{(k+1)}, z^{(k+1)}$;

步 5. 确定 $\mu_{k+1} < \mu_k$, 置 $k = k + 1$ 后转步 2.

对步长的选取也可以采用原始问题和对偶问题不同的步长, 即有

$$x^+ = x + \theta_P dx, \quad y^+ = y + \theta_D dy, \quad z^+ = z + \theta_D dz,$$

其中 θ_P 和 θ_D 的意义同上面一样. 这样做有助于提高算法的有效性和效率, 这尤其适用于下面要介绍的不可行内点算法.

上述算法要求用户提供一个可行的初始内点 (x, y, z) , 对大多数实际问题这是一个比较困难的问题. 对此, 我们在这里提供两种选择. 一种选择为通过引入人工变量, 形成具有明显严格初始可行内点的人工原始问题 and 人工对偶问题, 再用上述原始对偶内点算法来确定原来原始问题和对偶问题的最优解. 设 $x^{(0)} > 0, y^{(0)}, z^{(0)} > 0$ 为任意内点, 通常取 $x^{(0)} = z^{(0)} = e, y^{(0)} = 0$, 其中 e 同前面一

样是分量全为 1 的向量,则对于满足下述两不等式

$$> (b - Ax^{(0)})^T y^{(0)}, \quad > (A^T y^{(0)} + z^{(0)} - c)^T x^{(0)}$$

的参数 μ 和 ν , 下述人工原始问题

$$\begin{aligned} \min \quad & c^T x + x_{n+1}, \\ \text{s.t.} \quad & Ax + (b - Ax^{(0)}) x_{n+1} = b, \\ & (A^T y^{(0)} + z^{(0)} - c)^T x + x_{n+2} = \mu, \\ & x \geq 0, x_{n+1} \geq 0, x_{n+2} \geq 0, \end{aligned}$$

和人工对偶问题

$$\begin{aligned} \max \quad & b^T y + y_{m+1}, \\ \text{s.t.} \quad & A^T y + (A^T y^{(0)} + z^{(0)} - c) y_{m+1} + z = c, \\ & (b - Ax^{(0)})^T y + z_{n+1} = \nu, \\ & y_{m+1} + z_{n+2} = 0, \\ & z \geq 0, z_{n+1} \geq 0, z_{n+2} \geq 0, \end{aligned}$$

具有明显的严格可行内点, 其中 $x_{n+1}, x_{n+2}, y_{m+1}, z_{n+1}, z_{n+2}$ 为引入的人工变量.

这时可取 $(x^{(1)}, x_{n+1}^{(1)}, x_{n+2}^{(1)})$ 为人工原始问题的初始严格可行内点, 其中

$$x^{(1)} = x^{(0)}, x_{n+1}^{(1)} = 1, x_{n+2}^{(1)} = \mu - (A^T y^{(0)} + z^{(0)} - c)^T x^{(0)},$$

而取 $(y^{(1)}, y_{m+1}^{(1)}, z^{(1)}, z_{n+1}^{(1)}, z_{n+2}^{(1)})$ 为人工对偶问题的初始严格可行解, 其中

$$\begin{aligned} y^{(1)} &= y^{(0)}, y_{m+1}^{(1)} = -1, \\ z^{(1)} &= z^{(0)}, z_{n+1}^{(1)} = \nu - (b - Ax^{(0)})^T y^{(0)}, z_{n+2}^{(1)} = 1. \end{aligned}$$

以 $(x^{(1)}, x_{n+1}^{(1)}, x_{n+2}^{(1)})$ 和 $(y^{(1)}, y_{m+1}^{(1)}, z^{(1)}, z_{n+1}^{(1)}, z_{n+2}^{(1)})$ 为初始点用原始对偶内点算法求解, 设所得最优解为 $(x^*, x_{n+1}^*, x_{n+2}^*)$ 和 $(y^*, y_{m+1}^*, z^*, z_{n+1}^*, z_{n+2}^*)$, 如果 $x_{n+1}^* = 0, y_{m+1}^* = 0$, 则 x^* 和 (y^*, z^*) 分别是原来原始问题和对偶问题的最优解.

另一个方案是采用不可行的原始对偶内点算法, 该算法不要求初始点可行, 而只要求是严格内点, 也就是可以取任意 $x > 0, y, z > 0$ 作为初始点. 对一个不可行内点, 方法同样确定修正量 dx, dy 和 dz , 使得 $x + dx, y + dy, z + dz$ 满足

$$\begin{aligned} A(x + dx) &= b, \\ A^T(y + dy) + (z + dz) &= c, \\ (x_i + dx_i)(z_i + dz_i) &= \mu, i = 1, 2, \dots, n, \\ x + dx &> 0, z + dz > 0, \end{aligned} \tag{2.4.5}$$

再取

$$x^+ = x + \rho dx, y^+ = y + \rho dy, z^+ = z + \rho dz$$

作为新的迭代点, 其中步长 ρ 和 ρ 分别确保 $x + \rho dx > 0, z + \rho dz > 0$, 并尽可能增大新迭代点的可行性. 一旦某个迭代点成为可行的严格内点, 方法转而成为前面的原始对偶可行内点算法.

用同前面一样的表示方法,并略去高阶项,可把方程组(2.4.5)简化为

$$\begin{aligned} Zdx + Xdz &= (\mu I - XZ)e, \\ Adx &= b - Ax, \\ A^T dy + dz &= c - A^T y - z, \end{aligned} \quad (2.4.6)$$

将其用矩阵向量方式表示为

$$\begin{aligned} Z \quad 0 \quad X \quad dx &= (\mu I - XZ)e \\ A \quad 0 \quad 0 \quad dy &= r_p \\ 0 \quad A^T \quad I \quad dz &= r_d \end{aligned} \quad (2.4.7)$$

其中 $r_p = b - Ax$, $r_d = c - A^T y - z$ 分别称为原始残量和对偶残量,如果点 x 可行,则原始残量为零;同样如果点 (y, z) 可行,则对偶残量为零.用同前面一样的推导可得方程组的解为

$$\begin{aligned} dy &= (AWA^T)^{-1} [r_p + AWr_d - AZ^{-1}(\mu I - XZ)e], \\ dz &= r_d - A^T dy, \\ dx &= Z^{-1}[(\mu I - XZ)e - Xdz], \end{aligned}$$

其中 $W = Z^{-1}X$.

至于步长 p 的确定既依赖于对新点可行性的要求,又依赖于对新点严格正的要求.如果 $x + dx$ 能保持新点各分量的严格正性,则取步长 $p = 1$,这时新的迭代点为 $x^+ = x + p dx = x + dx$ 已经是一个严格的可行内点,而且其后的所有迭代点 x 都是严格的可行内点.如果 $x + dx > 0$ 不能成立,则步长 p 应是使 $x + dx > 0$ 成立的最大步长,即有 $p = p_{\min}$, $p_{\min} = \min_i -\frac{x_i}{dx_i} \mid dx_i < 0$. 因此步长 p 由下式确定,

$$p = \min\{1, p_{\min}\}, \quad p_{\min} = \min_i -\frac{x_i}{dx_i} \mid dx_i < 0, \quad .$$

同样可以得到对偶问题的步长 d :

$$d = \min\{1, d_{\min}\}, \quad d_{\min} = \min_i -\frac{z_i}{dz_i} \mid dz_i < 0, \quad .$$

从这里可以看出对原始问题和对偶问题采用不同的步长,可以分别尽快确定相应问题的严格可行内点.

线性规划问题的最优解必定在可行域的边界,而上述介绍的內点算法,不管求解的精度多高,所确定的只是一个接近最优解的严格內点.虽然我们可以直接取这样的—个內点作为最优解的一个近似,但是对于某些实际问题这样的—个解却没有什么实际意义,因此需要像单纯形方法那样确定一个最优的基本可行解.对此可利用最后所得迭代点提供的信息确定问题的—个最优基本可行解.下

列方法用于确定离迭代点 $x^{(k)}$ 最近的可行域的一个顶点. 记

$$X_k = \text{diag}(x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)}),$$

则已证明矩阵

$$D = X_k A^T (A X_k^2 A^T)^{-1} A X_k$$

的对角元素所提供的信息可用来确定离 $x^{(k)}$ 最近的基本可行解 x^* 的基变量. 具体说, 把上述矩阵 D 的对角元素记为向量

$$d = (d_1, d_2, \dots, d_n)^T, \quad d_i = d_{ii}, \quad i = 1, 2, \dots, n,$$

这里 d_{ii} 为矩阵 D 的第 i 个对角元素. 对于向量 d 的各个分量, 如果 d_i 的值接近 1, 则 x^* 的第 i 个变量是基变量, 如果 d_i 的值接近零, 则 x^* 的第 i 个变量是非基变量. 据此, 如果能从矩阵 D 的对角元素明确地确定 m 个基变量, 则通过解方程组 $Bx_B = b$ 得基变量的取值并设非基变量为零, 即得最优的基本可行解. 如果从向量 d 不能明确地确定 m 个基变量, 可以重置

$$X_k = \text{diag}(d_1, d_2, \dots, d_n),$$

再计算上述矩阵 D . 重复这一过程, 直至所得的向量 d 能明确地确定 m 个基变量.

最后简单提一下内点算法的计算复杂性问题. 自 1984 年 Karmarkar 的内点算法问世以来, 已经有大量的内点算法面世, 算法的计算复杂性大都是多项式时间的. 对于上述的原始对偶内点算法, 其运算工作量明显依赖于减小 μ 的具体方法. 对 dx, dy, dz 的计算方法. 初期的原始对偶内点方法一般需要 $O(nL)$ 阶的迭代次数, 每次迭代的算术运算为 $O(n^3)$ 阶, 因而其总的运算工作量是 $O(n^4 L)$ 阶, 其中 L 表示所论问题的输入规模, 它同问题的维数 n , 约束的个数 m , 以及问题中每个数据所占用的字节数有关. 经过对方法的不断改进, 近期的算法运算工作量已降为 $O(n^3 L)$ 阶. $O(nL)$ 阶的迭代次数, 每次 $O(n^{2.5})$ 阶的运算工作量.

小 结

第二章用于介绍在现实生活, 科学管理和科学实验中最常见也最有用的最优化技术和方法——线性规划, 全章共分四节分别介绍了线性规划的基本性质, 目前广泛使用的求解线性规划的方法: 单纯形法、对偶单纯形法和内点算法.

第 2.1 节给出了线性规划的标准型式, 以及它同各种不同形式的线性规划问题之间的关系和相互转换, 并就一个简单的线性规划问题通过直观的几何解释, 来引入线性规划问题的基本特性: 线性规划如果有最优解, 则必可在其可行

域的某个顶点上确定。据此,把线性规划基本性质的研究集中于可行域顶点性质和顶点代数表示的分析。由此引入了线性规划的基本解和基本可行解的概念和它们的代数表示,给出了基本可行解和可行域顶点的等价性,从而把求最优的顶点转化为求最优的基本可行解,最后从理论上严密地证明了线性规划如果有最优解,则必可在其可行域的某个顶点上确定这一基本性质。

第 2.2 节介绍线性规划最基本的也是使用最广泛的方法——单纯形方法。单纯形方法根据线性规划如果有最优解,则必可在其可行域的某个顶点上确定这一基本特性,在线性规划的基本可行解中确定最优解。方法从一个初始基本可行解开始,每次迭代产生一个新的使目标函数值得到改进的基本可行解,从而产生一个基本可行解的序列,对每一个作为基本可行解的迭代点,需要对其是否为一个最优解作出判定。如为最优解,停止进一步的迭代,否则确定另一更好的基本可行解,只要问题有有界的最优解,方法必经有限次迭代终止于最优解。为此介绍了同方法密切相关的基变量和非基变量的概念,用于判别最优性的简约价值系数,以及在确定新基本可行解时用到的入基变量和出基变量的概念和确定它们的方法,给出了单纯形方法的具体步骤,以及求解过程中使用的表格形式,并通过例子给出了用单纯形表求解的全过程。在本节的最后讲述了单纯形法同计算机实现有关的三个问题和解决办法。(1) 基矩阵的逆矩阵的计算,(2) 初始基本可行解的确定以及(3) 由于问题退化,迭代出现循环不能终止的问题。对问题(1)我们介绍了采用矩阵乘积表示的修正单纯形法,对于问题(2)我们介绍了两个常用的方法:a) 两阶段方法和 b) 大 M 方法;对于问题(3)的解决则可以采用 Bland 方法。

第 2.3 节论述线性规划问题的对偶理论和对偶单纯形方法,给出了标准型线性规划问题的对偶问题的形式,并由此给出了一般形式线性规划问题的对偶问题的形式,分析讨论了由原规划问题确定对偶问题的一般对偶关系和原则,给出了描述原问题和对偶问题的最优解之间关系的弱对偶定理和强对偶定理。根据对偶理论,线性规划问题可以用单纯形方法求解,也可以用对偶单纯形方法求解。如果说单纯形方法是在基本可行解中寻找满足最优性条件(简约价值系数非负)的最优解,对偶单纯形方法则是在所有满足最优性条件(简约价值系数非负)的基本解中寻找满足可行的最优解。对偶单纯形法从一个简约价值系数非负的基本解开始,产生一个简约价值系数全为非负的基本解序列。如果这样的一个基本解还是可行的,那么它就是问题的最优解。同样我们通过一个例子给出了对偶单纯形法的过程。在本节的最后简单提了一下有关线性规划的灵敏度分析的问题。

第 2.4 节用于介绍 20 世纪 80 年代后期发展起来,用于求解大规模线性规划问题的内点算法。由于有大量的求解线性规划的内点算法存在,我们在这一节只给出了应用比较广泛,效果相对好,又易于理解的原始对偶内点算法。该算法

通过用牛顿法求解由原始线性规划及其对偶规划所得到的原始对偶方程组,在可行域的内部产生一个收敛于最优解的严格可行内点序列.我们分析了迭代序列产生的基本原理和过程,给出了算法的步骤,介绍了确定初始可行内点的方法和算法的计算复杂性.



1. 用图解法确定下述线性规划问题的最优解.

$$\begin{aligned}
 & \min \quad 8x_1 + 5x_2, \\
 & \text{s.t.} \quad -x_1 + x_2 = 0, \\
 (1) \quad & 6x_1 + 11x_2 = 66, \\
 & 2x_1 + x_2 = 10, \\
 & x_1 = 0, x_2 = 0,
 \end{aligned}$$

$$\begin{aligned}
 & \max \quad 2x_1 + 3x_2, \\
 & \text{s.t.} \quad -x_1 + 2x_2 = 12, \\
 (2) \quad & -2x_1 + x_2 = 3, \\
 & x_1 = 0, x_2 = 0,
 \end{aligned}$$

$$\begin{aligned}
 & \min \quad 4x_1 - 3x_2, \\
 (3) \quad & \text{s.t.} \quad -11x_1 + 10x_2 = 20, \\
 & x_1 = 0, 0 \leq x_2 \leq 6,
 \end{aligned}$$

2. 确定由下述约束条件所形成的可行域的所有顶点,并计算函数 $f(x)$ 在这些顶点的函数值.

$$\begin{aligned}
 (1) \quad & f(x) = 3x_1 + 2x_2, \\
 & x_1 + x_2 = 3, \\
 & x_1 - 2x_2 = 4, \\
 & 0 \leq x_1 \leq 4, \\
 & 0 \leq x_2 \leq 1.
 \end{aligned}$$

$$\begin{aligned}
 (2) \quad & f(x) = -4x_1 - 3x_2, \\
 & 2x_1 + x_2 = 10, \\
 & x_1 + x_2 = 10, \\
 & 5x_1 + 3x_2 = 20, \\
 & x_1 = 0, x_2 = 0.
 \end{aligned}$$

3. 将下列线性规划问题转化成标准型.

$$\begin{aligned}
 & \max \quad 2x_1 - x_2 - 3x_3, \\
 & \text{s.t.} \quad 2x_1 - x_2 + 2x_3 = 2, \\
 & \quad \quad -x_1 + 2x_2 - 3x_3 = -2, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \\
 & \min \quad 2x_1 - x_2 - 3x_3 - x_4, \\
 & \text{s.t.} \quad x_1 + 3x_2 + 2x_3 + x_4 = 15, \\
 & \quad \quad 2x_2 + x_3 + x_4 = 10, \\
 & \quad \quad -x_1 + 5x_2 - 2x_3 + 2x_4 = 3, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0,
 \end{aligned}$$

$$\begin{aligned}
 & \max \quad 3x_1 - 13x_2, \\
 & \text{s.t.} \quad 5x_1 + 3x_2 = 18, \\
 & \quad \quad x_1 + 2x_2 = 9, \\
 & \quad \quad 2x_1 + 5x_2 = 3, \\
 & \quad \quad 2x_1 + x_2 = 2.5, \\
 & \quad \quad x_1 \geq 0,
 \end{aligned}$$

4. 用单纯形表的方法求下列线性规划问题的最优解.

$$\begin{aligned}
 & \max \quad 5x_1 + 6x_2 + 4x_3, \\
 & \text{s.t.} \quad 2x_1 + 2x_2 = 5, \\
 & \quad \quad 5x_1 + 3x_2 + 4x_3 = 15, \\
 & \quad \quad x_1 + x_2 = 10, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \\
 & \min \quad 3x_1 - 2x_2 + x_3, \\
 & \text{s.t.} \quad 2x_1 + 2x_2 - x_3 = 2, \\
 & \quad \quad -3x_1 - x_2 + 2x_3 = -2, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, \\
 & \max \quad 4x_1 + 2x_2 + 10x_3 + x_4 + 2x_5, \\
 & \text{s.t.} \quad 5x_1 + 3x_2 + 2x_3 + x_4 - x_5 = 32, \\
 & \quad \quad x_1 + 2x_2 + 3x_3 + 4x_4 + 4x_5 = 10, \\
 & \quad \quad x_i \geq 0, i=1,2,3,4,5.
 \end{aligned}$$

5. 用修正单纯形方法求下列问题的解.

$$\begin{aligned}
 & \min \quad -3x_1 + 2x_2 - x_3 - x_4, \\
 & \text{s.t.} \quad 2x_1 + x_2 + x_3 + 3x_4 = 20, \\
 & \quad \quad x_1 + x_3 + 2x_4 = 10, \\
 & \quad \quad -2x_1 - x_2 + 2x_3 + 5x_4 = 3, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0,
 \end{aligned}$$

$$\begin{aligned}
 & \max \quad 2x_1 - x_2 - 3x_3, \\
 & \text{s.t.} \quad 2x_1 - x_2 + 2x_3 = 2, \\
 & \quad \quad -x_1 + 2x_2 - 3x_3 = -2, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0,
 \end{aligned}
 \tag{2}$$

6. 考察下列线性规划问题

$$\begin{aligned}
 & \min \quad -\frac{3}{4}x_1 + 150x_2 - \frac{1}{50}x_3 + 6x_4, \\
 & \text{s.t.} \quad \frac{1}{4}x_1 - 60x_2 - \frac{1}{25}x_3 + 9x_4 = 0, \\
 & \quad \quad \frac{1}{2}x_1 - 90x_2 - \frac{1}{50}x_3 + 3x_4 = 0, \\
 & \quad \quad x_3 = 1, \\
 & \quad \quad x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0,
 \end{aligned}$$

(1) 引入松弛变量将其化成标准型后用单纯形方法进行迭代看是否出现循环的现象(在确定出基变量时如有多于一个指标使 $b_j^{(k)} / \alpha_{jp}^{(k)}$ 取得最小值,取其中最先出现的变量作为出基变量)。

(2) 把 Bland 方法结合进单纯形方法进行单纯形迭代看能否求得问题的最优解。

7. 给出下列线性规划问题的对偶

$$\begin{aligned}
 & \max \quad x_1, \\
 & \text{s.t.} \quad -4x_1 + 3x_3 = 0, \\
 & \quad \quad 2x_2 + 5x_3 = 0, \\
 & \quad \quad 6x_1 + 7x_2 = 0, \\
 & \quad \quad x_2 \geq 0, x_3 \geq 0, \\
 & \min \quad 2x_1 - 5x_2 + 3x_3 - 6x_4, \\
 & \text{s.t.} \quad -8x_1 + 7x_2 + 3x_3 + 4x_4 = 11, \\
 & \quad \quad 7x_1 + 6x_2 + 2x_3 + 3x_4 = 23, \\
 & \quad \quad 3x_1 + 2x_2 + 4x_3 + 7x_4 = 12, \\
 & \quad \quad x_1 \geq 0, x_3 \geq 0, x_4 \geq 0, \\
 & \max \quad c^T x, \\
 & \text{(3) s.t.} \quad Ax = b, \\
 & \quad \quad l \leq x \leq u,
 \end{aligned}$$

其中 l 和 u 是 x 的上下界向量。

$$\begin{aligned}
 & \min \quad c^T x, \\
 & \text{(4) s.t.} \quad b \leq Ax \leq \bar{b}, \\
 & \quad \quad x \geq 0,
 \end{aligned}$$

其中 b 和 \bar{b} 是 $Ax \in \mathbf{R}^m$ 的上下界向量。

8. 用对偶单纯形方法求下列线性规划问题的解。

$$\begin{aligned}
 & \min \quad 9x_1 + 5x_2 + 3x_3, \\
 & \text{s.t.} \quad 3x_1 + 2x_2 - 3x_3 = 3, \\
 & \quad \quad 2x_1 + \quad \quad x_3 = 5, \\
 & \quad \quad x_1 = 0, x_2 = 0, x_3 = 0,
 \end{aligned}
 \tag{1}$$

$$\begin{aligned}
 & \min \quad 3x_1 + 2x_2 + 4x_3, \\
 & \text{s.t.} \quad 2x_1 - x_2 = 5, \\
 & \quad \quad 2x_2 - x_3 = 10, \\
 & \quad \quad x_1 = 0, x_2 = 0, x_3 = 0,
 \end{aligned}
 \tag{2}$$

$$\begin{aligned}
 & \max \quad -6x_1 - 5x_2 - 7x_3 - 2x_4, \\
 & \text{s.t.} \quad -3x_1 - 2x_2 + 5x_3 + 4x_4 = 15, \\
 & \quad \quad 6x_1 - 2x_2 + 2x_3 + 7x_4 = 35, \\
 & \quad \quad -5x_1 - 4x_2 - 3x_3 + 2x_4 = 20, \\
 & \quad \quad x_1 = 0, x_2 = 0, x_3 = 0, x_4 = 0,
 \end{aligned}
 \tag{3}$$

9. 将 Klee - Minty 问题转换成标准型, 并对 $m=3$ 的情况用单纯形方法求解(要求取最负的价值系数的变量为入基变量). 如果在第一次迭代取第三个约束的松弛变量为入基变量, x_3 为入基变量, 结果又如何?

10. (1) 设 $\mu_{k+1} = 1 - \frac{1}{n} \mu_k, 0 < \mu_k < 1$ 给定 $\mu_0 > 0, n > 0$ 确定使 μ_k 成立所需要的迭代次数.

(2) 估计原始对偶内点算法的每次迭代的运算工作量.

11. 用原始对偶内点算法求下述线性规划问题的解

$$\begin{aligned}
 & \max \quad x_1 + 3x_2, \\
 & \text{s.t.} \quad x_1 + 4x_2 = 10, \\
 & \quad \quad x_1 + 2x_2 = 1, \\
 & \quad \quad x_1 = 0, x_2 = 0,
 \end{aligned}$$

第 3 章

线性搜索与信赖域方法

§ 3.1 线性搜索

如前所述,线性搜索是多变量函数最优化方法的基础.在多变量函数最优化中,迭代格式为

$$x_{k+1} = x_k + \alpha_k d_k, \quad (3.1.1)$$

其关键是构造搜索方向 d_k 和步长因子 α_k . 设

$$\phi(\alpha) = f(x_k + \alpha d_k). \quad (3.1.2)$$

从 x_k 出发,沿搜索方向 d_k ,确定步长因子 α_k ,使

$$\phi(\alpha_k) < \phi(0)$$

的问题就是关于 ϕ 的线性搜索问题.

理想的方法是使目标函数沿方向 d_k 达到极小,即使得

$$f(x_k + \alpha_k d_k) = \min_{\alpha > 0} f(x_k + \alpha d_k), \quad (3.1.3)$$

或者选取 $\alpha_k > 0$ 使得

$$\alpha_k = \min\{\alpha > 0 \mid f(x_k + \alpha d_k)^T d_k = 0\}. \quad (3.1.4)$$

这样的线性搜索称为精确线性搜索,所得到的 α_k 叫精确步长因子.我们在 § 3.2 和 § 3.3 分别介绍利用直接搜索法和插值法求精确步长因子 α_k .但是,一般地,精确线性搜索需要的计算量很大,而且在实际上也不必要.特别,对某些非光滑函数或导数表达式复杂的函数不能利用精确线性搜索.此外,在计算实践中人们也发现过分追求线性搜索的精度反而会降低整个方法的效率,因此人们提出了既花费较少的计算量,又能达到足够下降的不精确线性搜索方法.这些方法将在 § 3.5 介绍.

一般地,线性搜索算法分成两个阶段.第一阶段确定包含理想的步长因子(或问题最优解)的搜索区间;第二阶段采用某种分割技术或插值方法缩小这个区间.

确定初始搜索区间的一种简单方法叫进退法,其基本思想是从一点出发,按一定步长,试图确定出函数值呈现“高-低-高”的三点,即 (a) (c) (b) ,这

里 $a < c < b$. 具体地说, 就是给出初始点 $x_0 > 0$, 初始步长 $h_0 > 0$, 若

$$f(x_0 + h_0) < f(x_0),$$

则下一步从新点 $x_1 = x_0 + h_0$ 出发, 加大步长, 再向前搜索, 直到目标函数上升为止. 若

$$f(x_0 + h_0) > f(x_0),$$

则下一步仍以 x_0 为出发点, 沿反方向同样搜索, 直到目标函数上升就停止. 这样便得到一个搜索区间. 这种方法叫进退法.

利用 $f(x)$ 的导数, 也可以类似地确定搜索区间. 我们知道, 在包含极小点 x^* 的区间 $[a, b]$ 的端点处, $f'(a) < 0$, $f'(b) > 0$. 给定步长 $h > 0$, 取初始点 $x_0 = a$. 若 $f'(x_0) > 0$, 则取 $x_1 = x_0 + h$; 若 $f'(x_0) < 0$, 则取 $x_1 = x_0 - h$. 其余过程与上述方法类似.

算法 3.1.1 (进退法步骤)

- 步 1. 选取初始数据 $x_0 \in [0, \infty)$, $h_0 > 0$, 加倍系数 $t > 1$ (一般取 $t = 2$), 计算 $f(x_0)$, $k = 0$.
- 步 2. 比较目标函数值. 令 $x_{k+1} = x_k + h_k$, 计算 $f_{k+1} = f(x_{k+1})$, 若 $f_{k+1} < f_k$, 转步 3, 否则转步 4.
- 步 3. 加大搜索步长, 令 $h_{k+1} = th_k$, $x_{k+1} = x_k + h_{k+1}$, $f_{k+1} = f(x_{k+1})$, $k = k + 1$, 转步 2.
- 步 4. 反向搜索. 若 $k = 0$, 转换搜索方向, 令 $h_k = -h_k$, $x_{k+1} = x_k + h_{k+1}$, 转步 2; 否则, 停止迭代, 令

$$a = \min\{x_0, x_{k+1}\}, b = \max\{x_0, x_{k+1}\},$$

输出 $[a, b]$, 停止.

线性搜索方法根据是否采用导数信息分为无导数方法和导数方法. 由于没有利用导数信息, 无导数方法一般没有导数方法有效. 下面我们首先介绍典型的无导数方法 0.618 法和 Fibonacci 法, 然后介绍采用导数信息的插值方法, 最后讨论不精确线性搜索方法.

§ 3.2 0.618 法和 Fibonacci 法

0.618 法和 Fibonacci (斐波那契) 法都是分割方法. 其基本思想是通过取试探点和进行函数值的比较, 使包含极小点的搜索区间不断缩短, 当区间长度缩短到一定程度时, 区间上各点的函数值均接近极小值, 从而各点可以看作为极小点的近似. 这类方法仅需计算函数值, 不涉及导数, 又称直接法. 它们用途很广, 尤其适用于非光滑及导数表达式复杂或写不出的种种情形. 注意, 这些方法要求所

考虑区间上的目标函数是单峰函数, 如果这个条件不满足, 我们可以把所考虑的区间分成若干个小区间, 在每个小区间上函数是单峰的. 这样, 我们在每个小区间上求极小点, 然后选取其中的最小点.

3.2.1 0.618 法

设包含极小点 x^* 的初始搜索区间为 $[a, b]$, 设

$$f(x_k + d_k) \quad (3.2.1)$$

在 $[a, b]$ 上是凸函数. 0.618 法的基本思想是在搜索区间 $[a, b]$ 上选取两个对称点 x_k 和 μ_k , 且 $x_k < \mu_k$, 通过比较这两点处的函数值 $f(x_k)$ 和 $f(\mu_k)$ 的大小来决定删除左半区间 $[a, x_k]$, 还是删除右半区间 $(\mu_k, b]$. 删除后的新区间长度是原区间长度的0.618倍. 新区间包含原区间中两个对称点中的一点, 我们只要再选一个对称点, 并利用这两个新对称点处的函数值继续比较. 重复这个过程, 最后确定出极小点 x^* .

记 $a_0 = a, b_0 = b$, 区间 $[a_0, b_0]$ 经 k 次缩短后变为 $[a_k, b_k]$. 我们要求两个试探点 x_k 和 μ_k 满足下列条件:

$$b_k - x_k = \mu_k - a_k, \quad (3.2.2)$$

$$b_{k+1} - a_{k+1} = (b_k - a_k) \quad (3.2.3)$$

由(3.2.2)和(3.2.3)得到

$$x_k = a_k + (1 - \tau)(b_k - a_k), \quad (3.2.4)$$

$$\mu_k = a_k + \tau(b_k - a_k). \quad (3.2.5)$$

计算 $f(x_k)$ 和 $f(\mu_k)$. 如果 $f(x_k) < f(\mu_k)$, 则删掉右半区间 $(\mu_k, b_k]$, 保留 $[a_k, \mu_k]$, 从而新的搜索区间为

$$[a_{k+1}, b_{k+1}] = [a_k, \mu_k]. \quad (3.2.6)$$

为进一步缩短区间, 需取试探点 x_{k+1}, μ_{k+1} . 由(3.2.5),

$$\begin{aligned} \mu_{k+1} &= a_{k+1} + \tau(b_{k+1} - a_{k+1}) \\ &= a_k + \tau(\mu_k - a_k) \\ &= a_k + \tau(a_k + \tau(b_k - a_k) - a_k) \\ &= a_k + \tau^2(b_k - a_k). \end{aligned} \quad (3.2.7)$$

若令

$$\tau^2 = 1 - \tau, \quad (3.2.8)$$

则

$$\mu_{k+1} = a_k + (1 - \tau)(b_k - a_k) = x_k. \quad (3.2.9)$$

这样, 新的试探点 μ_{k+1} 不需要重新计算, 只要取 x_k 就行了, 从而在每次迭代中

(第一次迭代除外), 只需选取一个试探点即可.

类似地, 在 $(f_k) > (f_{\mu_k})$ 的情形, 新的试探点 $x_{k+1} = \mu_k$, 它也不需要重新计算. 在这种情形, 我们删去左半区间 $[a_k, x_k]$, 保留 $[x_k, b_k]$, 这时新的搜索区间为

$$[a_{k+1}, b_{k+1}] = [x_k, b_k].$$

令

$$x_k = \mu_k,$$

$$\mu_{k+1} = a_{k+1} + (b_{k+1} - a_{k+1}).$$

然后再比较 (f_{k+1}) 和 $(f_{\mu_{k+1}})$ 重复上述过程, 直到 $b_{k+1} - a_{k+1} < \epsilon$.

搜索区间长度缩短率 究竟是多少呢? 解方程(3.2.8)立得

$$= \frac{-1 \pm \sqrt{5}}{2}.$$

由于 $\epsilon > 0$, 故取

$$= \frac{5 - 1}{2} = 0.618.$$

这样(3.2.4)和(3.2.5)可分别写成

$$x_k = a_k + 0.382(b_k - a_k).$$

$$\mu_k = a_k + 0.618(b_k - a_k).$$

算法 3.2.1 (0.618 法计算步骤)

步 1. 选取初始数据. 确定初始搜索区间 $[a, b]$ 和精度要求 $\epsilon > 0$. 计算最初两个试探点 x_1, μ_1 ,

$$x_1 = a_1 + 0.382(b_1 - a_1),$$

$$\mu_1 = a_1 + 0.618(b_1 - a_1).$$

计算 (f_{x_1}) 和 (f_{μ_1}) , 令 $k=1$.

步 2. 比较目标函数值. 若 $(f_k) > (f_{\mu_k})$, 转步 3; 否则转步 4.

步 3. 若 $b_k - x_k < \epsilon$, 则停止计算, 输出 μ_k ; 否则, 令

$$a_{k+1} = x_k, b_{k+1} = b_k, x_{k+1} = \mu_k,$$

$$(f_{k+1}) = (f_{\mu_k}), \mu_{k+1} = a_{k+1} + 0.618(b_{k+1} - a_{k+1}).$$

计算 $(f_{\mu_{k+1}})$, 转步 2.

步 4. 若 $\mu_k - a_k < \epsilon$, 则停止计算, 输出 x_k ; 否则, 令

$$a_{k+1} = a_k, b_{k+1} = \mu_k, \mu_{k+1} = x_k,$$

$$(f_{\mu_{k+1}}) = (f_{x_k}), x_{k+1} = a_{k+1} + 0.382(b_{k+1} - a_{k+1}).$$

计算 $(f_{x_{k+1}})$, 转步 2.

3.2.2 Fibonacci 法

另一种与 0.618 法相类似的分割方法叫 Fibonacci 法. 它与 0.618 法的主要

区别之一在于:搜索区间长度的缩短率不是采用 0.618 而是采用 Fibonacci 数. Fibonacci 数列满足

$$\begin{aligned} F_0 &= F_1 = 1, \\ F_{k+1} &= F_k + F_{k-1}, k=1, 2, \dots \end{aligned} \quad (3.2.10)$$

Fibonacci 法中的计算公式为

$$\begin{aligned} \mu_k &= a_k + \frac{1}{F_{n-k+1}} (b_k - a_k) \\ &= a_k + \frac{F_{n-k-1}}{F_{n-k+1}} (b_k - a_k), k=1, \dots, n-1. \end{aligned} \quad (3.2.11)$$

$$\mu_k = a_k + \frac{F_{n-k}}{F_{n-k+1}} (b_k - a_k), k=1, \dots, n-1. \quad (3.2.12)$$

显然,这里 $\frac{F_{n-k}}{F_{n-k+1}}$ 相当于黄金分割法(3.2.4)-(3.2.5)中的 $\frac{1}{2}$, 每次缩短率满足

$$b_{k+1} - a_{k+1} = \frac{F_{n-k}}{F_{n-k+1}} (b_k - a_k). \quad (3.2.13)$$

这里 n 是计算函数值的次数,即要求经过 n 次计算函数值后,最后区间的长度不超过 ϵ , 即

$$b_n - a_n \leq \epsilon.$$

由于

$$\begin{aligned} b_n - a_n &= \frac{F_1}{F_2} (b_{n-1} - a_{n-1}) \\ &= \frac{F_1}{F_2} \cdot \frac{F_2}{F_3} \cdots \frac{F_{n-1}}{F_n} (b_1 - a_1) \\ &= \frac{1}{F_n} (b_1 - a_1), \end{aligned} \quad (3.2.14)$$

故有

$$\frac{1}{F_n} (b_1 - a_1) \leq \epsilon,$$

从而

$$F_n \geq \frac{b_1 - a_1}{\epsilon}. \quad (3.2.15)$$

给出最终区间长度的上界 ϵ , 由(3.2.15)求出 Fibonacci 数 F_n , 再根据 F_n 确定出 n , 从而搜索一直进行到第 n 个搜索点为止.

Fibonacci 算法与 0.618 法几乎完全相同, 请读者自己写出算法和程序. 可以证明

$$\lim_{k \rightarrow \infty} \frac{F_{k-1}}{F_k} = \frac{5-1}{2} = \frac{1}{2}. \quad (3.2.16)$$

这表明,当 n 时, Fibonacci 法与 0.618 法的区间缩短率相同,因而 Fibonacci 法也以收敛比 线性收敛. Fibonacci 法是分割方法求一维极小化问题的最优策略,而 0.618 法是近似最优的,但由于 0.618 法简单易行,因而得到广泛应用.

3.2.3 二分法

本节最后,我们稍微提一下一种最简单的分割方法——二分法,其基本思想是通过计算函数导数值来缩短搜索区间.设初始区间为 $[a_1, b]$,第 k 步时的搜索区间为 $[a_k, b_k]$,满足

$$(a_k) \leq 0, (b_k) \geq 0.$$

取中点 $\alpha = \frac{1}{2}(a_k + b_k)$,若 $(\alpha) \leq 0$,则令 $a_{k+1} = a_k, b_{k+1} = \alpha$;若 $(\alpha) \geq 0$,则令 $a_{k+1} = \alpha, b_{k+1} = b_k$,从而得到新的搜索区间 $[a_{k+1}, b_{k+1}]$.依此进行,直到搜索区间的长度小于预定的容限为止.二分法每次迭代都将区间缩短一半,故二分法的收敛速度也是线性的,收敛比为 $\frac{1}{2}$.

§ 3.3 逐次插值逼近法

插值法是一类重要的线性搜索方法,其基本思想是在搜索区间中不断用低次(通常不超过三次)插值多项式来近似目标函数,并逐步用插值多项式的极小点来逼近 $()$ 的极小点.当函数具有比较好的解析性质时,插值方法比直接方法(如 0.618 法或 Fibonacci 法)效果更好.在本节中,我们首先介绍三点二次插值法.然后介绍两种二点二次插值法.

现在我们先介绍采用三点二次插值的线性搜索方法.

在包含 $()$ 的极小点 * 的搜索区间 $[a, b]$ 中,给定三点 x_1, x_2, x_3 满足

$$x_1 < x_2 < x_3, \quad (3.3.1)$$

$$(x_1) > (x_2) < (x_3). \quad (3.3.2)$$

利用三点处的函数值 $(x_1), (x_2), (x_3)$ 构造二次函数,并要求插值条件满足

$$\begin{aligned} q(x_1) &= a x_1^2 + b x_1 + c = (x_1), \\ q(x_2) &= a x_2^2 + b x_2 + c = (x_2), \\ q(x_3) &= a x_3^2 + b x_3 + c = (x_3). \end{aligned} \quad (3.3.3)$$

令 $x_i = (x_i), i=1,2,3$.解上述方程组得

$$a = - \frac{(x_2 - x_3)x_1 + (x_3 - x_1)x_2 + (x_1 - x_2)x_3}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)},$$

$$b = \frac{\left(\frac{x_2}{2} - \frac{x_3}{3}\right)x_1 + \left(\frac{x_3}{3} - \frac{x_1}{1}\right)x_2 + \left(\frac{x_1}{1} - \frac{x_2}{2}\right)x_3}{(x_1 - x_2)(x_2 - x_3)(x_3 - x_1)}.$$

于是,二次函数 $q(x)$ 的极小点为

$$x^- = -\frac{b}{2a} = \frac{1}{2} \frac{\left(\frac{x_2}{2} - \frac{x_3}{3}\right)x_1 + \left(\frac{x_3}{3} - \frac{x_1}{1}\right)x_2 + \left(\frac{x_1}{1} - \frac{x_2}{2}\right)x_3}{(x_2 - x_3)x_1 + (x_3 - x_1)x_2 + (x_1 - x_2)x_3}. \quad (3.3.4)$$

上式称为三点二次插值公式. 这个公式也可以直接利用二次拉格朗日插值公式

$$L(x) = \frac{(x - x_2)(x - x_3)}{(x_1 - x_2)(x_1 - x_3)}x_1 + \frac{(x - x_1)(x - x_3)}{(x_2 - x_1)(x_2 - x_3)}x_2 + \frac{(x - x_1)(x - x_2)}{(x_3 - x_1)(x_3 - x_2)}x_3, \quad (3.3.5)$$

并令 $L(x) = 0$ 得到

求得 x^- 以后, 如果

$$|x_2 - x^-| \geq \frac{1}{2}|x_1 - x_2|, \text{ 当 } |x_2| > |x_1| \text{ 时,}$$

或者如果

$$|x_2 - x^-| \geq |x_1|, \text{ 当 } |x_2| < |x_1| \text{ 时,}$$

则我们认为收敛准则满足. 如果 $x^- < x_2$, 则极小点估计为 x^- , 否则为 x_2 . 这里通常取 $\epsilon_1 = 10^{-3}$, $\epsilon_2 = 10^{-5}$.

若终止准则不满足, 则利用 x^- 提供的信息, 从 x_1, x_2, x_3 和 x^- 中选出相邻的三个点, 将原来的搜索区间例如 $[x_1, x_3]$ 缩小. 下面我们给出选取新的三点的方法.

算法 3.3.1

1) 比较 x^- 和 x_2 的大小, 如果 $x^- > x_2$, 则转 2); 否则转 3).

2) 如果 $x^- < x_2$, 则

$$x_1 = x_2, x_2 = x^-, x_3 = x_2, x_2 = x^-,$$

转 4); 否则, $x_3 = x^-, x_3 = x^-, \text{ 转 4) }.$

3) 若 $x^- < x_2$, 则

$$x_3 = x_2, x_2 = x^-, x_3 = x_2, x_2 = x^-,$$

转 4); 否则, $x_1 = x^-, x_1 = x^-, \text{ 转 4) }.$

4) 区间收缩完毕, 然后在新的搜索区间 $[x_1, x_3]$ 上按公式 (3.3.4) 计算二次插值函数的极小点 x^- .

可以证明: 如果 $f(x)$ 四阶连续可微, x^* 满足 $f'(x^*) = 0$, $f''(x^*) > 0$, 则三点二次插值法 (3.3.4) 产生的序列 $\{x_k\}$ 的收敛阶约为 1.32.

例 3.3.2 用二次插值法求

$$\min_{t \in [0, 3]} f(t) = t^3 - 3t + 2$$

的近似最优解 (精确极小点为 $t^* = 1$). 设已确定其初始搜索区间为 $[0, 3]$, 取初始插值点 $t_0 = 2$, 终止误差 $\epsilon = 0.05$.

解 $a = a = 0, b = b = 3$.

第一次迭代:

$$(a) = 2, (b) = 20, (t_0) = 0.$$

代入公式(3.3.4)求得

$$t_1 = 0.9.$$

由于

$$(f_1) = 0.029, (f_0) = 4,$$

并且 $|f_1 - f_0| = 1.1 > \epsilon$, 故要继续迭代. 这时迭代点 t_1 比插值内点 t_0 好, 又 $t_1 = 0.9$, $t_0 = 2$, 令

$$a_1 = a_0 = 0, b_1 = b_0 = 2, t_1 = t_1 = 0.9.$$

第二次迭代:

$$(a) = 2, (b) = 4, (t_1) = 0.029.$$

代入公式(3.3.4)求得

$$t_2 = 0.82759.$$

由于

$$(f_2) = 0.08405, (f_1) = 0.029,$$

并且 $|f_2 - f_1| = 0.07241 > \epsilon$, 要继续迭代. 因 $t_2 = 0.82759$, $t_1 = 0.9$, 故令

$$a_2 = t_2 = 0.82759, b_2 = b_1 = 2, t_2 = t_1 = 0.9.$$

第三次迭代:

$$(a) = 0.08405, (b) = 4, (t_2) = 0.029.$$

代入公式(3.3.4)求得

$$t_3 = 0.96577.$$

由于

$$(f_3) = 0.00347, (f_2) = 0.82759,$$

并且 $|f_3 - f_2| = 0.06577 > \epsilon$, 要继续迭代. 因 $t_3 = 0.96577$, $t_2 = 0.82759$, 故令

$$a_3 = t_2 = 0.82759, b_3 = b_2 = 2, t_3 = t_2 = 0.82759.$$

第四次迭代:

$$(a) = 0.029, (b) = 4, (t_3) = 0.00347.$$

代入公式(3.3.4)求得

$$t_4 = 0.98308.$$

由于

$$(f_4) = 0.00086, (f_3) = 0.00347,$$

并且 $|f_4 - f_3| = 0.00261 < \epsilon$, 停止迭代输出近似最优解 $t_4 = 0.98308$.

类似地, 我们可以讨论逐次两点二次插值法.

两点二次插值法()

给出不同的两点 x_1, x_2 , 函数值 $f(x_1), f(x_2)$, 及导数值 $f'(x_1)$ 或 $f'(x_2)$, 构造二次插值多项式

$$q(x) = a^2 + b + c, \quad (3.3.6)$$

取 $q(x)$ 的极小点为 (x_1) 的极小点的近似值. 显然, 令

$$q'(x) = 2a + b = 0,$$

得 $q(x)$ 的极小点为

$$x_1 = -\frac{b}{2a}. \quad (3.3.7)$$

考虑插值条件

$$\begin{aligned} q(x_1) &= a_1^2 + b_1 + c = f(x_1), \\ q(x_2) &= a_2^2 + b_2 + c = f(x_2), \\ q'(x_1) &= 2a_1 + b = f'(x_1). \end{aligned} \quad (3.3.8)$$

记 $x_i = (x_i)$, $f_i = f(x_i)$, $i = 1, 2$. 解方程组(3.3.8), 得

$$\begin{aligned} a &= \frac{f_1 - f_2 - x_1(f_1 - f_2)}{(x_1 - x_2)^2}, \\ b &= f_1 + 2x_1 \frac{f_1 - f_2 - x_1(f_1 - f_2)}{(x_1 - x_2)^2}. \end{aligned}$$

从而,

$$\begin{aligned} x_1 &= -\frac{b}{2a} = x_1 + \frac{1}{2} \frac{x_1(x_1 - x_2)^2}{f_1 - f_2 - x_1(f_1 - f_2)} \\ &= x_1 + \frac{1}{2} \frac{x_1 - x_2}{\frac{f_1 - f_2}{x_1 - x_2} - 1} \\ &= x_1 - \frac{1}{2} \frac{(x_1 - x_2) \frac{1}{x_1 - x_2}}{\frac{f_1 - f_2}{x_1 - x_2} - 1}. \end{aligned} \quad (3.3.9)$$

于是, 我们得到如下迭代格式:

$$x_{k+1} = x_k - \frac{1}{2} \frac{(x_k - x_{k-1}) \frac{1}{x_k - x_{k-1}}}{\frac{f_k - f_{k-1}}{x_k - x_{k-1}} - 1}. \quad (3.3.10)$$

特别, 如果 $x_1 = 0$, $f_2 = 0$, 则(3.3.9)给出

$$x_1 = -\frac{1}{2} \frac{(0)^2}{f(0) - (0) - (0)f'(0)}. \quad (3.3.11)$$

两点二次插值法()

给出不同的两点 x_1, x_2 , 函数值 $f(x_1)$, 及导数值 $f'(x_1)$, $f'(x_2)$, 构造二次插值多项式. 要求插值多项式满足插值条件

$$\begin{aligned} q(x_1) &= a_1^2 + b_1 + c = f(x_1), \\ q'(x_1) &= 2a_1 + b = f'(x_1), \\ q'(x_2) &= 2a_2 + b = f'(x_2). \end{aligned} \quad (3.3.12)$$

令 $x_i = (x_i)$, $x_i = (x_i)$, $i = 1, 2$. 类似于前面的讨论可以得到

$$x_{i+1} = x_i - \frac{b}{2a} = x_i - \frac{1 - x_i^2}{1 - x_i^2} x_i. \quad (3.3.13)$$

上式可写成迭代格式

$$x_{k+1} = x_k - \frac{x_k - x_{k-1}}{x_k - x_{k-1}} x_k. \quad (3.3.14)$$

上式也称为割线公式.

下面的定理表明: 两点二次插值法(3.3.14)的收敛阶约为 1.618, 即这个方法具有超线性收敛速度.

定理 3.3.3 设 $f(x)$ 存在连续三阶导数, x^* 满足 $f(x^*) = 0$, $f'(x^*) \neq 0$, 则两点二次插值公式(3.3.14)产生的序列 $\{x_k\}$ 收敛到 x^* , 其收敛速度约为 1.618.

类似地, 两点二次插值法(3.3.10)产生的序列的收敛速度也是 1.618.

三次插值法

类似上面的讨论, 考虑利用 x_{k-1} , x_k 及 $f(x_{k-1})$, $f(x_{k-1})$, $f(x_k)$, $f(x_k)$ 构造三次多项式, 并求这个三次多项式的局部极小点可得

$$x_{k+1} = x_k - (x_k - x_{k-1}) \frac{(f(x_k) + u_1 - u_2)}{(f(x_k) - (f(x_{k-1}) + 2u_1))},$$

其中

$$u_1 = (f(x_{k-1}) + f(x_k) - 3) \frac{(x_{k-1}) - (x_k)}{x_{k-1} - x_k},$$

$$u_2 = [u_1^2 - (f(x_{k-1}) - f(x_k))]^{\frac{1}{2}}.$$

可以证明, 二点三次插值法的收敛速度是 2 阶.

§ 3.4 精确线性搜索方法的收敛性

通常, 采用精确线性搜索的无约束最优化算法的一般形式如下:

算法 3.4.1

步 1 给出 $x_0 \in \mathbf{R}^n$, $0 \leq n \leq 1$, $k = 0$.

步 2. 计算 $f(x_k)$ 如果 $f(x_k) = 0$, 停止.

步 3. 计算下降方向 d_k .

步 4. 计算步长因子 α_k , 使得

$$f(x_k + \alpha_k d_k) = \min_0 f(x_k + \alpha d_k). \quad (3.4.1)$$

步 5. 令 $x_{k+1} = x_k + \alpha_k d_k$, 转步 2.

令

$$\varphi(\alpha) = f(x_k + \alpha d_k), \quad (3.4.2)$$

则显然有

$$\varphi(0) = f(x_k), \quad \varphi'(\alpha) = 0.$$

如前所述,在精确线性搜索中,我们往往选取 $\varphi(\alpha)$ 的一个平稳点,即选取 α_k 使得

$$\alpha_k = \min\{\alpha \mid f(x_k + \alpha d_k)^T d_k = 0, \alpha > 0\}. \quad (3.4.3)$$

前面几节介绍的直接法(0.618 法和 Fibonacci 法)和逐次插值逼近法都是精确线性搜索方法.

现在,我们讨论精确线性搜索的性质以及采用精确线性搜索的极小化算法的收敛性.

设 $\theta_k = \angle d_k, -g_k$ 表示向量 d_k 与 $-g_k$ 之间的夹角,即

$$\cos \theta_k = -\frac{d_k^T g_k}{\|d_k\| \|g_k\|}. \quad (3.4.4)$$

定理 3.4.2 设 d_k 是下降方向, α_k 是精确线性搜索的步长因子.若存在常数 $M > 0$,使得对所有 $k \geq 0$,

$$\|d_k\| \leq M, \quad \forall k, \quad (3.4.5)$$

则

$$f(x_k) - f(x_k + \alpha_k d_k) \geq \frac{1}{2M} \|g_k\|^2 \cos^2 \theta_k. \quad (3.4.6)$$

证明 由假设可知对任意 $\alpha > 0$ 有

$$\begin{aligned} f(x_k + \alpha d_k) &= f(x_k) + g_k^T d_k + \frac{1}{2} \alpha^2 d_k^T d_k + o(\alpha^2) f(x_k + \alpha d_k) d_k, \quad (0 < \alpha < 1) \\ f(x_k) + g_k^T d_k + \frac{1}{2} \alpha^2 M^2 \|d_k\|^2 &\leq f(x_k + \alpha d_k). \end{aligned} \quad (3.4.7)$$

令

$$\alpha^- = -\frac{g_k^T d_k}{M^2 \|d_k\|^2}.$$

由于 α_k 是精确线性搜索步长因子,故有

$$\begin{aligned} f(x_k) - f(x_k + \alpha_k d_k) &= f(x_k) - f(x_k + \alpha^- d_k) \\ &= -g_k^T d_k - \frac{1}{2} \alpha_k^2 M^2 \|d_k\|^2 \\ &= \frac{1}{2M} \frac{(g_k^T d_k)^2}{\|d_k\|^2} \\ &= \frac{1}{2M} \|g_k\|^2 \frac{(g_k^T d_k)^2}{\|g_k\|^2 \|d_k\|^2} \\ &= \frac{1}{2M} \|g_k\|^2 \cos^2 \theta_k. \end{aligned}$$

定理 3.4.3 设梯度 $g(x)$ 在水平集 $L = \{x \in \mathbf{R}^n \mid f(x) = f(x_0)\}$ 上存在且一

致连续,采用精确线性搜索的极小化算法 3.4.1 产生的方向 d_k 与 $-g_k$ 的夹角 θ_k 满足

$$\theta_k \leq \frac{\mu}{2}, \text{对某个 } \mu > 0, \quad (3.4.8)$$

则或者对某个有限的 k 有 $g_k = 0$, 或者 $f(x_k) \rightarrow -\infty$, 或者 $g_k \rightarrow 0$.

证明 假定对所有 k , $g_k \neq 0$, $f(x_k)$ 下有界. 由于 $\{f(x_k)\}$ 单调下降, 故极限存在, 因而

$$f(x_k) - f(x_{k+1}) \rightarrow 0. \quad (3.4.9)$$

现在用反证法证明结论. 假定 $g_k \rightarrow 0$ 不成立, 则存在常数 $\delta > 0$ 和一个子序列使得 $g_k \geq \delta$. 从而

$$-\frac{g_k^T d_k}{d_k} = \|g_k\| \cos \theta_k \geq \sin \mu \stackrel{\text{def}}{=} \delta. \quad (3.4.10)$$

又

$$\begin{aligned} f(x_k + d_k) &= f(x_k) + g(x_k)^T d_k \\ &= f(x_k) + g_k^T d_k + [g(x_k) - g_k]^T d_k \\ &= f(x_k) + d_k \left(-\frac{g_k^T d_k}{d_k} + \|g(x_k) - g_k\| \right), \end{aligned} \quad (3.4.11)$$

其中 x_k 在 x_k 与 $x_k + d_k$ 之间. 由于 g 在水平集 L 上一致连续, 故存在 η , 使得当 $0 < \|d_k\| < \eta$ 时,

$$\|g(x_k) - g_k\| \leq \frac{1}{2} \delta. \quad (3.4.12)$$

依次利用(3.4.11), (3.4.12)和(3.4.10)得

$$\begin{aligned} f(x_k + d_k) &\leq f(x_k) + d_k \left(-\frac{g_k^T d_k}{d_k} + \frac{1}{2} \delta \right) \\ &\leq f(x_k) - \frac{1}{2} \delta \|d_k\|. \end{aligned}$$

从而由精确线性搜索可得

$$f(x_{k+1}) \leq f(x_k + d_k) \leq f(x_k) - \frac{1}{2} \delta \|d_k\|.$$

这与(3.4.9)矛盾. 从而有 $g_k \rightarrow 0$.

§ 3.5 不精确线性搜索方法

前面介绍的几种线性搜索方法求 $\hat{x} = f(x_k + d_k)$ 的精确极小点, 花费的计算量较大. 一般地, 在迭代过程中, 没有必要把线性搜索搞得十分精确. 特别是当迭代点离目标函数的最优解尚远时, 过分追求线性搜索的精度反而会降低整

个算法的效率.另外,一些最优化方法,例如牛顿法和拟牛顿法,其收敛速度并不依赖于精确的一维搜索过程.因此,我们可以放松对 α_k 的精确度要求,只要求目标函数在迭代的每一步都有充分的下降即可,这样可以大大节省工作量.

3.5.1 Goldstein 准则

Goldstein 准则为

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k g_k^T d_k, \quad (3.5.1)$$

$$f(x_k + \alpha_k d_k) \geq f(x_k) + (1 - \rho) \alpha_k g_k^T d_k, \quad (3.5.2)$$

其中, $0 < \rho < \frac{1}{2}$. 上式中第一个不等式是充分下降条件,第二个不等式保证了 α_k 不会取得太小,因为当 α_k 取得太小时,算法前进很慢.从图中可知,满足不等式 (3.5.1) 的 α_k 构成区间 $(0, c]$, 满足不等式 (3.5.2) 的 α_k 构成区间 $[b, a]$. 所以满足上面 Goldstein 准则的 α_k 构成区间 $[b, c]$.

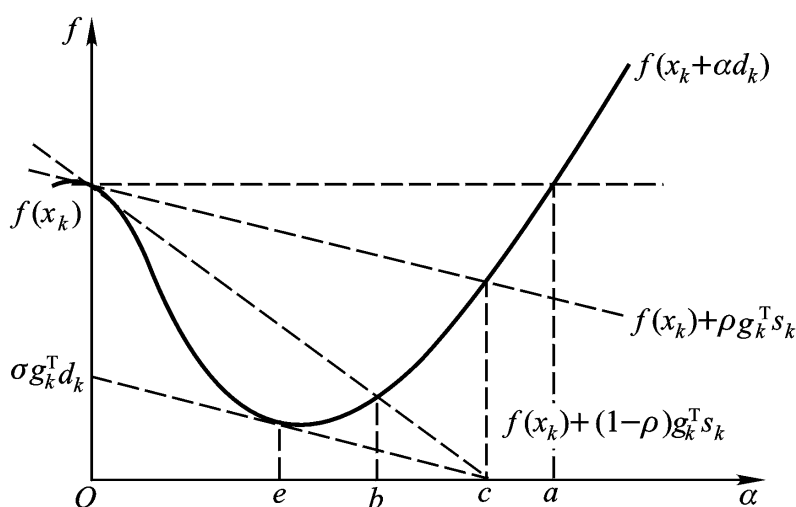


图 3.1

若设 $\phi(\alpha) = f(x_k + \alpha d_k)$, 则 (3.5.1) 和 (3.5.2) 可以分别写成

$$\phi(\alpha_k) \leq \phi(0) + \alpha_k \phi'(0), \quad (3.5.3)$$

$$\phi(\alpha_k) \geq \phi(0) + (1 - \rho) \alpha_k \phi'(0). \quad (3.5.4)$$

下面,我们给出 Goldstein 不精确线性搜索的步骤.

算法 3.5.1

步 1. 选取初始数据.在初始搜索区间 $[0, +\infty)$ (或 $[0, \alpha_{\max}]$) 中取定初始点 α_0 , 计算 $\phi(0)$, $\phi'(0)$, 给出 $\alpha_0, \frac{1}{2}$, $t > 1, k = 0$.

步 2. 检验准则 (3.5.3). 计算 $\phi(\alpha_k)$. 若

$$\phi(\alpha_k) \leq \phi(0) + \alpha_k \phi'(0),$$

转步 3; 否则, 令 $a_{k+1} = a_k, b_{k+1} = \alpha_k$, 转步 4.

步 3. 检验准则 (3.5.4). 若

$$f(x_k) = f(0) + (1 - \alpha_k) f(0),$$

停止迭代,输出 x_k ; 否则,令 $a_{k+1} = \alpha_k$, $b_{k+1} = b_k$.

若 $b_{k+1} < \alpha + \epsilon$ (或 ϵ_{\max}), 转步 4; 否则,令 $\alpha_{k+1} = t \alpha_k$, $k = k+1$, 转步 2.

步 4. $\alpha_{k+1} = \frac{a_{k+1} + b_{k+1}}{2}$, $k = k+1$, 转步 2.

3.5.2 Wolfe 准则

在 Goldstein 准则中, (3.5.2) 的一个缺点是可能把 $f(x_k + \alpha_k d_k)$ 的极小点排除在可接受区间之外. 为了克服这个缺点, 同时保证 α_k 不是太小, Wolfe 提出了下面的条件代替 (3.5.2):

$$g_{k+1}^T d_k \geq g_k^T d_k, \quad (\alpha, 1), \quad (3.5.5)$$

即

$$\begin{aligned} f(x_k + \alpha_k d_k) - g(x_k + \alpha_k d_k)^T d_k &\geq g_k^T d_k \\ &= f(0) > f(0), \end{aligned} \quad (3.5.6)$$

其几何解释是在可接受点处切线的斜率 $g(x_k + \alpha_k d_k)^T d_k$ 大于或等于初始斜率的 α 倍. 这个条件也叫做曲率条件. 这样, 充分下降条件和曲率条件一起构成了 Wolfe 准则:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k g_k^T d_k, \quad (3.5.7)$$

$$g(x_k + \alpha_k d_k)^T d_k \geq g_k^T d_k, \quad (3.5.8)$$

其中, $0 < \alpha < 1$.

应该指出, 不等式 (3.5.5) 是精确线性搜索所满足的正交条件

$$g_{k+1}^T d_k = 0$$

的近似. 但 (3.5.5) 的不足之处是即使在 $\alpha = 0$ 时也不能导致精确线性搜索. 因此, 下面的强 Wolfe 准则便提出来了:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \alpha_k g_k^T d_k \quad (3.5.9)$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leq |g_k^T d_k|, \quad (3.5.10)$$

其中, $0 < \alpha < 1$. 这样, 当 $\alpha = 0$ 时, (3.5.10) 的极限便是精确线性搜索.

一般地, α 愈小, 线性搜索愈精确. 不过, α 愈小, 工作量愈大. 而不精确线性搜索不要求过小的 α , 通常取 $\alpha = 0.1$, $\alpha \in [0.6, 0.8]$.

下面给出 Wolfe 不精确线性搜索方法的计算步骤.

算法 3.5.2

步 1. 选取初始数据. 给定初始搜索区间 $[0, \alpha_{\max}]$, $\alpha = 0, \frac{1}{2}$, $(\alpha, 1)$. 令

$\alpha_1 = 0$, $\alpha_2 = \alpha_{\max}$, 计算 $f_1 = f(x_k)$, $g_1 = g(x_k)^T d_k$, 取 (α_1, α_2) .

步 2. 计算 $\alpha = (\alpha_1, \alpha_2) = f(x_k + \alpha d_k)$. 若 $(\alpha) - \alpha_1 > \epsilon$, 转步 3; 否则, 由二次

插值公式(3.3.10)计算 \bar{x} :

$$\bar{x} = x_1 + \frac{x_2 - x_1}{2 + \frac{f(x_2) - f(x_1)}{f(x_1) - f(x_0)}}.$$

令 $x_2 = \bar{x}$, $x_1 = x_0$ 转步 2.

步 3. 计算 $\bar{x} = (x_1, x_2) = g(x_k + d_k)^T d_k$. 若 $\bar{x} \leq \epsilon$, 则令 $x_k = \bar{x}$, 输出 x_k , 停; 否则, 由二次插值公式(3.3.13)计算 \bar{x} :

$$\bar{x} = x_1 + \frac{(x_2 - x_1)(f(x_1) - f(x_0))}{f(x_1) - f(x_0) - (f(x_2) - f(x_1))}.$$

令 $x_1 = \bar{x}$, $x_2 = x_1$, $x_0 = x_1$, $x_1 = \bar{x}$, 转步 2.

3.5.3 Armijo 准则

设 d_k 是 $f(x)$ 在 x_k 处的下降方向, 给定 $(0, 1)$, $\alpha \in (0, \frac{1}{2})$, $\epsilon > 0$. 设 m_k 是使得下述不等式

$$f(x_k + \alpha^m d_k) \leq f(x_k) + \alpha^m g_k^T d_k. \quad (3.5.11)$$

成立的最小非负整数, 则令

$$x_{k+1} = x_k + \alpha^{m_k} d_k. \quad (3.5.12)$$

由于 d_k 是下降方向, 当 m 充分大时, 不等式(3.5.11)总是成立的, 因此上述 m_k 总是存在的. 由于 m_k 是使得不等式(3.5.11)成立的最小非负整数, 因而 x_{k+1} 不会太小, 从而保证了目标函数 $f(x)$ 的充分下降. 实际上, (3.5.11)就是充分下降条件

$$f(x_{k+1}) \leq f(x_k) + \alpha^{m_k} g_k^T d_k. \quad (3.5.13)$$

如果上式满足, 则终止搜索; 否则, 我们可以缩小 α , 或者在区间 $[0, \alpha]$ 上用二次插值公式(3.3.11)求近似极小点 $\bar{\alpha}$

$$\bar{\alpha} = \alpha - \frac{(f(x_k) - f(x_{k+1}))^2}{2(f(x_k) - f(x_{k+1})) - (f(x_k) - f(x_{k+1}))^2}. \quad (3.5.14)$$

将其作为新的 α , 这是一个插值法与充分下降条件组合起来的线性搜索方法.

3.5.4 不精确线性搜索的收敛性

本节最后, 我们讨论不精确线性搜索方法的收敛性定理. 为了保证方法的下降性, 我们要求避免搜索方向 $s_k = \alpha d_k$ 和负梯度方向 $-g_k$ 几乎直交的情形, 即要求 s_k 偏离 $-g_k$ 的正交方向远一点. 否则 $s_k^T g_k$ 接近于零, s_k 几乎不是下降方向. 为此, 我们假定 s_k 和 $-g_k$ 的夹角 θ_k 满足

$$\theta_k \leq \frac{\pi}{2} - \mu, \quad \mu > 0, \quad \forall k, \quad (3.5.15)$$

其中, $\mu > 0$, $\alpha_k \in (0, \frac{1}{2})$ 定义为

$$\cos \alpha_k = \frac{-g_k^T s_k}{\|g_k\| \|s_k\|}. \quad (3.5.16)$$

采用不精确线性搜索准则的一般下降算法的形式如下:

算法 3.5.3

步 1. 给出 $x_0 \in \mathbf{R}^n$, $0 < \mu < 1$, $k = 0$.

步 2. 如果 $\|g_k\| < \mu$, 则停止; 否则, 求出下降方向 d_k , 使其满足 $d_k^T g_k < 0$.

步 3. 利用 Goldstein 准则(3.5.1)-(3.5.2)或 Wolfe 准则(3.5.7)-(3.5.8)求出步长因子 α_k .

步 4. 令 $x_{k+1} = x_k + \alpha_k d_k$; $k = k + 1$, 转步 2.

类似于定理 3.4.2, 我们下面以 Wolfe 搜索为例, 给出不精确线性搜索在单步中至少下降的界.

引理 3.5.4 设函数 $f(x)$ 连续可微, 梯度 $g(x)$ 满足 Lipschitz 连续条件

$$\|g(y) - g(z)\| \leq M \|y - z\|, \quad (3.5.17)$$

如果 $f(x_k + \alpha d_k)$ 下有界, $\alpha > 0$, 则对满足 Wolfe 准则(3.5.7)-(3.5.8)的任何 $\alpha > 0$ 均有

$$f(x_k) - f(x_k + \alpha d_k) \geq \frac{(1 - \mu)}{M} \|g_k\|^2 \cos^2 \alpha, \quad \alpha \in (0, \frac{1}{2}). \quad (3.5.18)$$

证明 由 Lipschitz 条件和(3.5.8)得

$$\alpha M \|d_k\|^2 \geq d_k^T [g(x_k + \alpha d_k) - g_k] \geq (1 - \mu) \alpha d_k^T g_k,$$

即

$$\begin{aligned} \alpha \|d_k\| &\geq \frac{1 - \mu}{M \|d_k\|} \|d_k\| \|g_k\| \cos \alpha, \quad \alpha \in (0, \frac{1}{2}) \\ &= \frac{1 - \mu}{M} \|g_k\| \cos \alpha, \quad \alpha \in (0, \frac{1}{2}). \end{aligned} \quad (3.5.19)$$

利用(3.5.7)和(3.5.19), 有

$$\begin{aligned} f(x_k) - f(x_k + \alpha d_k) &\geq -\alpha d_k^T g_k \\ &= -\alpha \|d_k\| \|g_k\| \cos \alpha, \quad \alpha \in (0, \frac{1}{2}) \\ &= -\alpha \|g_k\| \cos \alpha \frac{1 - \mu}{M} \|g_k\| \cos \alpha, \quad \alpha \in (0, \frac{1}{2}) \\ &= -\frac{(1 - \mu)}{M} \|g_k\|^2 \cos^2 \alpha, \quad \alpha \in (0, \frac{1}{2}). \end{aligned}$$

下面两个定理分别讨论采用不精确线性搜索 Goldstein 准则和 Wolfe 准则的一般下降算法的总体收敛性.

定理 3.5.5 设在算法 3.5.3 中采用 Goldstein 准则(3.5.1)-(3.5.2)求步长因子 α_k , 并设夹角条件(3.5.15)满足. 如果 $g(x)$ 存在, 且在水平集 $\{x | f(x)$

$f(x_0)$ 上一致连续, 那么, 或者对某个 k , 有 $g_k = 0$, 或者 $f(x_k) - f(x_0) < -\epsilon$, 或者 $g_k = 0$.

证明 假定 $f(x_k)$ 下有界, 同时对所有的 k , $g_k = f(x_k) - f(x_0) > 0$. 由此得到 $s_k = 0$ 和 $f(x_k) - f(x_{k+1}) = 0$. 因此, 由 (3.5.1) 得到 $-g_k^T s_k = 0$.

今假定 $g_k = 0$ 不成立, 则存在 $\epsilon > 0$ 和一个子列使得 $g_{k_j} \neq 0$ 和 $s_{k_j} \neq 0$. 由于 $\|s_{k_j}\| = \frac{1}{2} - \mu$, 故

$$\cos \theta_{k_j} = \cos \frac{\pi}{2} - \mu = \sin \mu,$$

因而

$$-g_{k_j}^T s_{k_j} = \sin \mu \|g_{k_j}\| \|s_{k_j}\| = \sin \mu \|s_{k_j}\|.$$

但 Taylor 公式给出

$$f(x_{k+1}) = f(x_k) + g(x_k)^T s_k,$$

其中 θ_{k_j} 位于线段 (x_k, x_{k+1}) 上. 由 $g(x)$ 的一致连续性, 当 $s_k \rightarrow 0$ 时, $g(\theta_{k_j}) \rightarrow g_k$. 这样,

$$f(x_{k+1}) = f(x_k) + g_k^T s_k + o(\|s_k\|).$$

由此有

$$\frac{f(x_k) - f(x_{k+1})}{-g_k^T s_k} \rightarrow 1,$$

这与 (3.5.2) 矛盾. 因此 $g_k = 0$.

定理 3.5.6 设 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 在 \mathbf{R}^n 上连续可微和下有界. $g(x)$ 在水平集 $\mathcal{H} = \{x | f(x) = f(x_0)\}$ 上一致连续. 设不精确线性搜索方法采用 Wolfe 准则 (3.5.7)–(3.5.8), 则

$$\lim_k g_k \cos \theta_k = 0. \quad (3.5.20)$$

如果夹角条件 (3.5.15) 满足, 则

$$\lim_k g_k = 0. \quad (3.5.21)$$

证明 由于 $g_k^T s_k < 0$, 又由于 f 下有界, 因此序列 $\{x_k\}$ 是有定义的, 且在水平集 \mathcal{H} 中.

现在用反证法证明 (3.5.20). 假定 (3.5.20) 不成立, 则存在 $\epsilon > 0$ 和子序列, 其指标集为 K , 使得

$$-\frac{g_k^T s_k}{\|s_k\|} \geq \epsilon, \quad \forall k \in K.$$

于是, 由 (3.5.7),

$$f(x_k) - f(x_{k+1}) = \|s_k\| \left(-\frac{g_k^T s_k}{\|s_k\|} \right) \geq \epsilon \|s_k\|.$$

又由于 $\{f(x_k)\}$ 单调下降, 因而是收敛的, 故 $\{\|s_k\| / k \mid k \in K\}$ 收敛到零.

又由(3.5.8),

$$(1 - \alpha_k)(-g_k^T s_k) = (g(x_k + s_k) - g_k)^T s_k, \quad k = 0, 1, \dots$$

因此,

$$-\frac{g_k^T s_k}{s_k} = \frac{1}{1 - \alpha_k} (g(x_k + s_k) - g_k)^T s_k, \quad k = 0, 1, \dots, K.$$

由于 $\{s_k / k = K\}$ 收敛到零, 故由 $g(x)$ 在水平集上一致连续知上式右边趋于零, 从而产生矛盾. 这样, (3.5.20) 得证.

进一步, 若夹角条件(3.5.15)满足, 则存在一个正数 δ 使得

$$\cos \theta_k > 0, \quad \forall k. \quad (3.5.22)$$

(3.5.20) 和 (3.5.22) 立即给出 (3.5.21).

§ 3.6 信赖域方法的思想 and 算法框架

本章前面几节, 我们介绍了保证最优化方法总体收敛的线性搜索方法. 现在, 我们讨论另一种总体收敛策略——信赖域方法.

我们已经知道, 线性搜索方法先确定一个搜索方向 d_k , 然后沿着这个搜索方向 d_k 选择适当的步长因子 α_k , 新的迭代点定义为 $x_{k+1} = x_k + \alpha_k d_k \stackrel{\text{def}}{=} x_k + s_k$. 而信赖域方法是定义当前点的邻域 \mathcal{D}_k ,

$$\mathcal{D}_k = \{x \in \mathbf{R}^n \mid \|x - x_k\| \leq \rho_k\}, \quad (3.6.1)$$

这里 ρ_k 称为信赖域半径. 假定在这个邻域中, 二次模型是目标函数 $f(x)$ 的一个合适的近似, 则在这个邻域 (称为信赖域) 中极小化二次模型, 得到近似极小点 s_k , 并取 $x_{k+1} = x_k + s_k$, 其中 $\|s_k\| \leq \rho_k$. 这种方法不仅具有总体收敛性, 而且不要求目标函数的 Hesse 矩阵 (或其近似) 是正定的. 此外, 信赖域方法利用了二次模型求方向步 s_k , 使得目标函数的下降比线性搜索方法更有效.

信赖域方法的模型子问题是

$$\begin{aligned} \min \quad & q^{(k)}(s) = f(x_k) + g_k^T s + \frac{1}{2} s^T B_k s \\ \text{s.t.} \quad & \|s\| \leq \rho_k, \end{aligned} \quad (3.6.2)$$

其中, $s = x - x_k$, $g_k = \nabla f(x_k)$, B_k 是一个对称矩阵, 它是 Hesse 矩阵 $\nabla^2 f(x_k)$ 或其近似, $\rho_k > 0$ 为信赖域半径, $\|\cdot\|$ 为某一范数, 通常我们采用 l_2 范数.

如何选择 ρ_k ? 我们根据模型函数 $q^{(k)}(s)$ 对目标函数 $f(x)$ 的拟合程度来调整信赖域半径 ρ_k . 对于问题 (3.6.2) 的解 s_k , 设目标函数的下降量

$$Ared_k = f(x_k) - f(x_k + s_k) \quad (3.6.3)$$

为实际下降量, 设模型函数的下降量

$$Pred_k = q^{(k)}(0) - q^{(k)}(s_k) \quad (3.6.4)$$

为预测下降量. 定义比值

$$r_k = \frac{Ared_k}{Pred_k} = \frac{f(x_k) - f(x_k + s_k)}{q^{(k)}(0) - q^{(k)}(s_k)}, \quad (3.6.5)$$

它衡量模型函数 $q^{(k)}$ 与目标函数 f 的一致性程度. r_k 越接近 1, 表明模型函数 $q^{(k)}$ 与目标函数 f 的一致性程度越好. 这时可以增大 s_k 以扩大信赖域. 如果 $r_k > 0$ 但不接近 1, 我们保持信赖域半径 s_k 不变. 如果 r_k 接近零或取负值, 表明 $q^{(k)}$ 与 f 之间的一致性程度不理想, 我们就减小 s_k , 缩小信赖域. 下面的基本算法描述了这个过程.

算法 3.6.1 (信赖域算法)

步 1. 给出初始点 x_0 , 信赖域半径的上界 \bar{s} , $s_0 \in (0, \bar{s})$, $0, 0 < \alpha_1 < \alpha_2 < 1, 0 < \beta_1 < 1 < \beta_2, k = 0$.

步 2. 如果 $g_k = 0$, 停止.

步 3. (近似地) 求解子问题 (3.6.2) 得到 s_k .

步 4. 计算 $f(x_k + s_k)$ 和 r_k . 令

$$x_{k+1} = \begin{cases} x_k + s_k, & \text{如果 } r_k \geq \alpha_1, \\ x_k, & \text{否则.} \end{cases}$$

步 5. 校正信赖域半径. 令

$$\begin{aligned} s_{k+1} &= (0, \alpha_1 s_k], & \text{如果 } r_k < \alpha_1; \\ s_{k+1} &= [\alpha_1 s_k, s_k] & \text{如果 } r_k \in [\alpha_1, \alpha_2); \\ s_{k+1} &= [s_k, \min\{\beta_2 s_k, \bar{s}\}], & \text{如果 } r_k \geq \alpha_2. \end{aligned}$$

步 6. 产生 B_{k+1} , 校正 $q^{(k)}$, 令 $k = k+1$, 转步 2.

在上面的算法中, 如果 $r_k \geq \alpha_2$, 因而 $s_{k+1} = s_k$, 这种情形称为很成功迭代; 类似地, $r_k \in [\alpha_1, \alpha_2)$ 的情形称为成功迭代; $r_k < \alpha_1$ 的情形称为不成功迭代. 有时, 我们也把很成功迭代和成功迭代称为成功迭代. 上面算法中的一些参数选择可以建议如下: $\alpha_1 = 0.01$, $\alpha_2 = 0.75$, $\beta_1 = 0.5$, $\beta_2 = 2$, $s_0 = 1$ 或 $s_0 = \frac{1}{10} \|g_0\|$.

§ 3.7 信赖域方法的收敛性

我们首先给出如下假定:

假设 3.7.1 (A_0)

- 1) Hesse 近似 B_k 按范数一致有界, 即 $\|B_k\| \leq M$, M 是某个正常数;
- 2) 水平集 $L(x) = \{x \mid f(x) \leq f(x_0)\}$ 有界;

3) 函数 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 在水平集 $L(x)$ 上连续可微且下有界;

4) 对某个正的常数 α , $\|s_k\| \leq \alpha \|g_k\|$.

一般地,在信赖域方法的收敛性讨论中,我们总假设子问题(3.6.2)的(近似)解 s_k 满足

$$q^{(k)}(0) - q^{(k)}(s_k) \leq \frac{1}{2} \|g_k\| \min \left\{ \alpha, \frac{\|g_k\|}{B_k} \right\}, \quad (3.7.1)$$

其中 $\alpha \in (0, 1]$.

下面,我们指出,模型子问题的 Cauchy 点 s_k^c , 精确解和近似解均满足(3.7.1).

对于模型子问题(3.6.2),其 Cauchy 点可以表示为

$$s_k^c = - \frac{\|g_k\|}{B_k} g_k, \quad (3.7.2)$$

其中,

$$\alpha_k = \begin{cases} 1, & \text{如果 } g_k^T B_k g_k = 0 \\ \min \left\{ \frac{\|g_k\|^3}{g_k^T B_k g_k}, 1 \right\}, & \text{否则.} \end{cases} \quad (3.7.3)$$

引理 3.7.2 Cauchy 点满足

$$q^{(k)}(0) - q^{(k)}(s_k^c) \leq \frac{1}{2} \|g_k\| \min \left\{ \alpha_k, \frac{\|g_k\|}{B_k} \right\}. \quad (3.7.4)$$

证明 在 Cauchy 点的表达式(3.7.2)-(3.7.3)中, α_k 有三种可能性.下面,我们分三种情形来讨论.

先考虑 $g_k^T B_k g_k = 0$ 的情形.这时,由(3.7.3)可知 $\alpha_k = 1$,于是我们有

$$\begin{aligned} q^{(k)}(0) - q^{(k)}(s_k^c) &= q^{(k)}(0) - \left(q^{(k)}(0) - \frac{\|g_k\|}{B_k} g_k^T g_k \right) \\ &= \frac{\|g_k\|}{B_k} g_k^T g_k = \frac{1}{2} \frac{\|g_k\|^2 g_k^T B_k g_k}{g_k^T B_k g_k} \\ &= \frac{\|g_k\|}{B_k} \min \left\{ \alpha_k, \frac{\|g_k\|}{B_k} \right\}. \end{aligned} \quad (3.7.5)$$

在这种情形,(3.7.4)成立.

再考虑 $g_k^T B_k g_k > 0$ 且

$$\frac{\|g_k\|^3}{g_k^T B_k g_k} \leq 1 \quad (3.7.6)$$

的情形.这时,由(3.7.3)知 $\alpha_k = \frac{\|g_k\|^3}{g_k^T B_k g_k}$,于是我们有

$$q^{(k)}(0) - q^{(k)}(s_k^c) = \frac{\|g_k\|^4}{g_k^T B_k g_k} - \frac{1}{2} g_k^T B_k g_k \frac{\|g_k\|^4}{(g_k^T B_k g_k)^2}$$

$$\begin{aligned}
 &= \frac{g_k^4}{2 g_k^T B_k g_k} \\
 &\quad \frac{g_k^2}{2 B_k} \\
 &\quad \frac{1}{2} g_k \min_k, \frac{g_k}{B_k} .
 \end{aligned} \tag{3.7.7}$$

在这种情形, (3.7.4) 也成立.

最后, 考虑 $g_k^T B_k g_k > 0$ 且

$$\frac{g_k^3}{g_k^T B_k g_k} > 1 \tag{3.7.8}$$

的情形 这时, 由 (3.7.3) 可知 $\alpha_k = 1$ 利用 (3.7.8) 可得

$$\begin{aligned}
 q^{(k)}(0) - q^{(k)}(s_k^c) &= \frac{1}{2} g_k^2 - \frac{1}{2} \frac{g_k^2}{g_k^T B_k g_k} g_k^T B_k g_k \\
 &= \frac{1}{2} g_k^2 - \frac{1}{2} \frac{g_k^3}{g_k^T B_k g_k} \\
 &= \frac{1}{2} g_k \min_k, \frac{g_k}{B_k} .
 \end{aligned} \tag{3.7.9}$$

在这种情形, (3.7.4) 也成立.

综合上面三种情形的讨论, (3.7.4) 成立.

如果 s_k 是子问题 (3.6.2) 的精确解, 则显然有

$$q^{(k)}(s_k) = q^{(k)}(s_k^c), \tag{3.7.10}$$

从而, 我们立即得到

$$\begin{aligned}
 q^{(k)}(0) - q^{(k)}(s_k) &= q^{(k)}(0) - q^{(k)}(s_k^c) \\
 &= \frac{1}{2} g_k \min_k, \frac{g_k}{B_k} .
 \end{aligned} \tag{3.7.11}$$

如果 s_k 是子问题 (3.6.2) 的近似解, 满足

$$q^{(k)}(0) - q^{(k)}(s_k) \leq \eta (q^{(k)}(0) - q^{(k)}(s_k^e)), \tag{3.7.12}$$

其中 s_k^e 表示子问题 (3.6.2) 的精确解, $\eta \in (0, 1]$, 则我们也有

$$q^{(k)}(0) - q^{(k)}(s_k) \leq \eta (q^{(k)}(0) - q^{(k)}(s_k^c)). \tag{3.7.13}$$

从而, 我们立即得到: 近似解 s_k 满足

$$q^{(k)}(0) - q^{(k)}(s_k) \leq \frac{1}{2} g_k \min_k, \frac{g_k}{B_k} . \tag{3.7.14}$$

它表明近似解也满足 (3.7.1), 其中 $\alpha_k = \frac{1}{2}$. 这也告诉我们, 近似解达到的下降

仅是 Cauchy 点达到的下降的一个部分, 则下降性条件(3.7.1)就满足. 而下降性条件(3.7.1)对总体收敛性证明是十分重要的.

下面, 我们证明信赖域方法的总体收敛性定理. 为此, 我们先给出两个引理.

引理 3.7.3 设假设条件 A_0 成立, 则

$$|f(x_k + s_k) - q^{(k)}(s_k)| \leq \frac{1}{2} M \|s_k\|^2 + C(\|s_k\|) \|s_k\|, \quad (3.7.15)$$

其中, $C(\|s_k\|)$ 随着 s_k 的减小而任意减小.

证明 由 Taylor 定理,

$$f(x_k + s_k) = f(x_k) + g_k^T s_k + \int_0^1 [f(x_k + ts_k) - f(x_k)]^T s_k dt.$$

又

$$q^{(k)}(s_k) = f(x_k) + g_k^T s_k + \frac{1}{2} s_k^T B_k s_k.$$

于是,

$$|f(x_k + s_k) - q^{(k)}(s_k)| \leq \left| \frac{1}{2} s_k^T B_k s_k - \int_0^1 [f(x_k + ts_k) - f(x_k)]^T s_k dt \right| \leq \frac{1}{2} M \|s_k\|^2 + C(\|s_k\|) \|s_k\|.$$

引理 3.7.4 设假设条件 A_0 成立. 假定 $g_k \neq 0$ 和 $\|s_k\| \leq \delta_k$, δ_k 是某个小的容限. 则有 $\|s_{k+1}\| \leq \delta_{k+1}$.

证明 由假设条件和(3.7.14), 并设 $\|g_k\| > 0$, 有

$$\begin{aligned} Pred_k &= q^{(k)}(0) - q^{(k)}(s_k) \\ &\leq \frac{1}{2} \|g_k\|^2 \min\{\delta_k, \frac{\|g_k\|}{B_k}\} \\ &\leq \frac{1}{2} \min\{\delta_k, \frac{1}{M}\}. \end{aligned} \quad (3.7.16)$$

由算法 3.6.1, 并利用(3.7.15), (3.7.16)和假设 A_0 的 4), 得

$$\begin{aligned} |r_k - 1| &= \left| \frac{f(x_k) - f(x_k + s_k) - (q^{(k)}(0) - q^{(k)}(s_k))}{q^{(k)}(0) - q^{(k)}(s_k)} \right| \\ &= \left| \frac{f(x_k + s_k) - q^{(k)}(s_k)}{q^{(k)}(0) - q^{(k)}(s_k)} \right| \\ &\leq \frac{\frac{1}{2} M \|s_k\|^2 + C(\|s_k\|) \|s_k\|}{\frac{1}{2} \min\{\delta_k, 1/M\}} \\ &= \frac{\delta_k (M \delta_k + 2C(\delta_k))}{2 \min\{\delta_k, 1/M\}}. \end{aligned} \quad (3.7.17)$$

由于 $\delta_k \rightarrow 0$, 我们可以选择 δ_k 足够小, 使得

$$\frac{1}{M} \leq \frac{1}{M} + 2C(s_k) (1 - \frac{1}{2})^2, \quad (3.7.18)$$

从而得到 $|r_{k+1}| \leq 1 - \frac{1}{2}$, 这表明 $r_k > \frac{1}{2}$. 于是, 由算法 3.6.1 得到 $|r_{k+1}| \leq \frac{1}{2}$.

定理 3.7.5 设假设条件 A_0 成立. 则由算法 3.6.1 产生的序列满足

$$\liminf_k g_k = 0. \quad (3.7.19)$$

证明 用反证法. 假设定理不成立, 则存在 $\epsilon > 0$ 和某个正的指标 K , 使得

$$g_k \geq \epsilon, \quad \forall k \geq K. \quad (3.7.20)$$

先假定存在无限多个成功迭代, 则由算法 3.6.1 和 (3.7.14), 有

$$\begin{aligned} f(x_k) - f(x_{k+1}) &= \frac{1}{2} [q^{(k)}(0) - q^{(k)}(s_k)] \\ &\leq \frac{1}{2} (1 + \frac{1}{2}) g_k \min_k, \frac{g_k}{B_k} \\ &\leq \frac{1}{2} (1 + \frac{1}{2}) \min_k, \frac{g_k}{B_k}, \end{aligned}$$

其中, $\min_k = \max\{1 + \frac{1}{2}, B_k\}$. 由于 f 下有界, 上式意味着

$$\lim_k g_k = 0. \quad (3.7.21)$$

这与引理 3.7.4 的结果矛盾.

再假定只存在有限多个成功迭代, 即对于充分大的 k , 迭代都是不成功的, 信赖域收缩, 从而有 $|s_k| \rightarrow 0$. 这也与引理 3.7.4 的结果矛盾. 这些矛盾证明了结果 (3.7.19).

§ 3.8 解信赖域子问题

在信赖域算法 3.6.1 中最关键的是第 3 步: 解信赖域子问题. 这里我们介绍解信赖域子问题的折线法 (dog leg method), 它是由 Powell (1970) 提出来的.

为了近似地求解子问题 (3.6.2), 即求 $x_{k+1} = x_k + s_k$ 满足 $s_k^T g_k = -\frac{1}{2} g_k^T G_k g_k$, 折线法利用一条折线来近似 s . 其做法是连接 Cauchy 点 (即由最速下降法产生的极小点 $C.P.$) 和牛顿点 (即由牛顿法产生的极小点 x_{k+1}^N), 其连线与信赖域边界的交点取为 x_{k+1} . 显然, $|x_{k+1} - x_k| \leq |s_k|$. 当牛顿步 s_k^N 的长度 $|s_k^N| \leq |s_k|$ 时, x_{k+1} 就取为牛顿点 x_{k+1}^N .

对二次模型

$$q^{(k)}(x_k + g_k) = f(x_k) - \frac{1}{2} g_k^T G_k g_k + \frac{1}{2} g_k^T G_k g_k, \quad (3.8.1)$$

精确线性搜索因子 α_k 的表达式为

$$\alpha_k = \frac{g_k^T G_k g_k}{g_k^T G_k g_k}. \quad (3.8.2)$$

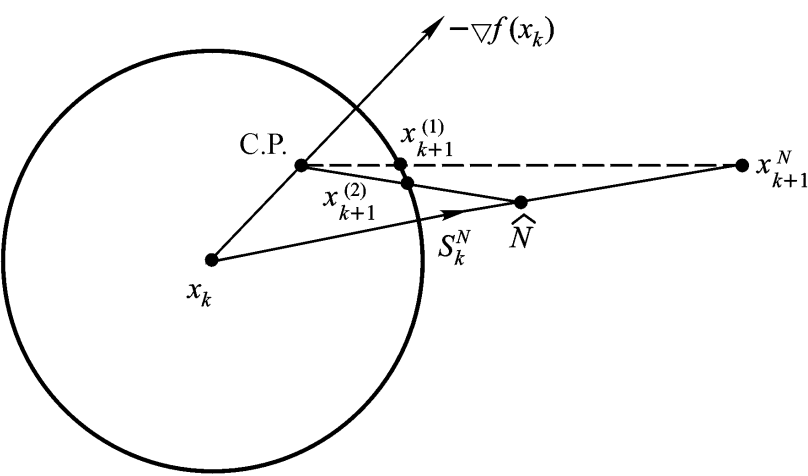


图 3 2 折线步方法和双折线步方法

Cauchy 步为

$$s_k^c = - \frac{g_k^T g_k}{g_k^T G_k g_k} g_k . \tag{3.8.3}$$

如果 $\|s_k^c\|_2 = \|g_k\|_2 \leq \tau_k$, 取

$$s_k = - \frac{\tau_k}{\|g_k\|_2} g_k , \tag{3.8.4}$$

这时,

$$x_{k+1} = x_k - \frac{\tau_k}{\|g_k\|_2} g_k . \tag{3.8.5}$$

如果 $\|s_k^c\|_2 = \|g_k\|_2 < \tau_k$, 我们再计算牛顿步 s_k^N ,

$$s_k^N = - G_k^{-1} g_k . \tag{3.8.6}$$

如果 $\|s_k^N\|_2 \leq \tau_k$, 则取

$$s_k = s_k^N = - G_k^{-1} g_k , \tag{3.8.7}$$

否则, 取

$$s_k = s_k^c + (\|s_k^N\|_2 - \|s_k^c\|_2) \frac{s_k^N}{\|s_k^N\|_2} , \tag{3.8.8}$$

其中 τ_k 的值由解方程

$$\|s_k^c\|_2 + (\|s_k^N\|_2 - \|s_k^c\|_2) = \tau_k \tag{3.8.9}$$

得到 .

综上所述, 我们得到

$$x_{k+1} = \begin{cases} x_k - \frac{\tau_k}{\|g_k\|_2} g_k , & \text{当 } \|s_k^c\|_2 \leq \tau_k \text{ 时,} \\ x_k + s_k^c + (\|s_k^N\|_2 - \|s_k^c\|_2) \frac{s_k^N}{\|s_k^N\|_2} , & \text{当 } \|s_k^c\|_2 < \tau_k \text{ 且 } \|s_k^N\|_2 > \tau_k \text{ 时,} \\ x_k - G_k^{-1} g_k , & \text{当 } \|s_k^c\|_2 < \tau_k \text{ 且 } \|s_k^N\|_2 \leq \tau_k \text{ 时.} \end{cases} \tag{3.8.10}$$

上述折线方法满足

- 1) 从点 x_k 到 Cauchy 点 C.P., 到牛顿点 x_{k+1}^N 的距离单调增加;
- 2) 从 Cauchy 点 C.P. 到牛顿点 x_{k+1}^N , 模型函数值单调减少.

在 Powell 单折线方法的基础上, Dennis 和 Mei(1979) 提出了双折线法. 他们发现, 如果让信赖域迭代中产生的点偏向牛顿方向, 则算法的性态会得到进一步改善. 于是, 他们把 Cauchy 点和牛顿方向上的 \hat{N} 点连接起来, 并将这条连线与信赖域边界的交点取为 x_{k+1} . 这就称为双折线. 在双折线情形, (3.8.10) 成为

$$\begin{aligned} x_k &= \frac{k}{g_k^T g_k} g_k, & \text{当 } s_k^c &= k \text{ 时,} \\ x_{k+1} &= x_k + s_k^c + (s_k^{\hat{N}} - s_k^c), & \text{当 } s_k^c < k \text{ 且 } s_k^{\hat{N}} > k \text{ 时,} \\ x_k &= G_k^{-1} g_k, & \text{当 } s_k^c < k \text{ 且 } s_k^{\hat{N}} < k \text{ 时.} \end{aligned} \quad (3.8.11)$$

其中,

$$s_k^{\hat{N}} = s_k^N, \quad (s_k^c, 1), \quad = \frac{g_k^T g_k}{(g_k^T G_k g_k)(g_k^T G_k^{-1} g_k)}. \quad (3.8.12)$$

如果 $\alpha = 1$, (3.8.11) 就是 Powell 单折线法 (3.8.10). 一般地, 取 $\alpha = 0.8 + 0.2$. 上述双折线法也满足

- 1) 从点 x_k 到 Cauchy 点 C.P., 到 $x_{k+1}^{\hat{N}}$ 的距离单调增加;
- 2) 从 Cauchy 点 C.P. 到点 $x_{k+1}^{\hat{N}}$, 模型函数值单调减少.

在产生 C.P. 点和 \hat{N} 点后, 所求的新点 $x_{k+1}(\cdot)$ 由 (3.8.11) 产生, 选择 α 使得

$$s_k^c + (s_k^{\hat{N}} - s_k^c) \frac{\alpha}{2} = \frac{\alpha}{k}. \quad (3.8.13)$$

如果所得到的 $x_{k+1}(\cdot)$ 满足下降性要求

$$f(x_{k+1}(\cdot)) \leq f(x_k) + g_k^T (x_{k+1}(\cdot) - x_k), \quad 0, \frac{1}{2}, \quad (3.8.14)$$

则接受 $x_{k+1}(\cdot)$ 为新点 x_{k+1} , 并根据信赖域算法 3.6.1 步 5 校正信赖域半径; 如果 $x_{k+1}(\cdot)$ 不满足 (3.8.14), 则令 $x_{k+1} = x_k$, 并缩小信赖域半径.

在折线法中, 可以考虑将牛顿点改为拟牛顿点或不定情形改进的牛顿点, 从而可以得到折线法的若干其他变形 (例如, 见 Zhang 和 Xu(1999)).

例 3.8.1 设 $f(x) = x_1^4 + x_1^2 + x_2^2$, 当前点 $x_k = (1, 1)^T$, $\alpha = \frac{1}{2}$, 试用双折线法求 x_{k+1} .

解 由于 $x_k = (1, 1)^T$, 由计算可得

$$g_k = (6, 2)^T, \quad G_k = \begin{pmatrix} 14 & 0 \\ 0 & 2 \end{pmatrix},$$

$$s_k^N = -G_k^{-1} g_k = -\frac{3}{7}, -1^T.$$

$$s_k^c = -\frac{g_k^T g_k}{g_k^T G_k g_k} g_k = -\frac{40}{512} \frac{6}{2} = -0.469.$$

由于 $s_k^c = -0.469 < -\frac{1}{2}$, 计算到 s_k^N 的步长 \hat{s}_k^N , 有

$$\begin{aligned} &= \frac{g_k^T g_k}{(g_k^T G_k g_k)(g_k^T G_k^{-1} g_k)} = \frac{(40)^2}{(512) \frac{32}{7}} = 0.684, \\ &= 0.8 + 0.2 = 0.747, \\ s_k^{\hat{N}} &= s_k^N = -0.320 \\ &\quad -0.747. \end{aligned}$$

由于 $s_k^{\hat{N}} = -0.813 > -\frac{1}{2}$, 故取双折线步长为

$$s_k = s_k^c + (\hat{s}_k^N - s_k^c), \quad (0, 1),$$

使得 $s_k = -\frac{1}{2}$ 解二次方程

$$s_k^c + (\hat{s}_k^N - s_k^c)^2 = -\frac{1}{2}$$

得

$$0.867.$$

因此

$$\begin{aligned} s_k &= s_k^c + 0.867(\hat{s}_k^N - s_k^c) = -0.340 \\ &\quad -0.669, \\ x_{k+1} &= x_k + s_k = -0.660 \\ &\quad -0.331. \end{aligned}$$

小 结

线性搜索方法与信赖域方法是保证最优化方法总体收敛的两类技术。

在线性搜索方法中,有精确线性搜索和不精确线性搜索.在六十年代中期以前,线性搜索方法是精确线性搜索的一统天下. § 3.2 和 § 3.3 主要描述了 0.618 法, Fibonacci 法和逐次插值逼近法, 这些方法属于精确线性搜索的范畴. § 3.4 建立了精确线性搜索的收敛性。

自从 Armijo(1966), Goldstein(1965)提出了不精确线性搜索方法以后,不精确线性搜索由于计算量小,效率高而成了现在流行的线性搜索方法.我们在 § 3.5 较详细地讨论了不精确线性搜索,介绍了三种最主要的不精确线性搜索准则: Goldstein 准则, Wolfe 准则(强 Wolfe 准则), Armijo 准则,并分别给出了算法. § 3.5.4 建立了不精确线性搜索方法的总体收敛性。

读者应至少掌握 0.618 法, 三点二次插值法, 以及 Goldstein 和 Wolfe 不精确搜索法的算法和程序. 精确线性搜索的收敛性定理 3.4.3, 不精确线性搜索的收敛性定理 3.5.5 和 3.5.6 都是很重要的, 它们在后面最优化方法收敛性的讨论中起重要作用, 希望读者加深理解.

信赖域方法是另一类保证最优化方法总体收敛的技术. 虽然信赖域方法的历史可以追溯到 Levenberg (1944), Marquardt (1963), Goldfeld, Quandt and Trotter (1966), 但现代信赖域方法是 Powell (1970) 提出来的. 他明确提出了信赖域子问题, 接受方向步 s_k 的准则, 校正信赖域半径 ρ_k 的准则, 及收敛性定理. 这些措施使信赖域方法比线性搜索方法具有更大的优越性. 应该指出, 引理 3.7.2 是重要的, 它指出了在信赖域方法中 Cauchy 点是充分下降的. 并且, 如果 s_k 是模型子问题的近似解, 那么它满足 (3.7.1). 这个结果也是 Powell (1975) 首先得到的. 本章最后讨论了解信赖域子问题的折线法和双折线法及若干修改. 其他的解子问题的方法可以参看 Moré and Sorensen (1983), Steihaug (1983) 等. 一个著名的关于信赖域方法的综合性评述见 Moré (1983), 一本专门论述信赖域方法的专著见 Conn, Gould, Toint (2000).



1. 分别用 0.618 法, 二次插值法, Goldstein 不精确线性搜索方法求函数

$$f(x) = (\sin x)^6 \tan(1-x)e^{30x}$$

在区间 $[0, 1]$ 上的极大值.

2. 写出 Fibonacci 法的计算过程和 C 程序 (或 MATLAB, FORTRAN 程序).

3. 设 $\phi(t) = e^{-t} + e^t$, 区间为 $[-1, 1]$.

- (1) 用 0.618 法极小化 $\phi(t)$.

- (2) 用 Fibonacci 法极小化 $\phi(t)$.

4. 设 $\phi(t) = 1 - te^{-t^2}$, 区间为 $[0, 1]$. 试用三点二次插值法极小化 $\phi(t)$.

5. 设 $\phi(t) = -2t^3 + 21t^2 - 60t + 50$.

- (1) 用 Goldstein 方法极小化 $\phi(t)$, $\rho_0 = 0.5$, $\epsilon = 0.1$.

- (2) 用 Wolfe 方法极小化 $\phi(t)$, $\rho_0 = 0.5$, $\epsilon = 0.1$, $\sigma = 0.8$.

6. 设 $f(x) = x_1^4 + x_1^2 + x_2^2$, 给定当前点 $x_k = (1, 1)^T$, 方向 $d_k = (-3, -1)^T$, 并设 $\epsilon = 1$, $t = 0.5$.

- (1) 试分别用 Goldstein 方法和 Wolfe 方法求新点 x_{k+1} .

- (2) 分别以 $\epsilon = 1$, $\epsilon = 0.5$, $\epsilon = 0.1$ 说明哪些 ρ_k 满足 Wolfe 准则, 哪些 ρ_k 不满足 Wolfe 准则?

7. 设 $f(x) = \frac{1}{2}x_1^2 + x_2^2$. 设初始点 $x_0 = (1, 1)^T$. 对于 (1) $\rho_0 = 1$, (2) $\rho_0 = \frac{5}{4}$,

- (1) 分别用折线法计算下一个迭代点 x_1 ;

- (2) 分别用双折线法计算下一个迭代点 x_2 .

第 4 章

无约束最优化方法

§ 4.1 最速下降法

最速下降法是以负梯度方向作为下降方向的极小化算法, 又称梯度法, 是 1874 年法国科学家 Cauchy 提出的. 最速下降法是无约束最优化中最简单的方法.

设目标函数 $f(x)$ 在 x_k 附近连续可微, 且 $g_k^T > f'(x_k) > 0$. 将 $f(x)$ 在 x_k 处 Taylor 展开,

$$f(x) = f(x_k) + g_k^T (x - x_k) + o(\|x - x_k\|). \quad (4.1.1)$$

记 $x - x_k = d_k$, 则上式可写为

$$f(x_k + d_k) = f(x_k) + g_k^T d_k + o(\|d_k\|). \quad (4.1.2)$$

显然, 若 d_k 满足 $g_k^T d_k < 0$, 则 d_k 是下降方向, 它使得 $f(x_k + d_k) < f(x_k)$. 当取定后, $d_k^T g_k$ 的值越小, 即 $-d_k^T g_k$ 的值越大, 函数 $f(x)$ 在 x_k 处下降量越大. 由 Cauchy Schwartz 不等式

$$|d_k^T g_k| \leq \|d_k\| \|g_k\| \quad (4.1.3)$$

可知, 当且仅当 $d_k = -g_k$ 时, $d_k^T g_k$ 最小, $-d_k^T g_k$ 最大, 从而 $-g_k$ 是最速下降方向. 以 $-g_k$ 为下降方向的方法叫最速下降法.

事实上, 最速下降方向也可以这样来考虑. 因为目标函数 f 沿方向 d 的变化率是 $g(x_k)^T d$, 故最速下降的单位方向 d 是问题

$$\min_d g_k^T d, \quad (4.1.4)$$

$$\text{s.t. } \|d\| = 1 \quad (4.1.5)$$

的解. 注意到

$$\begin{aligned} d^T g_k &= \|d\| \|g_k\| \cos \theta_k \\ &= \|g_k\| \cos \theta_k, \end{aligned} \quad (4.1.6)$$

其中, θ_k 是 g_k 与 d 之间的夹角. 极小化上式, 便得到当 $\theta_k = 0$, 即 $\cos \theta_k = 1$ 时, $d^T g_k$ 极小, 这时,

$$d = - \frac{g^k}{g^k} . \quad (4.1.7)$$

最速下降法的迭代格式为

$$x_{k+1} = x_k - \alpha_k g^k , \quad (4.1.8)$$

其中步长因子 α_k 由线性搜索策略确定 .

下面我们给出最速下降算法 .

算法 4.1.1 (最速下降法)

步 1. 给出 $x_0 \in \mathbf{R}^n, 0 \leq n \leq 1, k = 0$.

步 2. 计算 $d_k = -g^k$; 如果 $g^k = 0$, 停止 .

步 3. 由线性搜索求步长因子 α_k .

步 4. 计算 $x_{k+1} = x_k + \alpha_k d_k$.

步 5. $k = k+1$, 转步 2 .

最速下降法的收敛性理论在最优化中具有重要的理论意义 现在我们讨论最速下降法的总体收敛性 . 在 § 3.4 节中我们讨论了精确线性搜索方法的收敛性定理 对于最速下降法, $\alpha_k = 0$, 因而, 利用定理 3.4.3 立即可知最速下降法是总体收敛的 .

定理 4.1.2 设 $f(x)$ 在水平集 $L = \{x \in \mathbf{R}^n / f(x) = f(x_0)\}$ 上存在且一致连续, 则最速下降法产生的序列满足或者对某个 k 有 $g^k = 0$, 或者 $f(x_k) \rightarrow f^*$, $g^k \rightarrow 0$.

证明 利用定理 3.4.3 立得 .

下面的定理也论证了最速下降法的总体收敛性 .

定理 4.1.3 设函数 $f(x)$ 二次连续可微, 且 $\lambda_{\min}^2 f(x) \geq M$, 其中 M 是某个正常数 . 对任何给定的初始点 x_0 , 最速下降算法 4.1.1 或有限终止, 或 $\lim_{k \rightarrow \infty} f(x_k) = f^*$, 或 $\lim_{k \rightarrow \infty} g^k = 0$.

证明 考虑无限迭代下去的情形, 由定理 3.4.2, 有

$$f(x_k) - f(x_{k+1}) \geq \frac{1}{2M} \|g^k\|^2 . \quad (4.1.9)$$

于是,

$$\begin{aligned} f(x_0) - f(x_k) &= \sum_{i=0}^{k-1} [f(x_i) - f(x_{i+1})] \\ &\geq \frac{1}{2M} \sum_{i=0}^{k-1} \|g^i\|^2 . \end{aligned} \quad (4.1.10)$$

两边取极限, 于是, 或者 $\lim_{k \rightarrow \infty} f(x_k) = f^*$, 或者 $\lim_{k \rightarrow \infty} g^k = 0$. 从而定理成立 .

最速下降法具有程序设计简单, 计算工作量小, 存储量小, 对初始点没有特别要求等优点 . 但是, 最速下降方向仅是函数的局部性质, 对整体求解过程而言,

这个方法下降非常缓慢 数值试验表明,当目标函数的等值线接近于一个圆(球)时,最速下降法下降较快;而当目标函数的等值线是一个扁长的椭球时,最速下降法开始几步下降较快,后来就出现锯齿现象,下降十分缓慢 事实上,由于精确线性搜索满足 $g_{k+1}^T d_k = 0$, 则

$$g_{k+1}^T g_k = d_{k+1}^T d_k = 0, \tag{4.1.11}$$

这表明最速下降法中相邻两次的搜索方向是相互直交的,这就产生了锯齿形状. 越接近极小点,步长越小,前进越慢.

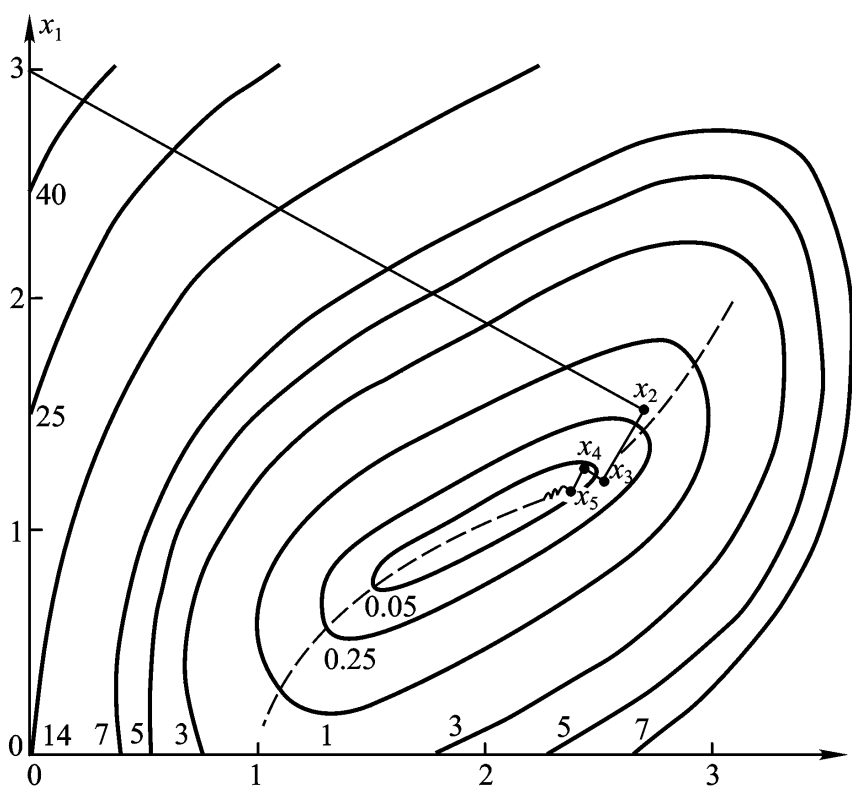


图 4.1 最速下降法中的锯齿形

可以证明,采用精确线性搜索的最速下降法的收敛速度是线性的.对于极小化正定二次函数 $\min f(x) = \frac{1}{2} x^T G x$,最速下降法产生的序列满足

$$\frac{f(x_{k+1}) - f(x^*)}{f(x_k) - f(x^*)} = \frac{1 - \lambda_1^2}{1 + \lambda_1^2} = \frac{-1}{+1}, \tag{4.1.12}$$

$$\frac{x_{k+1} - x^*}{x_k - x^*} = \frac{1}{n} \frac{1 - \lambda_1}{1 + \lambda_1} = \frac{-1}{+1}, \tag{4.1.13}$$

其中 λ_1 和 λ_n 分别是矩阵 G 的最大和最小特征值, $\kappa = \lambda_1 / \lambda_n$ 是矩阵 G 的条件数. 在非二次情形,如果 $f(x)$ 在 x^* 附近二次连续可微, $f(x^*) = 0$, $\lambda_1^2 f(x^*)$ 正定,则(4.1.12)也成立.

§ 4.2 牛 顿 法

牛顿法的基本思想是利用目标函数 $f(x)$ 在迭代点 x_k 处的二次 Taylor 展开作为模型函数,并用这个二次模型函数的极小点序列去逼近目标函数的极小点.

设 $f(x)$ 二次连续可微, $x_k \in \mathbf{R}^n$, Hesse 矩阵 $\nabla^2 f(x_k)$ 正定. 我们在 x_k 附近用二次 Taylor 展开近似 f ,

$$f(x_k + s) \approx q^{(k)}(s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s, \quad (4.2.1)$$

其中 $s = x - x_k$, $q^{(k)}(s)$ 为 $f(x)$ 的二次近似. 将上式右边极小化, 即令

$$q^{(k)}(s) = f(x_k) + \nabla f(x_k)^T s + \frac{1}{2} s^T \nabla^2 f(x_k) s = 0, \quad (4.2.2)$$

得

$$x_{k+1} = x_k - [\nabla^2 f(x_k)]^{-1} \nabla f(x_k), \quad (4.2.3)$$

这就是牛顿法迭代公式. 相应的算法称为牛顿法. 令 $G_k = \nabla^2 f(x_k)$, $g_k = \nabla f(x_k)$, 则 (4.2.3) 也可写成

$$x_{k+1} = x_k - G_k^{-1} g_k. \quad (4.2.4)$$

对于正定二次函数, 牛顿法一步即可达到最优解. 对于一般非二次函数, 牛顿法并不能保证经过有限次迭代求得最优解, 但如果初始点 x_0 充分靠近极小点, 牛顿法的收敛速度一般是快的. 下面的定理证明了牛顿法的局部收敛性和二阶收敛速度.

定理 4.2.1 (牛顿法收敛定理) 设 $f(x)$ 二阶连续可微, x^* 是 $f(x)$ 的局部极小点, $\nabla^2 f(x^*)$ 正定. 假定 $f(x)$ 的 Hesse 矩阵 $G_k = \nabla^2 f(x_k)$ 满足 Lipschitz 条件, 即存在 $L > 0$, 使得对于所有 $1 \leq i, j \leq n$ 有

$$|G_{ij}(x) - G_{ij}(y)| \leq L \|x - y\|, \quad x, y \in \mathbf{R}^n, \quad (4.2.5)$$

其中 $G_{ij}(x)$ 是 Hesse 矩阵 G_k 的 (i, j) 元素. 则当初始点 x_0 充分靠近 x^* 时, 对于一切 k , 牛顿迭代 (4.2.4) 有意义, 迭代序列 $\{x_k\}$ 收敛到 x^* , 并且具有二阶收敛速度.

证明 设 $h_k = x_k - x^*$, $g(x) = \nabla f(x)$, $G(x) = \nabla^2 f(x)$. 由于 $f(x)$ 二阶连续可微和 (4.2.5), 故利用 Taylor 公式得到

$$g(x_k + h) = g_k + G_k h + O(\|h\|^2).$$

令 $h = -h_k$, 得

$$0 = g(x^*) = g_k - G_k h_k + O(\|h_k\|^2).$$

由于 $f(x)$ 二次连续可微, $G^* = \nabla^2 f(x^*)$ 正定, 故当 x_k 充分靠近 x^* 时, G_k 也正

定且 G_k^{-1} 有界, 用 G_k^{-1} 乘以上式两边, 得

$$\begin{aligned} 0 &= G_k^{-1} g_k - h_k + O(\|h_k\|^2) \\ &= -s_k - h_k + O(\|h_k\|^2) \\ &= -h_{k+1} + O(\|h_k\|^2). \end{aligned}$$

由 $O(\cdot)$ 的定义, 存在常数 C , 使得

$$\|h_{k+1}\| \leq C \|h_k\|^2. \quad (4.2.6)$$

若 x_k 充分靠近 x^* 使得 $x_k \in \{x \mid \|x - x^*\| \leq 1/C\} =: D$, 则由 (4.2.6) 有

$$\|h_{k+1}\| \leq C \|h_k\|^2 / C < \|h_k\| / C, \quad (4.2.7)$$

这表明 x_{k+1} 也在这个领域 D 中. 由归纳法, 迭代对所有 k 有定义. 由于

$$\|h_k\| \leq \|h_{k-1}\| \leq \dots \leq \|h_0\|,$$

令 $k \rightarrow \infty$, 得 $\|h_k\| \rightarrow 0$. 因此迭代序列 $\{x_k\}$ 收敛. (4.2.6) 表明收敛速度是二阶的.

应该注意的是, 当初始点远离最优解时, G_k 不一定正定, 牛顿方向不一定是下降方向, 其收敛性不能保证. 为此, 在牛顿法中引进步长因子, 得到

$$\begin{aligned} d_k &= -G_k^{-1} g_k, \\ x_{k+1} &= x_k + \alpha_k d_k, \end{aligned}$$

其中 α_k 由线性搜索策略确定.

算法 4.2.2 (带步长因子的牛顿法)

步 1. 选取初始数据. 取初始点 x_0 , 终止误差 $\epsilon > 0$, 令 $k = 0$.

步 2. 计算 g_k . 如果 $\|g_k\| \leq \epsilon$, 停止迭代, 输出 x_k .

步 3. 解方程组构造牛顿方向. 解 $G_k d = -g_k$ 得 d_k .

步 4. 进行线性搜索求 α_k , 使得

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k) \quad (4.2.8)$$

步 5. 令 $x_{k+1} = x_k + \alpha_k d_k$, $k = k+1$, 转步 2.

下面我们证明上述带步长因子的牛顿法是总体收敛的.

定理 4.2.3 设 $f(x)$ 二阶连续可微, 又设对任意 $x_0 \in \mathbf{R}^n$, 存在常数 $m > 0$, 使得 $f(x)$ 在水平集 $L(x_0) = \{x \mid f(x) \leq f(x_0)\}$ 上满足

$$u^T \nabla^2 f(x) u \geq m \|u\|^2, \quad \forall u \in \mathbf{R}^n, x \in L(x_0). \quad (4.2.9)$$

则在精确线性搜索条件下, 带步长因子的牛顿法产生的迭代点列 $\{x_k\}$ 满足

(1) 当 $\{x_k\}$ 是有限点列时, 其最后一个点为 $f(x)$ 的唯一极小点.

(2) 当 $\{x_k\}$ 是无穷点列时, 它收敛到 $f(x)$ 的唯一极小点 \bar{x} .

证明 首先由 (4.2.9) 知 $f(x)$ 为 \mathbf{R}^n 上的严格凸函数, 从而其平稳点为总体极小点, 且是唯一的.

由 (4.2.9) 和定理 4.2.1 可知, 水平集 $L(x_0)$ 是有界闭凸集. 由于 $\{f(x_k)\}$ 单调下

降, 可知 $\{x_k\} \subset L(x_0)$, 故 $\{x_k\}$ 是有界点列, 于是存在 $\bar{x} \in L(x_0)$, 使得 $x_k \rightarrow \bar{x}$. 又 $\{f(x_k)\}$ 单调下降且有下界, 故 $f(x_k) \rightarrow f(\bar{x})$.

由于 $f(x)$ 连续及 $L(x_0)$ 是有界闭凸的, 故存在常数 $M > m > 0$ 使得

$$-d_k^T g_k \leq M, \quad \forall x \in L(x_0). \quad (4.2.10)$$

于是,

$$\begin{aligned} \cos \theta_k &= \frac{-d_k^T g_k}{\|d_k\| \|g_k\|} \\ &= \frac{d_k^T \nabla^2 f(x_k) d_k}{\|d_k\| \sqrt{d_k^T \nabla^2 f(x_k) d_k}} \\ &= \frac{m}{M}. \end{aligned} \quad (4.2.11)$$

再由定理 3.4.2 可得

$$\begin{aligned} f(x_k) - f(x_k + \alpha_k d_k) &\geq \frac{1}{2M} \nabla^2 f(x_k) \cos^2 \theta_k \\ &\geq \frac{1}{2M} \frac{m^2}{M^2} \nabla^2 f(x_k). \end{aligned} \quad (4.2.12)$$

令 $k \rightarrow \infty$, 并注意到上式左边趋于 0, 从而得到 $\nabla^2 f(x_k) \rightarrow 0$. 这表明 $\{x_k\}$ 收敛到唯一极小点 \bar{x} .

完全类似的, 如果线性搜索采用 Wolfe 不精确线性搜索, 根据引理 3.5.4, 我们有

$$f(x_k) - f(x_k + \alpha_k d_k) \geq \nabla^2 f(x_k) \cos^2 \theta_k, \quad (4.2.13)$$

在定理 4.2.3 的证明中只要将 (4.2.12) 式换成

$$\begin{aligned} f(x_k) - f(x_k + \alpha_k d_k) &\geq \nabla^2 f(x_k) \cos^2 \theta_k \\ &\geq \frac{m^2}{M^2} \nabla^2 f(x_k) \end{aligned} \quad (4.2.14)$$

同样可以推出 $\nabla^2 f(x_k) \rightarrow 0$. 于是我们有如下关于不精确线性搜索牛顿法的总体收敛性定理.

定理 4.2.4 设 $f(x)$ 二次连续可微, 又设对任意 $x_0 \in \mathbf{R}^n$, 存在常数 $m > 0$, 使得 $f(x)$ 在水平集 $L(x_0) = \{x \mid f(x) = f(x_0)\}$ 上满足 (4.2.9). 则在 Wolfe 不精确线性搜索条件下, 带步长因子的牛顿法产生的点列满足

$$\lim_{k \rightarrow \infty} \nabla^2 f(x_k) = 0, \quad (4.2.15)$$

且 $\{x_k\}$ 收敛到 $f(x)$ 的唯一极小点.

上述带步长因子的牛顿法告诉我们, 仅当目标函数的 Hesse 矩阵 $G(x)$ 正定时, 方法才是总体收敛的. 但是, 当初始点远离局部极小点时, $G(x_k)$ 可能不正定, 也可能奇异, 这样, 所产生的搜索方向 d_k 可能不是下降方向.

例 4.2.5 极小化 $f(x) = x_1^4 + x_1 x_2 + (1 + x_2)^2$.

解 设 $x^{(0)} = (1, 1)^T$, $g(x^{(0)}) = (0, 2)^T$, $G_0 = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}$.

显然, G_0 不是正定的.

$$G_0^{-1} = \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix}.$$

于是, $d_0 = -G_0^{-1}g_0 = (-2, 0)^T$. 沿方向 d_0 进行线性搜索, $f(x^{(0)} + d_0) = (-2)^4 + 1$, 其极小点 $\alpha_0 = 0$. 因此

$$x^{(1)} = x^{(0)} + \alpha_0 d_0 = x^{(0)}.$$

从而迭代不能继续进行下去.

为了克服上述缺点, 我们提出下列措施.

设 $g_k \neq 0$. 若 G_k 正定, 则

$$d_k = -G_k^{-1}g_k$$

是下降方向. 若 G_k 不正定, 但可逆, 且 $g_k^T G_k^{-1}g_k < 0$, 则

$$d_k = G_k^{-1}g_k$$

是下降方向. 若 G_k 奇异或 $G_k^{-1}g_k = 0$, 则取 $d_k = -g_k$.

当 G_k 不是正定时, 另一种有效的方法是 Gill Murray 修改牛顿法. 当 G_k 是正定时, 我们总有 Cholesky 分解

$$G_k = L_k D_k L_k^T,$$

其中 L_k 是单位下三角阵, D_k 是对角阵. 当 G_k 不是正定时, Gill Murray (1974) 提出了强迫正定的修改 Cholesky 分解, 使得

$$\bar{G}_k = L_k D_k L_k^T = G_k + E_k,$$

其中 E_k 是一个对角阵. 然后解

$$\bar{G}_k d = L_k D_k L_k^T d = -g_k,$$

求出 d_k .

牛顿法的另一个困难是当 $g_k = 0$ 且 G_k 不正定时, x_k 通常不是极小点, 而是鞍点. 事实上, 当 $g_k = 0$ 时, 我们有

$$\begin{aligned} f(x) &= f(x_k) + g_k^T d_k + \frac{1}{2} d_k^T G_k d_k + \dots \\ &= f(x_k) + \frac{1}{2} d_k^T G_k d_k + \dots \end{aligned}$$

这时若 G_k 正定, x_k 已经是极小点; 若 G_k 不正定, 就要继续寻找下降方向. 这时使 $d_k^T G_k d_k < 0$ 的负曲率方向 d_k 是使函数值下降的方向, 从而在这种情况下, 我们可以计算负曲率方向. 有关 Gill Murray 修改牛顿法的详细计算步骤, 有兴趣的读者可参见 Gill Murray, Wright (1981) 或邓乃杨 (1982).

§ 4.3 共轭梯度法

共轭梯度法是介于最速下降法与牛顿法之间的一个方法.它仅需利用一阶导数信息,但克服了最速下降法收敛慢的缺点,又避免了牛顿法需要存储和计算 Hesse 矩阵并求逆的缺点.共轭梯度法不仅是解大型线性方程组最有用的方法之一,也是解大型非线性最优化问题最有效的算法之一.

共轭梯度法最早是由 Hestenes 和 Stiefel(1952)提出来的,用于解正定系数矩阵的线性方程组.在这个基础上, Fletcher 和 Reeves(1964)首先提出了解非线性最优化问题的共轭梯度法.由于共轭梯度法不需要矩阵存储,且有较快的收敛速度和二次终止性等优点,现在共轭梯度法已经广泛地应用于实际问题中.

4.3.1 共轭方向法

共轭梯度法是共轭方向法的一种.所谓共轭方向法就是其所有的搜索方向都是互相共轭的方法.

定义 4.3.1 设 G 是 $n \times n$ 对称正定矩阵, d_1, d_2 是 n 维非零向量.如果

$$d_1^T G d_2 = 0, \quad (4.3.1)$$

则称向量 d_1 和 d_2 是 G -共轭的(或 G -直交的),简称共轭的.

设 d_1, d_2, \dots, d_m 是 \mathbf{R}^n 中任一组非零向量,如果

$$d_i^T G d_j = 0, \quad (i \neq j) \quad (4.3.2)$$

则称 d_1, d_2, \dots, d_m 是 G -共轭的,简称共轭的.

显然,如果 d_1, d_2, \dots, d_m 是 G -共轭的,则它们是线性无关的.如果 $G = I$,则共轭性就是通常的直交性.

一般共轭方向法步骤如下:

算法 4.3.2 (一般共轭方向法)

步 1. 给出初始点 x_0 , $\epsilon > 0$, $k = 0$.

计算 $g_0 = g(x_0)$ 和初始下降方向 d_0 , 使 $d_0^T g_0 < 0$.

步 2. 如果 $\|g_k\| < \epsilon$, 停止迭代.

步 3. 计算 α_k 和 x_{k+1} , 使得

$$f(x_k + \alpha_k d_k) = \min_0 f(x_k + \alpha d_k), \quad (4.3.3)$$

$$x_{k+1} = x_k + \alpha_k d_k. \quad (4.3.4)$$

步 4. 采用某种共轭方向法计算 d_{k+1} 使得

$$d_{k+1}^T G d_j = 0, \quad j = 0, 1, \dots, k.$$

步 5. 令 $k := k + 1$, 转步 2.

对于正定二次函数的极小化

$$\min_x f(x) = \frac{1}{2} x^T G x - b^T x, \quad (4.3.5)$$

它相当于解线性方程组

$$Gx = b, \quad (4.3.6)$$

其中 G 是 $n \times n$ 对称正定矩阵. 在这种情形, 精确线性搜索因子 α_k 的显式表示是

$$\alpha_k = - \frac{g_k^T d_k}{d_k^T G d_k}, \quad (4.3.7)$$

这里

$$g_k = Gx_k - b = r_k. \quad (4.3.8)$$

这表明, 在正定二次函数极小化的情形, 目标函数的梯度 $g(x)$ 与线性方程组 (4.3.6) 的残量 $r(x)$ 是一致的.

下面叙述共轭方向的基本定理, 它表明: 在精确线性搜索条件下, 共轭方向法具有二次终止性 (即对于正定二次函数, 方法是有限步终止的).

定理 4.3.3 设 $x_0 \in \mathbf{R}^n$ 是任意初始点. 对于极小化二次函数 (4.3.5), 共轭方向法至多经 n 步精确线性搜索终止; 且每一 x_{i+1} 都是 $f(x)$ 在 x_0 和方向 d_0, \dots, d_i 所张成的线性流形 $\{x \mid x = x_0 + \sum_{j=0}^i \alpha_j d_j, \alpha_j \in \mathbf{R}\}$ 中的极小点.

证明 因为 G 正定, 且共轭方向 d_0, d_1, \dots 线性无关, 故只需要证明对所有的 $i = n-1$, 有

$$g_{i+1}^T d_j = 0, \quad j = 0, \dots, i, \quad (4.3.9)$$

就可得出定理的两个结论.

事实上, 由于

$$g_{k+1} - g_k = G(x_{k+1} - x_k) = Gd_k \quad (4.3.10)$$

和在精确线性搜索 (4.3.5) 下,

$$g_{k+1}^T d_k = 0, \quad (4.3.11)$$

故有当 $j < i$ 时,

$$\begin{aligned} g_{i+1}^T d_j &= g_{j+1}^T d_j + \sum_{k=j+1}^i (g_{k+1} - g_k)^T d_j \\ &= g_{j+1}^T d_j + \sum_{k=j+1}^i \alpha_k d_k^T G d_j \\ &= 0. \end{aligned} \quad (4.3.12)$$

在 (4.3.12) 中两项为零分别由精确线性搜索和共轭性得到. 当 $j = i$ 时, 直接由精确线性搜索可知

$$g_{i+1}^T d_i = 0. \quad (4.3.13)$$

从而 (4.3.9) 成立.

4.3.2 共轭梯度法

共轭梯度法是一个典型的共轭方向法,它的每一个搜索方向是互相共轭的.而这些搜索方向 d_k 仅仅是负梯度方向 $-g_k$ 与上一次迭代的搜索方向 d_{k-1} 的组合.因此,存储量少,计算方便.

记

$$d_k = -g_k + \alpha_{k-1} d_{k-1}, \quad (4.3.14)$$

左乘 $d_{k-1}^T G$, 并使得 $d_{k-1}^T G d_k = 0$, 得

$$\alpha_{k-1} = \frac{g_k^T G d_{k-1}}{d_{k-1}^T G d_{k-1}}. \quad (\text{Hestenes Stiefel 公式}) \quad (4.3.15)$$

利用(4.3.10)和(4.3.11), 上式也可以写成

$$\alpha_{k-1} = \frac{g_k^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - g_{k-1})}, \quad (\text{Crowder Wolfe 公式}) \quad (4.3.16)$$

$$= \frac{g_k^T g_k}{g_{k-1}^T g_{k-1}}. \quad (\text{Fletcher Reeves 公式}) \quad (4.3.17)$$

另外三个常用的公式为

$$\alpha_{k-1} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T g_{k-1}}, \quad (\text{Polak Ribiere Polyak 公式}) \quad (4.3.18)$$

$$\alpha_{k-1} = -\frac{g_k^T g_k}{d_{k-1}^T g_{k-1}}, \quad (\text{Dixon 公式}) \quad (4.3.19)$$

$$\alpha_{k-1} = \frac{g_k^T g_k}{d_{k-1}^T (g_k - g_{k-1})}. \quad (\text{Dai - Yuan 公式}) \quad (4.3.20)$$

对于正定二次函数,若采用精确线性搜索,以上几个关于 α_k 的共轭梯度公式等价.但在实际计算中,FR 公式和 PRP 公式最常用.

注意到对于正定二次函数,

$$g_k = Gx_k - b \geq r_k, \quad (4.3.21)$$

其中 r_k 是方程组 $Gx_k = b$ 的残量, 以及

$$r_{k+1} - r_k = \alpha_k G d_k, \quad \alpha_k = -\frac{g_k^T d_k}{d_k^T G d_k} = \frac{r_k^T r_k}{d_k^T G d_k}, \quad (4.3.22)$$

我们可以给出如下关于正定二次函数极小化的共轭梯度法.

算法 4.3.4 (共轭梯度法)

步 1. 初始步: 给出 x_0 , $\epsilon > 0$, 计算 $r_0 = Gx_0 - b$, 令 $d_0 = -r_0$, $k = 0$.

步 2. 如果 $\|r_k\| < \epsilon$, 停止.

步 3. 计算

$$\alpha_k = \frac{r_k^T r_k}{d_k^T G d_k}, \quad (4.3.23)$$

$$x_{k+1} = x_k + \alpha_k d_k, \quad (4.3.24)$$

$$r_{k+1} = r_k + {}_k G d_k,$$

$$\alpha_k = \frac{r_{k+1}^T r_{k+1}}{r_k^T r_k}, \quad (4.3.25)$$

$$d_{k+1} = -r_{k+1} + \alpha_k d_k. \quad (4.3.26)$$

步 4. 令 $k = k+1$, 转步 2.

例 4.3.5 用 FR 共轭梯度法解极小化问题

$$\min f(x) = \frac{3}{2} x_1^2 + \frac{1}{2} x_2^2 - x_1 x_2 - 2 x_1.$$

解 将 $f(x)$ 写成 $f(x) = \frac{1}{2} x^T G x - b^T x$ 的形式, 有

$$G = \begin{pmatrix} 3 & -1 \\ -1 & 1 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, r(x) = Gx - b.$$

$$\text{设 } x^0 = (-2, 4)^T, \text{ 得 } r_0 = \begin{pmatrix} -12 \\ 6 \end{pmatrix}, d_0 = -r_0 = \begin{pmatrix} 12 \\ -6 \end{pmatrix},$$

$$\alpha_0 = \frac{r_0^T r_0}{d_0^T G d_0} = \frac{5}{17},$$

$$x^1 = x^0 + \alpha_0 d_0 = \begin{pmatrix} -2 \\ 4 \end{pmatrix} + \frac{5}{17} \begin{pmatrix} 12 \\ -6 \end{pmatrix} = \begin{pmatrix} \frac{26}{17} \\ \frac{38}{17} \end{pmatrix}.$$

$$r_1 = \begin{pmatrix} \frac{6}{17} \\ \frac{12}{17} \end{pmatrix}, \alpha_1 = \frac{r_1^T r_1}{r_1^T r_1} = \frac{1}{289},$$

$$d_1 = -r_1 + \alpha_1 d_0 = -\begin{pmatrix} \frac{6}{17} \\ \frac{12}{17} \end{pmatrix} + \frac{1}{289} \begin{pmatrix} 12 \\ -6 \end{pmatrix} = \begin{pmatrix} -\frac{90}{289} \\ -\frac{210}{289} \end{pmatrix}.$$

$$\alpha_1 = \frac{r_1^T r_1}{d_1^T G d_1} = \frac{17}{10},$$

$$x^2 = x^1 + \alpha_1 d_1 = \begin{pmatrix} \frac{26}{17} \\ \frac{38}{17} \end{pmatrix} + \frac{17}{10} \begin{pmatrix} -\frac{90}{289} \\ -\frac{210}{289} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

$$r_2 = Gx^2 - b = \begin{pmatrix} 2 \\ 0 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 0.$$

从而, $x^2 = (1, 1)^T$ 是所求的极小点.

下面的定理包含了共轭梯度法的基本性质. 为方便起见, 我们仍用 g_k 代替

r_k 叙述这个定理. 在这个定理中 m 是满足 $g_i \neq 0$ 的最大正整数.

定理 4.3.6 对于正定二次函数, 采用精确线性搜索的共轭梯度法在 $m \leq n$ 步后终止, 且对 $1 \leq i \leq m$ 成立下列关系式:

$$d_i^T G d_j = 0, \quad j = 0, 1, \dots, i-1, \quad (4.3.27)$$

$$g_i^T g_j = 0, \quad j = 0, 1, \dots, i-1, \quad (4.3.28)$$

$$d_i^T g_i = -g_i^T g_i, \quad (4.3.29)$$

$$[g_0, g_1, \dots, g_i] = [g_0, Gg_0, \dots, G^i g_0], \quad (4.3.30)$$

$$[d_0, d_1, \dots, d_i] = [g_0, Gg_0, \dots, G^i g_0], \quad (4.3.31)$$

其中 $[g_0, g_1, \dots, g_i]$ 和 $[d_0, d_1, \dots, d_i]$ 分别表示由 g_0, g_1, \dots, g_i 及 d_0, d_1, \dots, d_i 张成的子空间, $[g_0, Gg_0, \dots, G^i g_0]$ 表示 g_0 的 i 阶 Krylov 子空间.

证明 我们先用归纳法来证明 (4.3.27) - (4.3.29).

当 $i=1$ 时, (4.3.27) 和 (4.3.28) 成立; 当 $d_0 = -g_0$ 时, (4.3.29) 成立. 今设这些关系式对某个 $i < m$ 成立, 我们证明对于 $i+1$, 这些关系式也成立.

由于

$$g_{i+1} = g_i + \alpha_i G d_i \quad (4.3.32)$$

和

$$\alpha_i = -\frac{g_i^T d_i}{d_i^T G d_i} = -\frac{g_i^T g_i}{d_i^T G d_i} \neq 0, \quad (4.3.33)$$

利用 (4.3.32) 和 (4.3.14), 得

$$\begin{aligned} g_{i+1}^T g_j &= g_i^T g_j + \alpha_i d_i^T G g_j \\ &= g_i^T g_j - \alpha_i d_i^T G (d_j - \beta_{j-1} d_{j-1}). \end{aligned} \quad (4.3.34)$$

当 $j=i$ 时,

$$g_{i+1}^T g_i = g_i^T g_i - \frac{g_i^T g_i}{d_i^T G d_i} d_i^T G d_i = 0.$$

当 $j < i$ 时, (4.3.34) 直接由归纳法假设可知为零. 这样, (4.3.28) 得证.

再由 (4.3.14) 和 (4.3.32), 有

$$\begin{aligned} d_{i+1}^T G d_j &= -g_{i+1}^T G d_j + \alpha_i d_i^T G d_j \\ &= g_{i+1}^T (g_j - g_{j+1}) \beta_j + \alpha_i d_i^T G d_j. \end{aligned} \quad (4.3.35)$$

当 $j=i$ 时, 由 (4.3.28), (4.3.33) 和 (4.3.17), 上式成为

$$d_{i+1}^T G d_i = -\frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} d_i^T G d_i + \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} d_i^T G d_i = 0.$$

当 $j < i$ 时, 直接由归纳法假设可知 (4.3.35) 为零. 于是 (4.3.27) 得证.

由 (4.3.14) 和精确线性搜索, 有

$$\begin{aligned} d_{i+1}^T g_{i+1} &= -g_{i+1}^T g_{i+1} + \alpha_i d_i^T g_{i+1} \\ &= -g_{i+1}^T g_{i+1}, \end{aligned}$$

因此, (4.3.29) 成立.

下面我们用数学归纳法证明(4.3.30)和(4.3.31)成立.

显然, 当 $i=0$ 时, 结论成立. 今假定(4.3.30)和(4.3.31)对某个 i 成立, 我们证明这些结论对于 $i+1$ 也成立.

由归纳法假设可知,

$$g_i \in [g_0, Gg_0, \dots, G^i g_0], \quad Gd_i \in [Gg_0, \dots, G^{i+1} g_0].$$

所以

$$g_{i+1} = g_i + \alpha_i Gd_i \in [g_0, Gg_0, \dots, G^{i+1} g_0].$$

从而

$$[g_0, \dots, g_{i+1}] \subset [g_0, Gg_0, \dots, G^{i+1} g_0]. \quad (4.3.36)$$

反之, 由归纳法假设(4.3.31),

$$G^{i+1} g_0 = G(G^i g_0) \in [Gd_0, Gd_1, \dots, Gd_i].$$

利用(4.3.32)有 $Gd_j = (g_{j+1} - g_j)/\alpha_j, j=1, \dots, i$, 故

$$G^{i+1} g_0 \in [g_0, g_1, \dots, g_i, g_{i+1}].$$

再利用归纳法假设(4.3.30), 得

$$[g_0, Gg_0, \dots, G^{i+1} g_0] \subset [g_0, g_1, \dots, g_i, g_{i+1}]. \quad (4.3.37)$$

关系式(4.3.36)和(4.3.37)表明(4.3.30)对 $i+1$ 成立.

类似地, 可以证明(4.3.31)对 $i+1$ 也成立.

最后, 方法在 $m=n$ 步后终止可由(4.3.28)看出, 这里显然有 $g_n = 0$.

请注意, 定理的证明依赖于 $d_0 = -g_0$, 即初始搜索方向 d_0 取最速下降方向. 在这个定理中, (4.3.27)表示搜索方向的共轭性, (4.3.28)表示梯度的直交性, (4.3.29)表示下降条件, (4.3.30)和(4.3.31)表示方向向量和梯度向量之间的关系.

本小节结束, 我们提一下共轭梯度法的收敛速度. 定义

$$\kappa(G) = \frac{\lambda_{\max}(G)}{\lambda_{\min}(G)} = \frac{G^{-1}}{G} = 1/\mu_n.$$

可以证明

$$\|x_k - x^*\|_G \leq \frac{(\kappa(G) - 1)^k}{(\kappa(G) + 1)} \|x_0 - x^*\|_G. \quad (4.3.38)$$

与最速下降法的收敛速度(4.1.13)比较一下, 这时的收敛速度依赖于 $\kappa(G)$ 而不是 $\lambda(G)$.

(4.3.38)告诉我们, 共轭梯度法的收敛速度依赖于 $\kappa(G)$. 如果我们能够采取某种措施改善 G 的条件数, 则收敛可以加快. 下面的预处理技术提供了解决这个问题的一个办法.

设 C 是某个非奇异矩阵, 令

$$\hat{x} = Cx. \quad (4.3.39)$$

则二次目标函数(4.3.5)变成

$$\min \hat{f}(\hat{x}) = \frac{1}{2} \hat{x}^T (C^T G C^{-1}) \hat{x} - (C^T b)^T \hat{x}. \quad (4.3.40)$$

这等价于解线性方程组

$$(C^T G C^{-1}) \hat{x} = C^T b. \quad (4.3.41)$$

这时,方法的收敛速度依赖于 $C^T G C^{-1}$ 的条件数而不是 G 的条件数.如果我们能够选择非奇异矩阵 C 使得 $C^T G C^{-1}$ 的条件数明显好于 G 的条件数,则收敛速度将明显改善.根据这个思想,我们定义 $M = C^T C$, 并给出下面的预处理共轭梯度法.

算法 4.3.7 (预处理共轭梯度法)

步 1. 初始步: 给出 x_0 , $\epsilon > 0$, 预处理矩阵 M , 令 $r_0 = Gx_0 - b$. 解 $My = r_0$ 得 y_0 , 令 $d_0 = -r_0$, $k = 0$.

步 2. 如果 $\|r_k\| < \epsilon$, 停止.

步 3. 计算

$$\alpha_k = \frac{r_k^T y_k}{d_k^T G d_k}, \quad (4.3.42)$$

$$x_{k+1} = x_k + \alpha_k d_k, \quad (4.3.43)$$

$$r_{k+1} = r_k + \alpha_k G d_k,$$

解 $My = r_{k+1}$ 得 y_{k+1} ,

$$\alpha_k = \frac{r_{k+1}^T y_{k+1}}{r_k^T y_k}, \quad (4.3.44)$$

$$d_{k+1} = -y_{k+1} + \alpha_k d_k. \quad (4.3.45)$$

步 4. 令 $k = k+1$, 转步 2.

显然, 如果令 $M = I$, 则预处理共轭梯度算法 4.3.7 就是标准的共轭梯度算法 4.3.4.

4.3.3 对于非二次函数的共轭梯度法

在上一小节, 我们讨论了极小化正定二次函数的共轭梯度法. 现在, 我们把这个方法推广到处理一般非二次函数. Fletcher 和 Reeves (1964) 首先提出了极小化非二次函数的共轭梯度法.

算法 4.3.8 (FR - CG)

步 1. 初始步: 给出 x_0 , $\epsilon > 0$, 计算 $f_0 = f(x_0)$, $g_0 = \nabla f(x_0)$, 令 $d_0 = -g_0$, $k = 0$.

步 2. 如果 $\|g_k\| < \epsilon$, 停止.

步 3. 由线性搜索求步长因子 α_k , 并令

$$x_{k+1} = x_k + \alpha_k d_k.$$

步 4.

$$\alpha_k = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k},$$

$$d_{k+1} = -g_{k+1} + \alpha_k d_k.$$

步 5. 令 $k = k+1$, 转步 2.

这个算法由于程序简单, 计算量小, 没有矩阵存储与计算, 是解大型非线性规划的首选方法.

下面我们讨论 FR 共轭梯度法的总体收敛性.

定理 4.3.9 (FR 共轭梯度法总体收敛性定理)

假定 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 在有界水平集 $L = \{x \in \mathbf{R}^n \mid f(x) \leq f(x_0)\}$ 上连续可微, 且有下界, 那么采用精确线性搜索的 Fletcher-Reeves 共轭梯度法产生的序列 $\{x_k\}$ 至少有一个聚点是驻点, 即

(1) 当 $\{x_k\}$ 是有穷点列时, 其最后一个点是 $f(x)$ 的驻点.

(2) 当 $\{x_k\}$ 是无穷点列时, 它必有聚点, 且任一聚点都是 $f(x)$ 的驻点.

证明 (1) 当 $\{x_k\}$ 是有穷点列时, 由算法的终止性条件可知, 其最后一个点 x^* 满足 $g(x^*) = 0$, 故 x^* 是 $f(x)$ 的驻点.

(2) 当 $\{x_k\}$ 是无穷点列时, 对所有 k , $g_k \neq 0$. 由于 $d_k = -g_k + \alpha_{k-1} d_{k-1}$, 故

$$g_k^T d_k = -\|g_k\|^2 + \alpha_{k-1} g_k^T d_{k-1} = -\|g_k\|^2 < 0, \quad (4.3.46)$$

从而 d_k 是下降方向. 由于 $\{f(x_k)\}$ 是单调下降且有下界的序列, 故 $\lim_k f(x_k) = f^*$ 存在. 由 $\{x_k\}$ 水平集 L 及 L 有界可知 $\{x_k\}$ 是有界点列, 故存在收敛子列 $\{x_k\}_{K_1} \rightarrow x^*$, 这里 K_1 是子序列的指标集. 由于 $\{x_k\}_{K_1} \subset \{x_k\}$, 故 $\{f(x_k)\}_{K_1} \subset \{f(x_k)\}$, 从而由 f 的连续性可知, 对于 $k \in K_1$, 有

$$f(x^*) = f(\lim_{k \in K_1} x_k) = \lim_{k \in K_1} f(x_k) = f^*. \quad (4.3.47)$$

类似地, $\{x_{k+1}\}$ 也是有界点列, 故存在 $\{\bar{x}_{k+1}\}_{K_2} \rightarrow \bar{x}^*$, 这里 K_2 是 $\{x_{k+1}\}$ 的子序列的指标集. 故对于 $k+1 \in K_2$,

$$f(\bar{x}^*) = f(\lim_{k+1 \in K_2} x_{k+1}) = \lim_{k+1 \in K_2} f(x_{k+1}) = f^*. \quad (4.3.48)$$

于是

$$f(\bar{x}^*) = f(x^*). \quad (4.3.49)$$

下面用反证法证明 $g(x^*) = 0$. 假定 $g(x^*) \neq 0$, 则存在一个方向 d^* , 使得对于充分小的 α , 有

$$f(x^* + \alpha d^*) < f(x^*). \quad (4.3.50)$$

由于

$$f(x_{k+1}) = f(x_k + \alpha_k d_k) = f(x_k + d_k), \quad \alpha_k > 0,$$

故对于 $k+1 \leq K_2$, 令 k , 并利用(4.3.50)得

$$f(\bar{x}^*) = f(x^* + d^*) < f(x^*), \quad (4.3.51)$$

这与(4.3.49)矛盾 这证明了 $g(x^*) = 0$, 即 x^* 是 $f(x)$ 的驻点.

关于 FR 共轭梯度法的收敛速度, 我们有下面的结果: 在定理 4.3.9 的条件下, 采用精确线性搜索的 FR 共轭梯度法产生的迭代点列 $\{x_k\}$ 线性收敛到 $f(x)$ 的极小点 x^* .

在 FR 共轭梯度算法 4.3.8 中, 为了保证方向是下降的, 必须要求

$$g_k^T d_k = -\|g_k\|^2 + \alpha_{k-1} g_k^T d_{k-1} < 0.$$

在精确线性搜索情形, 由于 $g_k^T d_{k-1} = 0$, 故 $g_k^T d_k < 0$ 是显然的; 但在非精确线性搜索情形, 可能有 $g_k^T d_k > 0$ 的情形发生, 从而 d_k 不是下降的. 幸运的是, 如果采用强 Wolfe 不精确线性搜索准则(3.5.9)-(3.5.10)产生步长因子 α_k , 则方向 d_k 肯定是下降方向.

再开始共轭梯度法

对于一般非二次函数, 共轭梯度法常常采用再开始技术, 即每 n 步以后周期性地采用最速下降方向作为新的搜索方向. 尤其是当迭代从一个非二次区域进入 $f(x)$ 可由二次函数很好地逼近的区域时, 重新取最速下降方向作为搜索方向, 则其后 n 次迭代方向接近于共轭方向, 从而使方法有较快的收敛速度. 这一点也可以从共轭梯度法的二次终止性依赖于取最速下降方向作为初始搜索方向这个事实略见一斑. 对于大型问题, 常常更经常地进行再开始, 例如每隔 r 步迭代再开始, 这里 $r < n$ 或 $rn \leq n$.

主要的再开始策略除了每隔 n 步迭代再开始外, 还有当 $g_k^T d_k > 0$, 即 d_k 是上升方向时再开始. 另外, 从(4.3.28)可知, 对于二次函数, 相邻二次迭代的梯度互相直交, 因此, 如果它们偏离直交性较大, 例如

$$\frac{g_k^T g_{k-1}}{\|g_k\| \|g_{k-1}\|} \geq \nu, \quad (4.3.52)$$

这里可取 $\nu = 0.1$, 则进行再开始.

算法 4.3.10 (再开始 FR 共轭梯度法)

步 1. 初始步: 给出初始点 x_0 , 容限 $\epsilon > 0$, $k = 0$.

步 2. 计算 $g_0 = g(x_0)$. 如果 $\|g_0\| \leq \epsilon$, 停止迭代, 输出 $x^* = x_0$; 否则令 $d_0 = -g_0$.

步 3. 线性搜索求 α_k , 并令

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k d_k, \\ k &= k+1. \end{aligned}$$

步 4. 计算 $g_k = g(x_k)$.

若 $\frac{g_k^T g_{k-1}}{g_k^T g_{k-1}} \geq 0.1$, 令 $x_0 := x_k$, 转步 2;

如果 $g_k = 0$, 停止迭代.

步 5. 若 $k = n$, 令 $x_0 = x_k$, 转步 2.

步 6. 计算

$$d_k = -\frac{g_k^T g_{k-1}}{g_{k-1}^T g_{k-1}} g_{k-1} - g_k + d_{k-1}.$$

步 7. 如果 $d_k^T g_k > 0$, 令 $x_0 = x_k$, 转步 2; 否则转步 3.

与共轭梯度法一样, 再开始共轭梯度法仍然具有总体收敛性, 并且至少有线性收敛速度. 我们还可以证明一个更强的结果: 再开始共轭梯度法产生的迭代点列具有 n 步二次收敛速度, 即

$$\|x_{k+n} - x^*\| = O(\|x_k - x^*\|^2). \quad (4.3.53)$$

§ 4.4 拟牛顿法

牛顿法成功的关键是利用了 Hesse 矩阵提供的曲率信息, 但计算 Hesse 矩阵工作量大, 并且有的目标函数的 Hesse 矩阵很难计算, 甚至不好求出, 这就导致了一个想法: 能否仅利用目标函数值和一阶导数的信息, 构造出目标函数的曲率近似, 使方法具有类似牛顿法的收敛速度快的优点. 拟牛顿法就是这样的一类算法. 由于它不需要二阶导数, 拟牛顿法往往比牛顿法更有效.

4.4.1 拟牛顿条件

和牛顿法的推导一样, 考虑目标函数 $f(x)$ 在当前点 x_k 处的二次模型

$$m_k(d) = f(x_k) + g_k^T d + \frac{1}{2} d^T B_k d, \quad (4.4.1)$$

其中, B_k 是 $n \times n$ 对称正定矩阵, 是 Hesse 近似, 它将在每次迭代中进行校正. 极小化这个二次模型得到

$$d_k = -B_k^{-1} g_k, \quad (4.4.2)$$

从而新的迭代点为

$$x_{k+1} = x_k + \alpha_k d_k = x_k - \alpha_k B_k^{-1} g_k, \quad (4.4.3)$$

其中, α_k 是线性搜索步长因子, 上述迭代(4.4.3)称为拟牛顿迭代, 它与牛顿迭代的主要区别在于在(4.4.3)中我们用 Hesse 近似 B_k 代替了牛顿迭代中的 Hesse 矩阵 G_k .

设 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 在开集 $D \subset \mathbf{R}^n$ 上二次连续可微, $f(x)$ 在 x_{k+1} 附近的二次近似

为

$$f(x) = f(x_{k+1}) + g_{k+1}^T (x - x_{k+1}) + \frac{1}{2} (x - x_{k+1})^T G_{k+1} (x - x_{k+1}), \quad (4.4.4)$$

对上式两边求导,有

$$g(x) = g_{k+1} + G_{k+1} (x - x_{k+1}). \quad (4.4.5)$$

令 $x = x_k$, 得

$$G_{k+1} (x_{k+1} - x_k) = g_{k+1} - g_k. \quad (4.4.6)$$

令

$$s_k = x_{k+1} - x_k, y_k = g_{k+1} - g_k, \quad (4.4.7)$$

(4.4.7)成为

$$G_{k+1} s_k = y_k. \quad (4.4.8)$$

显然,如果 $f(x)$ 是正定二次函数,上述关系式(4.4.8)精确成立.现在,我们要求在拟牛顿法中构造出来的 Hesse 近似 B_{k+1} 满足这种关系,从而得到

$$B_{k+1} s_k = y_k. \quad (4.4.9)$$

上式称为拟牛顿条件或拟牛顿方程.

如果令 $H_k = B_k^{-1}$, 则拟牛顿条件为

$$H_{k+1} y_k = s_k, \quad (4.4.10)$$

拟牛顿迭代为

$$x_{k+1} = x_k + s_k d_k = x_k - s_k B_k^{-1} g_k \quad (4.4.11)$$

或

$$x_{k+1} = x_k + s_k d_k = x_k - s_k H_k g_k. \quad (4.4.12)$$

拟牛顿条件使二次模型具有如下插值性质: 如果 B_{k+1} 满足拟牛顿条件(4.4.9), 那么在 x_{k+1} 点的二次模型

$$m_{k+1}(x) = f(x_{k+1}) + g_{k+1}^T (x - x_{k+1}) + \frac{1}{2} (x - x_{k+1})^T B_{k+1} (x - x_{k+1}) \quad (4.4.13)$$

满足

$$m_{k+1}(x_{k+1}) = f(x_{k+1}), \quad m_{k+1}(x_{k+1}) = g_{k+1}, \quad m_{k+1}(x_k) = g_k. \quad (4.4.14)$$

上式中第一、第二个等式是显然的, 第三个等式是利用拟牛顿条件(4.4.9)得到的.

一般的拟牛顿算法如下:

算法 4.4.1

步 1. 给出初始点 $x_0 \in \mathbf{R}^n$, B_0 (或 H_0) $\in \mathbf{R}^{n \times n}$, $\alpha_0 > 0$, $k = 0$.

步 2. 如果 $\|g_k\|$ 足够小, 停止.

步 3. 解 $B_k d = -g_k$ 得搜索方向 d_k ; (或计算 $d_k = -H_k g_k$).

步 4. 由线性搜索求步长因子 α_k , 并令 $x_{k+1} = x_k + \alpha_k d_k$.

步 5. 校正 B_k 产生 B_{k+1} (或校正 H_k 产生 H_{k+1}), 使得拟牛顿条件 (4.4.9) (或 (4.4.10)) 成立.

步 6. $k = k+1$, 转步 2.

在上述拟牛顿法中, 初始 Hesse 近似 B_0 通常取为单位矩阵, 即 $B_0 = I$, 这样, 拟牛顿法的第一次迭代等价于一个最速下降迭代.

拟牛顿法有下列优点:

(1) 仅需一阶导数 (牛顿法需二阶导数).

(2) B_k (或 H_k) 保持正定, 使得方法具有下降性质 (在牛顿法中, G_k 可能不正定).

(3) 每次迭代需 $O(n^2)$ 次乘法运算 (牛顿法需 $O(n^3)$ 次乘法运算).

(4) 搜索方向是相互共轭的, 从而具有二次终止性.

(5) 具有超线性收敛性.

正如牛顿法是在椭球范数 $\|\cdot\|_{G_k}$ 意义下的最速下降法一样, 拟牛顿法是在椭球范数 $\|\cdot\|_{B_k}$ 意义下的最速下降法, 这里

$$\|x\|_{B_k} = (x^T B_k x)^{\frac{1}{2}}, \quad x \in \mathbf{R}^n. \quad (4.4.15)$$

事实上, 由

$$f(x_k + d) = f(x_k) + g_k^T d + O(\|d\|^2) \quad (4.4.16)$$

知 $g_k^T d$ 越小, 目标函数下降得越快, 故极小化问题

$$\begin{aligned} \min \quad & g_k^T d, \\ \text{s.t.} \quad & \|d\|_{B_k} = 1 \end{aligned} \quad (4.4.17)$$

的解就是 $f(x)$ 在椭球范数 $\|\cdot\|_{B_k}$ 意义下在 x_k 处下降最快的方向. 由广义 Cauchy Schwartz 不等式,

$$(g_k^T d)^2 \leq (g_k^T B_k^{-1} g_k) (d^T B_k d), \quad (4.4.18)$$

且当 $d = -B_k^{-1} g_k$ 时等式成立, 这时 $g_k^T d$ 取极小. 于是得到问题 (4.4.17) 的解为

$$d = -\frac{B_k^{-1} g_k}{\|B_k^{-1} g_k\|_{B_k}} = -\frac{B_k^{-1} g_k}{(g_k^T B_k^{-1} g_k)^{\frac{1}{2}}}. \quad (4.4.19)$$

因此, 拟牛顿方向

$$d_k = -B_k^{-1} g_k \quad (4.4.20)$$

是在椭球范数 $\|\cdot\|_{B_k}$ 意义下 $f(x)$ 在 x_k 处的最速下降方向. 由于在每一次迭代中尺度矩阵 B_k 总是变化的, 故方法也叫做变尺度方法.

4.4.2 DFP 校正和 BFGS 校正

DFP 校正是第一个拟牛顿校正, 是 1959 年由 Davidon 提出的, 后来由

Fletcher 和 Powell(1963)解释和发展的 .BFGS 校正是目前最流行的也是最有效的拟牛顿校正,它是由 Broyden, Fletcher, Goldfarb 和 Shanno 在 1970 年各自独立提出的拟牛顿法.下面我们介绍这两个典型的拟牛顿法的推导和性质.

考虑 Hesse 逆近似序列 $\{H_k\}$. 设对称秩二校正为

$$H_{k+1} = H_k + auu^T + bv v^T, \quad (4.4.21)$$

其中 a 和 b 为给定的数. 令拟牛顿条件(4.4.10)满足, 则

$$H_k y_k + auu^T y_k + bv v^T y_k = s_k, \quad (4.4.22)$$

这里 u 和 v 并不唯一确定, 但 u 和 v 的明显的选择是

$$u = s_k, \quad v = H_k s_k.$$

于是, 从(4.4.22)得

$$au^T y_k = 1, \quad bv^T y_k = -1,$$

从而

$$a = \frac{1}{u^T y_k} = \frac{1}{s_k^T y_k}, \quad b = -\frac{1}{v^T y_k} = -\frac{1}{y_k^T H_k y_k}.$$

因此

$$H_{k+1} = H_k + \frac{s_k s_k^T}{s_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k}. \quad (4.4.23)$$

上述公式称为 DFP 校正公式(关于 H_k). 从这个公式我们可以看出, 序列 $\{H_k\}$ 的生成并不是每次迭代中重复计算的, 而是通过一种简单的校正从 H_k 生成 H_{k+1} , 这使得拟牛顿法是节省的和有效的. 完全类似地, 利用拟牛顿条件(4.4.9), 我们可以得到

$$B_{k+1} = B_k + \frac{y_k y_k^T}{y_k^T s_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}. \quad (4.4.24)$$

上述公式称为 BFGS 校正公式(关于 B_k).

容易看出, 只要通过对(4.4.23)作简单变换 $H \rightarrow B$ 和 $s \rightarrow y$, 关于 B_k 的 BFGS 校正就可得到. 因此, 我们也把(4.4.24)称为互补 DFP 公式.

如果对(4.4.24)两次应用逆的秩一校正的 Sherman - Morrison 公式, 就得到关于 H_k 的 BFGS 校正:

$$\begin{aligned} H_{k+1}^{(\text{BFGS})} &= H_k + \frac{(s_k - H_k g_k) s_k^T + s_k (s_k - H_k y_k)^T}{s_k^T y_k} - \\ &\quad \frac{(s_k - H_k y_k)^T y_k}{(s_k^T y_k)^2} s_k s_k^T \end{aligned} \quad (4.4.25)$$

$$= I - \frac{s_k y_k^T}{s_k^T y_k} H_k I - \frac{y_k s_k^T}{s_k^T y_k} + \frac{s_k s_k^T}{s_k^T y_k}. \quad (4.4.26)$$

进一步, 若将上式中 $H \rightarrow B$ 和 $s \rightarrow y$ 互换, 便得到关于 B_k 的 DFP 校正:

$$B_{k+1}^{(\text{DFP})} = B_k + \frac{(y_k - B_k s_k)^T y_k + y_k (y_k - B_k s_k)^T}{y_k^T s_k} - \frac{(y_k - B_k s_k)^T s_k}{(y_k^T s_k)^2} y_k y_k^T \quad (4.4.27)$$

$$= I - \frac{y_k s_k^T}{y_k^T s_k} B_k I - \frac{s_k y_k^T}{y_k^T s_k} + \frac{y_k y_k^T}{y_k^T s_k}. \quad (4.4.28)$$

下面,我们讨论 DFP 和 BFGS 校正的几个主要性质.

第一,在所有满足拟牛顿条件的矩阵中, $B_{k+1}^{(\text{DFP})}$ (或 $H_{k+1}^{(\text{BFGS})}$) 在某种范数的意义上,是最靠近当前矩阵 B_k (或 H_k) 的.

具体地说,由(4.4.28)定义的 DFP 校正 $B_{k+1}^{(\text{DFP})}$ 是极小化问题

$$\min_B \|B - B_k\|_{W, F}, \quad (4.4.29)$$

$$\text{s.t. } B = B^T, B s_k = y_k \quad (4.4.30)$$

的唯一解,其中 W 是满足 $W^2 y_k = s_k$ 的任何非奇异对称矩阵,

$$\|B\|_{W, F} = \|WBW\|_F \quad (4.4.31)$$

代表加权 Frobenius 范数.

类似地,由(4.4.26)定义的 BFGS 校正 $H_{k+1}^{(\text{BFGS})}$ 是极小化问题

$$\min_H \|H - H_k\|_{W, F}, \quad (4.4.32)$$

$$\text{s.t. } H = H^T, H y_k = s_k \quad (4.4.33)$$

的唯一解,其中 W 是满足 $W^2 s_k = y_k$ 的任何非奇异对称矩阵.

第二,使校正保持正定性是非常重要的性质.如果 $\{B_k\}$ 或 $\{H_k\}$ 正定,则 f 的局部二次模型就有唯一的局部极小点,这时搜索方向 $d_k = -H_k g_k$ 就是下降方向.因此,我们通常要求校正保持正定性,即若 H_k 正定,则 H_{k+1} 也正定.DFP 和 BFGS 校正都具有这样的性质.

定理 4.4.2 (DFP 校正的正定性) 当且仅当 $s_k^T y_k > 0$ 时, DFP 校正 (4.4.23) 保持正定性.

证明 用归纳法证明

$$z^T H_k z > 0, \quad \forall z \neq 0. \quad (4.4.34)$$

由初始选择, H_0 显然正定.今假定对于某个 k , 结论成立,并记 $H_k = LL^T$ 为 H_k 的 Cholesky 分解.设

$$a = L^T z, b = L^T y_k. \quad (4.4.35)$$

则

$$\begin{aligned} z^T H_{k+1} z &= z^T H_k z - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} z + z^T \frac{s_k s_k^T}{s_k^T y_k} z \\ &= a^T a - \frac{(a^T b)^2}{b^T b} + \frac{(z^T s_k)^2}{s_k^T y_k}. \end{aligned} \quad (4.4.36)$$

由 Cauchy 不等式知

$$a^T a - \frac{(a^T b)^2}{b^T b} \geq 0, \quad (4.4.37)$$

又由题设 $s_k^T y_k > 0$, 故(4.4.36)中第二项非负, 从而

$$z^T H_{k+1} z \geq 0. \quad (4.4.38)$$

下面证明(4.4.36)右边二项中至少一项严格大于零. 假定(4.4.36)中第一项等于零, 由 Cauchy 不等式, 这等价于 a 与 b 平行, 即 z 与 y_k 平行. 设 $z = \alpha y_k$, $\alpha = 0$, 这时,

$$\frac{(z^T s_k)^2}{s_k^T y_k} = \alpha^2 s_k^T y_k > 0,$$

这表明当(4.4.36)中第一项为零时, 第二项严格大于零.

假定(4.4.36)中第二项等于零, 由于 $s_k^T y_k > 0$, 则有 $z^T s_k = 0$, 但在这种情形, z 与 y_k 必不平行, 从而第一项必大于零, 因为否则的话, $z = \alpha y_k$, $\alpha = 0$, 及题设 $s_k^T y_k > 0$ 将导致 $z^T s_k = \alpha y_k^T s_k = 0$, 这是矛盾的. 于是, 对于任何 $z \neq 0$, 总有 $z^T H_{k+1} z > 0$ 成立. 从而充分性得证.

必要性可以类似地证明.

这个定理对 BFGS 校正同样成立.

由上述定理可知, $s_k^T y_k > 0$ 对于保持 DFP 和 BFGS 校正的正定性是关键性的, 同时, 这个条件也是实际的, 并且是可以满足的. 对于正定二次函数, 我们有

$$s_k^T y_k = s_k^T G s_k > 0. \quad (4.4.39)$$

对于一般函数,

$$s_k^T y_k = g_{k+1}^T s_k - g_k^T s_k. \quad (4.4.40)$$

注意到 $g_k^T s_k < 0$ (s_k 是下降步) 和当采用精确线性搜索时, $g_{k+1}^T s_k = 0$, 从而 $s_k^T y_k > 0$.

当采用 Wolfe 不精确线性搜索准则(3.5.7)-(3.5.8)时, 有

$$g_{k+1}^T s_k = \rho g_k^T s_k, \quad (\rho < 1).$$

因此, 注意到 d_k 是下降方向, $g_k^T d_k < 0$, 则

$$y_k^T s_k = (-\rho - 1) g_k^T d_k > 0. \quad (4.4.41)$$

一般地, 适当提高线性搜索的精度, 就可以使得 $g_{k+1}^T s_k$ 在数量上小到所要求的程度, 从而使得 $s_k^T y_k > 0$.

第三, DFP 方法和 BFGS 方法具有二次终止性, 即, 对于正定二次函数, DFP 方法和 BFGS 方法产生的方向是共轭的, 方法至多 n 步终止. 这个结论的证明与 § 4.3 共轭梯度法的二次终止性定理类似.

下面给出用 BFGS 方法极小化函数的例子, 它也表明 BFGS 方法是目前最好的拟牛顿法.

例 4.4.3 用 BFGS 方法极小化函数

$$\min f(x_1, x_2) = (x_1 - 2)^4 + (x_1 - 2)^2 x_2^2 + (x_2 + 1)^2.$$
$$x_0 = (1, 1)^T, \quad x^* = (2, -1)^T.$$

解 采用 BFGS 方法

$$x_{k+1} = x_k - B_k^{-1} g_k,$$
$$s_k = x_{k+1} - x_k, \quad y_k = g_{k+1} - g_k,$$
$$B_{k+1} = B_k + \frac{y_k y_k^T}{s_k^T y_k} - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k}.$$

初始点 $x_0 = (1, 1)^T$, $B_0 = \frac{1}{2} f''(x_0)$. 计算结果如下:

x_k	$f(x_k)$
$x_0 = (1, 1)^T$	6.0
$x_1 = (1, -0.5)^T$	1.5
$x_2 = (1.45, -0.3875)^T$	5.12×10^{-1}
$x_3 = (1.5889290, -0.63729087)^T$	2.29×10^{-1}
$x_4 = (1.8254150, -0.97155747)^T$	3.05×10^{-2}
$x_5 = (1.9460278, -1.0705597)^T$	8.33×10^{-3}
$x_6 = (1.9641387, -1.0450875)^T$	3.44×10^{-3}
$x_7 = (1.9952140, -1.0017148)^T$	2.59×10^{-5}
$x_8 = (2.0000653, -1.0004294)^T$	1.89×10^{-7}
$x_9 = (1.9999853, -0.99995203)^T$	2.52×10^{-9}
$x_{10} = (2.0, -1.0)^T$	0

上面的计算结果表明,用 BFGS 方法经 10 步迭代达到极小点;同样地,从相同的初始点出发,DFP 方法需要 12 步迭代达到极小点;而牛顿法需要 6 次或 7 次迭代达到极小点.

上面我们讨论了 DFP 和 BFGS 校正,它们都是由 y_k 和 $B_k s_k$ 组成的秩二校正. 本小节最后我们提出由 DFP 和 BFGS 校正加权组合产生的一类校正族

$$B_{k+1} = \alpha_k B_{k+1}^{DFP} + (1 - \alpha_k) B_{k+1}^{BFGS} \tag{4.4.42}$$

$$= B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k} + \alpha_k (s_k^T B_k s_k) w_k w_k^T. \tag{4.4.43}$$

其中 α_k 为实参数,

$$w_k = \frac{y_k}{y_k^T s_k} - \frac{B_k s_k}{s_k^T B_k s_k}. \tag{4.4.44}$$

这类校正称为 Broyden 族.显然 $\alpha_k = 0$, 就是 BFGS 校正; $\alpha_k = 1$, 就是 DFP 校正. 由于 DFP 和 BFGS 校正都是保持正定的,因此,若 $y_k^T s_k > 0$, 只要 $\alpha_k \geq 0$, 则 Broyden 族校正保持正定.

4.4.3 拟牛顿法的收敛性*

本小节为选学内容. 本小节我们讨论拟牛顿法的总体收敛性和局部收敛性. 在定理 4.4.5 中, 我们给出总体收敛性定理, 它表明采用精确线性搜索或 Wolfe 不精确线性搜索的算法 4.4.1 产生的迭代序列 $\{x_k\}$ 收敛到极小点 x^* . 在定理 4.4.5 中, 我们给出了拟牛顿法超线性收敛的充分必要条件. 定理 4.4.6 指明 BFGS 方法满足这个充分必要条件, 因而具有超线性收敛性. 由于这个证明较长, 我们没有给出. 拟牛顿法的收敛性是重要的, 我们希望学生至少了解这些定理的内容, 不要求掌握这些定理的证明. 这里给出的证明仅供教师和感兴趣的同学参考.

设 x_0 是任意初始点, B_0 是对称正定的初始 Hesse 近似.

假设 4.4.4

(a) $f: \mathbf{R}^n \rightarrow \mathbf{R}^1$ 在开凸集 $D \subset \mathbf{R}^n$ 上二次连续可微;

(b) 水平集 $\{x \in \mathbf{R}^n / f(x) \leq f(x_0)\}$ 是凸的, 存在正的常数 m 和 M 使得 Hesse 矩阵 $G(x)$ 满足

$$m \|z\|^2 \leq z^T G(x) z \leq M \|z\|^2, \quad \forall z \in \mathbf{R}^n, x \in D. \quad (4.4.45)$$

(c) 在 x^* 的邻域 $N(x^*, \delta)$ 内, $G(x)$ 是 Lipschitz 连续的, 即

$$\|G(x) - G(\bar{x})\| \leq L \|x - \bar{x}\|, \quad \forall x, \bar{x} \in N(x^*, \delta). \quad (4.4.46)$$

上述假设条件(b)意味着 Hesse 矩阵 $G(x)$ 在 D 上是正定的, f 有唯一的极小点 x^* .

由 Taylor 定理,

$$g(x_k + s_k) = g(x_k) + \int_0^1 G(x_k + \theta s_k) d\theta,$$

令

$$\bar{G}_k = \int_0^1 G(x_k + \theta s_k) d\theta = \int_0^1 G(x_k + \theta d_k) d\theta, \quad (4.4.47)$$

则

$$y_k = g_{k+1} - g_k = \bar{G}_k s_k. \quad (4.4.48)$$

于是, 利用(4.4.45)和(4.4.48), 有

$$\frac{y_k^T s_k}{s_k^T s_k} = \frac{s_k^T \bar{G}_k s_k}{s_k^T s_k} \geq m, \quad (4.4.49)$$

令 $z_k = \bar{G}_k^{-\frac{1}{2}} s_k$, 则

$$\frac{y_k^T y_k}{y_k^T s_k} = \frac{z_k^T \bar{G}_k z_k}{z_k^T z_k} \leq M. \quad (4.4.50)$$

注意到矩阵 A 的迹 $\text{trace}(A)$ 是 A 的对角元的和, 也是 A 的特征值的和, 即

$$\text{trace}(\mathbf{A}) = \sum_{i=1}^n a_{ii}, \quad \text{trace}(\mathbf{A}) = \sum_{i=1}^n \lambda_i; \quad (4.4.51)$$

矩阵的行列式 $\det(\mathbf{A})$ 是 \mathbf{A} 的特征值的乘积, 即

$$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i. \quad (4.4.52)$$

在下面的定理的证明中, 我们利用了这两个概念来估计 Hesse 近似的最大和最小特征值的大小.

定理 4.4.5 设 B_0 是任意初始对称正定矩阵, x_0 是初始点, 使得假设 4.4.4(a)(b) 成立. 则由算法 4.4.1 产生的序列 $\{x_k\}$ 收敛到 f 的极小点 x^* .

证明 定义

$$m_k = \frac{y_k^T s_k}{s_k^T s_k}, \quad M_k = \frac{y_k^T y_k}{y_k^T s_k}. \quad (4.4.53)$$

由 (4.4.49) 和 (4.4.50) 得

$$m_k \leq m, \quad M_k \leq M. \quad (4.4.54)$$

由 BFGS 校正 (4.4.24), 计算其迹和行列式, 得

$$\text{Trace}(B_{k+1}) = \text{Trace}(B_k) - \frac{B_k s_k^2}{s_k^T B_k s_k} + \frac{y_k^2}{y_k^T s_k} \quad (4.4.55)$$

和

$$\det(B_{k+1}) = \det(B_k) \frac{y_k^T s_k}{s_k^T B_k s_k}. \quad (4.4.56)$$

定义

$$\cos \theta_k = \frac{s_k^T B_k s_k}{s_k^T s_k}, \quad q_k = \frac{s_k^T B_k s_k}{s_k^T s_k}. \quad (4.4.57)$$

这里 θ_k 是 s_k 和 $B_k s_k$ 之间的夹角. 于是,

$$\frac{B_k s_k^2}{s_k^T B_k s_k} = \frac{B_k s_k^2}{(s_k^T B_k s_k)^2} \frac{s_k^2}{s_k^2} \frac{s_k^T B_k s_k}{s_k^2} = \frac{q_k}{\cos^2 \theta_k}. \quad (4.4.58)$$

又由 (4.4.56) 和 (4.4.53), 有

$$\det(B_{k+1}) = \det(B_k) \frac{y_k^T s_k}{s_k^T s_k} \frac{s_k^T s_k}{s_k^T B_k s_k} = \det(B_k) \frac{m_k}{q_k}. \quad (4.4.59)$$

现在, 我们定义

$$\phi(B) = \text{Trace}(B) - \ln(\det(B)), \quad (4.4.60)$$

其中 $\ln(\cdot)$ 表示自然对数. 不难证明, $\phi(B) > 0$. 由 (4.4.60), (4.4.55) - (4.4.59), 得

$$\begin{aligned} \phi(B_{k+1}) &= \text{Trace}(B_{k+1}) - \ln(\det(B_{k+1})) \\ &= \text{Trace}(B_k) + M_k - \frac{q_k}{\cos^2 \theta_k} - \ln(\det(B_k)) - \ln m_k + \ln q_k \\ &= \phi(B_k) + (M_k - \ln m_k - 1) + \end{aligned}$$

$$1 - \frac{q_k}{\cos^2 \theta_k} + \ln \frac{q_k}{\cos^2 \theta_k} + \ln \cos^2 \theta_k. \quad (4.4.61)$$

注意到对所有 $t > 0$, $h(t) = 1 - t + \ln t \leq 0$, 上面方括号中的项是非正的, 因而由 (4.4.54), 并反复利用 (4.4.61), 得

$$0 < (B_{k+1}) - (B_1) + ck + \sum_{j=1}^k \ln \cos^2 \theta_j, \quad (4.4.62)$$

其中 $c = M - \ln m - 1 > 0$.

下面, 我们利用 Wolfe 不精确线性搜索的总体收敛性定理证明结果.

由于 $s_k = -B_k^{-1}g_k$, $B_k s_k = -g_k$, 故 $s_k^T B_k s_k = -g_k^T B_k^{-1}g_k = -g_k^T d_k$, 这表明由 (4.4.57) 定义的 s_k 也是最速下降方向 $-g_k$ 和拟牛顿搜索方向 $d_k = -B_k^{-1}g_k$ 之间的夹角. 于是, 由 Wolfe 不精确线性搜索的总体收敛性定理得

$$g_k^T s_k \cos \theta_k \geq 0. \quad (4.4.63)$$

为了要证明 $\liminf_k g_k^T s_k = 0$, 只要证明存在子序列 $\{k_j\}$, 使得 $\{\cos \theta_{k_j}\} \geq 0$.

假定 $\cos \theta_{k_j} > 0$. 则存在 $k_1 > 0$, 使得对所有 $j > k_1$, 有

$$\ln \cos^2 \theta_j < -2c, \quad (4.4.64)$$

其中 $c = M - \ln m - 1 > 0$ 是上面定义的常数. 利用 (4.4.62) 得到: 对所有 $k > k_1$,

$$\begin{aligned} 0 &< (B_1) - (B_k) + kc + \sum_{j=1}^{k_1} \ln \cos^2 \theta_j + \sum_{j=k_1+1}^k (-2c) \\ &= (B_1) - (B_k) + \sum_{j=1}^{k_1} \ln \cos^2 \theta_j + 2k_1 c - kc. \end{aligned} \quad (4.4.65)$$

在 (4.4.65) 中, 第一项和第三项是正的, 但有限, 第二项小于零, 第四项也小于零且与 k 有关, 故当 k 充分大时, 上式右边是负的, 从而给出矛盾.

这个矛盾表明存在子序列 $\{k_j\}$, 使得 $\{\cos \theta_{k_j}\} \geq 0$. 从而

$$\liminf_k g_k^T s_k = 0. \quad (4.4.66)$$

由假设 4.4.4(b), 问题是强凸的, 这表明 $\{x_k\} \rightarrow x^*$.

上述定理证明了: 采用 Wolfe 不精确线性搜索的 BFGS 拟牛顿算法是总体收敛的. 这个结果可推广到所有 $\alpha_k \in [0, 1)$ 的 Broyden 族, 即不包括 DFP 校正.

下面, 我们研究 BFGS 方法的局部超线性收敛性. 首先给出拟牛顿法超线收敛的充分必要条件.

定理 4.4.6 设 $f: \mathbf{R}^n \rightarrow \mathbf{R}$, 满足假设 4.4.4(a)(b) 成立. 考虑迭代 $x_{k+1} = x_k + d_k$, $d_k = -B_k^{-1}g_k$. 设 $\{x_k\}$ 收敛到解点 x^* . 则当且仅当

$$\lim_k \frac{(B_k - G(x^*))d_k}{d_k} = 0 \quad (4.4.67)$$

时, 序列 $\{x_k\}$ 超线性收敛到 x^* .

证明 设拟牛顿步为 $d_k = -B_k^{-1}g_k$, 牛顿步为 $d_k^N = -G_k^{-1}g_k$. 由于 $G(x^*)$ 是

正定的,故当 x_k 充分靠近 x^* 时, G_k^{-1} 是上有界的.我们先证明(4.4.67)等价于

$$d_k - d_k^N = o(\|d_k\|). \quad (4.4.68)$$

假定(4.4.67)成立,则

$$\begin{aligned} d_k - d_k^N &= G_k^{-1} (G_k d_k + g_k) \\ &= G_k^{-1} (G_k - B_k) d_k \\ &= O(\|G_k - B_k\| \|d_k\|) \\ &= o(\|d_k\|). \end{aligned} \quad (4.4.69)$$

最后一个等式来自(4.4.67).

反之,设(4.4.68)成立,则由 G_k 左乘以(4.4.68)两边,得

$$G_k d_k - G_k d_k^N = o(\|d_k\|). \quad (4.4.70)$$

注意到 $-G_k d_k^N = g_k = -B_k d_k$,从而(4.4.70)成为

$$(G_k - B_k) d_k = o(\|d_k\|),$$

此即(4.4.67).

下面,借助于牛顿法的二次收敛性结果来完成证明.

$$\begin{aligned} &\|x_k + d_k - x^*\| \\ &= O(\|x_k - x^*\|^2) + o(\|d_k\|) \\ &= O(\|x_k - x^*\|^2) + o(\|x_k - x^*\|) \end{aligned} \quad (4.4.71)$$

从上式易知 $\|d_k\| = O(\|x_k - x^*\|)$,代入(4.4.71),得到

$$\|x_{k+1} - x^*\| = \|x_k + d_k - x^*\| = o(\|x_k - x^*\|).$$

这表明 $\{x_k\}$ 是超线性收敛的.

这个定理告诉我们三点:

1) 超线性收敛的充要条件是(4.4.67)成立,即 B_k 只要沿搜索方向 d_k 收敛到 Hesse 矩阵 $G(x^*)$,则拟牛顿法超线性收敛.

2) (4.4.68)也是拟牛顿法超线性收敛的充要条件,即当且仅当拟牛顿步 d_k 在长度和方向上都趋向于牛顿步 d_k^N ,则拟牛顿法超线性收敛.

3) 如果将(4.4.67)用

$$\lim_k \frac{\|B_k - G_k\| \|d_k\|}{\|d_k\|} = 0$$

代替,定理仍然成立.

这个定理是基本的和一般的,当我们讨论每个具体的拟牛顿法的超线性收敛性时,都要验证充要条件(4.4.67).作为典型代表,我们不加证明地给出 BFGS 方法局部超线性收敛定理,感兴趣的读者请参看 Dennis 和 Moré(1977),Dennis 和 Schnabel(1983),以及袁亚湘和孙文瑜(1997).

定理 4.4.7 设 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 满足假设 4.4.4(a)(b)(c).那么,存在 $\delta > 0$ 和 $\epsilon > 0$,使得对于 $\|x_0 - x^*\|_2 < \delta$ 和 $\|B_0 - G(x^*)\|_2 < \epsilon$,BFGS 方法 $x_{k+1} = x_k -$

$B_k^{-1} g_k$ 有定义. 如果

$$\|x_k - x^*\| < \epsilon, \quad (4.4.72)$$

则序列 $\{x_k\}$ 超线性收敛到 x^* .

定理 4.4.8 设 $f: \mathbf{R}^n \rightarrow \mathbf{R}$ 满足假设 4.4.4(a)(b)(c). 假定采用 Wolfe 线性搜索的 BFGS 算法 $x_{k+1} = x_k - \alpha_k B_k^{-1} g_k$ 产生的迭代序列 $\{x_k\}$ 收敛到 x^* , 假定 (4.4.72) 成立, 则序列 $\{x_k\}$ 超线性收敛到 x^* .

小 结

本章是全书最主要的部分之一, 这不仅因为在科学工程实际中很多问题是无约束优化问题, 而且因为许多约束优化问题都是化成无约束优化问题来处理的.

本章介绍了最基本的无约束最优化方法, 其中, 最速下降法是最基本的方法, 牛顿法是最主要的方法, 共轭梯度法是解大型最优化问题的首选, 拟牛顿法尤其是其中的 BFGS 方法是目前最成功的方法.

关于拟牛顿法的收敛性, 我们打了 * 号, 仅供教师和感兴趣的学生选学.



1. 设 $f(x) = \frac{3}{2}x_1^2 + \frac{1}{2}x_2^2 - x_1x_2 - 2x_1$, 设初始点 $x^{(0)} = (-2, 4)^T$. 试用最速下降法和牛顿法极小化 $f(x)$.

2. 设 (1) $f(x) = \frac{1}{2}(x_1^2 + 9x_2^2)$, (2) $f(x) = \frac{1}{2}(x_1^2 + 10^4x_2^2)$. 试讨论最速下降法的收敛速度.

3. 设 $f(x) = \frac{1}{2}x^T x + \frac{1}{4}(x^T A x)$, 其中

$$A = \begin{pmatrix} 5 & 1 & 0 & \frac{1}{2} \\ 1 & 4 & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & 3 & 0 \\ \frac{1}{2} & 0 & 0 & 2 \end{pmatrix}$$

设 (1) $x_0 = (\cos 70^\circ, \sin 70^\circ, \cos 70^\circ, \sin 70^\circ)^T$,

(2) $x_0 = (\cos 50^\circ, \sin 50^\circ, \cos 50^\circ, \sin 50^\circ)^T$,

试在 $\epsilon = 1$ 和 $\epsilon = 10^4$ 的情况下讨论基本牛顿法、线性搜索牛顿法和牛顿型信赖域方法的数值结果和收敛性态.

4. 证明: 当极小化正定二次函数时, 共轭梯度法 FR 公式, PRP 公式和 Dixon 公式是等价的.

5. 设 $s \in \mathbf{R}^n, s^T s > 0, I$ 是 $n \times n$ 单位矩阵. 证明:

$$\left\| I - \frac{ss^T}{s^T s} \right\| = 1.$$

6. 证明逆的秩一校正公式 (Sherman Morrison 公式): 设 $A \in \mathbf{R}^{n \times n}$ 非奇异, $u, v \in \mathbf{R}^n$ 是任意向量, 若 $1 + v^T A u \neq 0$, 则 $A + uv^T$ 非奇异, 且其逆矩阵可以表示为

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u}.$$

7. 证明 Wolfe 条件 (3.5.5) 和强 Wolfe 条件 (3.5.9) 意味着 $s_k^T y_k > 0$.

8. (1) 证明公式 (4.4.23) 和公式 (4.4.24) 是互逆的.

(2) 证明公式 (4.4.26) 和公式 (4.4.27) 是互逆的.

9. 分别用牛顿法和拟牛顿法极小化 Rosenbrock 函数 $f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$, $x_0 = (-1.2, 1)^T, x^* = (1, 1)^T, f(x^*) = 0$.

10. 分别用共轭梯度法和拟牛顿法极小化 Powell 奇异函数 $f(x) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, x_0 = (3, -1, 0, 1)^T, x^* = (0, 0, 0, 0), f(x^*) = 0$.

第5章

线性与非线性最小二乘问题

实践中大量的最优化问题以最小二乘问题的形式出现,对于无约束最优化问题,最小二乘的形式为

$$\min_{x \in \mathbf{R}^n} f(x) = \frac{1}{2} r(x)^T r(x) = \frac{1}{2} \sum_{i=1}^m [r_i(x)]^2, \quad m \geq n, \quad (5.0.1)$$

其中 $r_i(x) (i=1, 2, \dots, m)$ 称为残量函数, $1/2$ 是为表达和运算方便加进去的, 有无 $1/2$ 对问题的最优解没有影响. 当 $r_i(x)$ 是 x 的线性函数时, (5.0.1) 称为线性最小二乘问题, 这时 $r_i(x)$ 可表示成

$$r_i(x) = a_i^T x - b_i,$$

其中 $a = (a_1^T, a_2^T, \dots, a_m^T)^T$ 为 n 维向量, $b \in \mathbf{R}^1$ 为纯量. 对线性最小二乘问题

$$\min \frac{1}{2} \sum_{i=1}^m (a_i^T x - b_i)^2,$$

通常用下述的矩阵向量表示

$$\min \frac{1}{2} \|Ax - b\|^2, \quad (5.0.2)$$

其中

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{pmatrix}.$$

当 $r_i(x) (i=1, 2, \dots, m)$ 为 x 的非线性函数时, (5.0.1) 称为非线性最小二乘问题.

最小二乘问题在科学实验、科学计算、预测预报、模拟、设计与工程技术等许多领域有广泛的应用. 最小二乘的一个典型例子是数据拟合. 设已知 m 组观测值

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \dots & \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{pmatrix}$$

且已确定 $b, i=1, 2, \dots, m$ 与 $a_i, i=1, 2, \dots, m$ 关系的一个估计式为

$$b \approx f(x, a_i), i=1, 2, \dots, m,$$

其中 $x \in \mathbf{R}^n, a_i = (a_{i1}, a_{i2}, \dots, a_{in})^T$. 问题为确定 n 维向量 x 的一个值使得 $f(x, a_i)$ 在 l_2 范数意义下最优拟合 $b, i=1, 2, \dots, m$, 即要求变量 $x \in \mathbf{R}^n$ 的值应是下述问题:

$$\min_{x \in \mathbf{R}^n} \sum_{i=1}^m [f(x, a_i) - b]^2 = \min_{x \in \mathbf{R}^n} \sum_{i=1}^m [r_i(x)]^2$$

的解. 除去少了系数 $1/2$ 外, 上述问题即为标准的最小二乘问题 (5.0.1). 而当 $f(x, a_i)$ 为线性函数 $f(x, a_i) = a_i^T x$ 时, 即为线性最小二乘问题. 另一个应用最小二乘的例子为科学计算中常见的方程组的求解,

$$r(x) = \begin{pmatrix} r_1(x) \\ r_2(x) \\ \vdots \\ r_m(x) \end{pmatrix} = 0.$$

当 $m > n$ 时称上述方程组为超定方程组; 当 $m = n$, 称为适定方程组; 当 $m < n$ 时为亚定方程组. 当所有 $r_i(x), i=1, 2, \dots, m$ 都是 x 的线性函数时, 为线性方程组. 对方程组的求解, 无论是线性的还是非线性的有许多方法可供选择, 其中最小二乘方法是比较常用的方法, 尤其对于超定方程组的求解. 本章我们将学习求解线性与非线性最小二乘问题的常用方法. 第 5.1 节学习线性最小二乘问题的解法, 第 5.2 节学习非线性最小二乘问题的 Gauss-Newton 法, 在第 5.3 节我们将介绍 $L-M$ 算法和信赖域方法, 本章的最后一节则介绍有关对非线性最小二乘问题 Gauss-Newton 法的拟牛顿修正.

§ 5.1 线性最小二乘问题的解法

5.1.1 解线性最小二乘问题

设 $A = (a_{ij})$ 为 $m \times n$ 阶矩阵, $b \in \mathbf{R}^m$ 为 m 维向量, 线性最小二乘问题为求下述最优化问题的最优解:

$$\min_{x \in \mathbf{R}^n} q(x) = \frac{1}{2} \|Ax - b\|^2, \quad (5.1.1)$$

其中 $\|\cdot\|$ 为向量的 l_2 范数. 将 $\|Ax - b\|^2$ 按范数展开得

$$\min_{x \in \mathbf{R}^n} q(x) = \frac{1}{2} x^T A^T A x - b^T A x + \frac{1}{2} b^T b.$$

这是一个以对称矩阵 $A^T A$ 为二阶 Hesse 矩阵的二次函数的无约束极小化问题. 由于矩阵 $A^T A$ 至少半正定, 问题的任何最优解也是问题的全局最优解. 设 x^* 为问题的最优解, 如果 $m = n$, 则必有 $Ax^* - b = 0$, 如果 $m > n$, 则 $Ax^* - b$ 不一定为零. 当向量 b 属于矩阵 A 的像空间 $R(A) = \{y / y = Ax, x \in \mathbf{R}^n\}$, 则存 $x^* \in \mathbf{R}^n$ 使 $\|Ax^* - b\|^2 = 0$, 否则问题的最优值不为零.

由于线性最小二乘问题(5.1.1)是一个凸二次函数的无约束极小化, 由

$$q(x) = A^T A x - A^T b,$$

根据无约束最优化问题最优解的一阶必要条件, 问题(5.1.1)的最优解是方程组

$$A^T A x = A^T b \quad (5.1.2)$$

的解. 方程组(5.1.2)称为法方程组或正规方程组. 如果矩阵 A 列满秩, 即 A 的列线性无关, 则矩阵 $A^T A$ 正定, 方程组(5.1.2)的解唯一且可表示为

$$x^* = (A^T A)^{-1} A^T b = A^+ b,$$

其中 $n \times m$ 阶矩阵 $A^+ = (A^T A)^{-1} A^T$ 称作 $m \times n$ 阶矩阵 A 的广义逆. 对于这样一个线性最小二乘问题我们可以用解一般凸二次函数的无约束最优化方法求解, 例如解方程组(5.1.2)的直接解法, 或迭代法, 如共轭梯度法等.

作为一个二次函数的无约束优化, 如果矩阵 $A^T A$ 正定, 对方程组(5.1.2)的求解, 一般采用矩阵 $A^T A$ 的 Cholesky 分解 LL^T , 其中 L 为下三角矩阵, 或 LDL^T 分解, 这里的 L 为对角线元素全为 1 的单位下三角矩阵, 然后再通过前代和回代求得问题的最优解 x^* . 然而对于线性最小二乘问题, 我们可以利用其问题的特殊结构, 设计更为有效的求解方法. 对于由线性最小二乘问题形成的矩阵 $A^T A$, 不宜采用先计算积矩阵 $A^T A$ 再对所得积矩阵进行分解的方法, 这样做既可以避免积矩阵的计算, 又可以避免计算过程中舍入误差的影响, 因为矩阵 $A^T A$ 的条件数是矩阵 A 的条件数的平方. 设矩阵 $A^T A$ 存在误差 E , 由误差 E 所引起的解的误差为 δx^* , 则由方程组(5.1.2)和

$$(A^T A + E)(x + \delta x^*) = b,$$

可得 δx^* 关于 x^* 的相对误差与 E 与 $A^T A$ 的相对误差之间有如下关系:

$$\frac{\|\delta x^*\|}{\|x^*\|} = \text{cond}(A^T A) \frac{\|E\|}{\|A^T A\|},$$

其中 $\text{cond}(A^T A) = \lambda_1^2 / \lambda_n^2$ 是矩阵 $A^T A$ 的条件数, λ_1 和 λ_n 分别是矩阵 A 的最大和最小奇异值. 下面的例子是 Gill 和 Murry(1991)给出的, 它说明矩阵 $A^T A$ 的条件数是矩阵 A 的条件数的平方. 设

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 10^{-4} \end{pmatrix},$$

则矩阵 A 的条件数为 $\text{cond}(A) = 2.8 \times 10^4$, 而相应的矩阵

$$A^T A = \begin{pmatrix} 2 & 2 \\ 2 & 2 + 10^{-8} \end{pmatrix}$$

的条件数为 $\text{cond}(A^T A) = 7.8 \times 10^8$, 可见, $A^T A$ 的条件数要大大的大于矩阵 A 的条件数. 因此计算积矩阵, 再分解后求解的方法不是一个可取的方法.

为避免计算积矩阵 $A^T A$ 再进行分解, 可以采用对增广矩阵 $[A \ b]$ 作 QR 正交分解的方法. 我们知道 l_2 范数在正交变换下具有不变性, 即

$$Q(Ax - b)^2_2 = QAx - Qb^2_2 = Ax - b^2_2.$$

设

$$A = QR = [Q \ Q_2] \begin{pmatrix} R_1 \\ 0 \end{pmatrix} = Q R_1$$

和

$$[A \ b] = Q[R \ \bar{b}],$$

其中 Q 为 $m \times m$ 阶正交矩阵, $Q_1 \in \mathbf{R}^{m \times n}$, $Q_2 \in \mathbf{R}^{m \times (m-n)}$, $\bar{b} = Q^T b$, R_1 为 $n \times n$ 上三角矩阵. 于是由 $A^T A x = A^T b$ 有

$$R_1^T R_1 x = R_1^T Q_1^T b = R_1^T \bar{b}_n,$$

因此最优解 x^* 可由三角方程组

$$R_1 x = \bar{b}_n$$

经回代确定, 其中 $\bar{b}_n = Q_1^T b$. 这时由于矩阵 R_1 的条件数即为矩阵 A 的条件数 $\text{cond}(R_1) = \lambda_1 / \lambda_n$, 通过上述方程组确定的最优解的相对误差要远小于通过先计算积矩阵再分解后所求得解的相对误差, 这种情况在 A 的条件数较大时尤为明显. 因此求线性最小二乘问题最优解的正交分解算法如下:

算法 5.1.1 (线性最小二乘的正交分解算法)

步 1. 对增广矩阵 $[A \ b]$ 作 QR 正交分解得 R_1 和向量 \bar{b} ;

步 2. 取 \bar{b}_n 为向量 \bar{b} 的前 n 个分量形成的向量;

步 3. 用回代解方程组 $R_1 x = \bar{b}_n$ 得解 x^* .

有关矩阵 A 的 QR 正交分解的方法可参阅有关矩阵计算的书籍.

在上述 QR 正交分解算法的分析过程中我们假定矩阵 A 的列线性无关. 当矩阵 A 的列线性相关, 即 A 不是列满秩时, 在上述矩阵 A 的 QR 正交分解式中, R_1 不再是一个 $n \times n$ 而是 $r \times n$ 阶的上三角矩阵, 而确定最优解的方程组成为

$$R_1 x = \bar{b}_r, \quad (5.1.3)$$

其中 $r = \text{rank}(A) < n$ 表示矩阵 A 的秩数, \bar{b}_r 表示由向量 \bar{b} 的前 r 个分量组成的向量. 方程组 (5.1.3) 有无穷多组解, 我们可依矩阵 A 的奇异值分解算法, 从中确定使 l_2 范数最小的解 x^* , 即

$$x^* = \min \{ \|x\|_2 \mid x \in X^* \},$$

其中 X^* 是最小二乘问题(5.1.1)所有解的集合. 设 $m \times n$ ($m > n$) 阶矩阵 A 的秩为 $r < n$, 则存在 $m \times m$ 阶正交矩阵 U 和 $n \times n$ 阶正交矩阵 V 使得

$$A = USV^T = U \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix} V^T,$$

其中 S 为 $m \times n$ 阶的块对角矩阵,

$$S = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix}$$

为 $r \times r$ 阶对角矩阵, $1 \quad 2 \quad \dots \quad r > 0$ 为矩阵 A 的奇异值. 满足上述要求使 l_2 范数取极小的最优解为

$$x^* = VS^+ U^T b = A^+ b,$$

这里 A^+ 是矩阵 A 的广义逆

$$A^+ = VS^+ U^T, S^+ = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 0 \end{pmatrix},$$

而

$$S^+ = \begin{pmatrix} 1^{-1} & & & \\ & \ddots & & \\ & & 1^{-1} & \\ & & & 0 \end{pmatrix}.$$

当 $r = \text{rank}(A) = n$ 时, 由奇异值分解所确定的矩阵 A^+ 即为前述定义的广义逆矩阵

$$A^+ = VS^+ U^T = (A^T A)^{-1} A^T.$$

5.1.2 解线性等式约束的线性最小二乘问题

下面讨论带有线性约束的线性最小二乘问题, 先考虑线性等式约束的线性最小二乘问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|Ax - b\|_2^2, \\ \text{s.t.} \quad & Cx = y, \end{aligned} \tag{5.1.4}$$

其中 $x \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $C \in \mathbf{R}^{p \times n}$ 和 $y \in \mathbf{R}^p$, $p < n$, 并假定 $\text{rank}(C) = p$, 即约束方程组相容.

求解问题(5.1.4)最直接的方法是用等式约束消去部分变量把问题转化为无约束线性最小二乘问题求解. 由于矩阵 C 行满秩, 可以把矩阵 C 剖分成

$$[G \quad C_2],$$

其中 G 为 $p \times p$ 阶非奇异矩阵, 对矩阵 A 和向量 x 作相应的剖分

$$A = \begin{bmatrix} A_1 & A_2 \end{bmatrix}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

其中 $x_1 \in \mathbf{R}^p$, $x_2 \in \mathbf{R}^{n-p}$, $A_1 \in \mathbf{R}^{m \times p}$, $A_2 \in \mathbf{R}^{m \times (n-p)}$. 于是约束方程成为

$$G_1 x_1 + G_2 x_2 = y,$$

由 G 非奇异, 可得

$$x_1 = G_1^{-1} (y - G_2 x_2). \quad (5.1.5)$$

将其代入问题(5.1.4)的目标函数可得下述仅关于变量 x_2 的无约束线性最小二乘问题

$$\min \frac{1}{2} \| (A_2 - A_1 G_1^{-1} G_2) x_2 - (b - A_1 G_1^{-1} y) \|^2. \quad (5.1.6)$$

如果 $m \times (n-p)$ 阶矩阵 $(A_2 - A_1 G_1^{-1} G_2)$ 列满秩, 即矩阵 $(A_2 - A_1 G_1^{-1} G_2)$ 的秩为 $n-p$, 则问题(5.1.6)的解唯一, 设为 x_2^* , 将其代入(5.1.5)得 x_1^* , 从而得到问题(5.1.4)的唯一解 $x^* = (x_1^*, x_2^*)$. 下面的定理给出了问题(5.1.4)存在唯一解的充分必要条件.

定理 5.1.2 问题(5.1.4)的解唯一的充分必要的条件是矩阵 $B = \begin{pmatrix} C \\ A \end{pmatrix}$ 的秩为 n .

证明 假定矩阵 B 的秩为 $q < n$, 并设 x^* 是问题(5.1.4)的一个解. 考虑齐次线性方程组

$$Bx = 0.$$

由于矩阵 B 的秩 q 小于 n , 上述齐次方程组有非零解. 设 \bar{x} 为这样的非零解, 则 $x^* + \bar{x}$ 为问题(5.1.4)的一个不同于 x^* 的解, 即问题(5.1.4)的解不唯一. 这就证明了条件的必要性.

再证充分性. 设矩阵 B 的秩为 n , 考虑矩阵

$$\begin{aligned} M &= \begin{pmatrix} C & I & -G_1^{-1}G_2 \\ A & 0 & I \end{pmatrix} \\ &= \begin{pmatrix} G_1 & G_2 & I & -G_1^{-1}G_2 \\ A_1 & A_2 & 0 & I \end{pmatrix} \\ &= \begin{pmatrix} G_1 & & 0 & \\ A_1 & A_2 - A_1 G_1^{-1}G_2 & & \end{pmatrix}. \end{aligned}$$

显然矩阵 M 的秩为 n , 即 M 的列线性无关, 由此得矩阵 $A_2 - A_1 G_1^{-1} G_2$ 的秩为 $n-p$. 这就证明了问题(5.1.4)的解是唯一的.

有许多不同的方法确定矩阵 C 的剖分, 并求解问题(5.1.6), 在这里我们介绍 Lawson 和 Hanson 的正交分解算法. 设

$$C = RV^T$$

是矩阵 C 的正交分解, 其中 V 为 $n \times n$ 阶正交矩阵,

$$R = [R_1 \ 0],$$

R_1 为 $p \times p$ 阶上三角矩阵, 把矩阵 V 剖分为

$$V = [V_1 \ V_2], V_1 \in \mathbf{R}^{n \times p}, V_2 \in \mathbf{R}^{n \times (n-p)},$$

则问题(5.1.4)的任何可行解可表示为

$$x = C^+ y + V_2 z, \quad z \in \mathbf{R}^{n-p}, \quad (5.1.7)$$

其中 $n \times p$ 阶矩阵 C^+ 是 $p \times n$ 阶矩阵 C 的广义逆

$$C^+ = VR^+, \quad R^+ = \begin{pmatrix} R_1^{-1} \\ 0 \end{pmatrix}.$$

把这样的 x 代入问题(5.1.4)得关于 $n-p$ 维变量 z 的无约束线性最小二乘问题

$$\min \frac{1}{2} \|AV_2 z - \bar{b}\|^2, \quad (5.1.8)$$

其中 $\bar{b} = b - AC^+ y$. 对问题(5.1.8)的求解可以用前面介绍的无约束线性最小二乘问题的 QR 正交分解算法, 设所得解为 z^* , 将 z^* 代入(5.1.7)即得问题(5.1.4)的最优解. 下面的正交分解算法给出了具体的计算步骤.

算法 5.1.3 (线性等式约束最小二乘问题的正交分解算法)

步 1. 将矩阵 C 正交分解成

$$C = RV^T = [R_1 \ 0] \begin{pmatrix} V_1^T \\ V_2^T \end{pmatrix};$$

步 2. 用回代解方程组 $R_1 z_1 = y$ 得解 z_1^* , 再计算 $\hat{x} = V_1 z_1^* (= C^+ y)$;

步 3. 计算 $\bar{b} = b - A\hat{x}$ 以及 $\bar{A} = AV_2$;

步 4. 对增广矩阵 $[\bar{A} \ \bar{b}]$ 作 QR 正交分解得

$$Q[\bar{A} \ \bar{b}] = \begin{pmatrix} \hat{R}_1 & \hat{b} \\ 0 & \hat{b} \end{pmatrix};$$

步 5. 解方程组 $\hat{R}_1 z = \hat{b}$ 得 z^* ;

步 6. 计算 $x^* = \hat{x} + V_2 z^*$;

步 7. 计算 $g^* = A^T (Ax^* - b)$ 后解方程组 $R^T = V_1^T g^*$ 得 Lagrange 乘子 λ^* .

求解线性等式约束线性最小二乘问题还可以用 Lagrange 乘子法, 其基本思想是根据等式约束最优化问题的最优解是其 Lagrange 函数的稳定点, 由最优解的一阶必要条件求 Lagrange 方程组的解. 问题(5.1.4)的 Lagrange 函数为

$$L(x, \lambda) = \frac{1}{2} \|Ax - b\|^2 - \lambda^T (Cx - y),$$

根据最优解的一阶必要条件, 最优解是下述 Lagrange 方程组的解:

$$\begin{pmatrix} A^T A & -C^T \\ C & 0 \end{pmatrix} \begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} A^T b \\ -A^T y \end{pmatrix},$$

$$L(x, \lambda) = -Cx + y = 0.$$

将其表示成矩阵向量形式

$$\begin{bmatrix} A^T A & -C^T \\ -C & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^T b \\ -y \end{bmatrix}. \quad (5.1.9)$$

这是一个含有 $n+p$ 个变量和 $n+p$ 个方程的方程组, 如果矩阵 A 的秩为 n , C 的秩为 p , 则可以证明方程组的系数矩阵非奇异, 因而有唯一的解

$$\lambda^* = -[C(A^T A)^{-1} C^T]^{-1} [C(A^T A)^{-1} A^T b - y], \quad (5.1.10)$$

$$x^* = (A^T A)^{-1} (A^T b + C^T \lambda^*). \quad (5.1.11)$$

为得到这样的解, 可以利用上述系数矩阵的特殊结构进行分解,

$$\begin{bmatrix} A^T A & -C^T \\ -C & 0 \end{bmatrix} = \begin{bmatrix} R^T & 0 \\ W^T & E^T \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & -I \end{bmatrix} \begin{bmatrix} R & W \\ 0 & E \end{bmatrix},$$

其中 R 和 E 分别是 $n \times n$ 和 $p \times p$ 阶上三角矩阵, W 为 $n \times p$ 阶矩阵, 并满足

$$R^T R = A^T A, \quad W^T = -CR^{-1},$$

$$E^T E = W^T W = C(A^T A)^{-1} C^T.$$

根据最优解和相应 Lagrange 乘子的表达式(5.1.10)和(5.1.11), 以及系数矩阵的上述分解可得下述用 Lagrange 乘子法求解问题(5.1.4)的具体步骤.

算法 5.1.4 (线性等式约束线性最小二乘问题的 Lagrange 乘子法)

步 1. 对增广矩阵 $[A \ b]$ 作 QR 正交分解

$$[A \ b] = Q \begin{bmatrix} R & \hat{b}_1 \\ 0 & \hat{b}_2 \end{bmatrix};$$

步 2. 用回代解方程组 $Rx = \hat{b}_1$ 得 x_1^* ;

步 3. 计算矩阵 $W^T = -CR^{-1}$ 以及 $\hat{y} = Cx_1^* - y$;

步 4. 对矩阵 W 作正交分解

$$W = U^T \begin{bmatrix} L \\ 0 \end{bmatrix};$$

步 5. 用前代和回代分别解方程组 $L^T z = -\hat{y}$ 和 $Lz = 0$ 得 z^* ;

步 6. 解方程组 $Rx = -W^T z^*$ 得 x_2^* , 置 $x^* = x_1^* + x_2^*$.

在这一算法中, 第(1)和第(2)步是通过矩阵 A 的正交分解求方程组 $A^T Ax = A^T b$ 的解 x_1^* , 而且分解因子 R 在第 6 步求解方程组 $A^T Ax = C^T \lambda^*$ 时还要用到. 第 4 步和第 5 步通过对矩阵 W 的正交分解求方程组

$$C(A^T A)^{-1} C^T \lambda^* = C(A^T A)^{-1} A^T b - y = Cx_1^* - y = \hat{y}$$

的解 λ^* . 第 6 步解方程组 $A^T Ax = C^T \lambda^*$ 得 x_2^* , 并置 $x^* = x_1^* + x_2^*$, 其中 $x_1^* = (A^T A)^{-1} A^T b$, $x_2^* = (A^T A)^{-1} C^T \lambda^*$. 由于 $A^T A = R^T R$, $C^T \lambda^* = -R^T W^T z^*$, 因此方程组 $A^T Ax = C^T \lambda^*$ 成为 $R^T Rx = -R^T W^T z^*$, 消去非奇异矩阵 R^T 即得第 6 步的

方程组 $Rx = -W^*$.

5.1.3 解线性不等式约束的线性最小二乘问题*

在这一节我们介绍线性不等式约束的线性最小二乘问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|Ax - b\|^2, \\ \text{s.t.} \quad & Cx \leq y \end{aligned} \quad (5.1.12)$$

的解法,其中 C 是 $p \times n$ 阶矩阵,但对 p 的大小不再像等式约束时有任何限制. 设已知问题在最优解 x^* 处的有效约束集合

$$I(x^*) = \{i \mid c_i^T x^* = y_i\},$$

则问题的最优解 x^* 可以通过解下述线性等式约束的最小二乘问题确定:

$$\begin{aligned} \min \quad & \frac{1}{2} \|Ax - b\|^2, \\ \text{s.t.} \quad & c_i^T x = y_i, \quad i \in I(x^*), \end{aligned}$$

其中 c_i 表示由矩阵 C 的第 i 行形成的列向量, y_i 是 $y \in \mathbf{R}^p$ 的第 i 个分量. 尽管我们事先并不知道有效集 $I(x^*)$ 而无法通过解上述问题得到最优解,但通过解等约束问题来获得不等式约束问题的最优解却是下面要介绍的有效集方法的出发点.

设可行点 $x^{(k)}$ 是问题最优解的一个近似,在该点相应的有效约束集合为

$$I(x^{(k)}) = \{i \mid c_i^T x^{(k)} = y_i\}.$$

有效集方法通过解线性等式约束最小二乘问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|Ax - b\|^2, \\ \text{s.t.} \quad & c_i^T x = y_i, \quad i \in I(x^{(k)}), \end{aligned} \quad (5.1.13)$$

得到最优解的一个新近似 $x^{(k+1)}$, 如果 $I(x^{(k)}) = I(x^*)$, 则 $x^{(k+1)}$ 即为问题的最优解; 如果 $I(x^{(k)}) \subsetneq I(x^*)$, 则可以确定点 $x^{(k+1)}$ 处的一个新的有效约束集合 $I(x^{(k+1)})$, 再求解 k 为 $k+1$ 时的问题 (5.1.13), 重复这一过程直至求得问题的最优解. 在迭代过程中, 只要每次都能使函数值下降, 则由于有效约束的集合数是有限的, 方法必经有限次迭代后终止于最优解.

通过令 $\tilde{x} = x - x^{(k)}$, $\tilde{b} = b - Ax^{(k)}$, 则由 $x^{(k)}$ 的可行性可以把问题 (5.1.13) 改写成

$$\begin{aligned} \min \quad & \frac{1}{2} \|\tilde{A}\tilde{x} - \tilde{b}\|^2, \\ \text{s.t.} \quad & \tilde{c}_i^T \tilde{x} = 0, \quad i \in I(x^{(k)}), \end{aligned} \quad (5.1.14)$$

前述线性等式约束最小二乘问题的求解方法可用于求得它的解, 设解为 $\tilde{x}^{(k)}$. 如果 $x^{(k)} + \tilde{x}^{(k)}$ 不是一个可行点, 即存在某一非有效的约束指标 $i \notin I(x^{(k)})$ 使得有

$\bar{c}_i^T (x^{(k)} + \lambda^{(k)}) < y_i$, 这时需要沿方向 $\lambda^{(k)}$ 进行线性搜索确定一个步长以得到可行点

$$x^{(k+1)} = x^{(k)} + \lambda^{(k)},$$

其中步长 $\lambda^{(k)}$ 由下述准则确定:

$$\lambda^{(k)} = \frac{y_p - \bar{c}_p^T x^{(k)}}{\bar{c}_p^T \lambda^{(k)}} = \min \left\{ \frac{y_i - \bar{c}_i^T x^{(k)}}{\bar{c}_i^T \lambda^{(k)}} \mid \bar{c}_i^T \lambda^{(k)} < 0, i \in I(x^{(k)}) \right\},$$

这时指标为 p 的不等式约束已经从原来的无效约束变为在点 $x^{(k+1)}$ 处的有效约束, 需把 p 加入到点 $x^{(k+1)}$ 处的有效约束指标集得 $I(x^{(k+1)})$, 并对 $k+1$ 求解问题 (5.1.14) .

如果 $x^{(k+1)} = x^{(k)} + \lambda^{(k)}$ 是可行的, 则 $x^{(k+1)}$ 就是问题 (5.1.13) 的最优解, 对此我们须要确定 $x^{(k+1)}$ 是否是原问题 (5.1.12) 的最优解, 即是否有 $I(x^{(k+1)}) = I(x^*)$ 成

这可以通过检验相应的 Lagrange 乘子 $\lambda^{(k+1)}$ 来进行. 根据约束优化问题最优解一阶必要条件, 对应于上述约束最优化问题最优解的 Lagrange 乘子应非负. 显除条件 $\lambda^{(k+1)} \geq 0$ 之外, $x^{(k+1)}$ 与 $\lambda^{(k+1)}$ 满足所有的一阶必要条件. 因此只要有 $\lambda^{(k+1)} \geq 0$ 成立, 则所有一阶必要条件已成立, 由问题的凸规划性质可以确定 $x^{(k+1)}$ 已经是问题的最优解, 应停止迭代. 如果有某一或某几个有效约束的 Lagrange 乘取负值, 表明 $x^{(k+1)}$ 不是最优解, 通过把相应的有效约束变为无效约束可以得到好解的近似, 因此需要把它们从有效约束指标集合中移走, 通常我们一次只移走一个相应于 Lagrange 乘子取最负值的指标, 即选取指标 q :

$$q^{(k+1)} = \min \left\{ i^{(k+1)} \mid \lambda_i^{(k+1)} < 0, i \in I(x^{(k+1)}) \right\}$$

并取

$$I(x^{(k+1)}) = I(x^{(k)}) \setminus \{q\}$$

作为点 $x^{(k+1)}$ 处的有效约束指标集进行新的迭代.

根据上述分析, 求解线性不等式约束最小二乘问题的有效集方法的步骤如下.

算法 5.1.5 (线性不等式约束线性最小二乘问题的有效集方法)

步 1. 确定一个初始可行点 $x^{(1)}$ 及相应的有效集 I , 置 $k=1$;

步 2. 求问题 (5.1.14) 的解 $\lambda^{(k)}$;

步 3. 如果 $x^{(k)} + \lambda^{(k)}$ 不可行, 转步 6;

步 4. 置 $x^{(k+1)} = x^{(k)} + \lambda^{(k)}$, 计算 $\lambda^{(k+1)}$ 及 $q^{(k+1)}$;

步 5. 如果 $\lambda_q^{(k+1)} \geq 0$, 停止迭代; 否则置 $I_{k+1} = I_k \setminus \{q\}$, 并转步 8;

步 6. 确定步长 $\lambda^{(k)}$ 及指标 p ;

步 7. 置 $x^{(k+1)} = x^{(k)} + \lambda^{(k)}$, $k+1 = k + \{p\}$;

步 8. 置 $k := k+1$ 后转步 2.

可以证明如果在每一次迭代都有 $\lambda^{(k)} \geq 0$, 则方法必定经有限次迭代后终止于

最优解 这是因为在 $k \rightarrow 0$ 时, 由迭代所确定的点列使目标函数值严格单调下降, 而在所产生的序列 $\{x^{(k)}\}$ 中必有一个子序列使得 $x^{(k)} + \delta^{(k)}$ 是可行的. 由于这样的 $\delta^{(k)}$ 是子问题 (5.1.14) 的解, 而子问题的数目是有限的, 每次迭代函数值单调下降确保在子问题之间的循环不会出现, 因而迭代必定经有限次后终止.

至于初始可行点 $x^{(1)}$ 可通过有效集方法解下述辅助问题:

$$\begin{aligned} \min_{i \in v} & (y_i - G^T x), \\ \text{s.t. } & G^T x \leq y_i, i \in v \end{aligned} \quad (5.1.15)$$

确定, 其中 v 是在点 x 处不等式约束不成立的指标集合, 即

$$v = \{i \mid G^T x > y_i\}.$$

在用有效集方法解问题 (5.1.15) 时可任取 n 个便于确定解的约束作为初始有效约束集合, 以及由这 n 个有效约束确定的解作为初始点. 只要问题 (5.1.12) 的可行域非空, 上述问题的求解必定产生一个可行点. 这个可行点以及相应的有效约束指标集合就可以用作上述有效集方法的初始可行点及初始有效约束集合.

§ 5.2 非线性最小二乘的 Gauss Newton 法

考虑非线性最小二乘问题

$$\min_{x \in \mathbf{R}^n} f(x) = \frac{1}{2} r(x)^T r(x) = \frac{1}{2} \sum_{i=1}^m [r_i(x)]^2, \quad m \geq n, \quad (5.2.1)$$

其中 $r(x) = (r_1(x), r_2(x), \dots, r_m(x))^T$ 为 $x \in \mathbf{R}^n$ 的非线性函数, 称为残量函数. 如果函数 $r(x)$ 二阶连续可微, 则 $f(x)$ 的一阶和二阶导数 (Hesse 矩阵) 分别为

$$f'(x) = A(x)^T r(x) = \sum_{i=1}^m r_i(x) \nabla r_i(x)$$

和

$$f''(x) = A(x)^T A(x) + \sum_{i=1}^m r_i(x) \nabla^2 r_i(x) = M(x) + S(x),$$

其中

$$A(x) = \begin{bmatrix} r_1(x) & r_2(x) & \dots & r_m(x) \end{bmatrix}^T = \begin{bmatrix} \frac{r_1}{x_1}(x) & \frac{r_1}{x_2}(x) & \dots & \frac{r_1}{x_n}(x) \\ \frac{r_2}{x_1}(x) & \frac{r_2}{x_2}(x) & \dots & \frac{r_2}{x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{r_m}{x_1}(x) & \frac{r_m}{x_2}(x) & \dots & \frac{r_m}{x_n}(x) \end{bmatrix}$$

为向量函数 $r(x)$ 的 $m \times n$ 阶 Jacobian 矩阵, $M(x) = A(x)^T A(x)$, $S(x) = \sum_{i=1}^m r_i(x)^2$.

作为一个无约束最优化问题, 对非线性最小二乘问题可以用一般的无约束最优化方法求解, 但由于问题结构的特殊性, 可以利用其特殊结构设计更有效的求解方法. 考虑无约束最优化问题的牛顿类算法, 设 $x^{(k)}$ 是最优解 x^* 的第 k 次近似, 则算法求下列最优化问题:

$$\min q_k(\cdot) = f(x^{(k)}) + f'(x^{(k)})^T (\cdot - x^{(k)}) + \frac{1}{2} (\cdot - x^{(k)})^T B_k (\cdot - x^{(k)}) \quad (5.2.2)$$

的最优解 $x^{(k+1)}$ 作为对 $x^{(k)}$ 的修正量或搜索方向, 其中 $\Delta x = x - x^{(k)}$. 当 $B_k = \nabla^2 f(x^{(k)})$ 时, 即为牛顿法或阻尼牛顿法, 方法如收敛, 则收敛度是二次的; 当 B_k 为 Hesse 矩阵 $\nabla^2 f(x^{(k)})$ 的近似并由拟牛顿修正公式产生时即为拟牛顿法, 为确保 (5.2.2) 的解 $x^{(k+1)}$ 是 $f(x)$ 在点 $x^{(k)}$ 处的一个下降方向, 要求近似矩阵 B_k 正定, 方法的收敛率则依赖于 B_k 对 $\nabla^2 f(x^{(k)})$ 的近似程度, 一般的拟牛顿法具有超线性收敛性.

对于非线性最小二乘问题, 由于是对 $r_i(x)$, $i = 1, 2, \dots, m$ 在平方和意义下取极小, 极有可能 $r_i(x)$, $i = 1, 2, \dots, m$ 在最优解 x^* 处的值为零 (零残量问题), 或取较小值 (小残量问题). 如果是这样, $f(x)$ 的 Hesse 矩阵 $\nabla^2 f(x)$ 中的非线性项 $S(x) = \sum_{i=1}^m r_i(x)^2$ 在最优解处或为零, 或同 $M(x)$ 相比相对较小. 因此在最优解 x^* 的某个邻域内, $\nabla^2 f(x)$ 的线性项 $M(x)$ 将是 $\nabla^2 f(x)$ 的一个很好的近似, 即二次近似模型 (5.2.2) 中的矩阵 B_k 可简单地取为 $M_k = M(x^{(k)}) = A(x^{(k)})^T A(x^{(k)}) = A_k^T A_k$, 这时 $x^{(k+1)}$ 可通过求解下述二次模型:

$$\min q_k(\cdot) = f(x^{(k)}) + r_k^T A_k (\cdot - x^{(k)}) + \frac{1}{2} (\cdot - x^{(k)})^T A_k^T A_k (\cdot - x^{(k)}) \quad (5.2.3)$$

来确定, 其中 $r_k = r(x^{(k)})$. 由于矩阵 $A_k^T A_k$ 至少半正定, (5.2.3) 中的二次函数 $q_k(\cdot)$ 是凸的, 因而其解必为全局最优解. 设 $x^{(k+1)}$ 为其最优解, 则根据最优性的一阶必要条件, $x^{(k+1)}$ 是下述正则方程组的解:

$$A_k^T A_k \Delta x = -A_k^T r_k. \quad (5.2.4)$$

同牛顿法一样如取

$$x^{(k+1)} = x^{(k)} + \Delta x^{(k)}$$

作为最优解 x^* 的新估计, 即得著名的 Gauss Newton 法.

算法 5.2.1 (Gauss Newton 法)

- 步 1. 给定解的初始估计 $x^{(1)}$, 置 $k = 1$;
- 步 2. 如果 $x^{(k)}$ 满足精度要求, 停止迭代;
- 步 3. 解方程组 (5.2.4) 得 $\Delta x^{(k)}$;

步 4. 置 $x^{(k+1)} = x^{(k)} + e^{(k)}$, $k := k+1$ 后转步 2.

Gauss Newton 法充分利用了最小二乘问题的结构特点, 仅仅利用函数的一阶导数 ($A(x)$) 的信息直接获得 Hesse 矩阵的近似, 而无须像拟牛顿方法那样利用函数一阶导数信息的逐步积累来获取 Hesse 矩阵的近似. 由于 Gauss Newton 法是从牛顿法通过忽略 ${}^2 f(x)$ 的非线性项 $S(x)$ 形成的, 可以预料 Gauss Newton 的性能和它的收敛性将受这一忽略的影响, 下面的定理证实了这一点.

定理 5.2.2 设 $f(x) = \frac{1}{2} \sum_{i=1}^m [r_i(x)]^2$ 在开凸集 $D \subset \mathbf{R}^n$ 上二阶连续可微; 存在 $x^* \in D$ 使 $f(x^*) = A(x^*)^T r(x^*) = 0$; 对所有 $x \in D$, $A(x)$ 满秩, 且存在常数 $\alpha > 0$, $M > 0$ 使成立

$$A(x) \quad , \quad H(x) = M(x)^{-1} \quad , \quad \forall x \in D; \quad (5.2.5)$$

又 $A(x)$ 和 ${}^2 f(x)$ 在 D 上 Lipschitz 连续, 即存在常数 $\alpha > 0$, $\beta > 0$ 使得

$$A(x) - A(y) \leq \alpha \|x - y\|, \quad \forall x, y \in D, \quad (5.2.6)$$

$${}^2 f(x) - {}^2 f(y) \leq \beta \|x - y\|, \quad \forall x, y \in D, \quad (5.2.7)$$

则 Gauss-Newton 迭代对所有的 $x \in D$ 有定义, 且有

$$e^{(k+1)} = H(x^*) S(x^*) e^{(k)} + O(\|e^{(k)}\|^2), \quad (5.2.8)$$

其中 $e^{(k)} = x^{(k)} - x^*$.

证明 由于 $A(x)$ 满秩, $A(x)^T A(x)$ 正定, 因此 Gauss Newton 迭代对所有 $x \in D$ 都有定义. 根据定理的条件我们分别有

$$\begin{aligned} M(x) - M(y) &= A(x)^T A(x) - A(x)^T A(y) + A(x)^T \\ &\quad A(y) - A(y)^T A(y) \\ &= A(x)^T A(x) - A(x)^T A(y) + A(x)^T \\ &\quad A(y) - A(y)^T A(y) \\ &= (A(x) + A(y))^T (A(x) - A(y)) \\ &\leq 2\alpha \|x - y\|, \quad \forall x, y \in D, \end{aligned} \quad (5.2.9)$$

$$\begin{aligned} H(x) - H(y) &= H(x)[M(y) - M(x)]H(y) \\ &= H(x)H(y)[M(x) - M(y)] \\ &\leq 2\beta \|x - y\|, \quad \forall x, y \in D, \end{aligned} \quad (5.2.10)$$

$$\begin{aligned} S(x) - S(y) &= {}^2 f(x) - M(x) - [{}^2 f(y) - M(y)] \\ &= {}^2 f(x) - {}^2 f(y) + M(x) - M(y) \\ &\leq (\beta + 2\alpha) \|x - y\|, \quad \forall x, y \in D. \end{aligned} \quad (5.2.11)$$

把 $f(x^*)$ 在点 $x^{(k)}$ 处展开得

$$0 = f(x^*) = f(x^{(k)}) - {}^2 f(x^{(k)}) e^{(k)} + O(\|e^{(k)}\|^2).$$

在上式两边左乘 H_k 有

$$0 = -e^{(k)} - e^{(k)} - H_k S_k e^{(k)} + O(\|e^{(k)}\|^2)$$

$$= -e^{(k+1)} - H(x^*) S(x^*) e^{(k)} - H_k(S_k - S(x^*)) e^{(k)} - (H_k - H(x^*)) S(x^*) e^{(k)} + O(e^{(k)2}). \quad (5.2.12)$$

利用定理的条件(5.2.5)以及(5.2.9)~(5.2.11),上式给出了定理的结论.

这一定理刻画了 Gauss Newton 法的三个基本特征.

(1) Gauss Newton 法收敛与否直接取决于 Hesse 矩阵的省略部分 $S(x)$ 的大小.如果同 $M(x^*)$ 相比, $S(x^*)$ 相对较大,即有 $H(x^*) S(x^*) > 1$,方法就不收敛.这种情况一般出现在 $f(x^*)$ 的值较大,即所谓的大残量问题,或者 $r_i(x)$, $i=1,2,\dots,m$ 的非线性程度较高的时候.

(2) 如果 $A(x^*)$ 满秩,即 $M(x^*)$ 正定,且 $H(x^*) S(x^*) < 1$, Gauss Newton 法局部收敛,也就是当初始点充分接近最优解时,方法收敛.这是因为 $M(x^*)$ 正定,存在 x^* 的一个邻域,在这个邻域内 $M(x)$ 正定,且定理的条件成立.于是由定理的结论以及 $O(\cdot)$ 的意义,存在常数 $c>0$ 使成立

$$e^{(k+1)} \leq (1+c) e^{(k)}.$$

如果初始点在这样的邻域内,并满足 $1+c e^{(1)} < 1$,则有 $e^{(2)} \leq e^{(1)}$,因此 $x^{(2)}$ 也在这个邻域内.用归纳法可以证明方法产生的序列 $\{x^{(k)}\}$ 均在这个邻域内,且有 $e^{(k)} \rightarrow 0$.因此方法收敛.

(3) 在方法收敛的时候,收敛率同样取决于 $H(x^*) S(x^*)$ 的大小.如果 Hesse 矩阵的非线性部分在点 x^* 为零,即 $S(x^*)=0$,则方法是相当令人满意的,即对于零残量问题,方法二次收敛.如果 $S(x^*) \neq 0$,收敛速度是线性的,而具体的收敛速度取决于问题的非线性程度或残量的大小.

由于矩阵 $M(x)$ 至少是半正定的,在大多数情况是正定的,因此由方程组(5.2.4)确定的方向 $s^{(k)}$ 一般是函数的一个下降方向(至少不是一个上升的方向),因此最常用的一个改进是把线性搜索结合进 Gauss Newton 法,形成所谓的阻尼 Gauss Newton 法:

算法 5.2.3 (阻尼 Gauss Newton 法)

步 1. 给定解的初始估计 $x^{(1)}$,置 $k=1$;

步 2. 如果 $x^{(k)}$ 满足精度要求,停止迭代;

步 3. 解方程组(5.2.4)得 $s^{(k)}$,并置 $s^{(k)} = s^{(k)}$;

步 4. 沿方向 $s^{(k)}$ 进行线性搜索,确定步长 α_k ;置 $x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$, $k=k+1$ 后转步 2.

线性搜索可以采用在第三章学习的线性搜索方法.由于采用线性搜索,只要方向 $s^{(k)}$ 是一个下降方向,线性搜索就可以确定一个步长,使得迭代产生的函数值序列单调下降,避免了 Gauss Newton 法在每次迭代都取步长 1 有可能使函数值不下降(这种情况主要出现在点 $x^{(k)}$ 离最优解较远的时候).因此可以预料阻尼 Gauss Newton 法的收敛域要大于 Gauss Newton 法的收敛域.

由 Jacobian 矩阵不满秩导致矩阵 $M(x)$ 不正定的情况在实际中存在, 由此会出现这样的情况, 阻尼 Gauss Newton 法或者收敛于非稳定点, 或者在还没有接近稳定点之前就已终止. 考察下述由 Powell 所提供的含有两个变量的例子:

$$r_1(x) = x_1, r_2(x) = 2x_2^2 + \frac{10x_1}{x_1 + 0.1}.$$

函数 $f(x) = \frac{1}{2}[r_1(x)^2 + r_2(x)^2]$ 在半平面 $x_1 > -0.1$ 内仅有一个稳定点 $x^* = (0, 0)^T$. 取 $x^{(0)} = (3, 1)$ 为初始点, 阻尼 Gauss Newton 法收敛于点 $x = (1.8016, 0.0000)^T$, 在这一点, Jacobian 矩阵 $A(x)$ 不满秩, 而且函数的导数不为零 $f'(x) \neq 0$.

从上述分析, 可以看到作为一个能利用非线性最小二乘问题结构特征的方法, 阻尼 Gauss Newton 法对于 Jacobian 矩阵处处都满秩的零残量和小残量非线性最小二乘问题有令人满意的收敛性(二次收敛或收敛较快的线性收敛), 但是对于 Jacobian 矩阵可能不满秩或残量较大或非线性程度较高的非线性最小二乘问题, 方法未必收敛.

最后需要指出的是: 对于方程组(5.2.4)的求解不能用先计算矩阵乘积, 再计算矩阵分解的方法, 而是要采用对增广矩阵 $[A_k \quad r_k]$ 进行 QR 正交分解的方法, 具体过程见本章第一节.

§ 5.3 信赖域方法

信赖域方法是求解最优化问题的另一类有效方法, 其最初的设计思想可追溯至 Levenberg 和 Marquardt 对 Gauss Newton 法的修正. 当 $x^{(k)}$ 离最优解较远时, 由方程组(5.2.4)确定的 $\delta^{(k)}$ 即使 $M_k = A_k^T A_k$ 正定也未必使由 $x^{(k+1)} = x^{(k)} + \delta^{(k)}$ 确定的点 $x^{(k+1)}$ 满足 $f(x^{(k+1)}) < f(x^{(k)})$, 即函数值不下降. 这是因为如此确定的 $\delta^{(k)}$ 是函数 $f(x)$ 的下述二次近似的极小解:

$$\min q_k(\delta) = f(x^{(k)}) + r_k^T \delta + \frac{1}{2} \delta^T A_k^T A_k \delta, \quad (5.3.1)$$

其中 $\delta = x - x^{(k)}$. 由于这样的近似只是在 $x^{(k)}$ 的一个比较小的邻域内才是好的, 而由方程组(5.2.4)确定的 $\delta^{(k)}$ 可能比较大而使得 $x^{(k)} + \delta^{(k)}$ 超出了这个有效的邻域. 为此, 通过选择参数 μ 的适当值, 使得由方程组

$$(A_k^T A_k + \mu I) \delta = -A_k^T r_k \quad (5.3.2)$$

确定的 $\delta^{(k)}$ 落在上述有效的邻域内. 由上式可以看出, 由于矩阵 $A_k^T A_k$ 至少半正定, 对任何非零正数 μ 矩阵 $A_k^T A_k + \mu I$ 非奇异(正定), 因而有唯一解, 而且对于越大的正数 μ , 解向量 $\delta^{(k)}$ 的长度越短, 因此, 必可找到适当的 μ 值, 使 $x^{(k)} + \delta^{(k)}$

在 $x^{(k)}$ 的一个较小的邻域内,从而有 $f(x^{(k)} + \delta^{(k)}) < f(x^{(k)})$.

如果称由方程组(5.2.4)确定的方向为 Gauss Newton 方向,则由方程组(5.3.2)确定的 $\delta^{(k)}$ 是位于 Gauss Newton 方向与最速下降方向之间的一个方向.而由于矩阵 $A_k^T A_k + \mu I$ 正定,这样的方向一定是 $f(x)$ 在点 $x^{(k)}$ 的一个下降方向.这是因为,在方程(5.3.2)中如果有 $\mu=0$,则(5.3.2)即为 Gauss Newton 方程组(5.2.4),所得方向为 Gauss Newton 方向,而对充分大的 μ ,由(5.3.2)确定的方向接近于最速下降方向.因此 μ 值的大小确定了所得方向由 Gauss Newton 方向向最速下降方向的偏移程度, μ 值越大,向最速下降方向的偏移越大,所得的迭代点越接近于最速下降方向上的点.

初期的 Levenberg Marquardt 方法,对 μ 值的确定采取逐步增大的方法,即给定 μ 的一个初始值(通常为零),解方程组(5.3.2)得 $\delta^{(k)}$,并检验 $x^{(k)} + \delta^{(k)}$ 是否使 $f(x)$ 的函数值取得预先指定的下降量,如果有,就取 $x^{(k+1)} = x^{(k)} + \delta^{(k)}$ 并进行终止性检验;如果没有,就要增大 μ 重新解方程组(5.3.2).重复这一过程,直至获得满足要求的 $\delta^{(k)}$ 为止.可以看到为确定 μ 的一个适当的值,需要对不同的 μ 值解方程组(5.3.2).而方程组的求解是相当费时的,其计算工作量是 $O(n^3)$ 阶.同时在这样的方法中,由于不能有效地利用计算过程中函数值和导数值的信息对 μ 的值进行估计,这又无形中增加了对不同 μ 值的试算次数,降低了算法的有效性.下面的定理为这一问题的解决提供了新的途径.

定理 5.3.1 设 $\delta^{(k)}$ 是方程组(5.3.2)对某一给定 $\mu > 0$ 的解,则 $\delta^{(k)}$ 也是下述信赖域问题的全局最优解:

$$\begin{aligned} \min \quad & q_k(\delta) = \frac{1}{2} \|A_k \delta + r_k\|^2, \\ \text{s.t.} \quad & \|\delta\| \leq \rho_k, \end{aligned} \quad (5.3.3)$$

其中 $\rho_k = \rho(x^{(k)})$.

证明 由于 $\delta^{(k)}$ 是方程组(5.3.2)的解,将 $\delta^{(k)}$ 代入函数 $q_k(\delta)$ 得

$$\begin{aligned} q_k(\delta^{(k)}) &= \frac{1}{2} (\delta^{(k)})^T A_k^T A_k \delta^{(k)} + r_k^T A_k \delta^{(k)} + \frac{1}{2} r_k^T r_k \\ &= \frac{1}{2} (\delta^{(k)})^T A_k^T A_k \delta^{(k)} - (\delta^{(k)})^T (A_k^T A_k + \mu I) \delta^{(k)} + \frac{1}{2} r_k^T r_k \\ &= \frac{1}{2} r_k^T r_k - \mu (\delta^{(k)})^T \delta^{(k)} - \frac{1}{2} (\delta^{(k)})^T A_k^T A_k \delta^{(k)}. \end{aligned}$$

另一方面,对任意 δ 我们有

$$\begin{aligned} q_k(\delta) &= \frac{1}{2} r_k^T r_k + r_k^T A_k \delta + \frac{1}{2} \delta^T A_k^T A_k \delta \\ &= \frac{1}{2} r_k^T r_k - \delta^T (A_k^T A_k + \mu I) \delta + \frac{1}{2} \delta^T A_k^T A_k \delta \end{aligned}$$

$$= \frac{1}{2} r_k^T r_k - \mu^T (x^{(k)} - x^*) - \frac{1}{2} (x^{(k)} - x^*)^T A_k^T A_k (x^{(k)} - x^*) + \frac{1}{2} (x^{(k)} - x^*)^T A_k^T A_k x^*.$$

因此对于任意满足 $\mu > 0$ 的 μ , 根据矩阵 $A_k^T A_k$ 的半正定性并利用 Cauchy Schwartz 不等式可以得到

$$\begin{aligned} q_k(x) - q_k(x^{(k)}) &= \frac{1}{2} (x^{(k)} - x^*)^T A_k^T A_k (x^{(k)} - x^*) + \mu (x^{(k)} - x^*)^T (x^{(k)} - x^*) \\ &\quad - \frac{1}{2} (x^{(k)} - x^*)^T A_k^T A_k (x^{(k)} - x^*) + \mu (x^{(k)} - x^*)^T (x^{(k)} - x^*) \\ &= 0. \end{aligned}$$

这就证明了 $x^{(k)}$ 是信赖域问题 (5.3.3) 的全局最优解.

这一定理表明, 对于一个给定的 $\mu > 0$, 方程组 (5.3.2) 的解, 也是对于某一 $\delta > 0$ 信赖域问题 (5.3.3) 的解. 因此控制参数 μ 的值以确定修正量 $\delta^{(k)}$ 的问题可以由控制参数 μ 的值来确定, 也就是通过选取 μ 的适当值求信赖域问题 (5.3.3) 的解 $x^{(k)}$ 确定对 $x^{(k)}$ 的修正量. 由此形成的方法就是近年来得到广泛使用的信赖域方法.

比较问题 (5.3.2) 和 (5.3.3) 也可以看出, 问题 (5.3.3) 直接反映了 $q_k(x)$ 是对 $f(x)$ 在点 $x^{(k)}$ 的一个邻域内的近似. 对于给定的 δ , 在由 δ 确定的邻域内 $q_k(x)$ 的解能否使目标函数取得满意的下降取决于 $q_k(x)$ 对 $f(x)$ 近似的满意程度. 而在这个邻域内 $q_k(x)$ 是不是 $f(x)$ 的满意的近似可以通过比较两个函数在点 $x^{(k)}$ 处的函数值来确定. 设 $x^{(k)}$ 为问题 (5.3.3) 对于某一给定 $\delta > 0$ 的解. 记

$$\text{Pred}(\delta^{(k)}) = f(x^{(k)}) - q_k(x^{(k)})$$

为由近似函数关于 $\delta^{(k)}$ 所取得的下降量. 通常称为预期下降量. 记

$$\text{Ared}(\delta^{(k)}) = f(x^{(k)}) - f(x^{(k)} + \delta^{(k)})$$

为函数 $f(x)$ 从修正量 $\delta^{(k)}$ 所取得的实际下降量. 如果 $q_k(x)$ 在由 δ 确定的邻域内是 $f(x)$ 的一个满意的近似, 那么 $\text{Ared}(\delta^{(k)})$ 的值应接近 $\text{Pred}(\delta^{(k)})$ 的值, 或者说有

$$r_k = \frac{\text{Ared}(\delta^{(k)})}{\text{Pred}(\delta^{(k)})} \approx 1.$$

如果 $q_k(x)$ 在由 δ 确定的邻域内不是 $f(x)$ 的一个满意的近似, 那么 $\text{Ared}(\delta^{(k)})$ 的值会远小于 $\text{Pred}(\delta^{(k)})$ 的值, 甚至取负值, 或者说有

$$r_k = \frac{\text{Ared}(\delta^{(k)})}{\text{Pred}(\delta^{(k)})} \approx 0.$$

因此比值 r_k (也就是函数值) 可以直接用来判定 $q_k(x)$ 在邻域 δ 内对 $f(x)$ 近似的程度, 并在近似程度不好时对信赖域半径作出调整. 由此给出下述信赖域型的 Levenberg Marquardt 方法.

算法 5.3.2 (解非线性最小二乘问题的信赖域方法)

步 1. 给定控制迭代的参数值 $0 < \alpha_1 < \alpha_2 < 1, 0 < \beta_1 < 1 < \beta_2, r_{\max} > 0$, 初始点 $x^{(1)}$, 以及初始信赖域半径 $\rho_1 < r_{\max}$, 置 $k=1$;

步 2. 如果 $x^{(k)}$ 满足终止条件, 停止迭代, 取 $x^{(k)}$ 作为最优解的近似;

步 3. 对 k 解信赖域子问题(5.3.3)得解 $\hat{x}^{(k)}$;

步 4. 计算 $\text{Pred}(\hat{x}^{(k)})$, $\text{Ared}(\hat{x}^{(k)})$ 和 r_k ;

步 5. 如果 $r_k < \alpha_1$, 置 $\rho_k := \alpha_1 \rho_{k-1}$, 转步 3;

步 6. 置 $x^{(k+1)} = x^{(k)} + \hat{x}^{(k)}$,

$$\rho_{k+1} = \begin{cases} \min\{\alpha_2 \rho_k, r_{\max}\}, & \text{如果 } r_k > \alpha_2 \text{ 且 } \hat{x}^{(k)} = 0, \\ \rho_k, & \text{否则,} \end{cases}$$

置 $k = k+1$, 转步 2.

在上述算法中参数 α_1 和 α_2 用于判定在邻域 ρ_k 内 $q_k(\cdot)$ 是否是 $f(x)$ 的好的近似, 因此 α_1 的值应接近 0, 通常有 $\alpha_1 \in (0, 0.1)$, 而 α_2 的值应接近 1, 通常有 $\alpha_2 \in (0.5, 0.9)$. 从算法可以看出, 如果有 $r_k < \alpha_1$, 则 $\hat{x}^{(k)}$ 不被接受, 减小信赖域半径后重解问题(5.3.3)以确定一个新的解 $\hat{x}^{(k)}$. 如果有 $r_k < \alpha_1$, $\hat{x}^{(k)}$ 就被接受, 并确定一个新的信赖域半径. 新的信赖域半径或者等于原信赖域半径, 或大于原信赖域半径. 如果有 $\alpha_1 < r_k < \alpha_2$, 表明在邻域 ρ_k 内 $q_k(\cdot)$ 是 $f(x)$ 的一个可以接受的近似, 但不是一个令人满意的近似, 因此信赖域半径可以保持不变; 如果虽然有 $r_k > \alpha_2$, 但 $\hat{x}^{(k)} \neq 0$, 这表明使 $f(x)$ 获得充分下降的 $\hat{x}^{(k)}$ 是在信赖域的内部取得的, 因此也没有必要增大信赖域半径; 只有在既有 $r_k > \alpha_2$, 又有 $\hat{x}^{(k)} = 0$ 时才有必要增大信赖域半径. 参数 α_1 和 α_2 分别用于控制减小和增大信赖域半径的程度, r_{\max} 则确定了信赖域半径的一个上界.

下述定理给出了上述信赖域方法的收敛性.

定理 5.3.3 设函数 $f(x) = \frac{1}{2} \sum_{i=1}^m r_i(x)^2$ 二次连续可微, 对于给定的初始点 $x^{(1)}$, 水平集

$$L(x^{(1)}) = \{x \mid f(x) = f(x^{(1)})\}$$

有界, 且存在常数 $M_1 > 0, M_2 > 0$ 使得

$$\|f'(x)\| \leq M_1, \quad \forall x \in L(x^{(1)}), \quad (5.3.4)$$

$$A(x)^T A(x) \leq M_2, \quad \forall x \in L(x^{(1)}), \quad (5.3.5)$$

则由上述算法产生的点列 $\{x^{(k)}\}$, 如果不终止于某一点, 则必有一个极限点 x^* 使得

$$f'(x^*) = 0.$$

证明 用反证法证明本定理. 设方法产生一个无限点列 $\{x^{(k)}\}$ 满足 $f'(x^{(k)}) \neq 0, \forall k=1, 2, \dots$. 先证对所有 $k \geq 1$ 成立

$$\text{Pred}(\hat{x}^{(k)}) = \frac{1}{2} \|A_k^T r_k\| \leq \min\{\alpha_2 \rho_k, r_{\max}\}, \quad \frac{A_k^T r_k}{M_2} > 0. \quad (5.3.6)$$

对任意 $x^{(k)}$, 考虑 $q_k(\cdot)$ 沿最速下降方向的下降量

$$\begin{aligned} q_k(\cdot) &= f(x^{(k)}) - \frac{1}{2} \frac{A_k^T A_k r_k}{A_k^T r_k} + \frac{1}{2} \frac{r_k^T A_k A_k^T A_k A_k^T r_k}{A_k^T r_k} \\ &= f(x^{(k)}) - \frac{1}{2} \frac{A_k^T r_k}{A_k^T r_k} + \frac{1}{2} \frac{r_k^T A_k A_k^T A_k A_k^T r_k}{A_k^T r_k} \end{aligned}$$

由于 $\alpha_k = -\frac{A_k^T r_k}{A_k^T r_k}$ 是问题 (5.3.3) 的一个可行解, 因而由 $x^{(k)}$ 的最优性有

$$\begin{aligned} \text{Pred}(x^{(k)}) &= f(x^{(k)}) - q_k(\alpha_k) = f(x^{(k)}) - \frac{1}{2} \frac{A_k^T r_k}{A_k^T r_k} \\ &= f(x^{(k)}) - \frac{1}{2} \frac{r_k^T A_k A_k^T A_k A_k^T r_k}{A_k^T r_k} \end{aligned}$$

由于矩阵 $A_k^T A_k$ 至少正半定, 我们有 $r_k^T A_k A_k^T A_k A_k^T r_k \geq 0$. 如果 $r_k^T A_k A_k^T A_k A_k^T r_k = 0$, 则直接有

$$\text{Pred}(x^{(k)}) = f(x^{(k)}) - \frac{1}{2} \frac{A_k^T r_k}{A_k^T r_k}.$$

如果有 $r_k^T A_k A_k^T A_k A_k^T r_k > 0$, 则 $q_k(\cdot) = \frac{A_k^T r_k}{A_k^T r_k}$ 在

$$\alpha_k^* = \frac{A_k^T r_k}{r_k^T A_k A_k^T A_k A_k^T r_k}$$

处取极小. 如果 $\alpha_k^* < \alpha_k$, 则有

$$\begin{aligned} \text{Pred}(x^{(k)}) &= f(x^{(k)}) - q_k(\alpha_k^*) = f(x^{(k)}) - \frac{1}{2} \frac{A_k^T r_k}{A_k^T r_k} \\ &= \frac{1}{2} \frac{A_k^T r_k}{r_k^T A_k A_k^T A_k A_k^T r_k} \\ &= \frac{1}{2} \frac{A_k^T r_k}{A_k^T A_k} - \frac{1}{2} \frac{A_k^T r_k}{M_k} \end{aligned}$$

如果有 $\alpha_k^* \geq \alpha_k$, 则有

$$\begin{aligned} \text{Pred}(x^{(k)}) &= f(x^{(k)}) - q_k(\alpha_k) = f(x^{(k)}) - \frac{1}{2} \frac{r_k^T A_k A_k^T A_k A_k^T r_k}{A_k^T r_k} \\ &= f(x^{(k)}) - \frac{1}{2} \frac{A_k^T r_k}{A_k^T r_k} \end{aligned}$$

这就完成了对 (5.3.6) 式的证明.

根据算法我们有

$$\text{Ared}(x^{(k)}) = f(x^{(k)}) - \text{Pred}(x^{(k)}) > 0.$$

由于 $f(x^{(k)})$ 单调下降, $f(x)$ 连续, 以及水平集 $L(x^{(1)})$ 的有界性, 有 $\text{Ared}(x^{(k)}) > 0$, 由此得

$$\text{Pred}(x^{(k)}) > 0. \quad (5.3.7)$$

现在再用反证法证明在序列 $\{x^{(k)}\}$ 中存在一个子序列, 设为 $\{x^{(k_j)}\}$ 使得

$$f(x^{(k_j)}) = A(x^{(k_j)})^T r(x^{(k_j)}) = 0. \quad (5.3.8)$$

假定这一结论不成立, 则存在常数 $\bar{\epsilon} > 0$ 使

$$f(x^{(k)}) \geq \bar{\epsilon} \quad (5.3.9)$$

对所有 $k \geq 1$ 成立. 对此, 可以证明对所有 $k \geq 1$ 信赖域半径 ρ_k 有界, 即存在常数 $\bar{\rho} > 0$ 使成立

$$\rho_k \leq \bar{\rho}, \quad \forall k \geq 1. \quad (5.3.10)$$

利用积分中值定理得

$$\begin{aligned} & |Ared(x^{(k)}) - Pred(x^{(k)})| \\ &= \left| f(x^{(k)}) + f(x^{(k)})^T \rho^{(k)} + \frac{1}{2} \rho^{(k)T} A_k^T A_k \rho^{(k)} - f(x^{(k)} + \rho^{(k)}) \right| \\ &= \left| \frac{1}{2} \rho^{(k)T} A_k^T A_k \rho^{(k)} - \int_0^1 (1-t) \rho^{(k)T} \nabla^2 f(x^{(k)} + t \rho^{(k)}) \rho^{(k)} dt \right| \\ &\leq (M_1 + M_2) \rho^{(k)2} \leq (M_1 + M_2) \rho_k^2. \end{aligned}$$

由此由(5.3.6)和(5.3.9)得

$$\begin{aligned} \left| \frac{Ared(x^{(k)})}{Pred(x^{(k)})} - 1 \right| &\leq \left| \frac{Ared(x^{(k)}) - Pred(x^{(k)})}{Pred(x^{(k)})} \right| \\ &\leq \frac{(M_1 + M_2) \rho_k^2}{\min \rho_k, \bar{M}_2}. \end{aligned}$$

因此当 $\rho_k \rightarrow 0$ 时我们可得下列极限

$$\lim_{k \rightarrow \infty} \frac{Ared(x^{(k)})}{Pred(x^{(k)})} = 1.$$

这表明存在一个 $\bar{\epsilon} > 0$, 使得当 $\rho_k \leq \bar{\epsilon}$ 时有 $r_k = Ared(x^{(k)}) / Pred(x^{(k)}) \geq 1$ 成立. 这就证明了 $\bar{\epsilon} = \bar{\epsilon}$ 可使(5.3.10)式成立. 于是我们可以进一步得到

$$Pred(x^{(k)}) \geq \frac{1}{2} \min \bar{\epsilon}, \bar{M}_2 > 0, \quad \forall k \geq 1.$$

这与(5.3.7)矛盾, 这一矛盾证明了必有子序列使(5.3.8)成立. 再由序列 $\{x^{(k)}\}$ 的有界性知子序列 $\{x^{(k_j)}\}$ 必有聚点, 记 x^* 为这样的聚点, 由 $f(x)$ 的连续性便有 $f(x^*) = A(x^*)^T r(x^*) = 0$. 这就完成了定理的证明.

这一定理表明 Gauss Newton 法的信赖域修正, 不仅克服了 Gauss Newton 法步长($\rho_k = 1$)过长或由于 A_k 不满秩等问题所引起的不收敛问题, 而且增大了收敛区域, 即信赖域 Gauss Newton 法具有全局收敛性. 下述定理则更进一步表明信赖域 Gauss Newton 法保留了 Gauss Newton 法的二次收敛性.

定理 5.3.4 设函数 $f(x)$ 二次连续可微, 对给定的初始点 $x^{(1)}$ 水平集 $L(x^{(1)})$ 有界. 设由信赖域 Gauss Newton 法产生的序列 $\{x^{(k)}\}$ 收敛于 x^* , 且

$f(x^*)$ 正定, $S(x^*) = \sum_{i=1}^m r_i(x^*)^2$, $r_i(x^*) = 0$, 则 $\{x^{(k)}\}$ 以二次收敛速度收敛于 x^* .

证明 由 $S(x^*) = 0$ 知 $f(x^*) = A(x^*)^T A(x^*)$, 因此 $A(x^*)^T A(x^*)$ 正定. 再由 $A(x)^T A(x)$ 的连续性, 以及 $x^{(k)}$ 对 x^* 的收敛性, 存在正数 $C_1, C_2 > 0$ 使得

$$C_1 \|y\|^2 \leq y^T A_k^T A_k y \leq C_2 \|y\|^2, \quad \forall y \neq 0$$

对所有充分大的 k 都成立. 考虑到 $(A_k^T A_k + \mu_k I)^{-1} r_k = -A_k^T r_k$, 由此得

$$\begin{aligned} \text{Pred}(x^{(k)}) &= x^{(k)T} A_k^T r_k - \frac{1}{2} x^{(k)T} A_k^T A_k x^{(k)} \\ &= x^{(k)T} (A_k^T A_k + \mu_k I)^{-1} r_k - \frac{1}{2} x^{(k)T} A_k^T A_k x^{(k)} \\ &= \frac{1}{2} x^{(k)T} A_k^T A_k x^{(k)} + \mu_k \|x^{(k)}\|^2 \\ &\geq \frac{1}{2} C_2 \|x^{(k)}\|^2. \end{aligned}$$

再利用 $f(x^{(k)} + \Delta x)$ 在点 $x^{(k)}$ 的 Taylor 展开式, $S(x^*) = 0$ 以及 $x^{(k)}$ 对 x^* 的收敛性有

$$\begin{aligned} \text{Ared}(x^{(k)}) &= \text{Pred}(x^{(k)}) \\ &= \frac{1}{2} x^{(k)T} (A_k^T A_k - \nabla^2 f(x^{(k)})) x^{(k)} + o(\|x^{(k)}\|^2) \\ &= -\frac{1}{2} x^{(k)T} \sum_{i=1}^m r_i(x^{(k)})^2 r_i(x^{(k)}) x^{(k)} + o(\|x^{(k)}\|^2) \\ &= o(\|x^{(k)}\|^2). \end{aligned}$$

由此两式, 我们得到

$$\left| \frac{\text{Ared}(x^{(k)})}{\text{Pred}(x^{(k)})} - 1 \right| \leq \frac{o(\|x^{(k)}\|^2)}{\frac{1}{2} C_2 \|x^{(k)}\|^2},$$

这表明在 k 充分大时有 $\text{Ared}(x^{(k)}) / \text{Pred}(x^{(k)}) \geq 1$. 因此存在一个正整数 $k_0 > 0$ 和信赖域半径的一个下界 $\bar{\rho} > 0$ 使得

$$\rho_k \geq \bar{\rho}, \quad \forall k \geq k_0.$$

又由于 $x^{(k)}$ 收敛于 x^* 且 $f(x^*) = A(x^*)^T r(x^*) = 0$, 因此

$$(A_k^T A_k)^{-1} A_k^T r_k \leq \bar{\rho}_k,$$

对所有充分大的 k 均成立, 即对所有充分大的 k , $x^{(k)}$ 即为 Gauss Newton 方程组 (5.2.3) 的解. 因此方法的收敛速度即为 Gauss Newton 法的局部收敛速度, 而我们知道, 在定理所给条件下 Gauss Newton 法是局部二次收敛的. 这就完成了定理的证明.

在定理的证明过程中我们利用了这样一个事实:对于信赖域问题(5.3.3)的解我们有下述结论(见第一章练习题16):如果 $A_k^T A_k$ 正定,且 $(A_k^T A_k)^{-1} A_k^T r_k$ 有界,则其解为 $x^{(k)} = - (A_k^T A_k)^{-1} A_k^T r_k$; 否则 $x^{(k)}$ 应是对某一 $\mu_k > 0$ 下述方程组的解

$$(A_k^T A_k + \mu_k I) x^{(k)} = - A_k^T r_k,$$

对信赖域子问题(5.3.3)的求解是信赖域方法的主要工作,也是最关键的工作.同线性搜索方法一样,也没有必要对子问题(5.3.3)进行精确求解.根据定理5.3.4的证明可以看出,只要对问题的近似求解所得的 $x^{(k)}$ 满足不等式(5.3.6),则方法的收敛性同样有保证,因此大量的信赖域方法都采用求问题(5.3.3)近似解的方法,并确保所得近似解满足不等式(5.3.6),有关对子问题(5.3.3)具体的解法见第三章.

§ 5.4 对 Gauss Newton 矩阵的拟牛顿修正

从5.2节的分析可以看出,Gauss Newton 法的收敛性主要取决于 Gauss Newton 矩阵 $M(x)$ 对 Hesse 矩阵 $^2 f(x)$ 的近似程度,信赖域方法也可以看作是加矩阵 μI 至 Gauss Newton 矩阵以改善 Gauss Newton 矩阵对 Hesse 矩阵的近似程度.

作为一个无约束最优化问题,也可以直接用拟牛顿修正公式来产生 Hesse 矩阵的近似,但这样做不能利用最小二乘问题的结构特点,即由函数的雅可比矩阵所提供的部分二阶导数的信息不能得到利用.因此,一个可能的方案是利用拟牛顿修正公式产生 Hesse 矩阵的非线性部分 $S(x)$ 的近似,再把它加到 Gauss Newton 矩阵以获取对 Hesse 矩阵好的近似.

设在迭代点 $x^{(k)}$,已有对矩阵 $S(x^{(k)})$ 的近似 C_k ,则对 Hesse 矩阵的近似可记为

$$B_k = M_k + C_k = A_k^T A_k + C_k.$$

第一个提出用拟牛顿修正产生对矩阵 $S(x^{(k)})$ 的近似是 Brown 和 Dennis (1971).由于

$$S(x^{(k)}) = \sum_{i=1}^m r_i(x^{(k)})^2 r_i(x^{(k)}),$$

$^2 r_i(x^{(k)})$ 是函数 $r_i(x)$ 的 Hesse 矩阵, Brown 和 Dennis(1971)建议用拟牛顿修正公式对每一 i 产生 $^2 r_i(x^{(k)})$ 的近似,设为 W_{ik} 再由

$$C_k = \sum_{i=1}^m r_i(x^{(k)}) W_{ik}$$

来取得对 $S(x^{(k)})$ 的近似. 不难明白这样的方法不具有实用价值, 这是因为每一次迭代除存放 $m \times n$ 阶的雅可比矩阵 A 之外, 还要存放 m 个 $n \times n$ 阶的矩阵 W_{ik} , $i=1, 2, \dots, m$, 随着 m 和 n 的增大, 这样做是不现实的. 但 Brown 和 Dennis 的想法为确定对 $S(x)$ 的近似开创了新的途径.

基于确定拟牛顿修正公式的同样的思想, 矩阵 $S(x^{(k+1)})$ 的近似 C_{k+1} 应满足方程

$$C_{k+1}^{(k)} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}, \quad (5.4.1)$$

其中 $\begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} = x^{(k+1)} - x^{(k)}$. 问题是如何确定向量 $\begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}$ 使得由方程组 (5.4.1) 确定的 C_{k+1} 是 $S(x^{(k+1)})$ 有合理解释的近似.

根据积分中值定理有

$$f(x^{(k+1)}) = f(x^{(k)}) + \int_0^1 f'(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} dt,$$

其中

$$\begin{aligned} \int_0^1 f'(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} dt &= \int_0^1 M(x^{(k)} + t \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}) dt + \\ &= \int_0^1 \sum_{i=1}^m r_i(x^{(k)} + t \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} dt \\ &= M_k(\bar{x}) + S_k(\bar{x}). \end{aligned}$$

由此得

$$S_k(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} = y^{(k)} - M_k(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}, \quad (5.4.2)$$

其中 $y^{(k)} = f(x^{(k+1)}) - f(x^{(k)})$. 比较 (5.4.1) 和 (5.4.2) 知 C_{k+1} 作为 $S_k(\bar{x})$ 的近似应满足方程组

$$C_{k+1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} = y^{(k)} - M_k(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} = \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}. \quad (5.4.3)$$

向量 $y^{(k)}$ 和 $M_k(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}$ 的不同选取可产生不同的向量 $\begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}$. 一旦向量 $y^{(k)}$ 和 $M_k(\bar{x}) \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)}$ 选定, 就可以像第四章的无约束最优化方法那样给出相应的确定矩阵 C_{k+1} 的修正公式.

Dennis, Gay 和 Welsch (1981) 在确定 C_k 的修正公式时取

$$\begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} = (A_{k+1} - A_k)^T r_{k+1},$$

这是通过取

$$A_{k+1}^T A_{k+1} \begin{pmatrix} \cdot \\ \cdot \\ \cdot \end{pmatrix}^{(k)} + (A_{k+1} - A_k)^T r_{k+1}$$

作为 $y^{(k)}$ 的近似和 $A_{k+1}^T A_{k+1}$ 为 $M_k(\bar{x})$ 的近似确定的. Dennis, Gay 和 Welsch 给出的 C_{k+1} 的修正公式为

$$C_{k+1} = C_k + \frac{y^{(k)} (y^{(k)})^T + (y^{(k)})^T y^{(k)}}{(y^{(k)})^T y^{(k)}} - \frac{(y^{(k)})^T (y^{(k)}) y^{(k)} y^{(k)T}}{((y^{(k)})^T y^{(k)}) (y^{(k)})^T y^{(k)}}, \quad (5.4.4)$$

其中

$$y^{(k)} = x^{(k)} - C_k^{-1} g^{(k)},$$

并称相应的方法为自适应算法.

Dennis, Gay 和 Welsch 的上述修正公式是通过下述约束最优化问题

$$\begin{aligned} \min \quad & C - C_k^T W, \\ \text{s.t.} \quad & C^{(k)} = y^{(k)}, \\ & C = C^T \end{aligned}$$

确定的, 其中 C_k 是对矩阵 $S(x)$ 的上一次近似, 矩阵 W 要求满足

$$W^{(k)} = y^{(k)},$$

但我们可以用通常的 DFP 修正公式来确定这样一个修正公式. 回顾确定函数 $f(x)$ 的 Hesse 矩阵近似的 DFP 修正公式,

$$B_{k+1} = B_k + \frac{1}{1 + \frac{(y^{(k)})^T B_k y^{(k)}}{(y^{(k)})^T y^{(k)}}} \frac{y^{(k)} y^{(k)T}}{(y^{(k)})^T y^{(k)}} - \frac{B_k y^{(k)} y^{(k)T} + y^{(k)} (y^{(k)})^T B_k}{(y^{(k)})^T y^{(k)}}, \quad (5.4.5)$$

其中 $y^{(k)} = x^{(k+1)} - x^{(k)}$. 在通常情况下 B_k 是对 Hesse 矩阵的上一次近似, 即应有 $B_k = A_k^T A_k + C_k$. 但是对于非线性最小二乘问题, 在点 $x^{(k+1)}$ 已知矩阵 $A_{k+1}^T A_{k+1}$, 在上述 DFP 修正公式中用 $B_k = A_{k+1}^T A_{k+1} + C_k$ 要优于取 $B_k = A_k^T A_k + C_k$. 为避免混淆, 记 $W_k = A_{k+1}^T A_{k+1} + C_k$, 并用它代替式(5.4.5)中的 B_k , 得

$$B_{k+1} = W_k + \frac{1}{1 + \frac{(y^{(k)})^T W_k y^{(k)}}{(y^{(k)})^T y^{(k)}}} \frac{y^{(k)} y^{(k)T}}{(y^{(k)})^T y^{(k)}} - \frac{W_k y^{(k)} y^{(k)T} + y^{(k)} (y^{(k)})^T W_k}{(y^{(k)})^T y^{(k)}},$$

为 Hesse 矩阵的新的近似, 从 B_{k+1} 减去 $A_{k+1}^T A_{k+1}$ 即得对 $S(x)$ 的新近似,

$$C_{k+1} = C_k + \frac{1}{1 + \frac{(y^{(k)})^T W_k y^{(k)}}{(y^{(k)})^T y^{(k)}}} \frac{y^{(k)} y^{(k)T}}{(y^{(k)})^T y^{(k)}} - \frac{W_k y^{(k)} y^{(k)T} + y^{(k)} (y^{(k)})^T W_k}{(y^{(k)})^T y^{(k)}},$$

根据 $y^{(k)}$ 及 W_k 的定义有

$$\begin{aligned} y^{(k)} + W_k y^{(k)} &= A_{k+1}^T A_{k+1} y^{(k)} + y^{(k)} + A_{k+1}^T A_{k+1} y^{(k)} + C_k y^{(k)} \\ &= 2 y^{(k)} + y^{(k)} + C_k y^{(k)} = 2 y^{(k)} + y^{(k)}. \end{aligned}$$

另一方面,

$$\begin{aligned} W_k y^{(k)} + y^{(k)T} W_k &= W_k y^{(k)} y^{(k)T} + y^{(k)} y^{(k)T} + \\ &\quad y^{(k)} (y^{(k)})^T W_k + y^{(k)} y^{(k)T} - 2 y^{(k)} y^{(k)T} \\ &= - (y^{(k)} (y^{(k)})^T + y^{(k)} y^{(k)T}) + 2 y^{(k)} y^{(k)T}. \end{aligned}$$

因此我们有

$$C_{k+1} = C_k + \frac{1}{1 + \frac{(y^{(k)})^T W_k y^{(k)}}{(y^{(k)})^T y^{(k)}}} \frac{y^{(k)} y^{(k)T}}{(y^{(k)})^T y^{(k)}} - \frac{W_k y^{(k)} y^{(k)T} + y^{(k)} (y^{(k)})^T W_k}{(y^{(k)})^T y^{(k)}}$$

$$\begin{aligned}
&= C_k + 2 - \frac{{}^{(k)T} {}^{(k)}}{{}^{(k)T} y^{(k)}} \frac{y^{(k)} y^{(k)T}}{{}^{(k)T} y^{(k)}} - \frac{2y^{(k)} y^{(k)T} - y^{(k)} {}^{(k)T} - {}^{(k)} y^{(k)T}}{{}^{(k)T} y^{(k)}} \\
&= C_k + \frac{{}^{(k)} y^{(k)T} + y^{(k)} {}^{(k)T}}{{}^{(k)T} y^{(k)}} - \frac{{}^{(k)T} {}^{(k)} y^{(k)} y^{(k)T}}{{}^{(k)T} y^{(k)} {}^{(k)T} y^{(k)}}.
\end{aligned}$$

这就是前面介绍的产生矩阵 $S(x)$ 的修正公式(5.4.4)。

类似地,如果把 BFGS 修正公式应用于上述定义的矩阵 W_k ,得矩阵 B_{k+1} 后再减去矩阵 $A_{k+1}^T A_{k+1}$ 可得产生矩阵 C_{k+1} 的另一个由 Gill 和 Murry 给出的修正公式

$$C_{k+1} = C_k + \frac{y^{(k)} y^{(k)T}}{{}^{(k)T} y^{(k)}} - \frac{W_k {}^{(k)} {}^{(k)T} W_k}{{}^{(k)T} W_k {}^{(k)}}. \quad (5.4.6)$$

然而当把上述产生的修正矩阵加至 $A_k^T A_k$ 形成 Hesse 矩阵的近似 $B_k = A_k^T A_k + C_k$,并用于实际计算时,效果并不如预期的那么好,究其原因在于如此产生的矩阵未必是 $S(x) = \sum_{i=1}^m r_i(x)^2$ 好的近似.考虑 $\sum_{i=1}^m r_i(x^*)^2$ 或较小的非线性最小二乘问题,作为 $S(x^{(k)})$ 的近似, C_k 应逐步收敛于零或变得相对较小.但是,随着迭代的进行,拟牛顿修正公式一般不会产生近似程度很高的矩阵,即 C_k 不会收敛于零.克服这一问题的一个可能方法为采用所谓的调比技术(Scaling),即在对矩阵 C_k 进行修正之前,先选择一个调比因子,设为 α_k ,然后在修正公式中用调比后的矩阵 $\alpha_k C_k$ 代替 C_k 得修正矩阵 C_{k+1} .以修正公式(5.4.4)为例,采用调比因子后的公式为

$$C_{k+1} = \alpha_k C_k + \frac{y^{(k)} {}^{(k)T} + {}^{(k)} y^{(k)T}}{{}^{(k)T} y^{(k)}} - \frac{{}^{(k)T} {}^{(k)}}{({}^{(k)T} y^{(k)})} \frac{y^{(k)} y^{(k)T}}{y^{(k)}}, \quad (5.4.7)$$

其中 ${}^{(k)} = y^{(k)} - \alpha_k C_k {}^{(k)}$.

对调比因子 α_k 的选取有许多不同的取法,这里推荐由 Dennis, Gay 和 Welsch 给出的调比因子

$$\alpha_k = \min \frac{{}^{(k)T} {}^{(k)}}{{}^{(k)T} C_k {}^{(k)}}, 1. \quad (5.4.8)$$

由于

$$\frac{{}^{(k)T} {}^{(k)}}{{}^{(k)T} C_k {}^{(k)}} = \frac{{}^{(k)T} \sum_{i=1}^m r_i(x^{(k+1)})^2 r_i(x^{(k+1)}) {}^{(k)}}{{}^{(k)T} {}^{(k)}} \cdot \frac{{}^{(k)T} {}^{(k)}}{{}^{(k)T} C_k {}^{(k)}},$$

由(5.4.8)式给出的调比因子 α_k 使得矩阵 $\alpha_k C_k$ 的谱覆盖矩阵 S_{k+1} 的谱,并避免矩阵 C_{k+1} 的元素过大.根据 ${}^{(k)}$ 的定义,如果有 $r_{k+1} = 0$,则我们有 ${}^{(k)} = (A_{k+1} - A_k)^T r_{k+1} = 0$,因此 $\alpha_k = 0$,由(5.4.7)式确定的矩阵 $C_{k+1} = 0$.可见调比因子(5.4.8)可以确保修正矩阵 C_{k+1} 的元素取值不会过大,而当 r_k 趋于零时,有 ${}^{(k)}$

趋于零,因此有 $\|C_k\| \rightarrow 0$ 从而确保采用调比后的修正矩阵序列 $\{C_k\}$ 收敛于零矩阵.

至此,可以给出解非线性最小二乘问题利用问题结构的拟牛顿型方法如下.

算法 5.4.1 (Gauss - Newton 拟牛顿型方法)

步 1. 给定初始点 $x^{(1)}$ 及初始矩阵 C_1 , 置 $k=1$;

步 2. 解方程组 $(A_k^T A_k + C_k)s = -A_k^T r_k$ 确定方向 $s^{(k)}$;

步 3. 沿方向 $s^{(k)}$ 进行搜索确定步长 α_k , 并置 $x^{(k+1)} = x^{(k)} + \alpha_k s^{(k)}$;

步 4. 如果 $x^{(k+1)}$ 满足终止条件, 停止迭代, 取 $x^{(k+1)}$ 为最优解的近似;

步 5. 计算 $\|r^{(k)}\|$, $\|y^{(k)}\|$, $\|r^{(k)}\|/\|y^{(k)}\|$, $\|r^{(k)}\|/\|y^{(k)}\|$, $\|r^{(k)}\|/\|y^{(k)}\|$, 对矩阵 C_k 进行修正得 C_{k+1} , 置 $k = k+1$ 后转步 2.

初始矩阵一般取 $C_1 = 0$, 即方法的第一步为通常的 Gauss - Newton 法, 也可以在初始的若干步内都采用 Gauss - Newton 法, 然后再采用拟牛顿修正公式产生 S_k 的近似 C_k 以形成 $B_k = A_k^T A_k + C_k$.

关于步 2 中方程组

$$(A_k^T A_k + C_k)s = -A_k^T r_k \quad (5.4.9)$$

的求解, 可以采用矩阵 A_k 的奇异值分解

$$A_k = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_l & \\ & & 0 \end{bmatrix} V^T, \quad (5.4.10)$$

其中 U, V 分别为 $m \times m$ 和 $n \times n$ 阶标准正交矩阵, $\sigma_i = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_n)$, $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n \geq 0$ 为矩阵 A_k 的奇异值. 于是方程组(5.4.9)可改写为

$$[\sigma^2 + V^T C_k V]z = -[0]U^T r_k, \quad (5.4.11)$$

其中 z 与 s 之间有下列关系

$$s = Vz. \quad (5.4.12)$$

如果矩阵 $\sigma^2 + V^T C_k V$ 呈良性, 即它的条件数不大, 则对矩阵 $\sigma^2 + V^T C_k V$ 进行 LDL^T 分解后, 用前代和回代确定(5.4.11)的解 z , 再代入(5.4.12)得所需的解 $s^{(k)}$. 如果 $\sigma^2 + V^T C_k V$ 由于条件数较大而不是良性矩阵时, 则可以采用分解迭代法确定方程组的解.

根据矩阵 A_k 的 n 个奇异值的大小可把矩阵 $\sigma^2 + V^T C_k V$ 分解成 $\sigma^2 + V^T C_k V = \text{diag}[\sigma_1^2, \dots, \sigma_l^2]$ 和 $\sigma^2 + V^T C_k V = \text{diag}[\sigma_{l+1}^2, \dots, \sigma_n^2]$ 两部分, 其中 l 为数值较大的奇异值个数. 对矩阵 V 作相应的分割

$$V = [V_1 \quad V_2],$$

其中 V_1 和 V_2 分别为 $n \times l$ 和 $n \times (n-l)$ 阶正交矩阵. 于是任意向量 $d \in \mathbf{R}^n$ 可表示成

$$d = V_1 u + V_2 v, \quad u \in \mathbf{R}^l, \quad v \in \mathbf{R}^{n-l}.$$

代入(5.4.11)得

$$\begin{bmatrix} I_l & 0 \\ 0 & I_{n-l} \end{bmatrix} u + \begin{bmatrix} V_1^T C_k V_1 & V_1^T C_k V_2 \\ V_2^T C_k V_1 & V_2^T C_k V_2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = - \begin{bmatrix} r_u \\ r_v \end{bmatrix} \quad (5.4.11)$$

将此方程组表示成两个低阶的方程组

$$\begin{bmatrix} I_l & 0 \\ 0 & I_{n-l} \end{bmatrix} u + V_1^T C_k (V_1 u + V_2 v) = - r_u, \quad (5.4.13)$$

$$V_2^T C_k V_1 u + (\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2) v = - r_v, \quad (5.4.14)$$

其中 r_u, r_v 分别为向量 $U^T r_k$ 的前 l 个分量和接下来的 $n-l$ 个分量形成的 l 和 $n-l$ 维向量. 从这两个方程组可以得到下列确定方程组(5.4.11)解的迭代法

$$\begin{aligned} \begin{bmatrix} I_l & 0 \\ 0 & I_{n-l} \end{bmatrix} u_{j+1} &= - r_u - V_1^T C_k s_j, \\ (\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2) v_{j+1} &= - r_v - V_2^T C_k V_2 u_{j+1}, \\ s_{j+1} &= V_1 u_{j+1} + V_2 v_{j+1}, \quad j=1, 2, \dots \end{aligned}$$

迭代可以从 $s_1 = 0$ 开始, 由上面第一式获得 u_2 , 代入第二式得 v_2 , 再把 u_2, v_2 代入第三式得 s_2 , 重复这一过程直至获得方程组(5.4.11)满足精度要求的解为止. 当

$$C_k \begin{bmatrix} I_l & 0 \\ 0 & I_{n-l} \end{bmatrix}^{-2} (1 + C_k (\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2)^{-1}) < 1$$

时, 上述迭代格式收敛. 对于矩阵 A_k 的一个给定的分解 V_1 和 V_2 , 如果序列 $\{s_j\}$ 不能充分快地收敛于方程组的解, 可以把 V_1 中部分小的奇异值移入矩阵 V_2 后对新的剖分重新开始迭代.

在利用上述迭代格式求方程组(5.4.11)解的过程中, 只需要解一个低阶的方程组

$$(\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2) v_{j+1} = - r_v - V_2^T C_k V_2 u_{j+1}.$$

当矩阵 $A_k^T A_k + C_k$ 正定时, 矩阵 $\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2$ 也正定, 由于矩阵 $\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix}$ 不包含矩阵中大的奇异值, 矩阵 $\begin{bmatrix} I_{n-l} & 0 \\ 0 & I_{n-l} \end{bmatrix} + V_2^T C_k V_2$ 具有好的性态, 因而用它的 LDL^T 分解可以稳定地求解.

由于矩阵 $A_k^T A_k$ 正定或半正定, 矩阵 $A_k^T A_k + C_k$ 的正定性由矩阵 C_k 的正定性保证. 然而由上述修正公式所产生的矩阵 C_k 并不能像无约束最优化的拟牛顿修正公式那样确保修正矩阵的正定性(见第四章). 因此有些方法在采用本节的技术时一般采用信赖域方法, 也就是把上述方法的第2步和第3步改成:

步 2. 确定一个合适的信赖域半径 ρ_k ,

步 3. 解下列信赖域问题

$$\min q_k(\delta) = f(x^{(k)}) + \delta^T A_k^T r_k + \frac{1}{2} \delta^T (A_k^T A_k + C_k) \delta,$$

$$\text{s.t.} \quad \|\delta\| \leq \rho_k,$$

得解 $\delta^{(k)}$, 并置 $x^{(k+1)} = x^{(k)} + \delta^{(k)}$.

有关这一信赖域问题的求解见第三章, 这里不再重复.

小 结

实际中出现的最优化问题的模型,大多数是具有若干函数平方和形式的线性或非线性最小二乘问题,本章主要讨论和介绍能有效地利用这些问题特殊结构的最小二乘求解方法,全章共分四节,分别介绍了线性最小二乘问题的解法、非线性最小二乘问题的 Gauss Newton 法、信赖域方法和对 Gauss Newton 矩阵的拟牛顿修正等.

在 5.1 节就线性最小二乘问题的求解介绍了有关的方法,它们是求解无约束线性最小二乘问题的正交分解算法,具有线性等式约束的线性最小二乘问题的正交分解算法和 Lagrange 乘子法,以及求解带有线性不等式约束的线性最小二乘问题的有效集方法. 5.1.3 节仅供感兴趣的同学选学.

从 5.2 节开始主要围绕非线性最小二乘问题的求解. 5.2 节介绍了著名的 Gauss Newton 法,该方法利用最小二乘问题的特殊结构由函数的一阶导数的信息形成 Gauss Newton 矩阵,直接构造函数二阶导数矩阵的近似,再用牛顿法的迭代格式求问题的最优解.我们在文中给出了算法的迭代格式,分析了算法的收敛性和算法的具体特征,方法适用问题的类型,以及方法的不足,并由此给出了方法的一种改进阻尼 Gauss Newton 法.

第 5.3 节主要介绍最优化问题的另一类重要方法——信赖域方法,由于这一类方法始于最小二乘问题,我们把它放在这一节加以介绍,但这一类方法早已广泛地应用于各种不同类型的最优化问题.可以这样说,线性搜索类方法把一个复杂的最优化问题转化成一系列简单的一维寻优问题,信赖域方法则是把优化问题转化成一系列相对简单的局部寻优问题.我们在这一节介绍了信赖域方法的基本原理在 Gauss Newton 矩阵的基础上再加上单位矩阵的一个适当的倍数以提高对二阶导数矩阵的近似,给出了方法的推导过程和方法的步骤,分析了方法的收敛性和收敛速度.

第 5.4 节进一步介绍对 Gauss Newton 法的修正,通过对 Gauss Newton 矩阵加上由拟牛顿修正公式确定的修正矩阵来提高对二阶导数矩阵的近似程度.在这一节里给出了若干这样的拟牛顿修正公式,分别给出了这些修正公式的推导原理和基本过程,还给出了利用这些公式求解非线性最小二乘问题的拟牛顿型方法的步骤.



1. 设 $f(x) = \frac{1}{2} \sum_{i=1}^5 r_i(x)^2$, 其中

$$r_1(x) = x_1^2 + x_2^2 + x_3^2 - 1,$$

$$r_2(x) = x_1^2 + x_2^2 + (x_3 - 2)^2 - 1,$$

$$r_3(x) = x_1 + x_2 + x_3 - 1,$$

$$r_4(x) = x_1 + x_2 - x_3 + 1,$$

$$r_5(x) = x_1^3 + 3x_2^2 + (5x_3 - x_1 + 1)^2 - 36,$$

通过计算 $r(x)$ 的雅可比矩阵确定 $f(x)$, $M(x) = A(x)^T A(x)$ 以及 $\nabla^2 f(x)$. 在点 $x = (0, 0, 1)^T$ 是否有 $M(x) = \nabla^2 f(x)$ 成立? 为什么?

2. 设 $f(x) = \frac{1}{2} \sum_{i=1}^4 r_i(x)^2$, $r_i(x) = x_1 e^{-x_2 t_i} - y_i$, $i = 1, 2, 3, 4$, 其中 $t_1 = -1$, $t_2 = 0$,

$t_3 = 1$, $t_4 = 2$, $y_1 = 2.7$, $y_2 = 1$, $y_3 = 0.4$, $y_4 = 0.1$.

(1) 以 $x^{(1)} = (1, 1)^T$ 为初始点计算进行一次 Gauss Newton 迭代所得的点 $x^{(2)}$;

(2) 在点 $x^{(2)}$ 还是用矩阵 $A_1^T A_1$ 进行一次近似的 Gauss Newton 迭代得点 $x^{(3)}$. 由这三个点估计方法对此问题的线性收敛速度.

3. 设 (1) 和 (2) 分别是方程组

$$(A^T A + \mu I) = -A^T r, i = 1, 2$$

对应于 μ 和 μ 的解, 其中 $\mu > \mu > 0$, $A \in \mathbf{R}^{m \times n}$, $r \in \mathbf{R}^m$. 证明有

$$q(x^{(2)}) < q(x^{(1)}),$$

这里 $q(x) = \frac{1}{2} \|A x + r\|^2$.

4. 设 $f(x) = x_1^4 + x_1 x_2 + (1 + x_2)^2$, $x^{(1)} = (0, 0)^T$, 确定参数 μ 的一个下界 $\bar{\mu}$ 使得 $\nabla^2 f(x^{(1)}) + \mu I$ 在 $\mu > \bar{\mu}$ 时是正定的. 令 $\mu = 1$, 确定方程组

$$(\nabla^2 f(x)^{(1)} + \mu I) = -\nabla f(x^{(0)})$$

的解 $x^{(1)}$, 并验证有 $f(x^{(1)} + x^{(1)}) < f(x^{(1)})$.

再验证只有当 $\mu = 0.9$ 时所得的解 $x^{(1)}$ 才能使 $f(x)$ 在 $x^{(1)} + x^{(1)}$ 处的函数值小于点 $x^{(1)}$ 处的函数值, 而最优的下降量在 $\mu = 1.2$ 时近似取得.

5. 设 $A \in \mathbf{R}^{m \times n}$, $f = b \in \mathbf{R}^n$, $\mu = 0$, $x^{(1)}$ 满足下述方程组

$$(A^T A + \mu I) = -b,$$

证明: (1) 如果 $\mu = 0$, 则 $x^{(1)}$ 是下述信赖域问题 (TRP) 的解

$$\min q(x) = f + x^T + \frac{1}{2} x^T A^T A x,$$

s.t. ;

(2) 如果 $\mu = 0$ 且有 $x^{(1)} = 0$, 则 $x^{(1)}$ 也是问题 (TRP) 的解;

(3) 如果 $\mu > 0$, 则 $x^{(1)}$ 还是 (TRP) 的唯一解.

6. 设 $d \in \mathbf{R}^n$ 为一满足

$$d^T f = c_1 d \quad f$$

的下降方向,其中 $c_1 \in (0,1]$. 又设 \bar{x} 是问题 5 中 (TRP) 的一个近似解,满足

$$\text{Pred}(\bar{x}) = f - q(\bar{x}) = f - \min \left\{ f + d^T \left(f + \frac{1}{2} d^T A^T A d \right) \mid d \right\},$$

证明

$$\text{Pred}(\bar{x}) = \frac{c_1}{2} \|f\|^2 \min_{\|d\|=1} \left(\frac{f^T A^T A d}{\|A^T A d\|} \right)^2.$$

7. 设

$$q(x) = \frac{1}{2} x^T G x + b^T x,$$

其中

$$G = \begin{pmatrix} 14 & 1 \\ 1 & 2 \end{pmatrix}, b = (6 \ 2)^T.$$

分别计算 $q(x)$ 沿最速下降方向的最优修正,记为 x_{cp} ,以及由方程组

$$Gx = -b$$

所确定的牛顿修正,记为 x_N . 验证有 $x_N > x_{cp}$. 令 $\phi(\alpha) = x_{cp} + \alpha(x_N - x_{cp})$, 确定 $(0,1)$ 的一个值使得 $\phi(\alpha) = 1$.

8. 考虑线性等式约束的线性最小二乘问题

$$\begin{aligned} \min \quad & \frac{1}{2} \|Ax - b\|_2^2, \\ \text{s.t.} \quad & Cx = y, \end{aligned}$$

其中 $x \in \mathbf{R}^n$, $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, $C \in \mathbf{R}^{p \times n}$ 和 $y \in \mathbf{R}^p$, $p < n$. 设 $\text{rank}(A) = n$, $\text{rank}(C) = p$, 证明 $x^* \in \mathbf{R}^n$ 是问题唯一解的充分必要条件为存在 Lagrange 乘子 λ^* 使得 (x^*, λ^*) 是下述方程组的解,并验证 x^*, λ^* 由 (5.1.10) 和 (5.1.11) 给出.

9. 对于问题 8,如果矩阵 A 不满秩,证明问题的解集合为

$$S = \{ (AE)^+ (b - AC^+ y) + C^+ y + (I - (AE)^+ A) N(C) \},$$

而目标函数的最优值为

$$\frac{1}{2} \|(I_A (AE)^+) (AC^+ y - b)\|_2^2,$$

其中 $E = I - C^+ C$, $N(C)$ 是矩阵 C 的零空间, C^+ 是矩阵 C 的广义逆, $(AE)^+$ 是矩阵 AE 的广义逆.

10. 设

$$C^T = QR = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \begin{matrix} \mathbf{R}_1^T \\ 0 \end{matrix}$$

其中 $Q = [V_1 \ V_2]$ 为一 $n \times n$ 正交矩阵, $V_1 \in \mathbf{R}^{n \times p}$, $V_2 \in \mathbf{R}^{n \times (n-p)}$. 假定 Lagrange 矩阵

$$\begin{aligned} A^T A - C^T \\ - C \quad 0 \end{aligned}$$

非奇异,证明:(1) 矩阵 C 满秩;(2) 矩阵 $V_2^T A^T A V_2$ 非奇异.

11. 设 $r_c = r(x^{(c)})$, $A_c = A(x^{(c)})$ 满秩, $S_c = S(x^{(c)}) = \sum_{i=1}^m r_i(x^{(c)})^2 r_i(x^{(c)})$,

$G_c = -(A_c^T A_c)^{-1} A_c^T r_c$, $N_c = -(A_c^T A_c + S_c)^{-1} A_c^T r_c$. 证明

$$x_c^G - x_c^N = (A_c^T A_c)^{-1} S_c^{(k)}.$$

设 x_c 充分接近 $f(x) = \frac{1}{2} \sum_{i=1}^m r_i(x)^2$ 的最优解 x^* , $f(x)$ 的 Hesse 矩阵 $H^2 f(x) = A(x)^T A(x) + S(x)$ 在点 x^* 非奇异, 在点 x^* 的邻域内 Lipschitz 连续, 利用上述关系式证明

$$x_c^G - x^* = (A_c^T A_c)^{-1} S_c (x_c - x^*) + O(\|x_c - x^*\|^2),$$

其中 $x_c^G = x_c + \frac{G}{c}$.

12. 已知 $r \in \mathbf{R}^m$, $A \in \mathbf{R}^{m \times n}$, $\mu > 0$. 证明 $x = -(A^T A + \mu I)^{-1} A^T r$ 是下述最小二乘问题的解

$$\min \|J + y\|^2,$$

其中

$$J = \begin{bmatrix} A \\ \mu^{\frac{1}{2}} I \end{bmatrix}, y = \begin{bmatrix} r \\ 0 \end{bmatrix}.$$

13. 考虑非线性最小二乘问题

$$\min f(x) = \frac{1}{2} \sum_{i=1}^m r_i(x)^2,$$

并假定 $r(x)$ 的雅可比矩阵 $A(x)$ 对所有 $x \in \mathbf{R}^n$ 都满秩, 用 G , $LM(\mu)$ 和 S 分别表示在点 x 的 Gauss Newton 方向, Levenberg Marquardt 方向和最速下降方向, 即有

$$\begin{aligned} G &= (A^T A)^{-1} A^T r, \\ LM(\mu) &= (A^T A + \mu I)^{-1} A^T r, \\ S &= -A^T r. \end{aligned}$$

证明

$$\begin{aligned} \lim_{\mu \rightarrow 0} LM(\mu) &= G, \\ \lim_{\mu} \frac{LM(\mu)}{LM(\mu)} &= \frac{S}{S}. \end{aligned}$$

第 6 章

二 次 规 划

§ 6.1 二 次 规 划

一个带有二次目标函数和线性约束的最优化问题称为二次规划.这类问题不仅自身有着重要作用,而且在通常的约束问题的求解中也有着重要应用,例如序列二次规划方法(SQP)和增广 Lagrange 方法等.

通常,二次规划(QP)问题叙述如下:

$$\min \quad q(x) = \frac{1}{2} x^T G x + d^T x, \quad (6.1.1)$$

$$\text{s.t.} \quad a_i^T x = b_i, \quad i \in I, \quad (6.1.2)$$

$$a_i^T x \leq b_i, \quad i \in J, \quad (6.1.3)$$

其中 G 是 $n \times n$ 对称矩阵, I 与 J 分别对应等式约束和不等式约束的有限的指标集, d, x 和 $\{a_i\}, i \in I \cup J$ 都是 n 维向量.二次规划能在有限次迭代中被解或者被证明是不可行的,但寻求解的效果非常依赖于目标函数的特性和不等式约束的数目.如果 Hesse 矩阵是正半定的,则称(6.1.1)——(6.1.3)式的问题是凸二次规划(QP),在这种情况下,此问题有时并不比解线性规划更困难.非凸二次规划,即目标函数的 Hesse 矩阵 G 是不定的,因为此规划可能有多个稳定点和局部极点,所以此规划更富有挑战性与更为困难.

下面给出具体的投资规划问题:每个投资者都认识到在投资中有风险和回报.为了增加在投资中期望的回报,投资者将必须承担更大风险.投资理论研究具有回报率 $r_i (i = 1, 2, \dots, n)$ 的 n 个可能投资如何投入的模型.此回报率 r_i 可能在未来通常并不知,然而通常假设是以正态分布的随机变量.我们使用期望值 $\mu_i = E[r_i]$ 与方差 $\sigma_i^2 = E[(r_i - \mu_i)^2]$ 构成此变量的特征,此方差测定变量 r_i 关于其均值的影响度,这样比值 σ_i^2 / μ_i^2 越大表示风险越大.

一个投资者构建以整个投资额的比率 x_i 投资于项目 $i (i = 1, 2, \dots, n)$ 的投资方案.假设所有投资额都投资并且不允许贷款,则约束是
$$\sum_{i=1}^n x_i = 1 \text{ 且 } x_i \geq 0$$

$(x_1, \dots, x_n)^T \geq 0$, 投资利润回报由下面给出

$$R = \sum_{i=1}^n x_i r_i. \quad (6.1.4)$$

为测定投资的合适利润, 需要获得期望回报与方差, 此期望回报简单记为:

$$E[R] = E \sum_{i=1}^n x_i r_i = \sum_{i=1}^n x_i E[r_i] = x^T \mu. \quad (6.1.5)$$

从统计的基本定律, 其方差也能计算. 它依赖于每对投资的相关协方差, 定义为

$$\sigma_{ij} = \frac{E[(r_i - \mu_i)(r_j - \mu_j)]}{\sigma_i \sigma_j}, \quad i, j = 1, 2, \dots, n. \quad (6.1.6)$$

相关系数测定了关于投资 i 和 j 项目的回报以相向方向的趋势. 两个项目投资回报一起上升和下降有正的协方差. σ_{ij} 越接近于 1, 两项投资互相越接近. 两项投资回报以反向运动具有负的协方差. 整个投资的方差 R 由下面给出:

$$E[(R - E(R))^2] = \sum_{i=1}^n \sum_{j=1}^n x_i x_j \sigma_{ij} = x^T G x, \quad (6.1.7)$$

其中定义 $n \times n$ 对称矩阵 G 的分量为

$$G_{ij} = \sigma_{ij} \sigma_i \sigma_j. \quad (6.1.8)$$

可以证明 G 是正半定的.

我们有兴趣研究在投资中所期望的回报 $x^T \mu$ 为最大而方差 $x^T G x$ 为最小, 在 Markowitz 提供的模型中, 组合上述两个目标加入风险允许参数 μ 后, 组成单目标函数, 求解下面问题获得最优组合投资.

$$\max \quad x^T \mu - x^T G x \quad (6.1.9)$$

$$\text{s.t.} \quad \sum_{i=1}^n x_i = 1 \quad (6.1.10)$$

$$x \geq 0, \quad (6.1.11)$$

这里参数 $\mu \in [0, \infty)$, 所选择的值依赖于个人投资的偏好, 协作投资更强调投资的风险, 所以他们将选择方差在目标函数中的权重, 而增加 μ 的值. 更冒险的投资者希望获得更大期望收益, 所以他们取 μ 接近于零. 应用组合投资优化问题的困难是在实际生活中如何确定期望回报、方差和协方差. 一种可能是选取历史数值来定义这些量 μ_i , σ_i 和 σ_{ij} . 当然, 不希望假定未来逊于过去. 进一步, 就是历史数值获得也可能不方便. 金融专家通常结合历史数值和自己的经验与期望产生值 μ_i , σ_i 和 σ_{ij} .

二次规划问题不仅本身具有相当广泛的实际应用背景, 而且正如我们在下一章中所看到的, 在求解其他类型的约束优化问题时, 常常需要求解一系列作为子问题出现的二次规划问题. 因此, 探讨有效求解问题 (6.1.1) — (6.1.3) 的方法是十分重要的.

作为非线性约束优化问题的特例, 自然可以用解约束非线性规划的某种方

法来求解二次规划问题.但由于其特殊性,我们一方面可以给出前述方法的简化与改进,使得算法更有效,如直接导出解析解,或者得到具有有限步终止性的好算法等;另一方面,我们还可探讨新型求解算法,如内点算法等.下面分别予以讨论.

§ 6.2 等式约束二次规划问题

等式约束二次规划问题可描述成下列形式

$$\min_{x \in \mathbf{R}^n} q(x) = \frac{1}{2} x^T G x + g^T x, \quad (6.2.1)$$

$$\text{s.t. } A^T x = b, \quad (6.2.2)$$

其中 $b \in \mathbf{R}^m$, $A \in \mathbf{R}^{n \times m}$.不失一般性,以下假设 A 为列满秩的,即其秩为 m .对于问题(6.2.1)–(6.2.2),可以应用各种消去法求解.又由于其特殊性,还可以给出其解析解.现在使用经典的直接消去法.设对等式约束(6.2.2)中矩阵 A 分块使得 A_B 为 $m \times m$ 非奇异方阵.此时等式约束(6.2.2)可改写成

$$A_B^T x_B + A_N^T x_N = b, \quad (6.2.3)$$

相关的量 x, g, A 与 G 作如下分块,

$$x = \begin{pmatrix} x_B \\ x_N \end{pmatrix}, \quad A = \begin{pmatrix} A_B \\ A_N \end{pmatrix}, \quad G = \begin{pmatrix} G_{BB} & G_{BN} \\ G_{NB} & G_{NN} \end{pmatrix}, \quad g = \begin{pmatrix} g_B \\ g_N \end{pmatrix},$$

其中 $x_B \in \mathbf{R}^m$, $x_N \in \mathbf{R}^{n-m}$, 其余类似.该分块使得 A_B 为 $m \times m$ 非奇异方阵.因 A_B^{-1} 存在,此时由方程(6.2.3)可得

$$x_B = A_B^{-T} (b - A_N^T x_N). \quad (6.2.4)$$

将此代入 $q(x)$,则可使问题(6.2.1)–(6.2.2)转化为下列无约束优化问题

$$\min_{x_N \in \mathbf{R}^{n-m}} \frac{1}{2} x_N^T \hat{G} x_N + \hat{g}^T x_N + \hat{c}, \quad (6.2.5)$$

其中

$$\hat{G} = G_{NN} - G_{NB} A_B^{-T} A_N^T - A_N A_B^{-1} G_{BN} + A_N A_B^{-1} G_{BB} A_B^{-T} A_N^T,$$

$$\hat{g} = g_N - A_N A_B^{-1} g_B + (G_{NB} - A_N A_B^{-1} G_{BB}) A_B^{-1} b,$$

$$\hat{c} = \frac{1}{2} b^T A_B^{-1} G_{BB} A_B^{-T} b + g_B^T A_B^{-T} b.$$

如果 \hat{G} 正定,则易知问题(6.2.5)的最优解为 $x_N^* = -\hat{G}^{-1} \hat{g}$.代入(6.2.4)即可确定对应的 x_B^* ,从而找到问题(6.2.1)–(6.2.2)的最优解

$$x^* = \begin{pmatrix} x_B^* \\ x_N^* \end{pmatrix} = \begin{pmatrix} A_B^{-T} b + A_B^{-T} A_N^T \hat{G}^{-1} \hat{g} \\ -\hat{G}^{-1} \hat{g} \end{pmatrix}, \quad (6.2.5a)$$

相应的最优 Lagrange 乘子 λ^* 可由下式确定,

$$A^T \lambda^* = g + Gx^* . \quad (6.2.6)$$

其实,只须考虑该方程组的前 m 行即给出 λ^* 的表达式如下,

$$\lambda^* = A_B^{-1} (g_B + G_{BB} x_B^* + G_{BN} x_N^*) .$$

如果 \hat{G} 正半定,则不难证明当

$$(I - \hat{G} \hat{G}^+) \hat{g} = 0 \quad (6.2.7)$$

时,问题(6.2.5)有界,且其最优解为

$$x_N^* = -\hat{G}^+ \hat{g} + (I - \hat{G}^+ \hat{G}) \tilde{x} ,$$

这里 $\tilde{x} \in \mathbf{R}^{n-m}$ 为任意向量, \hat{G}^+ 表示 \hat{G} 的广义逆矩阵.此时,原问题(6.2.1)—(6.2.2)的最优解与最优乘子则可通过上式与(6.2.3)类似地确定.若(6.2.7)不成立,则可以推出问题(6.2.5)无下界,从而原问题(6.2.1)—(6.2.2)亦无下界.

如果 \hat{G} 存在负的特征值,则问题(6.2.5)无下界,原问题(6.2.1)—(6.2.2)不存在有限最优解.

广义消去法

直接消去法简单明了,但因要计算 A_B 的逆涉及多个矩阵相乘,在 A_B 近于奇异时可能由于计算误差而不稳定.消去法的一个直接推广为广义消去法.基于 A 为列满秩,可以将全空间 \mathbf{R}^n 划分成两个互补的子空间:一个由 A 的列所生成的像空间,而另一个是 A^T 的零空间.设 Y 为由 A 的列所生成的像空间的一组基所构成的矩阵,而 Z 为由 A^T 的零空间的一组基所构成的矩阵. A 的秩为 m ,故 Y 和 Z 分别有 m 和 $n-m$ 个列,即 $Y \in \mathbf{R}^{n \times m}$ 与 $Z \in \mathbf{R}^{n \times (n-m)}$,并且满足矩阵 $[Y \ Z]$ 是正交矩阵,而 $A^T Z = 0$, $A^T Y$ 为非奇异矩阵.不失一般性,选择 Y 满足 $A^T Y = I$,令

$$x = Y \bar{x} + Z \hat{x} . \quad (6.2.8)$$

由约束条件 $A^T x = b$ 得

$$A^T Y \bar{x} + A^T Z \hat{x} = b ,$$

从而 $\bar{x} = b$.这样分解(6.2.8)可以写成

$$x = Yb + Z \hat{x} . \quad (6.2.9)$$

将上式代入(6.2.1),则可将原问题转化为下列无约束优化问题

$$\min \frac{1}{2} \hat{x}^T Z^T GZ \hat{x} + (g + GYb)^T Z \hat{x} + \frac{1}{2} (2g + GYb)^T Yb . \quad (6.2.10)$$

若矩阵 $Z^T GZ$ 正定,则此问题存在唯一最优解 \hat{x}^* ,它是下列方程组的唯一解,

$$(Z^T GZ) \hat{x} = -Z^T (g + GYb) , \quad (6.2.11)$$

则 $\hat{x}^* = -(Z^T GZ)^{-1} Z^T (g + GYb)$,代入(6.1.3)得问题的最优解为

$$x^* = Yb - Z(Z^T GZ)^{-1} Z^T (g + GYb), \quad (6.2.12)$$

而由(6.2.6)和(6.2.12)可以确定相应的最优 Lagrange 乘子 λ^* 为

$$\lambda^* = Y^T (Gx^* + g). \quad (6.2.13)$$

实际上,一般不会直接由上面两式来确定 x^* 与 λ^* ,而是通过计算矩阵 $Z^T GZ$ 的 LL^T 或 LDL^T 分解,用分解法来求方程组(6.2.11)的解 \hat{x}^* ,然后依(6.2.9)计算 x^* 以及 λ^* .这样做的好处是,它不仅可增加方法的数值稳定性,而且在分解过程中可直接检验 $Z^T GZ$ 是否正定.为控制误差传播与增大,可采用 QR 分解法确定所需的矩阵 Z .

通过选取不同的 Y 与 Z ,可得到计算 x^* 与 λ^* 的不同公式.可以证明,上述两种方法都是先确定最优解,再求出最优 Lagrange 乘子.其实,利用问题(6.2.1)—(6.2.2)的特殊性,存在同时确定其最优解与相应乘子的方法.这就是基于问题(6.2.1)—(6.2.2)的 Lagrange 函数而导出的所谓 Lagrange 方法.问题(6.2.1)—(6.2.2)的 Lagrange 函数为

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

二次规划问题的一阶必要条件,即 KKT 条件为

$$\begin{aligned} \frac{\partial L(x, \lambda)}{\partial x} &= Gx + g - A\lambda = 0, \\ \frac{\partial L(x, \lambda)}{\partial \lambda} &= A^T x - b = 0. \end{aligned}$$

其矩阵形式为

$$\begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} g \\ b \end{bmatrix}. \quad (6.2.14)$$

在上式中,系数矩阵称为 KKT 矩阵.

定理 6.2.1 假设 A 为列满秩矩阵,其秩为 m ,若投影 Hesse 矩阵 $Z^T GZ$ 正定,则问题(6.2.1)—(6.2.2)的 KKT 矩阵

$$K = \begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix}$$

是非奇异的,从而存在唯一的 KKT 对 (x^*, λ^*) 满足方程组(6.2.14).

证明 假设矩阵 K 是奇异的,则存在向量 $(p, v) \neq 0$ 使得

$$\begin{bmatrix} G & -A \\ -A^T & 0 \end{bmatrix} \begin{bmatrix} p \\ v \end{bmatrix} = 0.$$

因为 $A^T p = 0$,则从上式可得

$$0 = \begin{bmatrix} p^T & G & -A \\ v^T & -A^T & 0 \end{bmatrix} \begin{bmatrix} p \\ v \end{bmatrix} = p^T Gp.$$

因为 p 在 A^T 的零空间则可表示成 $p = Zu$,对某个 $u \in \mathbf{R}^{(n-m)}$.因此,

$$0 = p^T G p = u^T Z^T G Z u.$$

由矩阵 $Z^T G Z$ 正定, 则有 $u = 0$, 从而 $p = 0$. 又由 $A^T v = 0$ 和 A 为列满秩, 可得 $v = 0$. 这意味着与 $(p, v) \neq 0$ 矛盾. 由此定理的结论成立.

上述定理表明, 求解 KKT 方程组 (6.2.14) 即可同时得到最优解 x^* 和最优 Lagrange 乘子 λ^* .

例 6.2.2 考虑等式约束二次规划问题

$$\begin{aligned} \min \quad & q(x) = 3x_1^2 + 2x_1x_2 + x_1x_3 + 2.5x_2^2 + 2x_2x_3 + 2x_3^2 - 8x_1 - 3x_2 - 3x_3 \\ \text{s.t.} \quad & x_1 + x_3 = 3 \\ & x_2 + x_3 = 0. \end{aligned}$$

解 使用矩阵形式, 有

$$G = \begin{pmatrix} 6 & 2 & 1 \\ 2 & 5 & 2 \\ 1 & 2 & 4 \end{pmatrix}, \quad d = \begin{pmatrix} -8 \\ -3 \\ -3 \end{pmatrix}, \quad A^T = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

最优解 x^* 和相应的 Lagrange 乘子 λ^* 为

$$x^* = \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}, \quad \lambda^* = \begin{pmatrix} 3 \\ -2 \end{pmatrix}.$$

在此例中, 矩阵 G 是正定的, 而零空间的基矩阵定义为

$$Z = (-1, -1, 1)^T.$$

现在给出矩阵 K 的惯性指数定理, 其证明可参见其他参考书.

定理 6.2.3 假设 A 为列满秩, 其秩为 m , 若投影 Hesse 矩阵 $Z^T G Z$ 正定, 则问题 (6.2.1) — (6.2.2) 的 KKT 矩阵 K 有 n 个正的特征值, m 个负的特征值, 并且没有零特征值.

因为矩阵 K 是非奇异的, 则逆存在, 设

$$K^{-1} = \begin{pmatrix} G & -A^{-1} \\ -A^T & 0 \end{pmatrix} = \begin{pmatrix} H & -T \\ -T^T & U \end{pmatrix}$$

方程组 (6.2.14) 的唯一解可写成

$$x^* = -Hg + Tb, \quad (6.2.15)$$

$$\lambda^* = T^T g - Ub. \quad (6.2.16)$$

实际上, 若已知问题 (6.2.1) — (6.2.2) 的一个可行解 x_k , 则这两个方程可改为

$$x^* = x_k - Hg_k, \quad \lambda^* = T^T g_k,$$

其中 $g_k = Gx_k + g$. 这样, 我们唯一地同时确定了问题 (6.2.1) — (6.2.2) 的最优解 x^* 与最优 Lagrange 乘子 λ^* . 注意, 上述矩阵的非奇异性并不一定要求 G^{-1} 存在. 由于可用不同的方法给出分块矩阵 H, T, U 的不同表达式, 所以就有不同形式的计算公式 (6.2.15) — (6.2.16). 下面给出两种不同的形式.

首先,若 G 可逆,因 A 列满秩, $(A^T G^{-1} A)^{-1}$ 存在,这时可给出 H, T 与 U 的表达式如下:

$$\begin{aligned} H &= G^{-1} - G^{-1} A (A^T G^{-1} A)^{-1} A^T G^{-1}, \\ T &= G^{-1} A (A^T G^{-1} A)^{-1}, \\ U &= - (A^T G A)^{-1}. \end{aligned}$$

显然,应按先 U , 后 T , 再 H 的顺序计算这些矩阵. 由于这些公式中只出现 A , 故常称此方法为值空间方法.

如果我们已确定了矩阵 Y 与 Z , 则不难利用 Y 与 Z 应满足的条件得到 H, T, U 的另一组表达式

$$\begin{aligned} H &= Z (Z^T G Z)^{-1} Z^T, \\ T &= Y - Z (Z^T G Z)^{-1} Z^T G Y, \\ U &= Y^T G Z (Z^T G Z)^{-1} Z^T G Y - Y^T G Y, \end{aligned}$$

代入(6.2.15)与(6.2.16), 并与(6.2.12)和(6.2.13)比较可知二者此时是相等的. 与(6.2.11)的求解类似, 为避免矩阵直接求逆所具有的潜在不稳定性, 我们也可用分解法来求解 KKT 方程组(6.2.14). 不过这时可采用逐步分解的方法. 设 G 正定, 先求出矩阵 G 的 LDL^T 分解, 再用此分解因子消去(6.2.14)中系数矩阵的非对角元 $-A$ 与 $-A^T$, 此时对应的右下角对角块从 0 变为 $-A^T G^{-1} A$, 易知这是一个负定矩阵, 再把它分解成 $\hat{L} \hat{D} \hat{L}^T$ 的形式, 则可得到(6.2.14)系数矩阵的分解

$$K = \bar{L} \bar{D} \bar{L}^T,$$

其中

$$\bar{L} = \begin{bmatrix} L & 0 \\ B^T & \hat{L} \end{bmatrix}, \quad \bar{D} = \begin{bmatrix} D & \\ & \hat{D} \end{bmatrix},$$

其中 B 由 $LDB = -A$ 定义, 并易于通过回代确定. 为避免形成 $A^T G^{-1} A$ 时引起的计算损失, \hat{L} 与 \hat{D} 的计算最好采用先形成 $D^{-\frac{1}{2}} L^{-1} A$, 再由 QR 方法将其分解成 $Q \hat{D}^{\frac{1}{2}} \hat{L}^T$. 实际结果表明这种逐步分解法相当有效, 且特别适合于 G 有某种特殊结构即当约束数较小的情形. 下面给出使用这种逐步分解法解 KKT 系统的算法.

算法 6.2.4

步 1. 解 $Ly = \begin{bmatrix} g + Gx \\ Ax - b \end{bmatrix}$, 可得 y ;

步 2. 解 $B\hat{y} = y$, 得 \hat{y} ;

步 3. 解 $L\bar{y} = \hat{y}$, 可得 \bar{y} ;

步 4. 置 $\bar{x} = x^* - d^*$.

以上方法求解问题(6.2.1)—(6.2.2), 可得整体最优解.

定理 6.2.5 假设 A 为列满秩, 其秩为 m , 若投影 Hesse 矩阵 $Z^T GZ$ 正定, 则满足方程组(6.2.14)的 KKT 对 (x^*, λ^*) 中 x^* 是问题(6.2.1)—(6.2.2)的唯一整体最优解.

证明 假设 x 是问题(6.2.1)—(6.2.2)的任何可行解, 即 $A^T x = b$. 令 $d = x^* - x$. 因为 $A^T x^* = A^T x = b$, 则有 $A^T d = 0$. 代入(6.2.1)式的目标函数, 可得

$$\begin{aligned} q(x) &= \frac{1}{2} (x^* - d)^T G(x^* - d) + g^T (x^* - d) \\ &= \frac{1}{2} d^T Gd - d^T Gx^* - g^T d + q(x^*). \end{aligned}$$

由(6.2.14), KKT 对 (x^*, λ^*) 满足 $Gx^* = -g + A\lambda^*$, 又由 $A^T d = 0$, 得

$$d^T Gx^* = d^T (-g + A\lambda^*) = -g^T d.$$

代入前式可得

$$q(x) = \frac{1}{2} d^T Gd + q(x^*).$$

因为 d 在 A^T 的零空间, 则可表示成 $d = Zu$, 对某个 $u \in \mathbf{R}^{(n-m)}$. 因此,

$$q(x) = \frac{1}{2} u^T (Z^T GZ) u + q(x^*).$$

因为矩阵 $Z^T GZ$ 正定, 则当 $d \neq 0$, 即 $x^* \neq x$ 时, 有

$$q(x) > q(x^*).$$

因此, x^* 是问题(6.2.1)—(6.2.2)的唯一整体最优解.

如果矩阵 $Z^T GZ$ 不定, 设满足方程组(6.2.14)的 KKT 条件对 (x^*, λ^*) , 对某个 $u \in \mathbf{R}^{(n-m)}$ 使 $u^T Z^T GZ u < 0$, 并设 $p = Zu$, 则对任意 α 都有

$$A^T (x^* + \alpha p) = b,$$

所以 $x^* + \alpha p$ 为可行解, 而由 $p^T (Gx^* + g) = p^T (A\lambda^*) = u^T Z^T A\lambda^* = 0$, 可得

$$\begin{aligned} q(x^* + \alpha p) &= q(x^*) + \frac{1}{2} \alpha^2 p^T (Gx^* + g) + \frac{1}{2} \alpha^2 p^T Gp \\ &= q(x^*) + \frac{1}{2} \alpha^2 p^T Gp \\ &= q(x^*). \end{aligned}$$

从上所述, 对于满足方程组(6.2.14)的 KKT 对 (x^*, λ^*) , 可找到可行方向 $p = Zu$ 使目标函数 $q(\cdot)$ 并不增加. 从而不难证明原问题(6.2.1)—(6.2.2)亦无下界.

例 6.2.6 同例 6.2.2.

解 选取

$$Y = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \\ \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

和以前的 $z = (-1, -1, 1)^T$, 从 Y 的特殊选取, 则有 $A^T Y = I$. 假设问题中 $x = (0, 0, 0)^T$, 则

$$c = A^T x - b = -b, \quad g = d + Gx = d = (-8, -3, -3)^T.$$

简单的计算可得

$$P_y = (3, 0)^T, \quad P_z = 0.$$

从而

$$p = x^* - x = YP_y + zP_z = (2, -1, 1)^T.$$

由乘子 λ^* 可得

$$(A^T Y)^T \lambda^* = Y^T (g + Gp),$$

故 $\lambda^* = (3, -2)^T$, $x^* = (2, -1, 1)^T$.

§ 6.3 凸二次规划的有效集方法

本节讨论解不等式约束二次规划的有效集方法.

问题(6.1.1)——(6.1.3)的 Lagrange 函数为

$$L(x, \lambda) = \frac{1}{2} x^T Gx + x^T d - \sum_{i \in E} \lambda_i (a_i^T x - b). \quad (6.3.1)$$

另外, 定义在最优点 x^* 的有效约束集 $A(x^*)$ 为

$$A(x^*) = \{i \in I \setminus E \mid a_i^T x^* = b\}. \quad (6.3.2)$$

通过简单的计算可得一阶最优性必要条件为

$$Gx^* + d - \sum_{i \in A(x^*)} \lambda_i^* a_i = 0, \quad (6.3.3)$$

$$a_i^T x^* = b, \quad \forall i \in A(x^*), \quad (6.3.4)$$

$$a_i^T x^* < b, \quad \forall i \in I \setminus A(x^*), \quad (6.3.5)$$

$$\lambda_i^* = 0, \quad \forall i \in I \setminus A(x^*). \quad (6.3.6)$$

当(6.1.1)中矩阵 G 是正半定时, 因为由(6.1.2)——(6.1.3)所定义的约束可行域是凸集, 那么凸规划的任何局部最优解就是整体最优解.

如果 $A(x^*)$ 已知, 我们可以通过求解等式约束的二次规划问题

$$\min \quad q(x) = \frac{1}{2} x^T G x + x^T d,$$

$$\text{s.t.} \quad a_i^T x = b, \quad i \in (x^*).$$

求出原问题(6.1.1)–(6.1.3)的 KKT 点.有效约束集方法有三种形式,它们是原始有效集方法、对偶有效集方法和原始-对偶有效集方法.我们只限于讨论原始有效集方法.

原始有效集方法通常由计算初始可行点 x_0 开始,并且要求所有的迭代点列是可行的.通过求解由(6.1.2)–(6.1.3)中作为等式约束的二次规划子问题,获得下一次迭代步.如果等式约束二次规划子问题的解是原问题的可行点,则判别相应的 Lagrange 乘子 λ^* 是否满足

$$\lambda_i^* = 0, \quad i \in A(x^*),$$

其中 λ^* 满足

$$Gx^* + d = \sum_{i \in A(x^*)} \lambda_i^* a_i. \quad (6.3.7)$$

如果满足,则停止计算;否则可去掉一个约束重新求解约束问题.如果等式约束二次规划子问题的解不是原问题的可行点,则需要增加约束然后重新求解得到的等式约束二次规划子问题.设第 k 次迭代点 x_k 处的工作集为 W_k ,它由所有式约束和有效不等式约束组成.特别地,工作集 W_k 中所有约束的梯度 a_i 是线无关的.

给定一个迭代点 x_k 和工作集 W_k ,首先检验 x_k 是否是二次函数 $q(\cdot)$ 在由工作集所定义的子问题中的最优解.如果不是,则由等式约束二次规划子问题计步 p ,其中所有等式约束相应于工作集 W_k .定义

$$p = x - x_k, \quad g_k = Gx_k + d.$$

于是,

$$\begin{aligned} q(x) &= q(x_k + p) \\ &= \frac{1}{2} (x_k + p)^T G (x_k + p) + d^T (x_k + p) \\ &= \frac{1}{2} p^T G p + g_k^T p + c, \end{aligned}$$

这里 $c = \frac{1}{2} x_k^T G x_k + d^T x_k$ 是常数项.因为常数项不影响最优解,丢掉常数项 c ,可得二次规划子问题在第 k 次迭代中的表示为

$$\min \quad \frac{1}{2} p^T G p + g_k^T p, \quad (6.3.8)$$

$$\text{s.t.} \quad a_i^T p = 0, \quad i \in W_k, \quad (6.3.9)$$

当沿着方向 p_k 运动时,项 $a_i^T x$ 并不改变,因为 $a_i^T (x_k + p_k) = a_i^T x_k = b$.设 p_k 是

子问题(6.3.8)–(6.3.9)的解, $j_k(j_k)$ 是相应的 Lagrange 乘子. 如果 $p_k = 0$, 则知 x_k 是问题

$$\begin{aligned} \min \quad & \frac{1}{2} x_k^T G x_k + d^T x_k, \\ \text{s.t.} \quad & a_i^T x = b_i, \quad i \in k \end{aligned}$$

的 KKT 点.

如果 $p_k \neq 0$, 并且如果 $x_k + p_k$ 沿着所有的约束是可行的, 那么置 $x_{k+1} = x_k + p_k$. 否则, 置

$$x_{k+1} = x_k + \alpha_k p_k, \quad (6.3.10)$$

这里选择步长参数 $\alpha_k \in [0, 1]$ 为使 (6.3.10) 满足所有约束的最大值, 使得 $a_i^T (x_k + \alpha_k p_k) = b_i$. 如果 $\alpha_k < 1$, 则存在一个新的指标 j 使得 a_j 成为可行的, 于是将其加入约束. 这样, 我们取

$$\alpha_k = \min_{i \in k} \left\{ 1, \min_{i \in k} \frac{b_i - a_i^T x_k}{a_i^T p_k} \mid a_i^T p_k < 0 \right\}. \quad (6.3.11)$$

另外, 如果 $j_k \neq 0$ 对所有 $j \in k$ 成立, 则知 x_k 也是原问题 (6.1.1)–(6.1.3) 的 KKT 点. 否则, 选取最负的乘子 $j_k(j_k)$, 去掉第 j_k 个约束, 使目标函数值下降, 这里,

$$j_k = \min_{j \in k} \{ j_k \mid j_k < 0, j \in k \}.$$

然后, 新的工作集为 $W_{k+1} = k \setminus \{j_k\}$, 并重新求解子问题 (6.3.8)–(6.3.9).

下面, 我们给出有效集方法的主要步骤.

算法 6.3.1

步 0. 计算可行的初始点 x_0 . 置 x_0 处的有效约束集 W_0 ; $k = 0$.

步 1. 解子问题 (6.3.8)–(6.3.9) 得 p_k .

步 2. 若 $p_k = 0$, 则计算相应的满足 (6.3.7) 的 Lagrange 乘子 $\hat{\lambda}_i$, 置 $\hat{W} = k$; 若 $\hat{\lambda}_i = 0, \forall i \in k$, 则停止, 即得解 $x^* = x_k$; 否则, 置 $j = \arg \min_{j \in k} \hat{\lambda}_j$; $x_{k+1} = x_k$; $k+1 = k \setminus \{j\}$;

步 3. 若 $p_k \neq 0$, 计算 (6.3.11) 式得 α_k ; 置 $x_{k+1} = x_k + \alpha_k p_k$; 若存在回代约束 $a_j, j \notin k$, 加上其回代约束至 W_k 得 W_{k+1} ; 转步 4. 否则, 若不存在回代约束, 即 $k = 1$, 则置 $W_{k+1} = W_k$.

步 4. $k = k+1$, 转步 1.

若干方法可用于求初始可行点 x_0 , 例如线性规划阶段方法和“大 M ”方法.

例 6.3.2 应用算法 6.3.1 计算二维问题

(6.3.8)–(6.3.9)再次得 $p^{(2)} = 0$. 再从(6.3.12)得 Lagrange 乘子 $\hat{\lambda}_5 = -5$, 所以从工作集中去掉约束 5, 得 $W_3 =$.

第 3 次迭代是解无约束问题, 得 $p^{(3)} = (0, 2.5)^T$. 由公式(6.3.11)得到 $s_3 = 0.6$, 由此新的迭代点 $x^{(4)} = (1, 1.5)^T$. 只有一个回代约束, 即约束 1, 所以得 $W_4 = \{1\}$. 而当 $k=4$ 时, 子问题(6.3.8)–(6.3.9)的解为 $p^{(4)} = (0.4, 0.2)^T$, 并产生的步长为 1. 在此步仍然没有回代约束, 所以新的工作集没有变化, $s_5 = 1$. 新迭代点 $x^{(5)} = (1.4, 1.7)^T$.

最后, $k=5$ 时, 解子问题(6.3.8)–(6.3.9)得解 $p^{(5)} = 0$. 公式(6.3.11)产生 Lagrange 乘子 $\hat{\lambda}_1 = 1.25$, 所以找到了解 $x^* = (1.4, 1.7)^T$, 迭代终止.

小 结

本章研究二次规划的理论及算法. 第一节介绍了一些实际应用背景. 第二节讨论了解二次规划的直接消去法和广义消去法. 第三节讨论解不等式约束的有效集方法. 这些方法对解二次规划来说都是最基本的. 有关解二次规划的其他方法, 如内点法, 梯度投影法等, 可参阅有关参考文献.



1. 解下面二次规划问题并且作图解释几何意义.

$$\begin{aligned} \max \quad & f(x) = 2x_1 + 3x_2 + 4x_1^2 + 2x_1x_2 + x_2^2, \\ \text{s.t.} \quad & x_1 - x_2 \leq 0, \\ & x_1 + x_2 \leq 4, \\ & x_1 \leq 3. \end{aligned}$$

2. 从一点 x_0 的超平面 $\{x \mid Ax = b\}$ 寻找最短距离能构成二次规划形式, 这里 A 是行满秩矩阵

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}(x - x_0)^T(x - x_0), \\ \text{s.t.} \quad & Ax = b. \end{aligned}$$

证明最优乘子为

$$\lambda^* = (AA^T)^{-1}(b - Ax_0),$$

最优解为

$$x^* = x_0 - A^T(AA^T)^{-1}(b - Ax_0),$$

并且证明当 A 是行向量时, 从 x_0 至 $Ax = b$ 的可行解集的最短距离为

$$\frac{|b - Ax_0|}{|A|}.$$

3. 写出(6.2.1)–(6.2.2)的一阶必要性条件,并试证明之.
4. 写出下面既有等式约束又有不等式约束的二次凸规划的 KKT 条件

$$\min \quad q(x) = \frac{1}{2} x^T G x + x^T d,$$

$$\text{s.t.} \quad A^T x = b, \quad \bar{A}^T x = \bar{b},$$

这里 G 是对称正半定矩阵,使用这些条件导出类似的原始-对偶步.

5. 考虑二次规划

$$\max \quad q(x) = 6x_1 + 4x_2 - 13 - x_1^2 - x_2^2$$

$$\text{s.t.} \quad x_1 + x_2 = 3$$

$$x_1, x_2 \geq 0.$$

使用图解法求解,再使用有效约束集方法求解.

第7章

约束最优化的理论与方法

§ 7.1 约束最优化问题与最优性条件

本章中我们讨论解约束最优化问题的理论与方法,为清晰与方便起见,现将约束最优化问题表示如下:

$$\min f(x), \quad (7.1.1)$$

$$\text{s.t. } g_i(x) = 0, i = 1, 2, \dots, m, \quad (7.1.2)$$

$$g_i(x) \leq 0, i = m+1, m+2, \dots, m, \quad (7.1.3)$$

其中 $x \in \mathbf{R}^n$, $m \leq m$. 记可行域 D 为

$$D = \{x | g_i(x) = 0, i = 1, \dots, m; g_i(x) \leq 0, i = m+1, \dots, m\}. \quad (7.1.4)$$

记指标集 $I = \{1, 2, \dots, m, m+1, \dots, m\}$. 由可行域的定义可知,求解约束优化问题(7.1.1)——(7.1.3)就是在可行域 D 上寻求一点 x 使得目标函数 $f(x)$ 达到最小. 现在我们给出约束最优化问题解的精确定义.

定义 7.1.1 设 $x^* \in D$, 如果

$$f(x^*) \leq f(x), \quad \forall x \in D, \quad (7.1.5)$$

成立,则称 x^* 是问题(7.1.1)——(7.1.3)的全局极小点;进一步,如果

$$f(x^*) < f(x), \quad \forall x \in D, x \neq x^*, \quad (7.1.6)$$

则称 x^* 是严格全局极小点.

定义 7.1.2 设 $x^* \in D$, 如果对于某一常数 $\delta > 0$ 有

$$f(x^*) \leq f(x), \quad \forall x \in N(x^*, \delta) \cap D \quad (7.1.7)$$

成立,则称 x^* 是问题(7.1.1)——(7.1.3)的局部极小点,其中 $N(x^*, \delta)$ 是以 x^* 为中心以 δ 为半径的邻域(或为广义球)

$$N(x^*, \delta) = \{x | \|x - x^*\|_2 \leq \delta\}. \quad (7.1.8)$$

进一步,如果

$$f(x^*) < f(x), \quad \forall x \in N(x^*, \delta) \cap D, x \neq x^* \quad (7.1.9)$$

成立,则称 x^* 是严格局部极小点.

全局极小点通常也可以称为总(整)体极小点,显然一个全局(严格)极小点

也是一个局部(严格)极小点.假定 x^* 是问题(7.1.1)–(7.1.3)的一个局部极小点,如果存在 $i \in \{m+1, \dots, m\}$ 使得

$$c_i(x^*) > 0, \quad (7.1.10)$$

成立,则我们可将第 i 个约束条件去掉,且 x^* 仍然是去掉第 i 个约束条件所得到的问题的局部极小点.由于这一性质,我们称第 i 个约束在 x^* 处是无效约束.下面的定义将有效与非有效的概念推广到任何点处.记

$$E = \{1, 2, \dots, m\}, \quad (7.1.11)$$

$$I = \{m+1, \dots, m\}, \quad (7.1.12)$$

$$I(x) = \{i \mid c_i(x) = 0, i \in I\}. \quad (7.1.13)$$

定义 7.1.3 对任何 $x \in \mathbf{R}^n$, 我们称集合

$$A(x) = E \setminus I(x) \quad (7.1.14)$$

是在 x 点处的有效约束集合(或积极约束集合), $c_i(x) (i \in A(x))$ 是在 x 点处的有效约束(或积极约束), $c_i(x) (i \notin A(x))$ 是在 x 点处的非有效约束(或非积极约束).

假定我们已知问题(7.1.1)–(7.1.3)在解点处的有效约束 $A(x^*)$, 那么, 我们只需求解如下的等式约束优化问题

$$\begin{aligned} \min_{x \in \mathbf{R}^n} \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i \in A(x^*). \end{aligned} \quad (7.1.15)$$

因此,如何有效地求解上述等式约束问题将是非常重要的,本章将给出有效的算法并进行收敛性分析.

应该指出,实际上可行的只是求得最优化问题的一个局部(或严格局部)极小点,而非总体极小点.尽管我们可以考虑求总体极小点的可能性,但一般来说这是一个相当困难的任务.同时,在很多实际应用中,求局部极小点已能满足问题的要求.因此,如不作特别说明,求解一个非线性规划问题的涵义是求一个局部最优解向量 x^* , 而不是求可能存在的所有最优解.关于总体最优化的课题请参阅其他文献,如 Floudas 和 Pardalos(1992).一般情况下,直接利用定义 7.1.2 去判别某一给定点 x^* 是否为局部最优解是办不到的,一个显然的困难是 $N(x^*, D)$ 通常有无穷多个点,从而直接验证条件(7.1.7)是不可能的.因此,有必要给出只依赖于 x^* 处目标函数和约束函数的信息且与(7.1.7)等价的条通常,我们将这些条件称为最优性条件.

设 $x \in D \subset \mathbf{R}^n$ 并且实值函数 f 在这一点可微.我们知道如果一个实值函数 f 在内点 $x \in D$ 可微,则在 x 处存在一阶偏导数;此外,如果偏导数在 x 处还是连续,则称 f 在 x 处连续可微.同样地,如果 f 在 $x \in D$ 二次可微,则 f 在 x 处存在二阶偏导数;如果它们在 x 处还是连续的,则称 f 在 x 处是二次连续可微.由

于可行域上的点是否为局部极小点不仅取决于目标函数在该点的值,还与该点附近其他可行点上的值有关,因此可行方向在推导最优性条件时将起着十分重要的作用.下面我们就给出各种可行方向的定义.

定义 7.1.4 设 $x^* \in D, 0 \neq d \in \mathbf{R}^n$, 如果存在 $\delta > 0$ 使得

$$x^* + td \in D, \quad \forall t \in [0, \delta], \quad (7.1.16)$$

则称 d 是 D 在 x^* 处的可行方向. D 在 x^* 处的所有可行方向组成的集合记为 $FD(x^*, D)$.

定义 7.1.5 设 $x^* \in D, 0 \neq d \in \mathbf{R}^n$, 如果

$$\begin{aligned} d^T \nabla_i g(x^*) &= 0, \quad i \in E; \\ d^T \nabla_i g(x^*) &\leq 0, \quad i \in I(x^*); \end{aligned} \quad (7.1.17)$$

则称 d 是 D 在 x^* 处的线性化可行方向. D 在 x^* 处的所有线性化可行方向的集合记为 $LFD(x^*, D)$.

定义 7.1.6 设 $x^* \in D, 0 \neq d \in \mathbf{R}^n$, 如果存在序列 $\{z_k\}_{k=0}^\infty \subset \mathbf{R}^n$ 使得满足

- (i) $z_k \in D, \quad \forall k$;
- (ii) $\lim_{k \rightarrow +\infty} z_k = x^*$;
- (iii) $z_k - x^* \neq 0$, 对所有充分大的 k ,

则 z_k 称为 D 在 x^* 处的可行序列.记 $T(x^*)$ 是所有 $\lim_{k \rightarrow +\infty} z_k = x^*$ 的可行序列 $\{z_k\}$ 的集合.记 S_k 是序列 $\{z_k\}$ 的某一子序列, 因为可以证明存在 d 使得

$$\lim_{\substack{z_k \in S_k \\ k \rightarrow +\infty}} \frac{z_k - x^*}{\|z_k - x^*\|} = d, \quad (7.1.18)$$

由此,若序列 $\{z_k\}$ 有以(7.1.18)所定义的收敛极限方向 d , 则称 d 是 D 在 x^* 处的序列可行方向. D 在 x^* 处的所有序列可行方向的集合记为 $SFD(x^*, D)$.

根据定义,可以证明如果所有的约束函数都在 $x^* \in D$ 处可微, 则有

$$FD(x^*, D) \subset SFD(x^*, D) \subset LFD(x^*, D). \quad (7.1.19)$$

引理 7.1.7 设 $x^* \in D$ 是问题(7.1.1)–(7.1.3)的局部极小点, 如果 $f(x)$ 和 $g_i(x) (i=1, \dots, m)$ 都在 x^* 处可微, 则所有可行序列 $\{z_k\}$ 的序列可行方向 d 满足

$$d^T \nabla f(x^*) \leq 0, \quad \forall d \in SFD(x^*, D). \quad (7.1.20)$$

证明 假定存在可行序列 $\{z_k\}$ 的序列可行方向 d , 使得 $d^T \nabla f(x^*) < 0$, $\forall d \in SFD(x^*, D)$, 并且序列 $\lim_{k \rightarrow +\infty} z_k = x^*$. 由 Taylor 展开式, 并利用(7.1.18), 有

$$\begin{aligned} f(z_k) &= f(x^*) + (z_k - x^*)^T \nabla f(x^*) + o(\|z_k - x^*\|) \\ &= f(x^*) + \|z_k - x^*\| d^T \nabla f(x^*) + o(\|z_k - x^*\|). \end{aligned}$$

因为 $d^T \nabla f(x^*) < 0$, 所以上式的主部是第二项, 则对充分大的 k , 有

$$f(z^k) < f(x^*) + \frac{1}{2} \|z^k - x^*\|^T d^T f(x^*).$$

由此,给定任意以 x^* 为中心以 δ 为半径的邻域,都可选取充分大的 k 使得序列 z^k 在此邻域内,并且有比 $f(x^*)$ 更小的值.这与 $x^* \in D$ 是问题 (7.1.1) — (7.1.3) 的局部极小点相矛盾.

为获得约束优化问题的最优性条件,引入约束品性对约束优化问题是非常重要的.常见的容易验证的约束品性条件如下:

定义 7.1.8 给定可行点 x^* 和在 x^* 处的有效约束集合 $A(x^*) = I(x^*)$, 若 $c_i(x^*) (i \in A(x^*))$ 线性无关,则称在 x^* 点处满足线性独立约束品性条件.

下面我们讨论约束最优化问题的一阶最优性条件.

定理 7.1.9 设 x^* 是问题 (7.1.1) — (7.1.3) 的可行点,如果不存在 $d \in LFD(x^*, D)$ 满足 $d^T f(x^*) < 0$ 当且仅当存在 $i^* (i^* \in A(x^*))$ 使得

$$f(x^*) = \sum_{i^* \in A(x^*)} \lambda_{i^*}^* c_{i^*}(x^*), \quad \lambda_{i^*}^* = 0, \quad i^* \in I(x^*). \quad (7.1.21)$$

证明 由线性化可行方向的集合 $LFD(x^*, D)$ 的定义可知下面线性系统

$$\begin{aligned} d^T c_i(x^*) &= 0, \quad i \in A(x^*); \\ d^T c_i(x^*) &\leq 0, \quad i \in I(x^*); \\ d^T f(x^*) &< 0 \end{aligned} \quad (7.1.22)$$

无解,利用 Farkas 引理可得当且仅当存在 $i^*, \lambda_{i^*}^*$ 和 $\lambda_{i^*}^* = 0, (i^* \in I(x^*))$ 使得

$$f(x^*) = \sum_{i^* \in E} \lambda_{i^*}^* c_{i^*}(x^*) + \sum_{i^* \in I(x^*)} \lambda_{i^*}^* c_{i^*}(x^*). \quad (7.1.23)$$

利用以上定理和引理 7.1.7, 我们得到

定理 7.1.10 设 x^* 是问题 (7.1.1) — (7.1.3) 的局部极小点, 如果

$$SFD(x^*, D) = LFD(x^*, D), \quad (7.1.24)$$

则存在 $i^* (i = 1, 2, \dots, m)$ 使得

$$f(x^*) = \sum_{i=1}^m \lambda_{i^*}^* c_{i^*}(x^*), \quad \lambda_{i^*}^* = 0, \quad \lambda_{i^*}^* c_{i^*}(x^*) = 0, \quad i \in I(x^*). \quad (7.1.25)$$

证明 利用引理 7.1.7, 当 x^* 是问题 (7.1.1) — (7.1.3) 的局部极小点时, 可得

$$d^T f(x^*) = 0, \quad \forall d \in SFD(x^*, D).$$

从而利用定理 7.1.9, 并置 $\lambda_{i^*}^* = 0, i^* \in I \setminus I(x^*)$, (7.1.23) 式意味着 (7.1.25) 成立.

定理 7.1.10 是由 Kuhn 和 Tucker (1952) 给出的, 由于 Karush (1939) 也类似地考虑了约束最优化的最优性条件, 故它常称为 Karush-Kuhn-Tucker 定理,

简称 KKT 定理. 与 (7.1.25) 有着密切联系的一个函数是

$$L(x, \lambda) = f(x) - \sum_{i=1}^m \lambda_i c_i(x), \quad (7.1.26)$$

其中 $\lambda = (\lambda_1, \dots, \lambda_m)^T \in \mathbf{R}^m$, $c(x) = (c_1(x), \dots, c_m(x))^T$. 由于这一函数的思想可以追溯到 Lagrange, 故称之为 Lagrange 函数, 并称 λ_i ($i = 1, 2, \dots, m$) 为 Lagrange 乘子.

定义 7.1.11 如果 $x^* \in D$ 且存在向量 $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T \in \mathbf{R}^m$ 使得 (7.1.25) 成立, 则称 x^* 是问题 (7.1.1) — (7.1.3) 的 Karush Kuhn Tucker 点 (简称 KKT 点), 称 λ^* 是在 x^* 处的 Lagrange 乘子. 同时, 定理中使用的条件 (7.1.24) 通常称为约束品性.

由于上述条件不容易验证, 人们使用更强的, 容易验证的线性独立约束品性条件. 综合上面的定理可以得到如下定理.

定理 7.1.12 设 x^* 是问题 (7.1.1) — (7.1.3) 的一个局部极小点, 如果线性独立约束品性成立, 则必存在 $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T \in \mathbf{R}^m$ 使得 (λ^*, x^*) 满足

$$L(x^*, \lambda^*) = 0; \quad (7.1.27)$$

$$c_i(x^*) = 0, \quad \forall i \in I; \quad (7.1.28)$$

$$c_i(x^*) \leq 0, \quad \forall i \in I; \quad (7.1.29)$$

$$\lambda_i^* \geq 0, \quad \forall i \in I(x^*); \quad (7.1.30)$$

$$\lambda_i^* c_i(x^*) = 0, \quad \forall i \in A(x^*). \quad (7.1.31)$$

证明 由线性独立约束品性成立, 可得约束品性 (7.1.24) 成立. 由定理 7.1.10, 立得定理的结论.

因为线性独立约束品性易于检验, 所以定理 7.1.12 是一个最常见的并且也是非常有用的一阶最优性必要条件. 下面我们给出一阶最优性充分条件 (证明省略).

定理 7.1.13 设 $x^* \in D$, 如果 $f(x)$ 和 $c_i(x)$ ($i = 1, \dots, m$) 都在 x^* 处可微, 并且有

$$d^T \nabla f(x^*) > 0, \quad \forall 0 \neq d \in SFD(x^*, D) \quad (7.1.32)$$

成立, 则 x^* 是问题 (7.1.1) — (7.1.3) 的一个局部严格极小点.

有时利用一阶最优性条件来进行判断难以确定, 这样需要用二阶的条件来进行讨论.

定义 7.1.14 给定问题 (7.1.1) — (7.1.3) 的局部极小点. 存在 $\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)^T \in \mathbf{R}^m$ 满足 (7.1.31), 即 $\lambda_i^* c_i(x^*) = 0, i \in A(x^*)$ 成立, 这称为互补件. 进一步, 如果 λ_i^* 和 $c_i(x^*)$, $i \in A(x^*)$, 只有一个为零, 称为严格互补件.

因为现在讨论中要用到二阶导数矩阵 (Hesse 矩阵), 所以约定约束最优化问题中出现的函数 $f(x)$, $c_i(x)$ ($i = 1, 2, \dots, m$) 都是二次连续可微函数. 下面我们不加证明给出二阶最优性条件.

定义线性化可行方向集合 LFD 的子集 $E(x^*)$ 为

$$E(x^*) = \{d \in LFD \mid c_i(x^*)^T d = 0 \quad \forall i \in A(x^*) \text{ 且 } d^T \nabla c_i(x^*) > 0\}.$$

定理 7.1.15 (二阶必要性条件)

设 x^* 是问题 (7.1.1) — (7.1.3) 的局部极小点, λ^* 是相应的 Lagrange 乘子. 如果线性独立约束品性条件成立, 则

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d \geq 0, \quad \forall d \in E(x^*). \quad (7.1.33)$$

定理 7.1.16 (二阶充分性条件)

设 x^* 是一个 KKT 点, λ^* 是相应的 Lagrange 乘子. 如果

$$d^T \nabla_{xx}^2 L(x^*, \lambda^*) d > 0, \quad \forall d \in E(x^*), \quad d \neq 0, \quad (7.1.34)$$

则 x^* 是问题 (7.1.1) — (7.1.3) 的严格局部极小点.

在上面两个定理中,

$$\nabla_{xx}^2 L(x^*, \lambda^*) = \nabla_{xx}^2 f(x^*) - \sum_{i=1}^m \lambda_i^* \nabla_{xx}^2 c_i(x^*). \quad (7.1.35)$$

§ 7.2 二次罚函数方法

对于约束优化问题一类重要的求解方法就是通过解一系列无约束优化问题以获取原非线性约束问题解的罚函数方法. 用无约束最优化方法求解约束最优化问题时, 在考虑使目标函数值下降的同时还要求迭代点的可行性. 罚函数法在用无约束最优化方法时通过对不可行的迭代点施加惩罚, 并随着迭代的进展, 增大惩罚量, 迫使迭代点逐步向可行域靠近, 一旦迭代点成为一个可行点, 即为所求的原问题的最优解. 因此, 罚函数方法是一类不可行点的方法, 对初始点的选取没有可行性的限制. 对不可行约束所采用惩罚函数的不同, 形成许多不同的罚函数方法, 其中最简单, 也为大家乐于采用的是二次罚函数方法.

先考虑等式约束的最优化问题

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i = 1, \dots, m, \end{aligned} \quad (7.2.1)$$

二次罚函数 $Q(x; \mu)$ 定义为

$$Q(x; \mu) \stackrel{\text{def}}{=} f(x) + \frac{1}{2\mu} \sum_{i=1}^m c_i^2(x), \quad (7.2.2)$$

这里 $\mu > 0$ 是罚参数, 当 μ 趋于零时, 违反约束的罚项剧烈地增大. 直观地认为使

用罚序列 $\{\mu_k\}$,使得当 $k \rightarrow +\infty$ 时 $\mu_k \rightarrow 0$ 以获得 $Q(x; \mu_k)$ 的近似极小点 x_k .因为罚项是二次的,所以光滑可微,这样可以使用无约束优化技术来求解 x_k .近似点 x_k, x_{k-1} 等可作为第 $k+1$ 次迭代的好的初始点.

例 7.2.1 考虑问题

$$\begin{aligned} \min \quad & x_1 + x_2, \\ \text{s.t.} \quad & x_1^2 + x_2^2 - 2 = 0, \end{aligned}$$

二次罚函数为

$$Q(x; \mu) = x_1 + x_2 + \frac{1}{2\mu}(x_1^2 + x_2^2 - 2)^2,$$

当 $\mu=1$,则 $Q(x; \mu_k)$ 的极小点在 $(-1, -1)^T$.当 $\mu=0.1$, $Q(x; \mu_k)$ 的极小点非常接近于解 $(-1, -1)^T$.

对于通常的约束最优化问题(7.1.1)–(7.1.3),既有不等式约束又有等式约束,定义 $Q(x; \mu)$ 为

$$Q(x; \mu) \stackrel{\text{def}}{=} f(x) + \frac{1}{2\mu_{i \in E}} c_i^2(x) + \frac{1}{2\mu_{i \in I}} [c_i(x)]_+^2, \quad (7.2.3)$$

这里

$$[c_i(x)]_+ \stackrel{\text{def}}{=} \begin{cases} c_i(x), & i \in I, \\ \min(c_i(x), 0), & i \in E. \end{cases}$$

算法 7.2.2 (二次罚函数方法)

步 1. 给定 $\mu_0 > 0$,允许参数值 $\mu_0 > 0$,初始点 x_0 ; $k=0$.

步 2. 从 x_k 开始,求解 $Q(\cdot; \mu_k)$ 的近似极小点.

步 3. 当 $|c(x_k)| \leq \epsilon_k$ 时,终止,得近似解 x_k ;否则,选择新的罚参数 $\mu_{k+1} \in (0, \mu_k)$, $k = k+1$,转步 2.

注意,罚参数序列 $\{\mu_k\}$ 要合适地选择.当极小化 $Q(x; \mu_k)$ 的计算量很大时,可以选择适当缩小 μ_k ,例如 $\mu_{k+1} = 0.7\mu_k$.如果极小化 $Q(x; \mu_k)$ 计算量不大,可大大地缩小 μ_k ,例如 $\mu_{k+1} = 0.1\mu_k$.

现在讨论二次罚函数的收敛性.

定理 7.2.3 假设每个 x_k 都是罚函数算法 7.2.2 得到的 $Q(x; \mu_k)$ 的精确整体极小点,并且 $\mu_k \rightarrow 0$,则序列 $\{x_k\}$ 的每个极限点都是等式约束问题(7.2.1)的整体最优解.

证明 设 \bar{x} 是问题(7.2.1)的整体极小点,即

$$f(\bar{x}) \leq f(x), \quad \forall x \in \{x \mid c_i(x) = 0, i \in I\}.$$

因为 x_k 是第 k 次迭代 $Q(\cdot; \mu_k)$ 的极小点,则有 $Q(x_k; \mu_k) \leq Q(\bar{x}; \mu_k)$,从而有

$$f(x_k) + \frac{1}{2\mu_{k \in E}} c_i^2(x_k) \leq f(\bar{x}) + \frac{1}{2\mu_{k \in E}} c_i^2(\bar{x}) = f(\bar{x}). \quad (7.2.4)$$

整理得

$$c_i^2(x_k) = 2\mu_k [f(\bar{x}) - f(x_k)]. \quad (7.2.5)$$

假设 x^* 是 $\{x_k\}$ 的极限点, 使得存在一个无限子序列集 K 使

$$\lim_{k \in K} x_k = x^*.$$

对 (7.2.5) 两边取极限, $k \rightarrow \infty, k \in K$ 可得

$$c_i^2(x^*) = \lim_{k \in K} c_i^2(x_k) = \lim_{k \in K} 2\mu_k [f(\bar{x}) - f(x_k)] = 0,$$

这里最后的等式由 $\mu_k \rightarrow 0$ 得到. 因此有 $c_i(x^*) = 0, \forall i \in E$. 所以 x^* 是可行点. 进一步, 在 (7.2.4) 式中取极限 $k \rightarrow \infty, k \in K$ 由 μ_k 的非负性以及 $c_i(x)^2$ 的非负性, 有

$$f(x^*) = f(x^*) + \lim_{k \in K} \frac{1}{2\mu_k} c_i^2(x_k) = f(\bar{x}).$$

上式表明 x^* 也是整体极小点.

上述结果要求我们寻找每个子问题的精确整体最小解, 这是非常困难的. 事实上, 如果我们每一步都要求 $Q(\cdot; \mu_k)$ 的局部近似极小点, 上述定理也成立.

§ 7.3 内点障碍罚函数法

在本节我们考虑另一类称为内点障碍罚函数的罚函数方法, 同上一节的罚函数方法不同, 内点障碍罚函数法是一类保持严格可行性的方法, 即迭代点序列是严格可行内点, 因此这类方法只适用于只有不等式约束而无等式约束的非线性最优化问题. 内点障碍罚函数法通过在目标函数上引入一个关于约束的障碍项, 当迭代点由可行域的内部接近可行域的边界时, 障碍项将趋于无穷大来迫使迭代点返回可行域的内部, 从而保持迭代点的严格可行性.

考虑不等式约束最优化问题

$$\begin{aligned} \min_x \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i \in I. \end{aligned} \quad (7.3.1)$$

定义严格可行内点区域为

$$F^0 \stackrel{\text{def}}{=} \{x \in \mathbf{R}^n \mid c_i(x) > 0, \quad \forall i \in I\}, \quad (7.3.2)$$

并假设 F^0 是非空的. 对问题 (7.3.1), 障碍罚函数具有如下性质:

1. 在 F^0 外部的值或者无定义或者是无穷的;
2. 在 F^0 内部都是连续可微的;
3. 当 x 趋于 F^0 的边界, 其值趋于 ∞ .

常用的障碍函数是对数障碍函数

$$P(x; \mu) = f(x) - \mu \sum_{i=1}^l \log c_i(x) \quad (7.3.3)$$

和分数障碍罚函数

$$P(x, \mu) = f(x) + \mu \sum_{i=1}^l \frac{1}{c_i(x)}, \quad (7.3.4)$$

其中 $\log(\cdot)$ 为自然对数, $\mu > 0$ 为障碍参数. 下面的讨论限于对数障碍罚函数, 有关的结果和分析同样适用于分数障碍罚函数.

例 7.3.1 考虑非线性约束最优化问题

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

问题的最优解为 $x^* = (0, 1)^T$, 对数障碍罚函数为

$$P(x, \mu) = x_1 - 2x_2 - \mu[\log(1 + x_1 - x_2^2) + \log(x_2)].$$

对给定的 $\mu > 0$, 为确定 $P(x, \mu)$ 的无约束最优解, 考察一阶最优性条件

$$\begin{aligned} 1 - \frac{\mu}{1 + x_1 - x_2^2} &= 0, \\ -2 + \frac{2\mu x_2}{1 + x_1 - x_2^2} - \frac{\mu}{x_2} &= 0. \end{aligned} \quad (7.3.5)$$

求解这个方程组得

$$x_1(\mu) = \frac{1 + 2\mu + 3\mu - 1}{2}, \quad x_2(\mu) = \frac{1 + 2\mu + 1}{2}.$$

关于障碍参数 $\mu \rightarrow 0$ 取极限得

$$\begin{aligned} \lim_{\mu \rightarrow 0} x_1(\mu) &= \lim_{\mu \rightarrow 0} \frac{1 + 2\mu + 3\mu - 1}{2} = 0, \\ \lim_{\mu \rightarrow 0} x_2(\mu) &= \lim_{\mu \rightarrow 0} \frac{1 + 2\mu + 1}{2} = 1. \end{aligned}$$

即有 $\lim_{\mu \rightarrow 0} x(\mu) = x^*$.

上述例子很好地说明了对数障碍罚函数的一些有用的特征. 首先, $P(x, \mu)$ 的无约束最优解 $x(\mu)$ 如果存在, 那么在 $\mu \rightarrow 0$ 时确实收敛于原不等式约束问题的最优解; 其次, $x(\mu)$ 关于 μ 连续可微, 因而 $x(\mu)$ 在问题可行域的内部定义了一条以 μ 为参数的光滑曲线, (文献中一般称为障碍轨迹 Barrier Trajectory), 这条曲线在 $\mu \rightarrow 0$ 时的极限点即为问题的最优解, 事实上, 对数障碍罚函数法就是沿着这条轨迹确定问题的最优解.

现在来计算对数障碍函数的梯度与 Hesse 矩阵.

$$\nabla_x P(x; \mu) = \nabla f(x) - \sum_{i=1}^l \frac{\mu}{c_i(x)} \nabla c_i(x), \quad (7.3.6)$$

$$\frac{1}{2} P(x; \mu) = \frac{1}{2} f(x) - \sum_{i=1}^l \frac{\mu}{c_i(x)} \frac{1}{2} c_i(x) + \sum_{i=1}^l \frac{\mu}{c_i^2(x)} c_i(x) c_i(x)^T. \quad (7.3.7)$$

在(7.3.6)中令

$$\lambda_i^* = \frac{\mu}{c_i(x^*)}, \quad i = 1, \dots, l$$

代入(7.3.7)式,有

$$\frac{1}{2} P(x; \mu) = \frac{1}{2} L(x, \lambda^*) + \sum_{i=1}^l \frac{1}{\mu} (\lambda_i^*) c_i(x) c_i(x)^T. \quad (7.3.8)$$

与二次罚函数情形类似,可以使用牛顿法求解,即

$$\frac{1}{2} P(x; \mu) p = - \frac{1}{2} P(x; \mu). \quad (7.3.9)$$

下面给出障碍罚函数的算法.

算法 7.3.2 (对数障碍函数)

- 步 1. 给定 $\mu > 0$, 允许值 $\epsilon > 0$, 初始点 x_0 , $k = 0$.
- 步 2. 从 x_k 出发, 求 $P(x; \mu_k)$ 的近似最优解 x_{k+1} .
- 步 3. 如果收敛准则满足, 停止迭代, 并取 x_{k+1} 为最优解的近似.
- 步 4. 选择新的障碍参数 $\mu_{k+1} = (0, \mu_k)$; $k = k+1$, 转步 2.

一个好的初始点 x_k^s , 可从以前的近似最优点 x_{k-1}, x_{k-2}, \dots 进行外推得到. 另一种方法是微分 $\frac{1}{2} P(x; \mu)$ 得

$$\frac{1}{2} P(x; \mu) \dot{x} + \frac{1}{\mu} \frac{1}{2} P(x; \mu) = 0. \quad (7.3.10)$$

由计算第二项, 可得

$$\frac{1}{2} P(x; \mu) \dot{x} - \sum_{i=1}^l \frac{1}{c_i(x)} c_i(x) = 0. \quad (7.3.11)$$

用 $x = x_{k-1}$ 和 $\mu = \mu_{k-1}$ 代入此表示式, 得近似切线 \dot{x} , 利用 \dot{x} 可产生第 k 次迭代初始点 x_k^s 如下:

$$x_k^s = x_{k-1} + (\mu_k - \mu_{k-1}) \dot{x}. \quad (7.3.12)$$

可按以下方法选择新的障碍参数 μ_{k+1} : 如果求 $P(x; \mu_k)$ 的极小点不是太困难, 以及初始点 x_k^s 比较可靠, 那么可以大胆减小 μ_{k+1} , 例如取 $\mu_{k+1} = 0.2 \mu_k$, 或者 $\mu_{k+1} = 0.1 \mu_k$.

定理 7.3.3 假设 F^0 非空, x^* 是问题 (7.3.1) 的局部极小点, 相应的 Lagrange 乘子 λ^* 满足 KKT 条件. 进一步假设线性独立约束品性, 严格互补性条件及二阶充分性最优条件在点 (x^*, λ^*) 成立, 那么下列结论成立:

(1) 对充分小的 μ , 如果 $x(\mu)$ 是 x^* 邻域内 $P(x, \mu)$ 的局部极小点, 则存在连续可微的向量函数 $x(\mu)$, 使得

$$\lim_{\mu \rightarrow 0} x(\mu) = x^*. \quad (7.3.13)$$

(2) 对(1)中定义的函数 $x(\mu)$, 当 $\mu \rightarrow 0$ 时, 相应的 Lagrange 乘子估计

$$\lambda_i(\mu) = \frac{\mu}{c_i(x(\mu))} \text{ 收敛到 } \lambda_i^*.$$

(3) 对所有充分小的 μ , Hesse 矩阵 $\frac{\partial^2}{\partial x^2} P(x; \mu)$ 是正定的.

§ 7.4 序列二次规划方法

本节介绍求解约束最优化问题最有效的方法之一: 序列二次规划方法, 它是求解无约束最优化问题的牛顿法和拟牛顿方法对约束最优化问题的推广. 我们知道无约束最优化的牛顿法是通过求最优性的一阶必要条件方程组的解推导得出的. 考虑求光滑函数 $f(x)$ 的无约束最优解 x^* , 根据一阶最优性条件, x^* 是方程组

$$g(x) = 0$$

的解, 为得到这样的解, 用牛顿法解这个方程组, 得迭代格式

$$G_k s_k = -g_k, \quad x_{k+1} = x_k + d_k,$$

其中 $g_k = g(x_k)$, $G_k = G(x_k)$. 在 G_k 正定, d_k 是下降方向时, 沿方向 d_k 进行线性搜索, 确定一个步长 α_k 使 $f(x_k + \alpha_k d_k) < f(x_k)$, 并取 $x_{k+1} = x_k + \alpha_k d_k$ 以增大牛顿法的收敛域, 由此形成线性搜索方法. 在 G_k 不定, d_k 不一定是下降方向时, 在 G_k 上增加一个矩阵 $\mu_k I$, 使 $G_k + \mu_k I$ 正定, 再取 $x_{k+1} = x_k + d_k$, 这里 d_k 是方程组

$$(G_k + \mu_k I) d = -g_k \quad (7.4.1)$$

的解. μ_k 的选取同样要求 $f(x_k + d_k) < f(x_k)$ 以确保算法的收敛性. 由于方程组(7.4.1)的求解等价于下述问题的求解

$$\begin{aligned} \min \quad & q(d) = f(x_k) + g_k^T d + \frac{1}{2} d^T G_k d, \\ \text{s.t.} \quad & \|d\| \leq \Delta_k, \end{aligned} \quad (7.4.2)$$

且后者更易于通过控制 Δ_k 的值来确定(7.4.1)中的 μ_k 的大小, 因此目前都是通过解问题(7.4.2)来确定 d_k , 由此形成的算法称为信赖域方法, Δ_k 称为信赖域半径, $d = x - x_k$. 而所谓的拟牛顿方法只是用拟牛顿修正公式产生矩阵 G_k 的近似 B_k 代替上述问题中的矩阵 G_k , 以减少计算矩阵 G_k 的工作量.

把上述思想推广到非线性约束最优化问题, 就可以得到目前广泛使用的序列二次规划方法. 为简单起见, 还是先从等式约束最优化问题开始. 考虑

$$\begin{aligned} \min_x \quad & f(x) \\ \text{s.t.} \quad & c_i(x) = 0, \quad i = \{1, 2, \dots, m\}, \end{aligned} \quad (7.4.3)$$

其 Lagrange 函数为

$$L(x, \lambda) = f(x) - \lambda^T c(x),$$

一阶最优性必要条件为

$$L(x, \lambda) = 0, \quad (7.4.4)$$

或者为

$$\frac{\partial}{\partial x} L(x, \lambda) = g(x) - A(x)^T \lambda = 0,$$

$$L(x, \lambda) = f(x) - c(x) = 0,$$

其中 $A(x)$ 表示约束的 Jacobi 矩阵,

$$A(x)^T = [\alpha(x), \beta(x), \dots, \gamma_m(x)].$$

上述方程组是含有 $n+m$ 个方程和 $n+m$ 个未知量的非线性方程组, 用牛顿法解这个方程组可以得到迭代格式

$$x_{k+1} = x_k + d_k,$$

其中 d_k 和 d_k 是下述牛顿方程组的解

$$\frac{\partial^2}{\partial x^2} L(x_k, \lambda_k) \frac{dx}{d\lambda} = - \frac{\partial}{\partial \lambda} L(x_k, \lambda_k).$$

把这个方程组展开得

$$\begin{aligned} \frac{\partial^2}{\partial x^2} L(x_k, \lambda_k) \frac{dx}{d\lambda} &= - \frac{\partial}{\partial \lambda} L(x_k, \lambda_k) \\ &= - \frac{g(x_k) - A(x_k)^T \lambda_k}{0} = - \frac{g(x_k) - A(x_k)^T \lambda_k}{c(x_k)}. \end{aligned}$$

考虑到 $x_{k+1} = x_k + d_k$ 以及

$$\frac{\partial}{\partial x} L(x_k, \lambda_k) = g(x_k) - A(x_k)^T \lambda_k,$$

可以将上述方程组改写成

$$\begin{aligned} \frac{\partial^2}{\partial x^2} L(x_k, \lambda_k) \frac{dx}{d\lambda} &= - \frac{g(x_k) - A(x_k)^T \lambda_k}{c(x_k)} \\ &= - \frac{g(x_k) - A(x_k)^T \lambda_k}{c(x_k)}. \end{aligned}$$

所得的解为 d_k 和 λ_{k+1} . 而这样一个方程组恰好是下述二次规划问题

$$\begin{aligned} \min \quad & q(dx) = (dx)^T g(x_k) + \frac{1}{2} (dx)^T \frac{\partial^2}{\partial x^2} L(x_k, \lambda_k) dx, \\ \text{s.t.} \quad & A(x_k) dx + c(x_k) = 0 \end{aligned} \quad (7.4.5)$$

的一阶必要条件, λ_{k+1} 是其最优 Lagrange 乘子. 序列二次规划法就是通过解一系列这样的二次规划问题产生收敛于问题最优解和 Lagrange 乘子的迭代序列 $\{x_k\}$ 和 $\{\lambda_k\}$. 需要注意的是, 在上述二次规划问题中, 约束是问题 (7.4.3) 的约束的线性近似, 但目标函数却不是问题 (7.4.3) 中目标函数的二次近似. 在这个二次函数中其 Hesse 矩阵为 $\frac{\partial^2}{\partial x^2} L(x_k, \lambda_k)$, 它不仅包含目标函数 $f(x)$ 的二阶导数信息, 而且还包含了所有约束函数的二阶导数的信息. 下面为方便表示, 我们用 d 代替 dx .

例 7.4.1 考虑非线性约束最优化问题

$$\begin{aligned} \min \quad & f(x) = e^{3x_1 + 4x_2}, \\ \text{s.t.} \quad & a(x) = x_1^2 + x_2^2 = 0. \end{aligned}$$

其最优解为 $x^* = (-0.6, -0.8)^T$, 相应的 Lagrange 乘子为 $\lambda^* = -5e^{-5/2} - 0.016845$. 以点 $x_0 = (-0.7, -0.7)$ 为初始点, 逐次求下列二次规划问题

$$\begin{aligned} \min \quad & q(d) = \frac{1}{2} d^T (G(x_k) - \lambda_k^2 a(x_k)) d + g_k^T d, \\ \text{s.t.} \quad & a(x_k)^T d + a(x_k) = 0 \end{aligned}$$

的最优解及相应的 Lagrange 乘子, 其中

$$\begin{aligned} g(x) &= e^{3x_1 + 4x_2} \begin{pmatrix} 3 \\ 4 \end{pmatrix}, & a(x) &= 2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \\ G(x) &= e^{3x_1 + 4x_2} \begin{pmatrix} 9 & 12 \\ 12 & 16 \end{pmatrix}, & \lambda^2 a(x) &= \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \end{aligned}$$

在初始点, 解得

$$d_1 = \begin{pmatrix} 0.14196 \\ -0.15624 \end{pmatrix}, \quad \lambda_1^{(1)} = -0.014808.$$

由此得

$$x^{(1)} = \begin{pmatrix} -0.7 \\ -0.7 \end{pmatrix} + \begin{pmatrix} 0.14196 \\ -0.15624 \end{pmatrix} = \begin{pmatrix} -0.55804 \\ -0.85624 \end{pmatrix}.$$

经过五次迭代所得的迭代点为 $x^{(5)} = (-0.6, -0.8)^T$, Lagrange 乘子为 $\lambda_1^{(5)} = -0.016845$. 从这个例子可以看到序列二次规划方法具有很好的收敛性. 下述定理则从理论上阐明了方法的收敛性.

定理 7.4.2 假设 $f, a(i)$ 二阶连续可微, 问题 (7.4.3) 的最优解 x^* 存在, Jacobian 矩阵 $A(x^*)$ 满秩, 存在 λ^* 使一阶必要条件和二阶充分条件成立. 如点 x_0 充分接近 x^* , 则每次迭代二次规划子问题最优解存在, 迭代序列 $\{x_k\}$ 和 $\{\lambda_k\}$ 分别收敛于 x^* 和最优的 Lagrange 乘子 λ^* , 且收敛率是二次的 (证明略).

同无约束最优化的牛顿法相同, 在这个收敛性定理中同样要求 x_0 充分接近 x^* 以确保矩阵 $\frac{\partial^2}{\partial x \partial x} L(x, \lambda)$ 的正定性. 在初始点离 x^* 稍远的时候, 同样的问题 (方法不收敛, 矩阵 $\frac{\partial^2}{\partial x \partial x} L(x, \lambda)$ 不正定) 会出现, 克服这些问题的方法还是引入线性搜索或信赖域技术以扩大方法的收敛域.

在引入线性搜索时, 约束最优化问题同无约束最优化问题又有明显的不同. 对无约束最优化, 沿一个下降方向 d_k 确定步长 α_k , 只要求使目标函数在点 $x_k + \alpha_k d_k$ 有明显的下降. 对于约束最优化问题则不能同样地只要求目标函数下降. 从上面的例子我们看到, 序列二次规划方法不是一个可行点方法, 它每次迭代产生的点不一定可行. 对于一个不可行的迭代点, 衡量其好坏, 在考量其目标函数值的同时, 还要考量点的可行程度. 所谓的势函数 (Merit Function) 就是为

此而引入的用以衡量迭代点好坏的辅助函数,它同前面两节的罚函数类似,是把目标函数和约束函数结合在一起形成的.一个常用的势函数为 l_1 精确罚函数

$$(x, \mu) = f(x) + \frac{1}{\mu} \|c(x)\|_1,$$

这里 $\|c(x)\|_1$ 是向量 $c(x)$ 的 l_1 范数.对于这样的势函数可以证明,只要 $\frac{1}{2} L(x_k, \mu_k)$ 正定, μ 充分小,则问题(7.4.5)的解 d_k 是函数 (x, μ) 在点 x_k 的一个下降方向.沿方向 d_k 进行线性搜索确定步长 α_k 使得 $(x_k + \alpha_k d_k) < (x_k)$, 且有

$$\liminf_k (\|g_k - A(x_k)^T \alpha_k\|^2 + \alpha_k) = 0.$$

算法 7.4.3 (线性搜索序列二次规划方法)

- 步 1. 给定 $\mu > 0$, 允许值 $\epsilon > 0$, 初始点 x_0 , 置 $k=1$;
- 步 2. 解二次规划问题(7.4.5)得 d_k ;
- 步 3. 调正 μ_k 的值,使得 d_k 是函数 (x, μ_k) 在点 x_k 的下降方向,再沿方向 d_k 线性搜索确定步长 α_k ;
- 步 4. 置 $x_{k+1} = x_k + \alpha_k d_k$, 解方程组 $A_{k+1}^T \lambda = g_{k+1}$ 得 Lagrange 乘子 λ_{k+1} ;
- 步 5. 收敛性检验: 如果最后终止准则得到满足,停止迭代,取 x_{k+1} 和 λ_{k+1} 为最优解和 Lagrange 乘子的近似;否则,置 $k = k+1$ 后转步 2.

第三章有关的线性搜索方法可以用来确定步长.迭代的终止准则可以用

$$\|c_k\| + \|g_k - A_k^T \lambda_k\|$$

来判定点 x_k 和 λ_k 对最优解和相应的 Lagrange 乘子的接近程度.

需要注意的是上述的 l_1 精确罚函数并不处处可微,它在所有使至少有一个约束成等式的点不可微.由此产生的一个后果就是所谓的 Maratos 效应,也就是不管迭代点 x_k 离最优解 x^* 有多靠近,步长 $\alpha=1$ (即牛顿步)是不可接受的,从而影响了方法的快速收敛的特性.为克服这一困难,可以使用 Fletcher 的可微精确罚函数

$$(x, \mu) = f(x) - (x)^T c(x) + \frac{1}{2\mu} c(x)^T c(x),$$

其中

$$c(x) = [A(x) A(x)^T]^{-1} A(x) g(x).$$

同样可以证明在 $\frac{1}{2} L(x_k, \mu_k)$ 正定,对适当的 μ_k 值,方向 d_k 是 (x, μ_k) 在点 x_k 的下降方向.

上述方法中要求矩阵 $\frac{1}{2} L(x_k, \mu_k)$ 正定以确保搜索方向 d_k 的下降性,当 $\frac{1}{2} L(x_k, \mu_k)$ 非正定时,搜索方向的下降性得不到保证,乃至问题(7.4.5)无解.在这种情况下,如同无约束最优化的情形一样可以采用信赖域技术.然而对于问题(7.4.5),我们不能像无约束优化那样简单地加一个信赖域约束 $\|d\| \leq \Delta_k$, 这

是因为简单地加一个信赖域约束所形成的问题

$$\begin{aligned} \min \quad & q(d) = d^T g(x_k) + \frac{1}{2} d^T \frac{\partial^2 L(x_k, \lambda_k)}{\partial x x} d, \\ \text{s.t.} \quad & A(x_k) d + c(x_k) = 0, \\ & \|d\| \leq \Delta_k \end{aligned} \quad (7.4.6)$$

有可能由于两组约束的矛盾使可行域是空集.为此我们考虑问题(7.4.5)的一个变形.根据矩阵 $A_k = A(x_k)$,我们可以把方向向量 d 分解成互相正交的两部分 $d_k = h_k + v_k$,其中 $h_k \in N(A_k)$ (矩阵 A_k 的零空间), $v_k \in R(A_k)$ (矩阵 A_k 的值空间),要求 v_k 满足问题中的线性约束 $A_k v_k + c_k = 0$,而 h_k 是通过求一个同(7.4.5)有关的二次函数的极小解来确定的.对这样的 $d_k = h_k + v_k$,显然有 $A_k d_k + c_k = A_k(h_k + v_k) + c_k = A_k v_k + c_k = 0$.令

$$A_k = \begin{bmatrix} Y_k & Z_k \end{bmatrix} \begin{bmatrix} R_k \\ 0 \end{bmatrix}$$

是矩阵 A_k 的 QR 分解,这里 $Y_k \in \mathbf{R}^{n \times m}$, $Z_k \in \mathbf{R}^{n \times t}$ ($t = n - m$) 为正交矩阵,则矩阵 Y_k 和 Z_k 的列向量分别形成空间 $R(A_k)$ 和 $N(A_k)$ 的一组正交基,向量 h_k 可以表成 Z_k 的列向量的一个线性组合, $h_k = Z_k s_k$,而 $s_k \in \mathbf{R}^t$ 是下述无约束二次优化问题

$$\min \quad Q(s) = g_k^T Z_k s + \frac{1}{2} s^T Z_k^T \frac{\partial^2 L(x_k, \lambda_k)}{\partial x x} Z_k s \quad (7.4.7)$$

的最优解.据此,在上述线性搜索序列二次规划方法中,搜索方向可以如下确定:解方程组

$$A_k v + c_k = 0 \quad (7.4.8)$$

得 v_k ,解问题(7.4.7)得 s_k ,置 $d_k = v_k + Z_k s_k$.然后再沿 d_k 进行线性搜索.

由于问题(7.4.7)中无任何约束,对 s 就可以加上信赖域的限制得信赖域子问题

$$\begin{aligned} \min \quad & Q(s) = g_k^T Z_k s + \frac{1}{2} s^T Z_k^T \frac{\partial^2 L(x_k, \lambda_k)}{\partial x x} Z_k s, \\ \text{s.t.} \quad & \|s\| \leq \Delta_k. \end{aligned} \quad (7.4.9)$$

这个信赖域子问题的解存在,设为 s_k ,则取 $d_k = v_k + Z_k s_k$,计算势函数在点 $x_k + d_k$ 处的函数值,并同预计的函数下降量作比较以确定是否把 $x_k + d_k$ 取为新的迭代点或减小信赖域重新求解(7.4.9)以得到更好的 s_k 和 d_k .

算法 7.4.4 (信赖域序列二次规划方法)

- 步 1. 给定 $\epsilon > 0$, 控制参数 $\rho \in (0, 1/2)$, 初始点 x_1 , 置 $k = 1$;
- 步 2. 解方程组(7.4.8)得 v_k ;
- 步 3. 解二次规划问题(7.4.9)得 s_k , 置 $d_k = v_k + Z_k s_k$;
- 步 4. 调正 μ_k 的值, 计算 $\mu_k = (x_k) - (x_k + d_k)$, $\mu_k = -Q(s_k)$;

步 5. 如果 $\|g_k\| = \|d_k\| / \alpha_k < \epsilon$, 减小 α_k 后转步 3;

步 6. 置 $x_{k+1} = x_k + d_k$, 计算 α_{k+1} ;

步 7. 收敛性检验: 如果最后终止准则得到满足, 停止迭代, 取 x_{k+1} 和 α_{k+1} 为最优解和 Lagrange 乘子的近似; 否则, 取 $\alpha_{k+1} = \alpha_k$, 置 $k = k+1$ 后转步 2;

该算法步 5 中 $\alpha_k < \epsilon$ 表明, (7.4.9) 中的目标函数在给定的信赖域内不是原问题的一个好的近似, 需要减小信赖域来改善近似的程度, 并对减小的 α_k 重解问题 (7.4.9) 得新的解 s_k 和 d_k .

在上述无论是线性搜索序列二次规划法还是信赖域序列二次规划法中, 都要涉及 Lagrange 函数 $L(x, \lambda)$ 的 Hesse 矩阵的计算, 这不仅包含目标函数的 Hesse 阵, 还包含各约束函数的 Hesse 阵, 对规模稍微大一点的问题, 其计算工作量是相当大的. 一个有效的办法还是利用拟牛顿修正公式来产生 $^{-2}_{xx} L(x, \lambda)$ 的近似. 在利用相关的拟牛顿修正公式, 如 BFGS 公式产生 $^{-2}_{xx} L(x, \lambda)$ 的近似时, 公式中向量 y_k 的定义取为

$$y_k = -x L(x_{k+1}, \lambda_k) - x L(x_k, \lambda_k).$$

这时在采用线性搜索的序列二次规划算法中, 线性搜索的准则并不能保证有 $s_k^T y_k > 0$ 的条件成立, 从而不能保证有关修正矩阵的正定性, 这里 $s_k = x_{k+1} - x_k$. 一个值得推荐保证修正矩阵正定的替代方案为如果有 $s_k^T y_k > s_k^T B_k s_k$ 时就用 y_k 进行修正; 而在 $s_k^T y_k < s_k^T B_k s_k$ 时, 取 $\alpha_k y_k + (1 - \alpha_k) B_k s_k$ 作为 y_k 来进行矩阵的修正, 其中

$$\alpha_k = \frac{(1 - \alpha_k) s_k^T B_k s_k}{s_k^T B_k s_k - s_k^T y_k}.$$

可以看到对于如此选取的向量 y_k 有 $s_k^T y_k < s_k^T B_k s_k$, 可取为 (0, 0.5) 中的某一个值, 如 $\alpha_k = 0.2$. 对于问题 (7.4.7) 和 (7.4.9) 中采用矩阵修正公式时, 有两种方案可供选择, 一是上述办法用修正公式直接产生 $^{-2}_{xx} L(x, \lambda)$ 的近似; 另一个方案是用修正公式产生 $t \times t$ 阶矩阵 $Z^T \cdot^{-2}_{xx} L(x, \lambda) Z$ 的近似, 称此矩阵为简约 Hesse 矩阵 (Reduced Hesse). 这时有关修正公式中的向量 y_k 定义为

$$y_k = Z_k^T [-x L(x_{k+1}, \lambda_k) - x L(x_k, \lambda_k)],$$

而 s_k 定义为 $s_k s_k$.

当把序列二次规划方法推广至带有不等式约束的最优化问题

$$\begin{aligned} \min \quad & f(x), \\ \text{s.t.} \quad & c_i(x) = 0, \quad i = 1, \dots, m, \\ & c_i(x) \leq 0, \quad i = m+1, \dots, l \end{aligned}$$

时, 其推广是直接的, 每次迭代所要求解的二次规划子问题为

$$\begin{aligned} \min \quad & q(d) = d^T g(x_k) + \frac{1}{2} d^T W_k d, \\ \text{s.t.} \quad & c_i(x_k)^T d + c_i(x_k) = 0, \quad i = 1, \dots, m, \\ & c_i(x_k)^T d + c_i(x_k) \leq 0, \quad i = m+1, \dots, l \end{aligned} \quad (7.4.10)$$

其中 W_k 或为 $\frac{1}{2} L(x_k, \lambda_k)$ (牛顿法) 或为其近似 (拟牛顿法), Lagrange 函数为

$$L(x, \lambda) = f(x) - \sum_{i \in E} \lambda_i c_i(x) - \sum_{i \in I} \lambda_i c_i(x).$$

在进行线性搜索确定步长时所用的 l_1 精确罚函数为

$$P(x, \mu) = f(x) + \frac{1}{\mu} \sum_{i \in E} |c_i(x)| + \sum_{i \in I} \max\{0, -c_i(x)\}.$$

Fletcher 的可微精确罚函数为

$$P(x, \mu) = f(x) - (x)^T c(x) + \frac{1}{\mu} \sum_{i \in E} |c_i(x)| + \sum_{i \in I} \max\{0, -c_i(x)\}.$$

小 结

本章主要研究约束优化问题的理论与求解方法, 约束优化问题的理论与最优性条件是最优化算法的基础. KKT 条件在 1951 年由 Kuhn Tucker 提出, 而在 1939 年 Karush 也类似地提出了约束优化的最优性条件, 所以称为 KKT 条件. 本章主要简单介绍了约束优化的理论以及罚函数方法, 涉及二次罚函数, 内点障碍罚函数. 此外, 我们讨论了序列二次规划方法解一般约束最优化问题.



1. 对于 $x \in \mathbf{R}$, 证明函数 $\min(z, 0)^2$ 在 $z=0$ 处有二阶不连续导数.
2. 考虑问题

$$\begin{aligned} \min \quad & \frac{1}{1+x^2}, \\ \text{s.t.} \quad & x \leq 1, \end{aligned}$$

写出 $P(x; \mu)$.

3. 考虑问题

$$\begin{aligned} \min \quad & x, \\ \text{s.t.} \quad & x^2 \leq 0 \\ & x+1 \leq 0, \end{aligned}$$

解为 $x^* = -1$, 写出此问题的 $P(x; \mu)$, 寻找局部极小点.

4. 假设当前点 x_k 是 $P(x; \mu_k)$ 的精确极小点, 写出关于当前点 x_k 对于 $P(x; \mu_{k+1})$ 的牛顿步.

5. 对附录中试验函数 2.1 - 2.3 分别用罚函数法和序列二次规划法求解 (用 C 语言、MATLAB 等语言编制程序).

附录 试验函数

§ 1 无约束最优化问题的试验函数

问题 1.1 Rosenbrock 函数:

$$\begin{aligned} f(x) &= 100(x_2 - x_1^2)^2 + (1 - x_1)^2, \\ x_0 &= [-1.2, 1]^T, x^* = [1, 1]^T, f(x^*) = 0. \end{aligned} \quad (1.1)$$

问题 1.2 广义 Rosenbrock 函数:

$$\begin{aligned} f(x) &= \sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 + (1 - x_i)^2], \\ x_0 &= [-1.2, 1, \dots, -1.2, 1]^T, x^* = [1, 1, \dots, 1, 1]^T, f(x^*) = 0. \end{aligned} \quad (1.2)$$

问题 1.3 Wood 函数:

$$\begin{aligned} f(x) &= 100(x_1^2 - x_2)^2 + (x_1 - 1)^2 + (x_3 - 1)^2 + 90(x_3^2 - x_4)^2 + \\ &\quad 10.1[(x_2 - 1)^2 + (x_4 - 1)^2] + 19.8(x_2 - 1)(x_4 - 1), \\ x_0 &= [-3, -1, -3, -1]^T, x^* = [1, 1, 1, 1]^T, f(x^*) = 0. \end{aligned} \quad (1.3)$$

问题 1.4 Powell 奇异函数:

$$\begin{aligned} f(x) &= (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4, \\ x_0 &= [3, -1, 0, 1]^T, x^* = [0, 0, 0, 0]^T, f(x^*) = 0. \end{aligned} \quad (1.4)$$

问题 1.5 立方体函数

$$\begin{aligned} f(x) &= 100(x_2 - x_1^3)^2 + (1 - x_1)^2, \\ x_0 &= [-1.2, -1]^T, x^* = [1, 1]^T, f(x^*) = 0. \end{aligned} \quad (1.5)$$

问题 1.6 三角函数

$$\begin{aligned} f(x) &= \sum_{i=1}^n n + i(1 - \cos x_i) - \sin x_i - \sum_{j=1}^n \cos x_j^2, \\ x_0 &= \left[\frac{1}{5n}, \dots, \frac{1}{5n} \right]^T, x^* = [0, \dots, 0]^T, f(x^*) = 0. \end{aligned} \quad (1.6)$$

问题 1.7 螺旋形凹谷函数

$$f(x) = 100(x_3 - 10)^2 + x_1^2 + x_2^2 - 1^2 + x_3^2, \quad (1.7)$$

其中

$$\begin{aligned} 2 &= \arctan(x_1/x_2), \quad \text{if } x_1 > 0, \\ &+ \arctan(x_2/x_1), \quad \text{if } x_1 < 0, \\ x_0 &= [-1, 0, 0]^T, x^* = [1, 0, 0]^T, f(x^*) = 0. \end{aligned}$$

§ 2 约束最优化问题的试验函数

下面的约束最优化问题试验函数选自 Hock and Schittkowski(1981) .

问题 2.1

变量个数: $n=2$.

目标函数:

$$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 .$$

约束:

$$\begin{aligned} -0.25x_1^2 - x_2^2 + 1 &\leq 0, \\ x_1 - 2x_2 + 1 &= 0 . \end{aligned}$$

初始点: $x_0 = (2, 2)$, $f(x_0) = 1$.

解 $x^* = 0.5\sqrt{7} - 1, 0.25\sqrt{7} + 1$,

$$f(x^*) = 9 - 2.875\sqrt{7} .$$

问题 2.2

变量个数: $n=2$.

目标函数:

$$f(x) = (x_1 - 2)^2 + (x_2 - 1)^2 .$$

约束:

$$\begin{aligned} -x_1 - x_2 + 2 &\leq 0, \\ -x_1^2 + x_2 &\leq 0 . \end{aligned}$$

初始点: $x_0 = (2, 2)$, $f(x_0) = 1$.

解 $x^* = (1, 1)$, $f(x^*) = 1$.

问题 2.3

变量个数: $n=2$.

目标函数:

$$\begin{aligned} f(x) &= -75.196 + 3.8112x_1 + 0.0020567x_1^3 - 1.0345E-5x_1^4 + \\ &6.8306x_2 - 0.030234x_1x_2 + 1.28134E-3x_2x_1^2 + \\ &2.266E-7x_1^4x_2 - 0.25645x_2^2 + 0.0034604x_2^3 - 1.3514E-5x_2^4 + \\ &28.106'(x_2 + 1) + 5.2375E-6x_1^2x_2^2 + 6.3E-8x_1^3x_2^2 - \\ &7E-10x_1^3x_2^3 - 3.405E-4x_1x_2^2 + 1.6638E-6x_1x_2^3 + \end{aligned}$$

$$2.8673 \exp(0.0005 x_1 x_2) - 3.5256 E - 5 x_1^3 x_2.$$

约束:

$$\begin{aligned} x_1 x_2 - 700 &= 0, \\ x_2 - x_1^2 / 125 &= 0, \\ (x_2 - 50)^2 - 5(x_1 - 55) &= 0, \\ 0 \leq x_1 &\leq 75, \\ 0 \leq x_2 &\leq 65. \end{aligned}$$

初始点: $x_0 = (90, 10)$, $f(x_0) = 86.878639$.

解 $x^* = (13.55010424, 51.66018129)$, $f(x^*) = -7.804226324$.

问题 2.4

变量个数: $n = 3$.

目标函数:

$$f(x) = 1000 - x_1^2 - 2x_2^2 - x_3^2 - x_1 x_2 - x_1 x_3.$$

约束:

$$\begin{aligned} 8x_1 + 14x_2 + 7x_3 - 56 &= 0, \\ x_1^2 + x_2^2 + x_3^2 - 25 &= 0, \\ 0 \leq x_i, i &= 1, 2, 3. \end{aligned}$$

初始点: $x_0 = (2, 2, 2)$, $f(x_0) = 976$

解 $x^* = (3.512118414, 0.2169881741, 3.552174034)$,

$$f(x^*) = 961.7151721.$$

问题 2.5

变量个数: $n = 3$.

目标函数:

$$f(x) = 9 - 8x_1 - 6x_2 - 4x_3 + 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1 x_2 + 2x_1 x_3.$$

约束:

$$\begin{aligned} 3 - x_1 - x_2 - 2x_3 &= 0, \\ 0 \leq x_i, i &= 1, 2, 3. \end{aligned}$$

初始点: $x_0 = (0.5, 0.5, 0.5)$, $f(x_0) = 2.25$

解 $x^* = (4/3, 7/9, 4/9)$, $f(x^*) = 1/9$.

问题 2.6

变量个数: $n = 6$

目标函数:

$$\begin{aligned} f(x) &= 0.0204 x_1 x_4 (x_1 + x_2 + x_3) + 0.0187 x_2 x_3 (x_1 + 1.57 x_2 + x_4) + \\ &0.0607 x_1 x_4 x_5^2 (x_1 + x_2 + x_3) + \\ &0.0437 x_2 x_3 x_6^2 (x_1 + 1.57 x_2 + x_4) \end{aligned}$$

约束:

$$0.001 x_1 x_2 x_3 x_4 x_5 x_6 - 2.07 \quad 0$$

$$1 - 0.00062 x_1 x_4 x_5^2 (x_1 + x_2 + x_3) - 0.00058 x_2 x_3 x_6^2 (x_1 + 1.57 x_2 + x_4) \quad 0$$

$$0 \quad x_i, i = 1, \dots, 6$$

初始点: $x_0 = (5.54, 4.4, 12.02, 11.82, 0.702, 0.852)$, $f(x_0) = 137.066$

解

$$x^* = (5.332666, 4.656744, 10.43299, 12.08230, 0.7526074, 0.87865084),$$

$$f(x^*) = 135.075961$$

问题 2.7

变量个数: $n = 10$.

目标函数:

$$f(x) = \sum_{i=1}^{10} (\ln(x_i - 2))^2 + (\ln(10 - x_i))^2 - \sum_{i=1}^{10} x_i^2.$$

约束:

$$2.001 \leq x_i \leq 9.999, i = 1, \dots, 10.$$

初始点: $x_0 = (9, \dots, 9)$, $f(x_0) = -43.134337$

解 $x^* = (9.35025655, \dots, 9.35025655)$, $f(x^*) = -45.77846971$

问题 2.8

变量个数: $n = 10$.

目标函数:

$$f(x) = \sum_{j=1}^{10} \exp(x_j) \alpha_j + x_j - \ln \sum_{k=1}^{10} \exp(x_k),$$

其中,

$$\alpha_1 = -6.089, \alpha_2 = -17.164, \alpha_3 = -34.054,$$

$$\alpha_4 = -5.914, \alpha_5 = -24.721, \alpha_6 = -14.986,$$

$$\alpha_7 = -24.100, \alpha_8 = -10.708, \alpha_9 = -26.662, \alpha_{10} = -22.179$$

约束:

$$\exp(x_1) + 2\exp(x_2) + 2\exp(x_3) + \exp(x_6) + \exp(x_{10}) - 2 = 0$$

$$\exp(x_4) + 2\exp(x_5) + \exp(x_6) + \exp(x_7) - 1 = 0$$

$$\exp(x_3) + \exp(x_7) + \exp(x_8) + 2\exp(x_9) + \exp(x_{10}) - 1 = 0$$

$$-100 \leq x_i \leq 100, i = 1, \dots, 10.$$

初始点: $x_0 = (-2.3, \dots, -2.3)$, $f(x_0) = -21.015$

解

$$x^* = (-3.201212, -1.912060, -0.2444413, -6.537489, -0.7231524,$$

$$-7.267738, -3.596711, -4.017769, -3.287462, -2.335582),$$

$$f(x^*) = -47.76109026$$

参 考 文 献

- 1 邓乃扬 .无约束最优化计算方法 北京:科学出版社,1982
- 2 朱德通 .最优化模型与实验 同济大学出版社,2003
- 3 何旭初,孙文瑜 广义逆矩阵引论 江苏科技出版社,1991
- 4 席少霖 .非线性最优化方法 高等教育出版社,1992
- 5 胡毓达 .非线性规划 高等教育出版社,1990
- 6 徐成闲,陈志平,李乃成 .近代最优化方法 北京:科学出版社,2002
- 7 袁亚湘 .非线性规划数值方法 上海科学技术出版社,1993
- 8 袁亚湘,孙文瑜 最优化理论与方法 北京:科学出版社,1997
- 9 .Armijo .Minimization of functions having Lipschitz continuous partial derivatives .Pacific J .Mathematics, 1966,16:1 ~ 3
- 10 .S . Bazara and C .M . Shetty .Nonlinear Programming .Theory and Algorithms .New York: John Wiley and Sons,1979
- 11 E .M .L . Beale .A derivative of conjugate gradients,in F .A . Lootsma eds . Numerical Methods for Nonlinear Optimization .London: Academic Press, 1972:39 ~ 43
- 12 D .P . Bertsekas .Nonlinear Programming .Athena Scientific, Belmont, Mass ., 1995
- 13 M .C . Bartholomew Biggs .Recursive quadratic programming methods based on the augmented Lagrangian .Mathematical Programming Study, 1987,31:21 ~ 24
- 14 C .G .Broyden, J .E .Dennis, Jr , and J .J .Moré .On the local superlinear convergence of quasi Newton methods .J .Institute of Mathematics and Applications, 1973, 12:223 ~ 246
- 15 R .H . Byrd and J .Nocedal .A tool for the analysis of quasi Newton methods with application to unconstrained minimization .SIAM Journal on Numerical Analysis, 1989, 26:727 ~ 739
- 16 R .H .Byrd, R .B .Schnabel, and G .A .Schultz .Approximate solution of the trust region problem by minimization over two dimensional subspaces . Mathematical Programming, 1988, 40:247 ~ 263

- 17 A .Cohen .Rate of convergence of several conjugate gradient algorithms .
SIAM J .Numer Anal ., 1972, 9: 248 ~ 259
- 18 A . R . Conn, N . I . M . Gould and Ph . L . Toint . Trust Region
Methods .SIAM, 2000
- 19 W . C . Davidon, Variable metric methods for minimization, Argonne
National Labs Report , ANL - 5990
- 20 J .E .Dennis Jr ., and J .J .Moré .Quasi Newton Methods .motivation and
theory .SIAM Review, 1977, 19: 46 ~ 89
- 21 R .Fletcher .Practical Methods of Optimization(second edition) . Chiches-
ter: John Wiley and Sons, 1987
- 22 R .Fletcher and M .J .D . Powell .A rapid convergent descent method for
minimization .Computer Journal, 1963, 6: 163 ~ 168
- 23 C .Floudas, P .Pardalos, des ., Recent Advances in Global Optimization .
Princeton University Press .Princeton, NJ, 1992
- 24 P .E . Gill and W .Murray .Newton type methods for unconstrained and linearly
constrained optimization .Mathematical Programming, 1974, 28: 311 ~ 350
- 25 S .M . Goldfeld, R .E . Quandt and H .F . Trotter .Maximization by quad-
ratic hill climbing climbing, Econometrica, 1966, 34: 541 ~ 551
- 26 A .A . Goldstein .On steepest descent SIAM J .Control, 1965, 3: 147 ~ 151
- 27 A .A .Golstein and J .Price .An effective algorithm for minimization .Num-
er . Math ., 1967, 10: 184 ~ 189
- 28 G .H . Golub and C .F .Van Loan .Matrix Computations .The Johns Hop-
kins University Press .Baltimore, 3rd ed ., 1996
- 29 W .Karush .Minima of functions of several variables with inequalities as side
conditions, Master s thesis .University of chicago Chicago, Illinois, 1939
- 30 H .W .Kuhn and A .W .Tucker, Nonlinear Programming, in: J .Neyman
ed ., Proceeding of the Second Burkeley Symposium on Mathematical Sta-
tistics and Probability, University of Califorlia Press, Berkey, Califorlia,
1951, 481 ~ 492 .
- 31 D .G .Luenberger Introduction to Linear and Nonlinear Programming(2nd
Edition) .Massachusetts: Addison Wesley, 1984
- 32 G . P . McCormick .Nonlinear Programming: Theory, Algorithms, and
Applications .New York: John Wiley and Sons, 1983
- 33 J .J . Moré .Recent developments in algorithms and software for trust
region methods .in: A .Bachem, M .Gr tschel and B .Korte, eds ., Mathe-

- mathematical Programming: The State of the Art, Berlin: Springer Verlag, 1983, 258 ~ 287
- 34 J .J . Moré and D .C . Sorensen .Computing a trust region step .SIAM J . Sci . Stat .Comp ., 1983, 4: 553 ~ 572
 - 35 J .Nocedal and S .J . Wright .Numerical Optimization .New York: Springer, 1999
 - 36 J .Nocedal and Y .Yuan .Analysis of a self scaling quasi Newton method . Math .Prog ., 1993, 61: 19 ~ 37
 - 37 S .S . Oren .Self scaling variable metric algorithm : Implementation and experiments .Management Science, 1974, 20: 863 ~ 874
 - 38 S .S .Oren and D .G .Luenberber .Self scaling variable metric (SSVM) algorithm I: Criteria and sufficient conditions for scaling a class of algorithms .Management Science, 1974, 20: 845 ~ 862
 - 39 J .M .Ortega and W .C .Rheinboldt .Iterative Solution of Nonlinear Equations in Several Variables . New York: Academic Press, 1970
 - 40 M .J .D .Powell .A new algorithm for unconstrained optimization .in: J .B . Rosen, O .L .Mangasarian and K .Ritter, eds . .Nonlinear Programming . New York: Academic Press, 1970, 31 ~ 66
 - 41 M .J .D .Powell .Convergence properties of a class of minimization algorithms, in: O .L .Mangasarian .R .R .Meyer and S .M . Robinson, eds . . Nonlinear Programming 2 New York: Acadmic Press, 1975, 1 ~ 27
 - 42 M .J .D . Powell and Y .Yuan .A .recursive quadratic programming algorithm that use differentiable exact penalty function .Mathematical Programming, 1986, 35: 265 ~ 278
 - 43 L .Qi .Convergence analysis of some algorithms for solving nonsmoth equations . Mathematics of Operations Research, 1993, 18: 227 ~ 244
 - 44 L .Qi and J .Sun .A nonsmooth version of Newton 's method .Mathematical Programming, 1993, 58: 353 ~ 367
 - 45 R .B . Schnabel and P .D .Frank .Tensor methods for nonlinear equations . SIAM J .Numer .Anal ., 1984, 21: 815 ~ 843
 - 46 D .C . Sorensen .The q superlinear convergence of a collinear scaling algorithm for unconstrained optimization .SIAM J .Numer .Anal ., 1980, 17: 84 ~ 114
 - 47 D .C .Sorensen .Newton 's method with a model trust region modification . SIAM J .Numer .Anal ., 1982, 20: 409 ~ 426

- 48 T .Steihaug .The conjugate gradient and trust regions in large scale optimization .SIAM Journal on Numerical Analysis, 1983, 20: 626 ~ 637
- 49 J .Stoer and R .Bulirsch .Introduction to Numerical Analysis . New York: Springer Verlag, 1993
- 50 W .Sun and Z .Wu Numerical research on self scaling variable metric algorithm .Numerical Mathematics, A Journal of Chinese Universities, 1989, 11: 145 ~ 158
- 51 W .Sun . On nonquadratic model optimization methods . Asia and Pacific Journal of Operations Research, 1996, 13: 43 ~ 63
- 52 W .Sun .Newton 's method and quasi Newton SQP method for general LC^1 constrained optimization .Applied Mathematics and Computation, 1998, 92: 69 ~ 84
- 53 W .Sun and Y .Wei .Triple reverse order law for weighted generalized inverse .Applied Mathematics and Computation, 2002, 125: 221 ~ 229
- 54 W .Sun, J .Yuan and Y .Yuan .Conic trust region method for linearly constrained optimization .Journal of Computational Mathematics, 2003, 21: 295 ~ 304
- 55 Y .Yuan .On the superlinear convergence of a trust region algorithm for nonsmooth optimization .Mathematical Programming, 1985, 31: 269 ~ 285 . (1985b)
- 56 Y .Yuan and R .Byrd .Non quasi Newton updates for unconstrained optimization .J .Comp .Math ., 1995, 13: 95 ~ 107
- 57 Y .Yuan .On the truncated conjugate gradient method .Mathematical Programming, 2000, 87: 561 ~ 573
- 58 J .Z .Zhang and C .X .Xu .A class of indefinite dogleg path methods for unconstrained minimization SIAM J .on Optimization, 1994, 9: 646 ~ 667