

分类号_____

密级_____

UDC _____

编号_____

中国科学院研究生院 博士学位论文

特殊非线性最小二乘与 L_1 范数极小化问题的计算方法

刘歆

指导教师

袁亚湘(研究员、博士)

中国科学院数学与系统科学研究院

申请学位级别

博士

学科专业名称

计算数学

论文提交日期

2009年4月13日

论文答辩日期

2009年5月19日

培养单位

中国科学院数学与系统科学研究院

学位授予单位

中国科学院研究生院

答辩委员会主席

胡晓东

摘 要

非线性最小二乘问题是最优化领域里非常重要的一类问题。工程计算中大量的实际问题最终都会归结为非线性的参数估计或数据拟合问题,而非线性最小二乘问题正是这两类问题的数学模型。由于来源于实际问题的非线性最小二乘问题一般都是有特殊形式的。考虑和分析这些具有特殊形式的非线性最小二乘问题,往往可以取得事半功倍的效果,易于设计出一些高效、精确、鲁棒的计算方法,这些结果无论对最优化理论的发展还是对实际的工程计算都是有重大意义的。

我们首先研究了可分离变量的非线性最小二乘问题。我们的研究突破了传统的变量投影的算法框架,提出了结构型变量不分离的算法框架,并据此设计了三种结构型变量不分离算法。我们的算法继承了变量投影方法的优势,同时可与信赖域方法很好结合,解决一般约束优化问题。我们给出了算法的各种收敛性结果。

我们还对二次最小二乘问题的全局极小算法设计问题进行了研究。工程中应用广泛的距离几何模型是这类问题的一个特例。我们提出了全局下降方向的判定方法,根据问题的特殊结构。我们还给出了全局下降方向在二维子空间中的计算方法及全局下降二维子空间的概念和判定准则,并以之构造了随机子空间算法。我们的算法应用在无线传感器定位问题上有很好的数值结果,速度快,稳定性更好,比现有算法有显著的改进。

在实际工程问题中,由于误差往往不一定服从高斯分布,因此通过求解最小二乘模型并不一定能够得到很好的参数估计结果,此时就需要使用其它范数极小化模型。 L_1 范数极小化问题(也称为“最小一乘”)是一种非常好的模型。但是非线性 L_1 范数极小化问题除了具有非凸性外,还属于非光滑优化问题,目前还没有什么非常有效可行的方法能够快速、有效、稳定地求解这类问题。因此在工程应用界,很少使用非线性 L_1 范数极小化模型。

我们对非线性 L_1 范数极小化问题进行了研究。证明了求解此类问题的主要工具:信赖域方法的信赖域子问题是NP-难的。并且对此类非光滑的信赖域子问题提出了一种序列二维子空间算法,能够快速找到信赖域子问题的近似解。我们还给出了当信赖域子问题是凸的时候,算法的收敛性分析。

此外,在本篇论文中,我们还介绍了对最小二乘参数估计与观测次数的关系;提出了凸二次规划的不可行原始对偶积极集方法;给出了序列二维子空间极小方法的收敛性分析;并对无约束问题的随机性算法进行了讨论。

关键词: 非线性最小二乘问题, L_1 范数极小问题, 变量投影方法, 全局下降方向, 随机多初始方法, 序列二维子空间极小方法, 信赖域方法

Numerical Methods for Special Nonlinear Least Squares Problems and L_1 Norm Minimization Problems

Liu Xin (majored in computational mathematics)

Directed by Prof. Yuan Ya-xiang

Nonlinear least squares problem is a class of important optimization problem. Many practical problems in engineer computation can be reduced to nonlinear parameter-estimation or data-fitting problems, both of which exactly belong to nonlinear least squares problems. Since the nonlinear least squares problems derived from different practical problems are usually of special properties, considering and analyzing nonlinear least squares problems with special properties can be beneficial for us to design efficient, accurate and robust computational methods. And these results can be meaningful for both the theory development of optimization and practical engineer computation.

Firstly, we consider nonlinear least squares problems with separable variables. Most existing algorithms for this kind of problems are derived from the variable projection method, which utilizes the separability under a separate framework. We bring forward a new unseparated framework, based on which we proposed three algorithms. Our method maintains all the advantages of variable projection based methods, and moreover it can be combined with trust region methods easily and can be applied to general constrained separable nonlinear problems.

Secondly, We consider the global minimization of quadratic least squares(QLS) problems, which is a widely used model appearing in various areas such as communication engineering, computational biology and so on. For those QLS problems with a large number of variables, there may exist quite a few local optima, and stochastic method is the only possible choice in such cases. Here we propose a subspace sampling technique which is superior to general multistart methods for QLS problems and tested to be very efficient in solving distance geometry problems.

Taking the errors as Gaussian random variables, as we usually do for sake of simplicity, can be false in certain practical problems, in which cases the least square models can be problematic. Thus sometimes it is necessary to take other norm minimization models other than least squares model(L_2 norm minimization model). L_1 norm minimization can be a proper model. However, due to its non-convexity and non-smoothness, feasible algorithms which can solve L_1 minimization problems with

efficiency, accuracy, and stability are almost absent, which partly leads to the situation that nonlinear L_1 norm minimization method is rarely applied to engineer area.

Considering that trust region based method take an important role in solving L_1 norm minimization problem. We focus our research on the trust region subproblem for nonlinear L_1 norm minimization problem. We prove this kind of nonsmooth trust region subproblem is NP-hard. Special direct algorithms for one dimensional and two dimensional cases are introduced, which are of polynomial time complexity. A sequential 2-dimensional subspace minimization method are proposed for the trust region subproblem. We give the convergence analysis to the algorithm in the case of the trust region subspace.

In addition, we construct the relationship between least square estimation and the observation number; propose a infeasible prime-dual active set approach for convex quadratic programming; put forward the convergence analysis of sequential two dimensional minimization method; and discuss some random methods for unconstrained optimization problems.

Keywords: nonlinear least squares problem, L_1 norm minimization problem, variable projection method, global descent direction, stochastic multistart method, sequential two dimensional subspace method, trust region method

目 录

摘 要	i
第一章 引言	1
§1.1 可分非线性最小二乘问题	2
§1.2 二次最小二乘问题的全局优化算法	3
§1.3 L_1 范数极小化问题	3
§1.4 其它相关工作	4
第二章 可分非线性最小二乘问题	7
§2.1 非线性最小二乘基础	7
§2.2 可分非线性最小二乘问题与变量投影方法概述	9
§2.2.1 可分非线性最小二乘问题	9
§2.2.2 变量投影方法	10
§2.3 结构型变量不分离算法	15
§2.3.1 方法构想与框架	15
§2.3.2 算法描述	17
§2.4 新算法的收敛性分析	19
§2.4.1 局部收敛速度	19
§2.4.2 渐近收敛速度	20
§2.5 数值结果	23
第三章 二次最小二乘问题的全局优化算法	27
§3.1 全局优化简述	27
§3.2 多项式优化与距离几何问题	29
§3.2.1 多项式优化问题	30
§3.2.2 距离几何模型	32
§3.3 随机二维子空间方法	36
§3.3.1 随机多初始算法	36
§3.3.2 全局下降方向的判定	38
§3.3.3 全局下降二维子空间的判定	41
§3.3.4 算法构造	43
§3.4 数值结果	44

§3.4.1	算法参数选取	44
§3.4.2	数值结果I	45
§3.4.3	数值结果II	47
第四章	L_1范数极小化问题	49
§4.1	非光滑优化问题概述	49
§4.2	非线性 L_1 范数极小化问题的信赖域算法	52
§4.3	非光滑信赖域子问题分析	53
§4.4	低维非光滑信赖域子问题的直接算法	56
§4.4.1	假设和定义	56
§4.4.2	一维情形	59
§4.4.3	二维情形	63
§4.5	序列二维子空间轮换搜索算法及收敛性分析	73
§4.5.1	方法概述	73
§4.5.2	收敛性分析	76
第五章	其它相关工作	93
§5.1	最小二乘参数估计与观测次数的关系	93
§5.1.1	最小二乘估计	93
§5.1.2	几个线性代数引理	94
§5.1.3	估计与观测次数的关系定理	95
§5.2	凸二次规划的不可行原始对偶积极集方法	97
§5.2.1	问题描述	97
§5.2.2	算法描述	98
§5.2.3	数值实验	100
§5.3	序列二维子空间方法的收敛性分析	104
§5.3.1	算法的收敛性分析	106
§5.3.2	拉格朗日牛顿型子空间取法	110
§5.3.3	小结	119
§5.4	无约束优化的随机性算法	120
§5.4.1	数值比较	122
§5.4.2	随机水平集算法的收敛性	124
第六章	结论与展望	129

参考文献	131
------	-----

发表文章目录	143
--------	-----

致 谢	145
-----	-----

第一章 引言

在工程应用领域中有大量的问题可以归结为参数估计或者数据拟合问题:

$$h(x, t_i) = 0; \quad \forall i = 1, 2, \dots, m. \quad (1.0.1)$$

这里 $h: \mathbf{R}^{n+l} \mapsto \mathbf{R}$; $t_i \in \mathbf{R}^l$ 是观测数据; $x \in \mathbf{R}^n$ 是待定参数。我们需要通过 m 次观测得到的 $t_i (i = 1, 2, \dots, m)$, 来确定 x 。问题(1.0.1)广泛存在于物理、化学、生物、经济、信息科学等各个工程应用领域内。由于实际的观测都会存在误差, 为了保证能够得到较为准确的参数 x , 一般观测次数 m 都会大于 n 。

如果我们记:

$$f_i(x) = h(x, t_i); \quad \forall i = 1, 2, \dots, m. \quad (1.0.2)$$

并令:

$$f(x) = (f_1(x), f_2(x), \dots, f_m(x))^T. \quad (1.0.3)$$

则可以通过求解等价于(1.0.1)的优化问题:

$$\min_{x \in \mathbf{R}^n} F(x) = \|f(x)\|_p \quad (1.0.4)$$

来确定 x 。这里 $\|\cdot\|_p (p \geq 1)$ 是某种范数。

通常来说, 我们都使用 L_2 范数 ($p = 2$), 也即熟知的最小二乘模型。这是因为当 $f_i(x) (i = 1, 2, \dots, m)$ 都是线性的, 且观测误差是白噪声, 即观测数据的误差 $e_i = t_i - t_i^{real}$ 是服从高斯分布的, 则最小二乘估计是无偏估计 [7], [129], 也就是说(1.0.4)的最优解 x^* 和真实参数 x^{real} 满足:

$$\mathbf{E}(x^* - x^{real}) = 0. \quad (1.0.5)$$

当 $f_i(x) (i = 1, 2, \dots, m)$ 是非线性的时候, 工程上也一般使用最小二乘模型, 这是因为求解最小二乘问题已经有一套成熟的算法和理论体系, 参见 Levenberg [52], Marquadt [69], Nocedal 和 Wright [76], Sun 和 Yuan [124], 及 Yuan [166] 等文献。

但是实际应用问题中, 也存在观测噪声和白噪声相去甚远的情形, 这个时候 L_2 范数极小化模型就无法满足求解(1.0.1)的需要了。于是 L_1 范数极小化模型就应运而生。

由于非线性 L_1 范数极小化问题是非常困难的, 因此至今还没有非常有效成熟的算法。

本篇论文将始终围绕着优化问题(1.0.4)的求解。考虑到实际问题往往具有特殊性, 而利用问题的特殊性设计的具有针对性的算法往往有比直接套用一般算法有效, 我们讨论具有特殊性的非线性最小二乘问题和 L_1 范数极小化问题。

§1.1 可分非线性最小二乘问题

可分非线性最小二乘问题的提出是源于原子计算中的参数估计(数据拟合)问题。1972年, Golub和Pereyra在原子物理中粒子半衰期公式的参数估计中发现这类问题都具有如下形式:

$$\phi(a, \alpha; t) = a_1 + a_2 e^{-\alpha_1 t} + a_3 e^{-\alpha_2 t}. \quad (1.1.1)$$

这里 $a \in \mathbf{R}^3$, $\alpha \in \mathbf{R}^2$ 都是待定参数, 根据不同时刻 $t_i (i = 1, 2, \dots, m)$ 的观测, 就可以得到如下的指数拟合问题:

$$\min_{a \in \mathbf{R}^3, \alpha \in \mathbf{R}^2} \frac{1}{2} \sum_{i=1}^m [y_i - \phi(a, \alpha; t_i)]^2. \quad (1.1.2)$$

考虑到观测项 ϕ 具有关于一部分变量 a 是线性的, 关于另一部份变量 α 是非线性的这一特殊性, Golub和Pereyra [37]提出了变量投影方法, 其主要思想是利用线性最小二乘的显示表达式, 把线性部分变量用非线性部分变量表出, 并求降维的非线性最小二乘问题。Ruhe和Wedin [113]证明了变量投影方法的渐近收敛速度比直接使用一般的变量不分离算法要快。

[38]中提到具有这种特性的参数估计(数据拟合)问题是实际工程计算中很多问题的数学模型, 如: 反问题、信号处理、医学和生物成像、神经网络、机器人及可视化、通信、电气与电子工程、微分方程动力系统。可见实际工程应用领域对这类问题的高效算法是具有一定需求的。

不过从 [37]之后, 关于可分非线性最小二乘问题的研究进展, 都是基于变量投影思想。比如, Böckmann [9]介绍了基于有限差分方法来近似计算降维的非线性最小二乘问题的梯度, 并使用信赖域方法求解降维问题。Kaufman和Pereyra [40]将变量投影方法推广到约束具有和目标函数相同的可分离变量形式的带约束可分非线性最小二乘问题。

然而在实际工程计算中, 可分非线性最小二乘问题不一定具有变量可分离的约束。比如说: 界约束或者球约束都是很常见的。但是在这种约束下, 无论显式变量投影算法或隐式变量投影算法都不容易实现。事实上, 变量投影思想是很难推广到求解约束不具有和目标函数相同的可分离变量形式的约束可分非线性最小二乘问题上去的。

我们的工作是通过提出两种对可分非线性最小二乘问题Hessian矩阵的近似矩阵的构造方法, 提出了两组结构型变量不分离的算法框架。这个框架可以直接推广到一般约束优化问题中去。基于这种结构型变量不分离算法框架, 我们提出

了几种格式的算法。数值实验和理论分析都表明我们提出的结构型变量不分离算法具有和变量投影方法一样的渐近收敛速度。

§1.2 二次最小二乘问题的全局优化算法

二次最小二乘问题是(1.0.2)中每个 $f_i(x)$ 都是 x 的二次函数。通常情况下,二次最小二乘问题是非凸的,要求解其全局最优解是非常困难的(参见 [47], [78])。我们容易证明二次最小二乘问题是NP-难问题。

我们研究二次最小二乘问题的全局极小问题,一方面是因为它是欧几里德距离几何问题的推广,而后者在无线通讯网络、计算生物学等领域有着广泛的应用;另一方面是因为它是一类形式最简单的非线性最小二乘问题,研究它的全局极小化算法对设计其它非线性最小二乘问题的全局极小算法有一定的指导意义。

二次最小二乘问题可以看作为特殊的多项式优化问题。多项式优化问题是近年来比较热门的研究课题,很多著名的学者已经提出了不少成熟的确定性或非确定性算法来求解多项式优化问题的全局极小点。多项式优化的主要技巧有:代数几何方法 [18]; 同伦方法 [53], [54], [55]; SOS(平方和)松弛化方法 [116], [121], [79], [81], [75]等。Qi [109]提出了一种求解正定的四次多项式全局极小的方法。

然而通常来说上面介绍的多项式优化方法效率都很低,除了SOS松弛化方法外,其它方法都是只能求解很小规模的问题,基本上都是属于理论算法。而SOS方法,也并不是对所有问题都有效,且算法效率也不是很高。

如果考虑二次最小二乘问题的特例:欧几里德距离几何问题,根据问题来源的不同背景, Ye [132, 122]及Moré和Wu [74]分别提出了半定规划松弛化方法和整体光滑化方法。

我们发展了 [109]中全局下降方向的概念,针对二次最小二乘问题的特殊性,提出了全局下降二维子空间的概念。并且创新地提出了随机二维子空间搜索算法。数值实验表明我们设计的算法在求解来源于无线传感器网络问题的距离几何问题时,优于半定规划松弛化方法、整体光滑化方法和SOS松弛化方法。

§1.3 L_1 范数极小化问题

L_1 范数极小化问题也即是(1.0.4)中 p 取1的优化问题。首先 L_1 范数极小化问题属于非光滑优化问题 [11, 12], 非光滑优化问题是优化中一类很难而且很重要的课题。目前主要的算法有次梯度(subgradient)方法(Shor [114, 115, 116], Eremin [31], Polyak [82, 83, 84]等); 共轭次梯度方法(Wolfe [130]等); 捆集(bundle)法(Lemaréchal

[49, 50, 51], Zowe [151]等); 求解凸规划的割平面法(Kelley [41], Cheney和Goldstein [17]等)以及信赖域方法等。

L_1 范数极小化问题还属于复合非光滑优化问题。Fletcher([33])1981年给出了一个求解复合NDO的信赖域方法。这个方法的主要思想是每步迭代都求解一个非光滑的二阶近似的信赖域子问题。上个世纪八十年代一些国际知名优化学家建立了一系列关于用信赖域方法求解复合非光滑优化问题的重要收敛性结果(如: Fletcher [34], Powell [98]及Yuan [136, 137, 138, 139, 140]等)。此外求解复合非光滑优化问题还有Hald [44, 45]提出的两阶段方法。

由于目前求解 L_1 范数极小化问题的最有效的算法主要是信赖域方法,于是我们对非光滑的信赖域方法进行了研究。我们不是着眼于信赖域框架在非光滑优化时的收敛性,而是主要讨论每一步迭代都需要求解的非光滑信赖域子问题。

我们考虑到非光滑信赖域子问题本身是非凸、非光滑的,于是我们分析了它的计算复杂度,并证明了求解 L_1 范数极小化问题的非光滑信赖域子问题本身也是NP-难问题。于是我们就考虑构造求解这种信赖域子问题的算法。

首先我们给出了低维($n = 1$ 或 $n = 2$)情形下,求解这类非光滑信赖域子问题的全局极小点的多项式时间复杂度的直接算法。并推广了Powell [108]中提出的作为求解一般信赖域子问题的截断共轭梯度方法后处理技巧的序列二维子空间方法,提出了非光滑情形下的序列二维子空间方法。我们还对方法的收敛性进行了一定的分析。

§1.4 其它相关工作

在解决来自实际工程应用问题的参数估计或数据拟合问题时,我们发现增加观测次数,通常能够得到更好的解。于是我们在一些基本假设下,建立最小二乘参数估计与观测次数的关系。并发现在假设条件下,随着观测次数的增加,最小二乘估计的误差的期望是单调下降的。

我们提出了一种求解凸二次规划的不可行原始对偶积极集方法;我们注意到Kunisch和Rendl [43]及Voglis和Lagaris [127]都提出了类似的算法思想,不过前者是针对单向界约束,后者是针对双向界约束。而我们独立提出的算法适用范围更广。此外 [43]对单向界约束的问题所作的收敛性分析并不实用,因为文中的分析是基于 A 矩阵满足两个极其苛刻的条件。实验表明我们提出的算法有很好的数值表现。

Powell [108]提出的序列二维子空间极小技巧,作为用截断共轭梯度法([13]; [143])求解一般信赖域子问题的后处理技巧,可以使算法有很好的数值表现,然而

却没有做收敛性分析。我们给出了直接使用序列二维子空间极小方法求解信赖域子问题的收敛性和局部收敛速度。在探索如何选取能够高效收敛的子空间的问题中，我们提出了拉格朗日牛顿型子空间选取方法。并证明了以此构造的序列二维子空间算法具有局部二阶收敛速度。

设计和分析无约束优化问题的高效、低计算量的算法，一直是无约束优化算法研究的重要课题。我们比较了通过梯度方向结合随机步长、随机方向结合精确搜索步长、梯度方向加随机方向的序列二维子空间方法及随机水平集方法的数值表现，并给出了随机水平集方法的收敛性分析。

第二章 可分非线性最小二乘问题

可分非线性最小二乘问题是一类特殊的非线性最小二乘问题。本章将首先介绍非线性最小二乘问题及其一般算法；其次对可分非线性最小二乘问题的应用、特性和已有方法作一个简单的综述；最后介绍我们在这个问题上所作的工作。

§2.1 非线性最小二乘基础

非线性最小二乘问题的一般形式为：

$$\min_{x \in \mathbf{R}^n} F(x) = \frac{1}{2} \|f(x)\|_2^2 = \frac{1}{2} \sum_{i=1}^m f_i^2(x), \quad (2.1.1)$$

这里 $f: \mathbf{R}^n \mapsto \mathbf{R}^m$ 是关于 x 的非线性函数, $(f(x))_i = f_i(x)$ 。我们可以将(2.1.1)视为一般的无约束优化问题, 使用一般的无约束优化算法来求解。但是由于其目标函数 s 具有特殊形式, 所以我们可以充分利用这些特殊形式来构造求解非线性最小二乘的特殊方法。

最基本的求解非线性最小二乘问题的方法是高斯牛顿方法 ([28, 77])。高斯牛顿方法的思想非常直观: 设当前迭代点为 x_k , 用 $f(x_k + d)$ 在 x_k 处的线性近似 $f(x_k + d) \approx f(x_k) + J(x_k)d$ 来代替 $f(x_k + d)$, 通过极小化如下线性最小二乘模型:

$$\min_{d \in \mathbf{R}^n} \frac{1}{2} \|f(x_k) + J(x_k)d\|_2^2. \quad (2.1.2)$$

来得到一个迭代步 $d_k^{(GN)}$, 这里 $J(x)$ 是 $f(x)$ 的 Jacobian 矩阵。当 $J(x_k)$ 的秩为 n 时, (2.1.2) 的解 (一般称之为高斯牛顿步) $d_k^{(GN)}$ 有如下显式表示:

$$d_k^{(GN)} = -(J(x_k))^+ f(x_k) = -(J(x_k)^T J(x_k))^{-1} J(x_k)^T f(x_k). \quad (2.1.3)$$

于是, 下一个迭代点为:

$$x_{k+1} = x_k + d_k^{(GN)}. \quad (2.1.4)$$

如果我们比较(2.1.3)与(2.1.1)的牛顿步

$$\begin{aligned} d_k^{(N)} &= -\nabla^2 F(x_k)^{-1} \nabla F(x_k) \\ &= -(J(x_k)^T J(x_k) + \sum_{i=1}^m \nabla^2 f_i(x_k) f_i(x_k))^{-1} J(x_k)^T f(x_k), \end{aligned} \quad (2.1.5)$$

就可以发现对非线性最小二乘模型, 高斯牛顿法和牛顿法的区别在于, 高斯牛顿法仅用一阶信息 $J(x_k)^T J(x_k)$ (称为高斯牛顿矩阵) 来近似牛顿法中 $F(x)$ 的 Hessian 矩

阵 $\nabla^2 F(x)$, 而省略了二阶信息项 $\sum_{i=1}^m \nabla^2 f_i(x) f_i(x)$ 。因此容易证明, 当问题(2.1.1)在解处的函数值为零时, 高斯牛顿法有着和牛顿法一样的局部二阶收敛速度; 当省略的二阶信息项相对于高斯牛顿矩阵是无穷小量时, 高斯牛顿法是线性收敛的; 而一旦二阶信息项太大时, 高斯牛顿法就可能不收敛。

为了保证高斯牛顿法的全局收敛性, 我们往往会在(2.1.4)式的迭代方向 $d_k^{(GN)}$ 前加上一个步长因子 α_k , 使得新的迭代点

$$x_{k+1} = x_k + \alpha_k d_k^{(GN)}, \quad (2.1.6)$$

满足某种步长搜索准则, 这类方法我们称之为阻尼高斯牛顿法。

然而在实际问题中, 往往会碰到Jacobian矩阵 $J(x_k)$ 的秩小于 n 的情况。这时, 无论是高斯牛顿法还是阻尼高斯牛顿法都是没有定义的, 于是Levenberg [52]和Marquardt [69]提出了Levenberg-Marquardt方法。Levenberg-Marquardt方法的主要思想是使用一个正的标量矩阵来近似Hessian阵中的二阶信息项:

$$d_k^{(LM)} = -(J(x_k)^T J(x_k) + \lambda_k \cdot I)^{-1} J(x_k)^T f(x_k). \quad (2.1.7)$$

关于正则化参数 λ 的取法有很多种, 可以参见 [76, 124, 166]等教材的详细介绍。其中最著名的是Moré [70]在MINPACK软件包中使用的用信赖域技巧来控制迭代的Levenberg-Marquardt-Moré算法。2000年后, Yamashita和Fukushima [135], Fan和Yuan [32]分别给出了在弱于非奇异性条件的局部误差界条件下正则化参数 λ 的两种不同取法, 并证明了采用在相应取法下的Levenberg-Marquardt方法具有局部二阶收敛速度。

当问题(2.1.1)在解处的函数值很大(大残量问题)或者目标函数的非线性程度很高时, 如果仍旧采用高斯牛顿法, 忽略二阶信息项; 或者采用Levenberg-Marquardt方法用数量矩阵来代替二阶信息项, 都不能很好地近似目标函数的Hessian矩阵, 从而导致收敛速度很慢。而另一方面如果直接计算二阶信息项, 则需要求得所有 $f_i(x)$ 的Hessian矩阵, 代价又太大。所以Broyden和Dennis [25], Biggs [4], 及Dennis, Gay和Welsch [26] 分别提出了求解非线性最小二乘问题的结构型拟牛顿算法, 其主要思想是采用如下的结构型拟牛顿步来代替牛顿步(2.1.5):

$$d_k^{(SQN)} = -(J(x_k)^T J(x_k) + A_k)^{-1} J(x_k)^T f(x_k). \quad (2.1.8)$$

这里 A_k 是二阶信息项 $\sum_{i=1}^m \nabla^2 f_i(x) f_i(x)$ 的割线近似, 需要满足如下割线方程:

$$A_{k+1} s_k = u_k, \quad u_k = y_k - J(x_{k+1})^T J(x_{k+1}) s_k; \quad (2.1.9)$$

其中

$$s_k = x_{k+1} - x_k, \quad y_k = J(x_{k+1})^T f(x_{k+1}) - J(x_k)^T f(x_k). \quad (2.1.10)$$

之所以称之为结构型拟牛顿算法，是因为根据非线性最小二乘问题的特性，拟牛顿步(2.1.8)保留了高斯牛顿阵的部分，仅用割线近似二阶信息项。将结构型拟牛顿算法和信赖域思想结合，Dennis、Gay和Welsch [26]给出了求解非线性最小二乘的算法程序NL2SOL。对于大残量问题，NL2SOL有较大优越性，对于小残量问题，NL2SOL与Levenberg-Marquardt-Moré算法差不多，而Levenberg-Marquardt-Moré更简单。这两种算法都是目前求解非线性最小二乘问题的最流行的算法。

如果适当选取 A_k 的校正公式，结构型拟牛顿算法对零残量和非零残量的问题都有局部超线性收敛性 [27]。当运用线搜索技巧时，一般我们都希望结构型拟牛顿矩阵 $J(x_k)^T J(x_k) + A_k$ 是正定的，从而 $d_k^{(SQN)}$ 是 $F(x)$ 在当前迭代点 x_k 处的下降方向。但是大多数结构型拟牛顿法并不能保证这一点。为克服这种困难，Yabe和Takahashi [134]提出了分解结构型拟牛顿算法，其思想是在每步迭代计算：

$$d_k^{(FSQN)} = -((J(x_k) + L_k)^T (J(x_k) + L_k))^{-1} J(x_k)^T f(x_k), \quad (2.1.11)$$

其中 $L_k \in \mathbf{R}^{m \times n}$ ， $L_k^T L_k + J_k^T L_k + L_k^T J_k$ 是二阶信息项 $\sum_{i=1}^m \nabla^2 f_i(x) f_i(x)$ 的近似。因此 L_{k+1} 满足如下的拟牛顿条件：

$$(J(x_k) + L_k)^T (J(x_k) + L_k) s_k = z_k, \quad (2.1.12)$$

其中

$$z_k = y_k, \quad \text{或} \quad z_k = y_k + J_{k+1}^T J_{k+1} s_k. \quad (2.1.13)$$

关于非线性最小二乘的工作，还有很多，如Fletcher和Xu [35]提出的杂交方法等。有兴趣的读者可以查阅 [156]， [76]等。

§2.2 可分非线性最小二乘问题与变量投影方法概述

§2.2.1 可分非线性最小二乘问题

可分非线性最小二乘问题的提出是源于原子计算中的参数估计(数据拟合)问题。1972年，Golub和Pereyra在原子物理粒子半衰期公式的参数估计中发现，这类问题都具有如下形式：

$$y_i = \sum_{j=1}^p a_j \phi_j(b, t_i), \quad (i = 1, 2, \dots, m), \quad (2.2.1)$$

这里 $\phi_j(b, t)$ ($j = 1, \dots, p$)是定义在 \mathbf{R}^{q+1} 上的实值非线性函数; t_i 和 y_i ($i = 1, \dots, m$)是观测数据; $a \in \mathbf{R}^p$ 和 $b \in \mathbf{R}^q$ 是待估计的参数; 正整数 p 和 q 分别是两个参数的维数, 我们记 $n = p + q$ 。在实际观测中, 观测次数 m 往往远大于问题的维数 n 。

我们可以把参数估计(数据拟合)问题(2.2.1)写成如下的非线性方程组问题:

$$f_i(a, b) = 0, \quad i = 1, 2, \dots, m, \quad (2.2.2)$$

这里

$$f_i(a, b) = y_i - \sum_{j=1}^p a_j \phi_j(b, t_i), \quad i = 1, 2, \dots, m. \quad (2.2.3)$$

我们容易发现(2.2.2)中每一个观测项方程(2.2.3)都是关于一部分待估参数 a 的线性函数, 关于另一部分变量 b 的非线性函数。具有这种特性的参数估计(数据拟合)问题是实际工程计算中很多问题的数学模型, 如: 反问题、信号处理、医学和生物成像、神经网络、机器人及可视化、通信、电气与电子工程、微分方程动力系统, 等等 [37, 38]。

我们知道超定的非线性方程组问题(2.2.2), 一般都是转化为如下的非线性最小二乘问题来求解:

$$\min_{a \in \mathbf{R}^p, b \in \mathbf{R}^q} \frac{1}{2} \sum_{i=1}^m (f_i(a, b))^2. \quad (2.2.4)$$

因为当(2.2.2)有解时, (2.2.2)与(2.2.4)是等价的。(2.2.4)的数值算法即可用来求解(2.2.2)。可参阅 [28], [77], [124], [165], [161]等。

当 y_i ($i = 1, \dots, m$)还是 b 的函数时, 参数估计(数据拟合)问题(2.2.1)就可转化为标准的可分非线性最小二乘问题:

$$\min_{a \in \mathbf{R}^p, b \in \mathbf{R}^q} \psi(a, b) = \frac{1}{2} \|y(b) - \Phi(b)a\|_2^2, \quad (2.2.5)$$

这里 $y: \mathbf{R}^q \mapsto \mathbf{R}^m$; $\Phi: \mathbf{R}^q \mapsto \mathbf{R}^{m \times p}$, 且 $(\Phi(b))_{ij} = \phi_j(b, t_i)$, $i = 1, 2, \dots, m$, $j = 1, 2, \dots, p$ 。

§2.2.2 变量投影方法

Golub和Pereyra [37]早在1973年就提出了求解(2.2.4)(也即标准可分非线性最小二乘问题(2.2.5)中 $y(b) = y_0$ 的情形)的显示变量投影方法。下面我们来介绍显示变量投影方法的主要思想。对任何一个固定的 $b \in \mathbf{R}^q$, (2.2.5)就成为一个线性最小二乘问题, 并且我们可以得到其最小范数解:

$$\hat{a}(b) = \Phi^+(b)y_0, \quad (2.2.6)$$

这里算子 $\Phi^+(b)$ 表示 $\Phi(b)$ 的Moore-Penrose广义逆。我们将(2.2.6)代入(2.2.4), 就可以得到:

$$\begin{aligned}\min_{a \in \mathbf{R}^p, b \in \mathbf{R}^q} \psi(a, b) &= \min_{b \in \mathbf{R}^q} \frac{1}{2} \|y_0 - \Phi(b)\Phi^+(b)y_0\|_2^2 \\ &= \min_{b \in \mathbf{R}^q} \frac{1}{2} \|P_{\Phi(b)}^\perp y_0\|_2^2.\end{aligned}\quad (2.2.7)$$

这里算子 $P_{\Phi(b)}^\perp$ 是从 \mathbf{R}^m 到 $\Phi(b)^T$ 的零空间的正交投影算子。这样我们就得到了一个降维的非线性最小二乘问题(2.2.7)。为了求解这个问题, 还必须知道(2.2.7)中目标函数的Jacobian矩阵的计算方法。在Golub和Pereyra的工作 [37]中, 给出了正交投影算子 $P_{\Phi(b)}^\perp$ 的Fréchet广义导数的计算方法:

$$D(P_{\Phi(b)}^\perp) = -P_{\Phi(b)}^\perp \Phi'(b)\Phi^+(b) - (P_{\Phi(b)}^\perp \Phi'(b)\Phi^+(b))^T. \quad (2.2.8)$$

根据(2.2.8)我们就可以得到(2.2.7)中目标函数的Jacobian矩阵, 并根据前一节介绍的求解非线性最小二乘问题的各种算法来求解(2.2.7)。我们假设求得了(2.2.7)的一个最优解 \hat{b} , 根据(2.2.6)的定义, 我们可以相应地得到 $\hat{a}(\hat{b})$, 于是

$$\begin{pmatrix} \hat{a}(\hat{b}) \\ \hat{b} \end{pmatrix} \quad (2.2.9)$$

是(2.2.4)的一个最优解。根据不同的算法求解(2.2.7), 我们就可以得到了不同的显式变量投影算法。

[37]还证明了当 $\Phi(b)$ 的秩恒定时, 若 \hat{b} 是(2.2.7)的一个局部最优解, 则(2.2.9)是(2.2.4)的一个局部最优解; 若 \hat{b} 是(2.2.7)的一个全局最优解, 则(2.2.9)是(2.2.4)的一个全局最优解。

显式变量投影方法的主要思想在于显式消去线性部分变量, 这使得基于变量投影的算法与直接求解(2.2.4)的相应算法相比, 具有三大优势: 更少的迭代步数; 更少的初始点猜测; 当原问题是病态时, 可以降低病态程度。

Kaufman [39]将Golub和Pereyra提出的显式变量投影算法中正交投影算子的Jacobian矩阵简化为

$$D(P_{\Phi(b)}^\perp) = -P_{\Phi(b)}^\perp \Phi'(b)\Phi^+(b); \quad (2.2.10)$$

并论证了, 使用基于(2.2.10)构造的Jacobian矩阵的显式变量投影算法在每步迭代中都可以节省25%的计算量。

Ruhe和Wedin在1980年的工作 [113]中通过推广变量投影到一般的非线性无约束优化, 得到了可分非线性最小二乘问题(2.2.5)的隐式变量投影方法。下面我们介绍隐式变量投影方法的主要思想。

首先考虑极小化问题:

$$\min_{x \in \mathbf{R}^n} \psi(x). \quad (2.2.11)$$

这里 $\psi(x)$ 为一般的非线性函数, 且

$$x = \begin{pmatrix} y \\ z \end{pmatrix}, \quad y \in \mathbf{R}^p, \quad z \in \mathbf{R}^q, \quad p + q = n. \quad (2.2.12)$$

假设当 y 固定时,

$$\min_{y \in \mathbf{R}^p} \psi(y, z) \quad (2.2.13)$$

的最优解比较容易求得。显然可分非线性最小二乘问题是满足这个假设的。我们记 $y(z)$ 为(2.2.13)的最优解, 我们知道 $(y(\hat{z}), \hat{z})$ 是(2.2.5)的解, 当且仅当 \hat{z} 是

$$\min_{z \in \mathbf{R}^q} \nu(z) = \psi(y(z), z) \quad (2.2.14)$$

的解。隐式变量投影方法并不是直接求解(2.2.14)。而是在每个迭代步, 将 y 看作是关于 z 的隐函数, 来计算(2.2.14)的一阶信息项(导数)和二阶信息项($\nu(z)$ 的Hessian矩阵)的近似。

我们可以把 ν 看作是目标函数 ψ 在如下流形上的限制:

$$\mathfrak{S} = \left\{ \begin{pmatrix} y \\ z \end{pmatrix}, y = y(z) \text{ 是(2.2.13)的解} \right\}. \quad (2.2.15)$$

引入如下坐标变换:

$$y = w + y(z), \quad (2.2.16)$$

于是流形 \mathfrak{S} 变为:

$$\mathfrak{S}_2 = \left\{ \begin{pmatrix} w \\ z \end{pmatrix}, w = 0 \right\}. \quad (2.2.17)$$

为了得到目标函数(2.2.14)的导数, 我们先来计算坐标变换(2.2.16)的Jacobian矩阵。容易求得:

$$\frac{\partial(w, z)}{\partial(y, z)} = \begin{pmatrix} \frac{\partial w}{\partial y} & \frac{\partial w}{\partial z} \\ \frac{\partial z}{\partial y} & \frac{\partial z}{\partial z} \end{pmatrix} = \begin{pmatrix} I & \frac{\partial w}{\partial z} \\ 0 & I \end{pmatrix}. \quad (2.2.18)$$

由 $\psi_y = 0$, 我们得到

$$\frac{\partial w}{\partial z} = \frac{d(y - y(z))}{dz} = -\frac{dy(z)}{dz}. \quad (2.2.19)$$

又由于 $\psi_{yy}dy + \psi_{yz}dz = 0$, 知

$$\frac{dy}{dz} = -\psi_{yy}^{-1}\psi_{yz}. \quad (2.2.20)$$

将(2.2.20)代入(2.2.18), 我们可得:

$$\frac{\partial(y, z)}{\partial(w, z)} = \left(\frac{\partial(w, z)}{\partial(y, z)} \right)^{-1} = \begin{pmatrix} I & -\varphi_{yy}^{-1}\varphi_{yz} \\ 0 & I \end{pmatrix}. \quad (2.2.21)$$

若记坐标变换(2.2.16)后的目标函数 ψ 为 $\tilde{\psi}$, 根据(2.2.21)可得

$$\nabla \tilde{\psi} = \left(\frac{\partial(y, z)}{\partial(w, z)} \right)^T \nabla \psi = \begin{pmatrix} \psi_y \\ -\psi_{yz}^T \psi_{yy}^{-1} \psi_y + \psi_z \end{pmatrix}; \quad (2.2.22)$$

及

$$\nabla^2 \tilde{\psi} = \left(\frac{\partial(y, z)}{\partial(w, z)} \right)^T \nabla^2 \psi \left(\frac{\partial(y, z)}{\partial(w, z)} \right) = \begin{pmatrix} \psi_{yy} & 0 \\ 0 & \psi_{zz} - \psi_{yz}^T \psi_{yy}^{-1} \psi_{yz} \end{pmatrix}. \quad (2.2.23)$$

再次根据 $\psi_y = 0$, 我们有:

$$\nabla \nu = \psi_z; \quad (2.2.24)$$

$$\nabla^2 \nu = \psi_{zz} - \psi_{yz}^T \psi_{yy}^{-1} \psi_{yz}. \quad (2.2.25)$$

回到我们的可分非线性最小二乘问题, 也即(2.2.9)具有(2.2.5)的形式。我们设

$$\varphi(a, b) = y(b) - \Psi(b)a. \quad (2.2.26)$$

并通过直接计算得到:

$$\varphi_a(a, b) = \Psi(b), \quad \varphi_b(a, b) = y'(b) - \Psi'(b)a, \quad (2.2.27)$$

且

$$\varphi_{ab}(a, b) = \Psi'(b)^T, \quad \varphi_{ba}(a, b) = \Psi'(b), \quad \varphi_{aa}(a, b) = 0. \quad (2.2.28)$$

我们发现 φ 的二阶信息中仅有 $\varphi_{bb}(a, b)$ 一项是无法通过一阶信息得到的。

根据(2.2.25), 我们有

$$\nabla^2 \nu = (\varphi_b^T \varphi_b + \varphi_{bb}) \quad (2.2.29)$$

$$-(\varphi_a^T \varphi_b + \varphi_{ab})^T (\varphi_a^T \varphi_a + \varphi_{aa})^{-1} (\varphi_a^T \varphi_b + \varphi_{ab}); \quad (2.2.30)$$

这样用(2.2.31)近似(2.2.29), 我们就可以给出求解可分非线性最小二乘问题(2.2.5)的隐式变量投影的具体算法。

算法 2.2.1. (可分非线性最小二乘的隐式变量投影算法)

步一 给定 b_0 , $a_0 = 0$; $\epsilon > 0$ 是给定的充分小的正数; 并置 $k := 0$;

步二 计算 $\varphi^{(k)} = \varphi(a_k, b_k)$, 及 $\varphi_a^{(k)} = \Phi(b_k)$,

$$a_{k+1} = a_k - (\varphi_a^{(k)})^+ \varphi^{(k)}. \quad (2.2.31)$$

步三 计算

$$\varphi_b^{(k)} = y'(b_k) - \Phi'(b_k) a_k; \quad (2.2.32)$$

通过类Kaufman步更新 b_{k+1} :

$$b_{k+1} = b_k - (P_{\varphi_a^{(k)}}^\perp \varphi_b^{(k)})^+ [I - \varphi_a^{(k)} (\varphi_a^{(k)})^+] \varphi^{(k)} = b_k - (P_{\varphi_a^{(k)}}^\perp \varphi_b^{(k)})^+ \varphi^{(k)}, \quad (2.2.33)$$

或者通过类Golub-Pereyra步更新 b_{k+1} :

$$b_{k+1} = b_k - (P_{\varphi_a^{(k)}}^\perp \varphi_b^{(k)} + P_{\varphi_a^{(k)}} \varphi_a^{(k)})^+ \Phi'(b_k) \varphi^{(k)}. \quad (2.2.34)$$

步四 置 $k := k + 1$; 如果

$$\|\varphi_a^{(k)T} \varphi^{(k)} + \varphi_b^{(k)T} \varphi^{(k)}\| < \epsilon; \quad (2.2.35)$$

则停; 否则, 转步2。

根据 [113]的分析, 由步长(2.2.34)或步长(2.2.33)定义的隐式变量投影算法2.2.1和由Golub-Pereyra提出的(2.2.7)式或Kaufman提出的(2.2.10)式定义的显式变量投影算法有相同的渐近收敛速度, 但是每一个迭代步的计算量有所减少。[113]还证明了使用隐式变量投影算法和对原问题(2.2.5)直接使用高斯牛顿算法在每一个迭代步中都具有相同的计算量, 但是隐式变量投影算法具有更好的渐近收敛速度。

此后关于可分非线性最小二乘问题的研究进展, 都是基于变量投影思想。比如, Böckmann [9]基于有限差分技术, 提出了对降维的非线性最小二乘问题(2.2.7)

的梯度计算方法, 进而通过信赖域算法求解降维问题。Kaufman和Pereyra [40]将变量投影方法推广到约束具有和目标函数相同的可分离变量形式的带约束可分非线性最小二乘问题的情形。

然而在实际工程计算中, 可分非线性最小二乘问题不一定具有变量可分离的约束。比如界约束和球约束都在应用问题中非常常见, 但是在这些约束下, 无论显式变量投影算法或隐式变量投影算法都不容易实现。事实上, 变量投影思想是很难推广到求解约束不具有和目标函数相同的可分离变量形式的约束可分非线性最小二乘问题上去的。为了避免这个困难, 我们提出了一个变量不分离的算法框架。

§2.3 结构型变量不分离算法

§2.3.1 方法构想与框架

根据可分非线性最小二乘问题的特性, 变量投影方法消去了线性部分变量, 将原问题(2.2.5) 转化为仅关于非线性部分变量的降维函数, 然后只需用一般的非线性最小二乘算法来求解降维问题即可。无论理论分析和数值实验都表明基于某种非线性最小二乘问题的变量投影算法, 比相应的直接求解原问题的算法收敛要快 [37, 38]。一种比较易于理解的解释是, 传统的变量不分离算法框架, 忽视了可分非线性最小二乘问题的特性。

但是当我们对原问题(2.2.5)加一些约束后, 变量投影方法就失效了。因为无论是显式还是隐式的变量投影算法都在很大程度上依赖于当非线性部分变量固定时, 线性部分变量显式最优解的存在性。

考虑到传统变量不分离框架效率不高, 及变量投影方法的兼容性太差, 我们期望设计一种变量不分离的算法框架, 既能够充分利用可分非线性最小二乘问题的特性来提高算法的效率; 又能够很好地兼容一般约束问题。

我们按照(2.2.31)定义 φ , 则原问题(2.2.5)的目标函数 $\psi(a, b)$ 的Hessian矩阵可以写为:

$$\nabla^2\psi = \begin{pmatrix} \varphi_a^T\varphi_a + \varphi_{aa}^T\varphi & \varphi_a^T\varphi_b + \varphi_{ab}^T\varphi \\ \varphi_b^T\varphi_a + \varphi_{ba}^T\varphi & \varphi_b^T\varphi_b + \varphi_{bb}^T\varphi \end{pmatrix}. \quad (2.3.1)$$

根据问题的特殊结构, 即由(2.2.34)式, 我们可以得到:

$$\nabla^2\psi = \begin{pmatrix} \varphi_a^T\varphi_a & \varphi_a^T\varphi_b + \Phi'(b)\varphi \\ \varphi_b^T\varphi_a + (\Phi'(b)\varphi)^T & \varphi_b^T\varphi_b + \varphi_{bb}^T\varphi \end{pmatrix}. \quad (2.3.2)$$

由于牛顿方法具有二阶收敛速度,所以如果我们需要设计(2.2.5)的高效算法,就应该去寻找目标函数Hessian矩阵(2.3.2)的好的近似。在一般的变量不分离框架下,高斯牛顿法是通过 φ 的一阶信息来近似Hessian阵的:

$$H^{GN} = \begin{pmatrix} \varphi_a^T \varphi_a & \varphi_a^T \varphi_b \\ \varphi_b^T \varphi_a & \varphi_b^T \varphi_b \end{pmatrix}. \quad (2.3.3)$$

通过比较(2.3.3)与(2.3.2),我们发现 $\varphi_{ab} = \Phi'(b)\varphi$ 这个二阶信息项其实通过一阶信息的计算就可以得到,而高斯牛顿法(2.3.3)却对之弃之不用。而我们当然希望通过使用这个特殊的二阶信息项来构造一个更好的Hessian近似。比较直观的做法是用矩阵:

$$H^{(0)} = \begin{pmatrix} \varphi_a^T \varphi_a & \varphi_a^T \varphi_b + \Phi'(b)\varphi \\ \varphi_b^T \varphi_a + (\Phi'(b)\varphi)^T & \varphi_b^T \varphi_b \end{pmatrix}, \quad (2.3.4)$$

来近似Hessian矩阵(2.3.2)。然而,这个简单的方法并不可行,因为即使目标函数(2.2.5)的真实的Hessian矩阵 $\nabla^2\psi$ 是正定的,我们也不能保证简易近似(2.3.4)中的 $H^{(0)}$ 是正定的,所以以 $H^{(0)}$ 设计算法,将增加算法产生非下降方向的可能。事实上我们在后一节中的分析中可以证明用 $H^{(0)}$ 作为Hessian矩阵(2.3.2)近似的算法在大多数情况下还不如用高斯牛顿矩阵(2.3.3)。

下面我们就要想法来修正(2.3.4)以使得修正后的Hessian矩阵的近似阵满足正定性。首先我们令

$$\varphi_C = (\varphi_a^+)^T \Phi'(b)\varphi, \quad (2.3.5)$$

并用 $\varphi_a^T(\varphi_b + \varphi_C)$ 来代替 $\varphi_a^T \varphi_b + \Phi'(b)\varphi$ 。

如果我们用 $(\varphi_b + \varphi_C)^T(\varphi_b + \varphi_C)$ 来代替(2.3.4)中矩阵 $H^{(0)}$ 右下角的 $\varphi_b^T \varphi_b$,我们就可以得到如下的近似矩阵:

$$H^{(1)} = \begin{pmatrix} \varphi_a^T \varphi_a & \varphi_a^T(\varphi_b + \varphi_C) \\ (\varphi_b + \varphi_C)^T \varphi_a & (\varphi_b + \varphi_C)^T(\varphi_b + \varphi_C) \end{pmatrix}. \quad (2.3.6)$$

这样定义的矩阵 $H^{(1)}$ 总是具有半正定性的,并且和高斯牛顿矩阵(2.3.3)相比,它具有更多的二阶信息。若 $\Phi(b)$ 具有恒秩 p ,那么我们就有:

$$\varphi_a^T \varphi_b + \Phi'(b)\varphi = \varphi_a^T(\varphi_b + \varphi_C) \quad (2.3.7)$$

恒成立。

在这种情况下，我们有：

$$H^{(1)} = \nabla^2 \psi + \begin{pmatrix} 0 & 0 \\ 0 & \varphi_b^T \varphi_C + \varphi_C^T \varphi_b + \varphi_C^T \varphi_C - \varphi_{bb}^T \varphi \end{pmatrix}. \quad (2.3.8)$$

且误差矩阵 $H^{(1)} - \nabla^2 \psi$ 的秩为 q 。下一节我们将证明(2.3.6)定义的矩阵 $H^{(1)}$ 和高斯牛顿阵 H^{GN} 相比，确实是目标函数Hessian矩阵 $\nabla^2 \psi$ 的一个更好的近似。如果考虑在近似矩阵 $H^{(1)}$ 的右下方再加上一项 $\varphi_C^T \varphi_C$ ，我们就可以得到：

$$H^{(2)} = \begin{pmatrix} \varphi_a^T \varphi_a & \varphi_a^T (\varphi_b + \varphi_c) \\ (\varphi_b + \varphi_c)^T \varphi_a & (\varphi_b + \varphi_c)^T (\varphi_b + \varphi_c) + \varphi_C^T \varphi_C \end{pmatrix}. \quad (2.3.9)$$

我们也将证明 $H^{(2)}$ 是 $\nabla^2 \psi$ 的另一个好的近似。我们将基于(2.3.6)定义的近似矩阵 $H^{(1)}$ 和(2.3.9)定义的近似矩阵 $H^{(2)}$ 的算法称为可分非线性最小二乘问题的结构型变量不分离算法。

§2.3.2 算法描述

采用(2.3.6)与(2.3.9)定义的近似Hessian矩阵 $H^{(1)}$ 和 $H^{(2)}$ ，结合线搜索技巧、Levenberg-Marquardt技巧和信赖域策略，我们设计了三种结构型变量不分离算法。下面我们给出这三种算法的具体描述。

算法 2.3.1. (结构型变量不分离的线搜索算法)

步一 给定初始点 $x_0 = (a_0^T, b_0^T)^T$ ；给定 $\rho \in (0, 0.5)$ 及充分小的正数 $\epsilon > 0$ ；

置 $k := 0$ 。

步二 若 $\|\nabla \psi(x_k)\| < \epsilon$ ，则停止。

步三 根据式(2.3.6)或(2.3.9)计算 H_k ，令

$$d_k = -(H_k)^{-1} \nabla \psi(x_k). \quad (2.3.10)$$

步四 搜索 $\alpha > 0$ 满足：

$$\psi(x_k + \alpha d_k) \leq \psi(x_k) + \rho \alpha \nabla \psi(x_k)^T d_k, \quad (2.3.11)$$

$$\psi(x_k + \alpha d_k) \geq \psi(x_k) + (1 - \rho) \alpha \nabla \psi(x_k)^T d_k. \quad (2.3.12)$$

步五 更新 $x_{k+1} = x_k + \alpha d_k$ ；置 $k := k + 1$ ；转步二。

通常来说, 线搜索型算法在近似的Hessian矩阵 H_k 接近奇异时会产生数值困难。为了克服这个问题, 我们可以结合Levenberg-Marquardt技巧或信赖域策略来实现结构型变量不分离算法。

算法 2.3.2. (结构型变量不分离的Levenberg-Marquardt算法)

步一 给定初始点 x_0 ; 给定常数 $c_1 > 1$, $c_2 > 1$, $0 < p_1 < p_2 < 1$ 和充分小的正数 $\epsilon > 0$; 初始正则化参数 $\lambda_0 = \|\psi(x_0)\|$; 置 $k := 0$ 。

步二 若 $\|\nabla\psi(x_k)\| < \epsilon$, 则停止。

步三 根据式(2.3.6)或(2.3.9)计算 H_k , 令:

$$d_k = -(H_k + \lambda_k I)^{-1} \nabla\psi(x_k). \quad (2.3.13)$$

步四 计算实际下降量与预估下降量的比值:

$$r_k = \frac{\psi(x_k) - \psi(x_k + d_k)}{-\nabla\psi(x_k)^T d_k - \frac{1}{2} d_k^T H_k d_k}, \quad (2.3.14)$$

并以之更新正则化参数 λ_{k+1} 和下一个迭代点 x_{k+1} :

$$\lambda_{k+1} = \begin{cases} c_1 \lambda_k & \text{如果 } r_k \leq p_1, \\ \frac{1}{c_2} \lambda_k & \text{如果 } r_k \geq p_2, \\ \lambda_k & \text{否则;} \end{cases} \quad (2.3.15)$$

$$x_{k+1} = \begin{cases} x_k + d_k & \text{如果 } r_k > 0, \\ x_k & \text{否则.} \end{cases} \quad (2.3.16)$$

步五 置 $k := k + 1$, 转步二。

算法 2.3.3. (结构型变量不分离的信赖域算法)

步一 给定初始点 x_0 ; 给定初始信赖域半径 $\Delta_0 > 0$; 给定常数 $c_1 > 1$, $c_2 > 1$, $0 < p_1 < p_2 < 1$; 和充分小的正常数 $\epsilon > 0$; 置 $k := 0$ 。

步二 若 $\|\nabla\psi(x_k)\| < \epsilon$, 则停止。

步三 根据式(2.3.6)或(2.3.9)计算 H_k ; 并求解信赖域子问题:

$$\min_{d \in \mathbb{R}^n} \quad m_k(d) = \nabla\psi(x_k)^T d + \frac{1}{2} d^T H_k d, \quad (2.3.17)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta_k. \quad (2.3.18)$$

得到信赖域试探步 d_k 。

步四 计算实际下降量与预估下降量的比值:

$$r_k = \frac{\psi(x_k) - \psi(x_k + d_k)}{m(0) - m(d_k)}; \quad (2.3.19)$$

并以之更新信赖域半径 Δ_{k+1} 和下一个迭代点 x_{k+1} :

$$\Delta_{k+1} = \begin{cases} \frac{1}{c_1} \|d_k\|_2 & \text{如果 } r_k \leq p_1, \\ \max[c_2 \|d_k\|_2, \Delta_k] & \text{如果 } r_k \geq p_2, \\ \Delta_k & \text{否则;} \end{cases} \quad (2.3.20)$$

$$x_{k+1} = \begin{cases} x_k + d_k & \text{如果 } r_k > 0, \\ x_k & \text{否则.} \end{cases} \quad (2.3.21)$$

步五 置 $k := k + 1$, 转步二。

在算法2.3.3的步三中, 信赖域试探步 d_k 既可以是根据 [70]来精确计算信赖域子问题(2.3.17)-(2.3.18)所得到的精确解; 也可以是根据截断共轭梯度法 [125, 120] 来近似求解信赖域子问题(2.3.17)-(2.3.18)所得到的近似解。Yuan([143])证明了根据截断共轭梯度法求得的信赖域子问题近似解 d_k 将使得目标函数的下降量至少达到精确解所对应的下降量的一半。这个优美的结论使我们可以放心的使用截断共轭梯度法来求解信赖域子问题, 以提高算法的效率。我们还可以使用Powell在NEWUOA [108]中提出的序列二维子空间极小方法作为后处理技巧来修正截断共轭梯度法所得到的信赖域子问题的近似解。

§2.4 新算法的收敛性分析

本节中, 我们将研究我们提出的结构型变量不分离算法的收敛性。由于我们的算法2.3.1、算法2.3.2、算法2.3.3和求解非线性最小二乘问题的阻尼高斯牛顿法、Levenberg-Marquardt方法和Levenberg-Marquardt方法的信赖域实现的区别仅仅在近似Hessian矩阵的构造上, 而这并不影响我们把后三种算法的全局收敛性推广到我们的三种结构型变量不分离算法上。因此, 这一节中, 我们将仅讨论我们算法的局部收敛性质。首先我们将证明采用(2.3.6)与(2.3.9)定义的近似Hessian矩阵 $H^{(1)}$ 和 $H^{(2)}$ 都是目标函数Hessian矩阵 $\nabla^2\psi$ 的精确逼近, 并以之得到局部收敛性质。

§2.4.1 局部收敛速度

首先, 类似高斯牛顿法的局部收敛性分析, 我们的算法对于零残量的问题也具有局部二阶收敛速度。不失一般性, 下面给出我们的算法(2.3.1)的局部二阶收敛性的证明。

定理 2.4.1. 记 x^* 是(2.2.5)的局部极小点, 且满足 $\varphi(x^*) = 0$; 若假设由(2.3.6)和(2.3.9)给出的矩阵 $H^{(i)}(x^*) (i = 1, 2)$ 是正定的, 且 $\nabla^2\psi(x)$ 和 $H^{(i)}(x)^{-1}$ 在 x^* 的邻域内Lipschitz连续; 则存在 $\epsilon > 0$, 对任何 $x_k \in N(x^*, \epsilon)$, 都有:

$$\|x_k + d_k - x^*\| = O(\|x_k - x^*\|^2) \quad (2.4.1)$$

成立。

证明: 不失一般性, 我们仅考虑 H_k 是由(2.3.6)定义的情况。记 $S_k = \nabla^2\psi(x_k) - H_k$, 通过 φ_C 的定义(2.3.8)、(2.3.5)和已知条件, 我们可以推得:

$$\|S_k\| = O(\|\varphi(x_k)\|) = O(\|x_k - x^*\|). \quad (2.4.2)$$

记

$$d_k^N = -\nabla^2\psi(x_k)\nabla\psi(x_k) \quad (2.4.3)$$

为牛顿步。我们熟知牛顿步满足:

$$x_k + d_k^N - x^* = O(\|x_k - x^*\|^2), \quad (2.4.4)$$

于是可推得:

$$\|d_k^N\| = O(\|x_k - x^*\|). \quad (2.4.5)$$

进而有:

$$\begin{aligned} \|d_k - d_k^N\| &= \|(\nabla^2\psi(x_k)^{-1} - H_k^{-1})\nabla\psi(x_k)\| \\ &= \|(I - H_k^{-1}\nabla^2\psi(x_k))d_k^N\| \\ &\leq \|I - H_k^{-1}\nabla^2\psi(x_k)\|\|d_k^N\| \\ &= \|H_k^{-1}(H_k - \nabla^2\psi(x_k))\|\|d_k^N\| \\ &\leq \|H_k^{-1}\|\|S_k\|\|d_k^N\| = O(\|x_k - x^*\|^2). \end{aligned} \quad (2.4.6)$$

最后, 根据(2.4.4)和(2.4.6), 就可推得我们的结论(2.4.1)。证毕。 \square

§2.4.2 渐近收敛速度

前面我们通过分析, 知道我们的结构型变量不分离算法的局部收敛速度与传统的高斯牛顿法一致。现在我们来分析我们的结构型变量不分离算法、由Golub-Pereyra提出的(2.2.7)式及由Kaufman提出的(2.2.10)式定义的显式变量投影算法, 和传统的高斯牛顿方法的渐近收敛速度。这里我们的分析思想类似于Ruhe和Wedin [113]中提出的分析方法。[113]是第一篇分析变量投影方法和传统的高斯牛顿法的渐近收敛速度的工作, 同时第一次提出了变量投影方法优于传统的高斯牛顿法的理论依据。

定义 2.4.1. 对于迭代算法

$$x_{k+1} = h(x_k), \quad \lim_{k \rightarrow \infty} x_k = x^*, \quad (2.4.7)$$

我们定义渐近收敛速度为:

$$R = -\ln \rho(h'(x^*)). \quad (2.4.8)$$

一般来说, 我们知道算法迭代算子在解处的Jacobian矩阵 $h'(x^*)$ 的谱半径越小, 方法收敛地越快。这里

$$\rho(h'(x^*)) = \max\{-\alpha, \beta\}, \quad (2.4.9)$$

其中 α 和 β 分别为如下广义特征值问题的最小和最大特征值。

$$(H - \nabla^2 \psi)z = \lambda H z. \quad (2.4.10)$$

由于(2.4.10)的所有特征值都是如下Rayleigh商问题的稳定点:

$$\mu(z) = \frac{z^T (H - \nabla^2 \psi) z}{z^T H z}, \quad (2.4.11)$$

所以要估计算法迭代算子在解处的Jacobian矩阵 $h'(x^*)$ 的谱半径, 等价于估计(2.4.11)的极点 $\mu(z)$ 。

Ruhe和Wedin [113]的理论分析论证了如下引理是成立的:

引理 2.4.1. 我们设 μ^{GN} 、 $\mu^{(G.&P.)}(z)$ 和 $\mu^{(K.)}$ 分别为由(2.4.11)式中Rayleigh商的极点定义的 $\mu(z)$, 其中所用算法分别为传统变量不分离的高斯牛顿算法、由Golub-Pereyra提出的(2.2.7)式及由Kaufman提出的(2.2.10)式定义的显式变量投影算法。则 μ^{GN} 、

$\mu^{(G.&P.)}(z)$ 和 $\mu^{(K.)}$ 应分别由如下函数的极点定义:

$$\mu^{GN}(z) = \{r(\mu, z) = q(z) - \mu + (\mu + \mu^{-1} - 2)c(z) = 0 \text{ 的解}\}; \quad (2.4.12)$$

$$\mu^{(G.&P.)}(z) = q(z); \quad (2.4.13)$$

$$\mu^{(K.)} = \frac{q(z) - c(z)}{1 - c(z)}. \quad (2.4.14)$$

这里函数 q 和 c 分别定义如下:

$$q(z) = -\|\varphi\|_2 \lambda_{\max}(K_2); \quad (2.4.15)$$

$$c(z) = \frac{\|\varphi_a^+ \Phi'(b) \varphi z\|_2^2}{\|P_{\varphi_a}^+ \varphi_b z\|_2^2 + \|\varphi_a^+ \Phi'(b) \varphi z\|_2^2}. \quad (2.4.16)$$

其中 K_2 是算子 φ 限制在由(2.2.17)定义的流形 \mathcal{S} 上的曲率。

根据引理2.4.1, 我们有如下命题成立:

命题 2.4.1. 当目标函数的Hessian近似阵 H 分别取 H^M 时, 记 α^M 和 β^M 分别为广义特征值问题(2.4.10)的最小和最大特征值。这里 M 分别代表传统变量不分离的高斯牛顿算法、由Golub-Pereyra提出的(2.2.7)式及由Kaufman提出的(2.2.10)式定义的显式变量投影算法。则我们有如下结论成立:

$$\alpha^{GN} \leq \alpha^{(K.)} \leq \alpha^{(G.&P.)}, \quad \beta^{(K.)} \leq \beta^{(G.&P.)} \leq \beta^{GN}. \quad (2.4.17)$$

根据2.4.1, 我们可以发现传统变量不分离的高斯牛顿算法的算法迭代算子的谱半径总是比变量投影方法的要大。因此这个命题也解释了为什么传统变量不分离的高斯牛顿算法往往是这三种方法中收敛最慢的。

下面我们使用引理2.4.1的结论, 来证明我们设计的结构型变量不分离算法的渐近收敛性:

定理 2.4.2. 当目标函数的Hessian近似阵 H 分别取 $H^{(i)}$ ($i = 1, 2$)时, 记 $\alpha^{(i)}$ 和 $\beta^{(i)}$ 分别为广义特征值问题(2.4.10)的最小和最大特征值。这里 $H^{(1)}$ 和 $H^{(2)}$ 分别由(2.3.6)与(2.3.9)定义。则我们有如下结论成立:

$$\alpha^{(1)} = \alpha^{(K.)}, \quad \beta^{(1)} = \beta^{(K.)}; \quad \alpha^{(2)} = \alpha^{(G.&P.)}, \quad \beta^{(2)} = \beta^{(G.&P.)}. \quad (2.4.18)$$

证明: 首先考虑 $H = H^{(2)}$ 的情况。由于 $\mu^{(2)}$ 是广义特征值问题 $(H^{(2)} - \nabla^2\psi)z = \lambda H^{(2)}z$ 的极值, 故而它应为如下方程的一个零点:

$$\det[H^{(2)} - \nabla^2\psi - \mu^{(2)}H^{(2)}] = 0. \quad (2.4.19)$$

假设 X_{11} 非奇异, 可得:

$$\det \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix} = \det[X_{11}] \cdot \det[X_{22} - X_{21}X_{11}^{-1}X_{12}]. \quad (2.4.20)$$

于是

$$\begin{aligned} & \det[H^{(2)} - \nabla^2\psi - \mu^{(2)}H^{(2)}] \\ &= \det \begin{bmatrix} -\mu^{(2)}\varphi_a^T\varphi_a & -\mu^{(2)}\varphi_a^T\varphi_b + \varphi_{ab}^T\varphi \\ -\mu^{(2)}\varphi_b^T\varphi_a + \varphi_{ba}^T\varphi & (1 - \mu^{(2)})(\varphi_b^T\varphi_C + \varphi_C^T\varphi_b + 2\varphi_C^T\varphi_C) - \varphi_{bb}^T\varphi - \mu^{(2)}\varphi_b^T\varphi_b \end{bmatrix} \\ &= (-\mu^{(2)})^{p+q} \det[\varphi_a^T\varphi_a] \det[(H^{(G.&P.)} - \psi''_{\mathfrak{S}}) - \mu^{(2)}H^{(G.&P.)}]. \end{aligned} \quad (2.4.21)$$

这里

$$H^{(G.&P.)} = \varphi_b^T P_{\varphi_a}^\perp \varphi_b + (\varphi_a + \Phi'(b)\varphi)^T P_{\varphi_a} \varphi_a + \Phi'(b)\varphi, \quad (2.4.22)$$

是由Golub-Pereyra提出的(2.2.7)式定义的显式投影梯度算法的迭代矩阵, 而

$$\psi''_{\mathfrak{S}} = \varphi_{bb} - \varphi_{ba}\varphi_{aa}^{-1}\varphi_{ab}, \quad (2.4.23)$$

是目标函数 ψ 限制在由(2.2.17)定义的流形 \mathfrak{S} 上的Hessian矩阵。最后一个等式成立的依据是 $\varphi_C = \varphi_a^T(\varphi_b + \Phi'(b)\varphi)$ 。于是, 由(2.4.21), 可知

$$H^{(2)} - \nabla^2\psi - \mu^{(2)}H^{(2)} \quad (2.4.24)$$

的任何一个非零特征值都应为

$$\det[(H^{(G.\&P.)} - \psi''_{\mathfrak{S}} - \mu^{(2)}H^{(G.\&P.)})] \quad (2.4.25)$$

的一个零点, 这就证明了 $\mu^{(2)}(z)$ 和 $\mu^{(G.\&P.)}(z)$ 具有相同的极值。

以上证明可以类似地推广到由(2.3.6)定义的 $H = H^{(1)}$ 和由Kaufman提出的(2.2.10)式定义的显式投影梯度算法的迭代矩阵 $H^{(K.)} = \varphi_b^T P_{\varphi_a}^\perp \varphi_b$ 的情形。证毕。□

根据定理2.4.2, 我们知道采用我们设计的由(2.3.6)与(2.3.9)定义的近似Hessian矩阵 $H^{(2)}$ 和 $H^{(1)}$, 根据算法2.3.1所设计的结构型变量不分离算法分别和由Golub-Pereyra [37]提出的(2.2.7)式及由Kaufman [39]提出的(2.2.10)式定义的显式变量投影算法具有完全相同的渐近收敛速度。

类似地, 我们还可以估计由(2.3.4)定义的Hessian阵的近似矩阵 $H = H^{(0)}$ 作为迭代矩阵的算法的渐近收敛速度。当目标函数的Hessian近似阵 H 取 $H^{(0)}$ 时, 我们设 $\alpha^{(0)}$ 和 $\beta^{(0)}$ 分别为广义特征值问题(2.4.10)的最小和最大特征值。则 $\alpha^{(0)}$ 和 $\beta^{(0)}$ 应分别为如下函数的最小值和最大值:

$$\mu(z) = \frac{p(z)}{1 + p(z) - q(z)}, \quad (2.4.26)$$

这里

$$p(z) = z^T \varphi_{bb}^T \varphi z / (\|P_{\varphi_a}^\perp \varphi_b z\|_2^2 + \|\varphi_C z\|_2^2), \quad (2.4.27)$$

$q(z)$ 是当迭代矩阵为 $H = H^{(G.\&P.)}$ 时, 由(2.4.10)定义的Rayleigh商。(2.4.26)显示, 当非线性部分变量的二阶信息 $\varphi_{bb}^T \varphi$ 充分小的时候, 使用 $H^{(0)}$ 可以得到较快的渐近收敛速。但是当 $\varphi_{bb}^T \varphi$ 和 $\varphi_b^T \varphi_C$ 相比不可忽略时, 那么使用 $H^{(0)}$ 就会得到很坏的渐近收敛速度。

§2.5 数值结果

这一节中, 我们将报告我们提出的结构型变量不分离算法的数值表现。参与和我们算法2.3.1、算法2.3.2、算法2.3.3比较的为求解非线性最小二乘问题的一般

阻尼高斯牛顿法、Levenberg-Marquardt方法, 及隐式变量投影算法2.2.1。测试问题来自Golub和Pereyra [37] 及Ruhe和Wedin [113]。

本节数值结果的测试环境是Matlab 6.5; 使用CPU为1.60GHz, 内存504MB的HP Compaq nx6120型计算机。

问题一: 指数数据拟合问题

$$\phi(a, b; t) = a_1 + a_2 e^{-b_1 t} + a_3 e^{-b_2 t}, \quad (2.5.1)$$

这里观测项个数为 $m = 33$; 最优解为 $a_1^* = 0.37531$, $a_2^* = 1.9305$, $a_3^* = -1.4592$, $b_1^* = 0.012867$, $b_2^* = 0.022123$ 。

测试问题的构造根据如下规则:

- $t_i = i$ 时, $y_i = (1 + nf \cdot \text{randn}(1))\phi(a^*, b^*; t_i)$, 这里 nf 是观测噪音;
- 初始点取为 $a_i = (1 + rd \cdot (2\text{rand}(1) - 1))a_i^*$, $i = 1, 2, 3$; $b_j = (1 + rd \cdot (2\text{rand}(1) - 1))b_j^*$, $j = 1, 2$, 这里 rd 是衡量初始点到最优解远近的参数。

问题二: 分数数据拟合

$$\begin{aligned} \phi(a, b; t) = & a_1 + a_2 t + a_3 t^2 \\ & - a_4 \left[\frac{1}{1 + ((b_1 + 0.5b_2 - t)/b_3)^2} + \frac{1}{1 + ((b_1 - 0.5b_2 - t)/b_3)^2} \right] \\ & - a_5 \left[\frac{1}{1 + ((b_4 + 0.5b_5 - t)/b_6)^2} + \frac{1}{1 + ((b_4 - 0.5b_5 - t)/b_6)^2} \right] \\ & - a_6 \left[\frac{1}{1 + ((b_7 - t)/b_8)^2} \right], \end{aligned} \quad (2.5.2)$$

这里观测项个数为 $m = 188$; 最优解为 $a_1^* = 1.0$, $a_2^* = 0.2$, $a_3^* = 0.1$, $a_4^* = 0.9$, $a_5^* = 0.7$, $a_6^* = 0.3$, $b_1^* = 0.2$, $b_2^* = 0.8$, $b_3^* = 3.0$, $b_4^* = 0.3$, $b_5^* = 0.7$, $b_6^* = 3.0$, $b_7^* = 0.5$, $b_8^* = 2.0$ 。

测试问题的构造根据如下规则:

- $t_i = i/m$, $y_i = (1 + nf \cdot \text{randn}(1))\phi(a^*, b^*; t_i)$, 这里 nf 是观测噪音;
- 初始点取为 $a_i = (1 + rd \cdot (2\text{rand}(1) - 1))a_i^*$, $i = 1, 2, \dots, 6$; $b_j = (1 + rd \cdot (2\text{rand}(1) - 1))b_j^*$, $j = 1, 2, \dots, 8$, 这里 rd 是衡量初始点到最优解远近的参数。

我们分别对两个问题在无噪音($nf = 0$)和有噪音($nf = 0.1$)两种情况下作测试。对所有的测试问题, $\epsilon = 10^{-10}$, 近似Hessian矩阵由(2.3.6)定义。

表2.5.1中列出的结果是在无观测噪音情况下, 由传统的变量不分离的高斯牛顿算法、采用(2.2.33)定义的隐式变量投影算法和我们的算法2.3.1($\rho = 0.01$) 求解测试问题的迭代步数和运行时间(秒)。

问题	高斯牛顿算法	变量投影方法	算法2.3.1
P1, $rd = 0.01$	6/0.0160	4/0.0140	4/0.0160
P2, $rd = 0.01$	16/0.7540	9/0.4400	10/0.4970

表2.5.1 无观测噪音的测试结果

通过表2.5.1, 我们可以发现, 当初始点靠近解时, 使用算法2.3.1和隐式变量投影算法的表现较为一致, 迭代步数都少于传统的变量不分离的高斯牛顿法。这和我们的渐近收敛性分析吻合。

我们也测试了带有观测噪音的问题。比较的是传统的变量不分离的Levenberg-Marquadt方法、在(2.2.33)定义的迭代步中加入Levenberg-Marquadt修正的隐式变量投影方法, 以及我们的算法2.3.2和算法2.3.3。在我们的算法2.3.2和算法2.3.3中, 我们使用如下的参数选取方法: $c_1 = c_2 = 2$, $p_1 = 0.1$ 及 $p_2 = 0.4$ 。数值结果列在表2.5.2中。

问题	L-M方法	变量投影方法	算法2.3.2	算法2.3.3
P1, $rd = 0.1$	19/0.0310	13/0.0250	14/0.0340	17/0.0410
P1, $rd = 0.2$	21/0.0340	14/0.0270	16/0.0370	18/0.0430
P1, $rd = 0.5$	25/0.0410	16/0.0310	19/0.0430	20/0.0470
P2, $rd = 0.1$	162/0.7660	142/0.7410	154/1.0160	62/0.4370
P2, $rd = 0.2$	169/0.8030	144/0.7880	160/1.0790	72/0.5000
P2, $rd = 0.5$	195/0.9540	158/0.9370	182/1.2910	76/0.5310

表2.5.2 观测噪音为 $nf = 0.1$ 时的测试结果

根据上面的数值结果, 可以发现我们的算法2.3.2的表现非常类似于变量投影方法, 且都比传统的Levenberg-Marquadt方法要好。而我们的算法2.3.3对问题二, 比变量投影方法好; 但是对问题一, 表现略微不尽如人意。

第三章 二次最小二乘问题的全局优化算法

二次最小二乘问题是具有如下形式的非线性最小二乘问题:

$$\min_{x \in \mathbf{R}^n} f(x) = \sum_{i=1}^m f_i^2(x) = \sum_{i=1}^m \left(\frac{1}{2} x^T A_i x + g_i^T x + c_i \right)^2, \quad (3.0.1)$$

这里 $A_i \in \mathbf{R}^{n \times n}$, $g_i \in \mathbf{R}^n$, $c_i \in \mathbf{R}$ 。在计算生物学、通讯工程、计算机图形学、化学工程等工程领域应用甚广的超定二次系统(方程组):

$$\begin{cases} \frac{1}{2} x^T A_1 x + g_1^T x + c_1 = 0; \\ \frac{1}{2} x^T A_2 x + g_2^T x + c_2 = 0; \\ \dots\dots\dots \\ \frac{1}{2} x^T A_m x + g_m^T x + c_m = 0, \end{cases} \quad (3.0.2)$$

的求解问题即可转化为(3.0.1)。

尽管二次最小二乘问题是形式最简单的非线性最小二乘问题之一,但由于其具有非凸性,求解其全局最优解仍然是非常困难的。容易证明二次最小二乘问题是NP-难问题。于是我们的工作就是希望设计一种高效的算法可以稳定地求得(3.0.1)全局极小点的很好的近似。

本章首先对全局优化进行简要概述;然后讨论对于二次最小二乘问题的一些已有的进展;最后给出我们设计的随机子空间方法及其数值表现。

§3.1 全局优化简述

考虑一般的优化问题:

$$\min_{x \in \mathbf{R}^n} f(x) \quad (3.1.1)$$

$$\text{s.t. } x \in \Omega, \quad (3.1.2)$$

其中 Ω 是 \mathbf{R}^n 中的一个子集。

我们可以定义:

定义 3.1.1. 对于一个可行点 $\bar{x} \in \Omega$, 如果存在一个正数 δ , 满足:

$$f(\bar{x}) \leq f(x), \quad \forall x \in \Omega \cap \{x \mid |x - \bar{x}| \leq \delta\}. \quad (3.1.3)$$

那么我们就称 \bar{x} 为(3.1.1)-(3.1.2)的局部极小点。

定义 3.1.2. 对于一个可行点 $x^* \in \Omega$, 如果:

$$f(x^*) \leq f(x), \forall x \in \Omega. \quad (3.1.4)$$

那么我们就称 x^* 为(3.0.1)-(3.1.2)的全局极小点。

全局优化就是求解一个优化问题(3.0.1)-(3.1.2)的目标函数在约束限定的可行域范围(Ω)内的全局极小点(定义3.1.2中的 x^*)。而一般我们所熟知的优化算法, 一般来说都是求解一个优化问题(3.0.1)的目标函数在约束限定的可行域范围内的局部极小点(定义3.1.1中的 \bar{x})。

我们知道对于局部优化问题, 有局部最优性条件, 根据局部最优性条件, 我们可以设计迭代算法使得局部最优性条件在有限步迭代后, 在可以容忍的误差范围内被满足。

然而对于全局优化问题一般没有全局最优性条件。也就是说, 在一般情况下, 当我们找到一个优化问题的局部极小点时, 我们往往不知道它是否是问题的全局极小点。然而一般情况下我们很难甚至不可能求得问题的所有局部极小点, 即使求得问题的所有局部极小点, 当它的个数是变量个数的指数函数时, 那么比较这些局部极小点上目标函数值的算法就已经是指数时间复杂度的了。

关于全局优化的困难性, 有个著名的论断, 见 [47]。下面我们简要地给出这个著名的论断:

定理 3.1.1. 给定一个优化问题(3.0.1)-(3.1.2), 如果仅仅知道目标函数 f 是 k 次连续可微的(k 是给定的正整数或 $+\infty$), 那么通过有限次的函数值估计, 我们无法给出目标函数值的一个下界。

此论断的证明并不困难, 假设论断不成立, 通过有限次观测得到了下界 L 。我们知道有限次的函数值估计所包含点集 \mathcal{P} 的测度为零, 因此必存在一个点 $y^* \in \Omega$ 及一个充分小的正数 $\epsilon > 0$, 使得 y^* 的邻域 $\mathcal{U} := \{x \mid |x - y^*| \leq \epsilon\}$ 满足 $\mathcal{U} \cap \mathcal{P} = \emptyset$ 。因此我们可以构造一个新的函数 \bar{f} , 使之满足:

$$\bar{f}(y^*) = L - 1; \quad (3.1.5)$$

$$\bar{f}(x) = f(x), \quad \forall x \notin \mathcal{U}; \quad (3.1.6)$$

$$\bar{f} \text{ 是 } k \text{ 次连续可微的。} \quad (3.1.7)$$

(3.1.7)通过 \bar{f} 在 \mathcal{U} 内的定义达到。因此仅通过 \mathcal{P} 集中的函数值的观测, 无法得到 f 的一个下界。

事实上,也存在在可行域内只有一个局部极小点的问题,比如说凸规划问题,此时求解问题的全局极小和局部极小是等价的。但是在绝大多数情况下,全局优化比局部优化要困难得多。在很多应用领域,比如计算生物学、通讯工程、芯片设计、数据挖掘、交通问题中的优化问题往往具有很多的局部极小点,要求全局极小几乎是不可能的。

全局优化的算法分为两大类:确定性算法和不确定性算法。可参阅 [47], [78]等。

确定性算法,顾名思义就是通过算法的有限次迭代,可以精确求得问题的全局极小(在误差范围内)。根据前面的介绍,很自然地,我们知道确定性算法只适用于特殊类型的问题。比如说对于Lipschitz连续的目标函数,有Lipschitz全局优化算法;对于在任意可行区域内容易估计上下界的目标函数,有分枝定界算法;对于非凸二次规划、D.C.规划问题,有割平面算法;其它确定性算法还有同伦算法、填充(隧道)函数算法等。

不确定性算法,就是通过算法的有限次迭代,只能求得问题在一定意义下的近似解。不确定性算法主要分两类:一类是随机算法;一类是近似算法。

随机算法有随机多初始算法、随机函数算法、模拟退火算法、遗传算法、神经网络算法、蚁群杂交算法等。随机算法的特点是:算法是根据某种启发式的经验设计的;而且一般来说可以在概率意义下证明,当迭代步数趋于无穷时,算法将以概率1收敛于问题的全局极小。

近似算法一般通过某种松弛技巧,将非凸优化问题转化为凸优化问题,进行求解。一般来说,这种转化导致的最优值的差别,也就是近似比,是可以估计的。这种估计可以保证在最坏情况下,近似算法得到的近似解在一定意义下是可以接受的。比较成熟的松弛技巧有线性松弛、半定规划松弛等。

确定性算法除了只能针对特殊的全局优化问题外,还有一些其它的不足之处:算法效率不高;算法复杂度很高;甚至有些只是理论算法,无法付诸实施等等。因此对于实际工程应用问题,人们往往更倾向于采用不确定性算法。

§3.2 多项式优化与距离几何问题

这一节我们主要介绍一些和我们的二次最小二乘问题相关的两类问题:多项式优化问题和距离几何问题。二次最小二乘问题是多项式优化问题的一类特殊情况;而距离几何模型又是二次最小二乘问题的一种特例。

§3.2.1 多项式优化问题

本小节我们考虑如下多项式优化问题:

$$\min_{x \in \mathbf{R}^n} f(x), \quad (3.2.1)$$

这里 $f(x)$ 是次数为 $2d$ (d 为正整数)的多项式。我们假设(3.2.1)具有全局极小点,且全局极小点到原点的距离是有限的。考虑如下两类问题:

$$f(x_1, x_2) = x_1^2 - x_2^2; \quad (3.2.2)$$

$$f(x_1, x_2) = x_1^2 + (x_1x_2 - 1)^2. \quad (3.2.3)$$

(3.2.2)不具有全局极小点; (3.2.3)全局极小在无穷远处,都不在考虑范围内。

多项式优化问题近年来是比较热门的研究课题,很多著名的学者已经提出了不少成熟的确定性或非确定性算法来求解多项式优化问题的全局极小点。参见 [121], [79], [81]等。

多项式优化的理论基础基于代数几何学([16], [8])的结论。我们考虑由 f 的 n 个偏导数构成的理想:

$$I = \left\langle \frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right\rangle \in R[x_1, x_2, \dots, x_n] := R[x]. \quad (3.2.4)$$

定义 3.2.1. 定义理想 I 在 n 维实(复)空间中的零点集合为 $\nu_{\mathbf{R}}(I)$ ($\nu_{\mathbf{C}}(I)$)。

我们知道目标函数 $f(x)$ 的所有稳定点都对应于理想 I 在 n 维实空间 \mathbf{R}^n 中的零点。根据 [16], [8]中的结论,理想 I 在 n 维复空间 \mathbf{C}^n 中零点的个数等于其在 \mathbf{R} 上剩余环的维数:

$$|\nu_{\mathbf{C}}(I)| = \dim_{\mathbf{R}} \mathbf{R}[x]/I. \quad (3.2.5)$$

如果我们假设 $\nu_{\mathbf{C}}(I)$ 是有限的,则根据Bézout's定理,我们有:

$$|\nu_{\mathbf{C}}(I)| = (2d - 1)^n. \quad (3.2.6)$$

大多数求解(3.2.1)的代数几何算法都是不遗余力地要找出零点集合 $\nu_{\mathbf{C}}(I)$,有的甚至采用枚举的办法。我们感兴趣的其实是 $\nu_{\mathbf{R}}(I)$,尽管通常来说

$$|\nu_{\mathbf{R}}(I)| \leq |\nu_{\mathbf{C}}(I)|, \quad (3.2.7)$$

但 $\nu_{\mathbf{R}}(I)$ 的结构要更为复杂,可参阅 [54, 55]。

多项式优化的主要技巧有：代数几何方法；同伦方法；SOS(平方和)松弛化方法等。

代数几何方法主要有利用Gröbner基的方法、结式方法等([18])。其主要思想都是利用代数几何的理论将求理想的零点问题等价转化为求一个一维高次多项式的根的问题，然后通过求解一维多项式的根，约简原来的理想，并重复这个过程。

同伦算法(可参阅 [53], [54], [55])要求问题具有一定特性，即可以引入一个参数满足：当参数取1时等价于原问题，当参数取0时是一个非常易于求解的问题，然后通过一条由0到1的合适的路径，慢慢逼近原问题，求得原问题的所有零点。

SOS方法是一种近似方法，由Shor [116]于1983年提出，它是一种近似方法。它的主要思想是找到最大的 λ 使得 $f(x) - \lambda$ 可以被表示为多项式平方和的形式，我们记这样的 λ 为 f^{SOS} 。显然 f^{SOS} 是 $f(x)$ 的一个下界。

记 X 为每个元素都是由 $\{x_1, x_2, \dots, x_n\}$ 组成的次数最多为 d 的单项式的列向量。于是 X 的长度为 $N = C_{n+d}^d$ 。记 L_f 为所有满足

$$f(x) = X^T \cdot A \cdot X, \quad (3.2.8)$$

的实对称矩阵 A 组成的集合。 L_f 是 $N \times N$ 维矩阵空间中的一个仿射空间。

我们设 X 的一个元素是1。记 E_{11} 为仅有左上角元素为1，其它元素均为0的矩阵，于是我们有：

定理 3.2.1. 对任何实数 λ ，下面两个命题是等价的：

- 多项式 $f(x) - \lambda$ 能表示成实系数多项式平方和的形式；
- 存在矩阵 $A \in L_f$ 使得 $A - \lambda E_{11}$ 是半正定的。

于是我们可以通过求解如下半定规划问题来求 f^{SOS} ：

$$\max_{\lambda \in \mathbf{R}} \quad \lambda, \quad (3.2.9)$$

$$\text{s.t.} \quad A - \lambda E_{11} \succeq 0, \quad (3.2.10)$$

$$A \in L_f. \quad (3.2.11)$$

SOS方法的问题在于存在不能表示为多项式平方和的恒正多项式，最简单的例子是Motzkin's多项式：

$$f(x, y) = x^4 y^2 + x^2 y^4 - 3x^2 y^2 \geq 0, \text{ 但是 } f^{SOS} = -\infty. \quad (3.2.12)$$

这就是说下界 f^{SOS} 和 f 的全局极小值 f^* 之间的差距可以很大。

著名的希尔伯特第十七问题(Hilbert's 17th problem)给出的结论更是揭示了这种问题是普遍存在的(可参阅 [110,111]):

定理 3.2.2 (David Hilbert 1888). 总存在不能表示为多项式平方和形式的 n 维 $2d$ 次非负多项式, 除非 $n = 2$, 或者 $d = 1$, 或者 $n = 3, d = 2$.

还有许多工作([75])是围绕着怎样结合代数几何知识, 得到更好的SOS下界。

通常来说上面介绍的多项式优化方法效率都很低, 除了最后介绍的SOS方法外, 其它方法都只能求解规模很小的问题, 基本上都属于理论算法。而且它们对类似于 $x^2 = 0$ 和 $x^2 + \epsilon = 0$ ($\epsilon > 0$)这样的小扰动很敏感(前者有解, 后者无解)。而SOS方法, 正如前面的分析, 并不对所有问题都有效; 其次算法效率也不是很高, 我们后面还将介绍 [75]中最先进的SOS方法求解实际问题的效率。

Qi [109]提出了一种求解正定的四次多项式全局极小的方法。这种方法也只能求解很小规模的问题。不过在本章下一节中我们提出的一个算法的设计思想可以看作是对 [109]部分工作的推广。

§3.2.2 距离几何模型

我们先来介绍无线传感器网络定位问题(Sensor Network Localization Problem)。

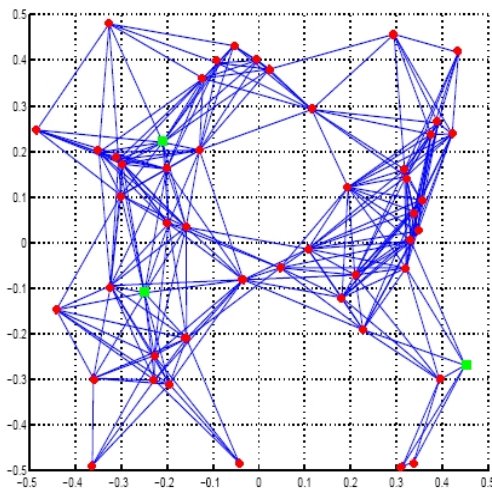


图 3.1

如图3.1在一块方形区域内有 m 个给定的点, 称为基站(绿色), 它们的坐标位

置是确定的；还有 n 个位置不确定的点，称为无线传感器(红色)，它们的坐标位置是待确定的。无线传感技术可以测量出所有小于接收半径 rd 的无线传感器与基站间的或者两个无线传感器间的距离(蓝色线段)。现在问题就是如何通过这些距离和给定的基站坐标，来确定所有无线传感器的坐标。

刻画上述问题的欧几里得距离几何模型(Euclidean Distance Geometry Problem)如下：

$$\|x_i - x_j\|^2 = d_{ij}, \forall (i, j) \in N_x, i < j; \quad (3.2.13)$$

$$\|a_k - x_j\|^2 = s_{kj}, \forall (k, j) \in N_a. \quad (3.2.14)$$

这里集合 $N_x = \{ (i, j) \mid d_{ij} \leq rd \}$ ；集合 $N_a = \{ (k, j) \mid s_{kj} \leq rd \}$ ； d_{ij} ， s_{kj} ， rd ，和 $a_k \in \mathbf{R}^2$ ， $(k = 1, \dots, l)$ 是距离几何问题(3.2.13)-(3.2.14)的参数； $x_j \in \mathbf{R}^2$ ， $(j = 1, \dots, n)$ 是变量。

这等价于如下的二次最小二乘问题：

$$\min_{x_i \in \mathbf{R}^{2 \times n}} \sum_{(i, j) \in N_x, i < j} (\|x_i - x_j\|^2 - d_{ij}^2)^2 + \sum_{(k, j) \in N_a} (\|a_k - x_j\|^2 - s_{kj}^2)^2. \quad (3.2.15)$$

对于无线传感器网络定位问题，Ye的研究小组提出了半定规划松弛化方法[122]，[132]。下面我们介绍半定规划松弛化方法的主要思想。

记传感器坐标为 $2 \times n$ 维矩阵 $X = [x_1, x_2, \dots, x_n]$ 。于是：

$$\|x_i - x_j\|^2 = e_{ij}^T X^T X e_{ij}; \quad (3.2.16)$$

$$\|a_k - x_j\|^2 = (a_k; e_j)^T [I \ X]^T [I \ X] (a_k; e_j). \quad (3.2.17)$$

这里 e_{ij} 为第 i 个元素为1，第 j 个元素为-1，其它元素为0的向量； e_j 为第 j 个元素为1，其它元素为0的向量。我们有：

$$e_{ij}^T Y e_{ij} = d_{ij}^2, \forall (i, j) \in N_x, i < j; \quad (3.2.18)$$

$$(a_k; e_j)^T \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} (a_k; e_j) = d_{kj}^2, \forall (k, j) \in N_a; \quad (3.2.19)$$

$$Y = X^T X. \quad (3.2.20)$$

我们用凸约束

$$Y \succeq X^T X \quad (3.2.21)$$

来松弛非凸约束

$$Y = X^T X. \quad (3.2.22)$$

(3.2.21)等价于

$$Z := \begin{pmatrix} I & X \\ X^T & Y \end{pmatrix} \succeq 0. \quad (3.2.23)$$

于是原问题(3.2.13)-(3.2.14)就转化为松弛的半定规划问题:

$$Z_{1:2,1:2} = I \quad (3.2.24)$$

$$(0; e_{ij})(0; e_{ij})^T \bullet Z = d_{ij}^2, \forall (i, j) \in N_x, i < j; \quad (3.2.25)$$

$$(a_k; e_j)(a_k; e_j)^T \bullet Z = d_{kj}^2, \forall (k, j) \in N_a; \quad (3.2.26)$$

$$Z \in \mathbf{R}^{(n+2) \times (n+2)} \succeq 0, \quad (3.2.27)$$

这里

$$A \bullet B = \text{tr}(A^T B). \quad (3.2.28)$$

我们后面还将介绍半定规划松弛化方法求解无线传感器网络定位问题的数值表现。

在计算生物学中,确定蛋白质分子空间结构是一项很重要的课题。迄今为

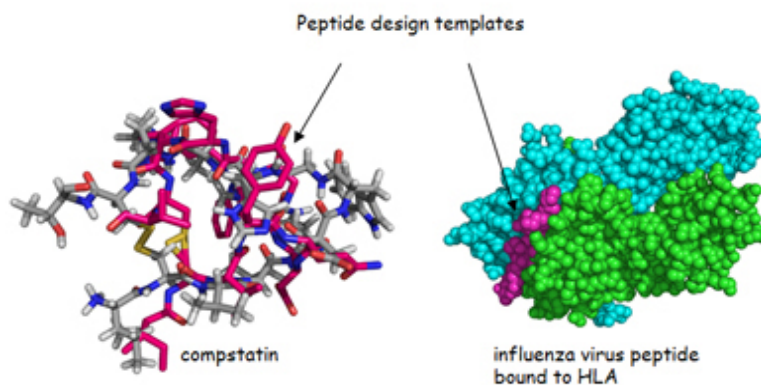


图 3.2

止,确定蛋白质分子构型的手段主要有三种: X光结晶学实验、核磁共振实验和能量动力学方程计算。如果采用核磁共振的实验手段,可以测得蛋白质部分分子之间的距离。于是下面的工作就是根据这些距离数据求解蛋白质各原子的空间坐标,于是就归结为如下的欧几里得距离几何问题:

$$\|x_i - x_j\|_2^2 = r_{ij}^2, \forall (i, j) \in \mathcal{S}; \quad (3.2.29)$$

这里 $\mathcal{S} \subset \{(i, j) \mid 1 \leq i < j \leq n\}$ (如: 可取 \mathcal{S} 为 $\{(i, j) \mid r_{ij} \leq md\}$); r_{ij} 是核磁共振实验的观测数据; $x_i \in \mathbf{R}^3$, ($i = 1, \dots, n$) 是变量。当然在蛋白质分子空间构型问题中, 我们先要固定一些原子的位置, 来消除平移和旋转不变性造成的冗余解。

Moré和Wu [74]提出的整体光滑化方法(Global Continuation Method)是求解距离几何问题(3.2.29)的基础性工作。下面简单介绍整体光滑化算法的思想。

距离几何模型(3.2.29)等价于如下的二次最小二乘问题:

$$\min_{x \in \mathbf{R}^{3 \times n}} f(x) = \sum_{(i,j) \in \mathcal{S}} (\|x_i - x_j\|^2 - d_{ij}^2)^2, \quad (3.2.30)$$

引入高斯变换作为光滑化算子:

$$\langle f \rangle_\lambda(x) = \int_{\mathbf{R}^n} f(y) \frac{1}{(2\pi)^{n/2} \lambda^n} e^{-\frac{\|y-x\|^2}{\lambda^2}} dy; \quad (3.2.31)$$

光滑化算子(3.2.31)可以磨平多峰函数, 当 λ 足够大时, $\langle f \rangle_\lambda(x)$ 将成为凸函数。所以整体光滑化算法的思想, 就是取一个足够大的 λ , 求得光滑化问题的唯一极小点; 然后适当减小 λ , 以该极小点为初始点, 求解新的光滑化函数的局部极小点; 依次类推, 最后当 λ 取零时, 就等价于原问题。整体光滑化算法和同伦算法在思想上是有类似之处的。

当 f 的具体形式如距离几何问题(3.2.30)中定义时, 光滑化算子(3.2.31)有如下算术表达式:

$$\langle f \rangle_\lambda(x) = \sum_{(i,j) \in \mathcal{S}} ((\|x_i - x_j\|^2 - \delta_{ij}^2)^2 + 10\lambda^2 \|x_i - x_j\|^2) + \gamma. \quad (3.2.32)$$

下面给出整体光滑化算法的步骤:

算法 3.2.1. (整体光滑化算法)

步一 设定初始光滑化参数 λ_0 ; 光滑次数 $GCNum$ 和多初始次数 $OpMax$; 最优点 x^{record} , 最优函数值 f^{record} 。

步二 置 $t := 1$, 随机选定初始点 x_{in} , 置 $k := 1$ 。

步三 若 $k = 1$, 置 $x_0 := x_{in}$; 否则 $x_0 = x_*^{(k-1)}$ 。

步四 以 x_0 为初始点, 采用某种局部极小算法求解

$$\min_{x \in \mathbf{R}^{3 \times n}} \langle f \rangle_{\lambda_k}(x), \quad (3.2.33)$$

并得到局部极小点 $x_*^{(k)}$ 。这里(3.2.33)中的 $\langle f \rangle_{\lambda_k}(x)$ 由(3.2.32)定义。

步五 若 $k \leq GCNum$, 则置 $k := k + 1$, 置

$$\lambda_k := \left(1 - \frac{k}{GCNum}\right)\lambda_0, \quad (3.2.34)$$

并转步三; 否则转步六。

步六 若 $f(x_*^{(k)}) < f^{record}$, 则更新 $x^{record} := x_*^{(k)}$, $f^{record} := f(x_*^{(k)})$ 。

步七 若 $t < OpMax$, 则置 $k := 0$, $t := t + 1$, 随机选定初始点 x_{in} , 转步三; 否则, 算法停止, 返回 x^{record} 。

整体光滑化算法有一个明显的不足之处。它的关键思想是希望通过一系列光滑化函数逐渐逼近原函数, 以前一个的局部极小作为后一个的初始点。最理想的情形是这条局部极小路径指向全局最优的吸引域, 这自然就要求第一个局部极小应包含某些全局极小的信息。最直观的想法是如果第一个光滑化函数是凸函数, 自然它的局部极小也是它的全局极小。我们发现当 $\lambda > \frac{1}{\sqrt{5}}\delta$ 时, $h(r) = (r^2 - \delta^2)^2 + 10\lambda^2 r^2$ 是凸的递增函数。因此当

$$\lambda > \frac{1}{\sqrt{5}} \max\{\delta_{i,j} : (i,j) \in \mathcal{S}(:= \lambda^C)\} \quad (3.2.35)$$

时, 光滑化函数 $\langle f \rangle_{\lambda_k}(x)$ 是凸的。但是此时, (3.2.33)的解为 $x_i = x_j, \forall 1 \leq i < j \leq n$, 显然这个解和问题的信息 $\delta_{i,j}$ 毫无关系, 因此这个凸的光滑函数不能反映出原函数的全局信息。

事实上 [74]也提到 λ_0 的取法很关键, 首要地, 应当有 $\lambda_0 \ll \lambda^C$ 。数值结果也印证了这个观点。然而即使 λ_0 很小, 我们也难以求得第一个光滑化函数的好的局部极小点。但从我们后面将汇报的数值结果上来看, 整体光滑算法对三维的距离几何问题(3.2.30)还是很有效的。

§3.3 随机二维子空间方法

本节我们介绍我们设计的随机二维子空间方法。首先我们给出一个一般的随机多初始方法的框架; 然后我们给出二次最小二乘问题全局下降方向的判定方法、讨论全局下降方向在二维子空间中的计算方法及全局下降二维子空间的概念和判定准则, 并以之构造我们的随机二维子空间算法。

§3.3.1 随机多初始算法

求解全局优化问题, 使用最广泛的是随机多初始算法。随机多初始算法的主要步骤是: 在问题的可行域内采集随机初始点(或点集); 根据某些特定的规则,

来决定是否以该点(或点集的一个子集)为初始点进行局部优化,以得到新的局部极小点;评判新的局部极小点,更新最优点记录;最后在若干次随机采点后,以当前最优点作为全局极小的近似。

一般来说只要优化问题的可行域有界,目标函数的全局极小的吸引域测度不为零,那么随机多初始算法就是以概率1收敛到问题全局极小点的。

随机采点策略、进行局部优化的条件及算法终止准则都会影响随机多初始算法的效率。

[146]中给出了一个非常有效的随机多初始算法的终止准则。下面根据这个终止准则,我们给出一个随机多初始算法的一般框架:

算法 3.3.1. (随机多初始算法框架)

初始赋值阶段

步一 采点次数上限 $SampleMax$, 卫冕次数上限 $ViMax$,
操作次数上限 $OpMax$;

步二 $Sample := 0$, $Vi := 0$, $Op := 0$; $x^{record} = 0$, $f^{record} = +\infty$ 。

随机采点阶段 在可行域内随机采点 x^{sample} ; 置 $Sample := Sample + 1$ 。

随机点评判阶段

步一 根据某特定规则,评判 x^{sample} 是否合格;若合格转步二;
否则转步三。

步二 置 $Sample := 0$; 转局部极小化阶段。

步三 如果 $Sample < SampleMax$, 转随机采点阶段;
否则置 $OpMax := OpMax + 1$, 置 $Sample := 0$, 转随机采点阶段。

局部极小化阶段 根据 x^{sample} , 选定初始点 $\overline{x^{sample}}$;

进行局部优化; 得到局部最优点 x^* ; 置 $OpMax := OpMax + 1$ 。

最优更新阶段 如果 $f(x^*) < f^{record}$, 则更新最优纪录: $x^{record} = x^*$,
 $f^{record} = f(x^*)$, 置 $Vi := 0$; 否则置 $Vi := Vi + 1$ 。

终止判断阶段 如果 $Vi < ViMax$ 且 $Op < OpMax$, 转随机采点阶段;
否则算法停止, 返回 x^{record} 。

根据 [47]和 [78]中介绍的随机采点方法,我们就可以得到如下两种最简单的随机多初始算法:

算法 3.3.2. (简单随机多初始算法I)

采用随机多初始算法框架3.3.1, 其中随机点评判阶段采用的规则是:

$$\forall x^{sample}, \text{ 通过}; \quad (3.3.1)$$

局部极小化阶段的初始点选取规则为:

$$\overline{x^{sample}} = x^{sample}. \quad (3.3.2)$$

算法 3.3.3. (简单随机多初始算法II)

采用随机多初始算法框架3.3.1, 其中随机点评判阶段采用的规则是:

$$\forall x^{sample} \in \{ x \mid f(x) < f(x^{record}) \}, \text{ 通过}; \quad (3.3.3)$$

局部极小化阶段的初始点选取规则为:

$$\overline{x^{sample}} = x^{sample}. \quad (3.3.4)$$

我们知道简单随机多初始算法I(算法3.3.2)相对而言较为保守, 因为怕错失任何一个在全局极小吸引域内的点, 所以对每个随机点都要作局部搜索。当局部搜索代价较大时, 算法效率很低; 而简单随机多初始算法II(算法3.3.3)相对而言较为激进, 仅对比最优纪录好的随机点作局部搜索, 这样算法比较容易错失好的随机点, 也会影响算法效率。

下面我们根据二次最小二乘问题(3.0.1)的特性, 通过在以上两种随机多初始算法的简单策略中间寻求一个权衡点, 构造新的随机点评价规则, 从而大大提高随机多初始算法的效率。

§3.3.2 全局下降方向的判定

首先我们考虑(3.0.1)的一维情形:

$$\min_{x \in \mathbf{R}} \varphi(x) = ax^4 + 2bx^3 + cx^2 + dx + e, \quad (a > 0); \quad (3.3.5)$$

假设 $x = 0$ 是 $\varphi(x)$ 的一个局部极小点, 我们有 $d = 0$ 。于是我们可以通过判断

$$\varphi'(x) = 2x(2ax^2 + 3bx + c) \quad (3.3.6)$$

的括号中二次多项式的判别式

$$\Delta_1 = 9b^2 - 8ac \quad (3.3.7)$$

的正负来确定(3.3.5)还有没有其它局部极小值点；我们可以通过判断

$$\varphi(x) = x^2(ax^2 + 2bx + c) + e \quad (3.3.8)$$

的括号中二次多项式的判别式

$$\Delta_2 = 4b^2 - 4ac \quad (3.3.9)$$

的正负来确定(3.3.5)还有没有比 $\varphi(0)$ 函数值更小的局部极小值点。从而我们有如下命题：

命题 3.3.1. 我们假设 $x = 0$ 是(3.3.5)的一个局部极小点，则如下命题成立：

- 1) 如果 $\Delta_1 < 0$ ，则 $x = 0$ 是 $\varphi(x)$ 唯一的稳定点；
- 2) 如果 $\Delta_1 = 0$ ，则 $x = 0$ 是 $\varphi(x)$ 唯一的极小点， $x = -\frac{3b}{4a}$ 是一个鞍点；
- 3) 如果 $\Delta_1 > 0$ ，且 $\Delta_2 < 0$ ，则 $x = 0$ 是 $\varphi(x)$ 的全局极小点， $x = -\frac{3b + \text{sign}(b)\sqrt{\Delta_1}}{4a}$ 是另一个局部极小点；
- 4) 如果 $\Delta_2 = 0$ ，则 $x = 0$ 和 $x = -\frac{b}{a}$ 是 $\varphi(x)$ 的两个函数值相等的局部极小点；
- 5) 如果 $\Delta_2 > 0$ ，则 $x = 0$ 是 $\varphi(x)$ 的局部极小点， $x = -\frac{3b + \text{sign}(b)\sqrt{\Delta_1}}{4a}$ 是 $\varphi(x)$ 的全局极小点。

证明： $\varphi(x)$ 的所有稳定点都是 $\varphi'(x) = 0$ 的根，显然除了 $x = 0$ 外还有没有其它稳定点，取决于判别式 Δ_1 是否大于等于零。如果判别式等于零，则还有一个鞍点；如果判别式大于零，则 $2ax^2 + 3bx + c = 0$ 距 $x = 0$ 远的根是另一个局部极小。

$\varphi(x)$ 有没有比 $x = 0$ 处函数值更小的点，被 $ax^2 + 2bx + c$ 的正负决定。如果判别式 Δ_2 小于零，则 $ax^2 + 2bx + c$ 恒正，于是 $\varphi(0)$ 是全局极小值；如果判别式 Δ_2 等于零，则还有一个等值的局部极小点；如果判别式 Δ_2 大于零，则在 $ax^2 + 2bx + c$ 的两根之间，都是函数值比 $x = 0$ 更小的点，此时另一局部极小点是 $\varphi(x)$ 的全局极小点。证毕。 □

对于多维问题，我们定义全局极小方向如下：

定义 3.3.1. 任意给定一点 \bar{x} ，如果存在 $\lambda \neq 0$ 满足

$$f(\bar{x} + \lambda \bar{d}) < f(\bar{x}), \quad (3.3.10)$$

我们称方向 \bar{d} 为 $f(x)$ 在点 \bar{x} 处的一个全局下降方向。

假设我们已经通过局部算法求得了问题的一个局部极小点 x_0 , 并给定一个方向 d , 我们如何判断它是否是目标函数在当前局部极小点 x_0 处的全局下降方向呢?

我们考虑(3.0.1)中 $f(x)$ 在 x_0 处沿 d 方向的展开:

$$\psi_d(\alpha) = f(x_0 + \alpha d) = P(d)\alpha^4 + 2R(d)\alpha^3 + S(d)\alpha^2 + f(x_0); \quad (3.3.11)$$

这里

$$\begin{cases} P(d) = \sum_{i=1}^m (\frac{1}{2}d^T A_i d)^2; \\ R(d) = \sum_{i=1}^m (\frac{1}{2}d^T A_i d)(A_i x_0 + g_i)^T d; \\ S(d) = \sum_{i=1}^m [2(\frac{1}{2}d^T A_i d)f_i + ((A_i x_0 + g_i)^T d)^2]. \end{cases} \quad (3.3.12)$$

考虑到 x_0 是 $f(x)$ 的一个局部极小点, 因此 $\alpha = 0$ 是 $\psi_d(\alpha)$ 的局部极小值点。根据命题3.3.1, 我们有:

命题 3.3.2. 如(3.3.11)定义的 $\psi_d(\alpha)$ 具有另一个局部极小点, 当且仅当:

$$\Gamma_{\psi}^{(1)}(d) = 8P(d)S(d) - 9R(d)^2 < 0; \quad (3.3.13)$$

$\psi_d(\alpha)$ 具有一个函数值比 $\psi_d(0)$ 更小的局部极小点, 当且仅当:

$$\Gamma_{\psi}^{(2)}(d) = P(d)S(d) - R(d)^2 < 0. \quad (3.3.14)$$

证明: 命题3.3.1的直接推论。□

由命题3.3.2, 我们可以得到如下全局下降方向的判定定理:

定理 3.3.1. 假设 x_0 是 $f(x)$ 的一个局部极小点。如果存在 d 满足(3.3.14), 则 d 是 $f(x)$ 在 x_0 处的一个全局下降方向; 且以

$$x^{in}(d) = x_0 - \frac{3R(d) + \text{sign}(R(d))\sqrt{9R(d)^2 - 8P(d)S(d)}}{4P(d)}, \quad (3.3.15)$$

为初始点, 采用某种单调的局部优化算法求解(3.0.1), 可以得到 $f(x)$ 的一个局部极小点 x^* , 该局部极小点必满足 $f(x^*) < f(x_0)$;

如果

$$\Gamma_{\psi}^{(2)}(d) \geq 0, \quad \forall d \in \mathbf{R}^n, \quad (3.3.16)$$

则 x_0 是 $f(x)$ 的全局极小点。

证明: 定理的前半部分可由命题3.3.2直接推出。下面证明后半部分。我们用反证法, 假设 x_0 不是 $f(x)$ 的局部极小点, 则必存在 \bar{x} 满足 $f(\bar{x}) < f(x_0)$, 于是我们令 $\bar{d} = \bar{x} - x_0$, 根据命题3.3.2, 我们有 $\Gamma_\psi^{(2)}(\bar{d}) < 0$, 这与(3.3.16)矛盾。因此假设不成立, 定理得证。□

剩下的问题是在一个局部极小点 x_0 处, 我们怎么才能求得全局下降方向或者判断出 x_0 是问题的全局极小呢? 我们考虑如下的六次齐次多项式优化问题:

$$\min_{d \in \mathbf{R}^n} \quad \Gamma_\psi^{(2)}(d) = P(d)S(d) - R(d)^2, \quad (3.3.17)$$

$$\text{s.t.} \quad \|d\|_2 = 1. \quad (3.3.18)$$

这里球约束(3.3.18)包含了 $d \neq 0$ 的所有方向(对于齐次多项式优化问题, 函数值大小只和方向相关, 与大小无关)。

如果我们能求得(3.3.17)-(3.3.18)的全局极小点 d^* , 则根据定理3.3.1, $\Gamma^{(2)}(d^*)$ 的正负号决定了 x_0 是否为 $f(x)$ 全局极小点, 并且如果答案是负, d^* 就是 $f(x)$ 在 x_0 处的一个全局下降方向。

§3.3.3 全局下降二维子空间的判定

尽管我们发现(3.3.17)-(3.3.18)的全局极小并不比原问题(3.0.1)的好求, 但是对于问题在二维子空间中的情形, 我们稍加转换, 就可以把它变为一个等价的一维多项式极小问题。

考虑空间中任意两个线性无关的单位向量 u 和 v , 定义:

定义 3.3.2. $\Upsilon(u, v) = \{u\} \cup \{\lambda u + v \mid \forall \lambda \in \mathbf{R}\}$.

于是在二维子空间 $\text{span}\{u, v\}$ 中, 我们可以用 $d \in \Upsilon$ 代替约束(3.3.18)来表示所有的非零方向。我们引入如下命题:

命题 3.3.3. 假设 u 和 v 是 \mathbf{R}^n 中两个线性无关的单位向量, 则:

$$\min_{d \in \text{span}\{u, v\}, \|d\|_2=1} \Gamma_\psi^{(2)}(d) = \min\{\Gamma_\psi^{(2)}(u), \min_{\lambda \in \mathbf{R}} \Gamma_\psi^{(2)}(\lambda u + v) / \|\lambda u + v\|_2^6\}. \quad (3.3.19)$$

证明: 任意 $d \in \text{span}\{u, v\}$, 都能表示为

$$d = (\lambda u + \varrho v), \quad \forall \lambda, \varrho \in \mathbf{R} \quad (3.3.20)$$

的形式, 而上式等价于

$$d = \varrho \left(\frac{\lambda}{\varrho} u + v \right), \quad \forall \lambda \in \mathbf{R}, \varrho \neq 0, \quad \text{或者} \quad d = \lambda u, \quad \forall \lambda \in \mathbf{R}. \quad (3.3.21)$$

结合约束 $\|d\|_2 = 1$, 我们有

$$\{d \mid d \in \text{span}\{u, v\}\} \cup \{d \mid \|d\|_2 = 1\} = \{u\} \cup \left\{ \frac{\lambda u + v}{\|\lambda u + v\|_2} \mid \forall \lambda \in \mathbf{R} \right\}. \quad (3.3.22)$$

由于 $\Gamma_\psi^{(2)}(d)$ 是六次齐次多项式, 并根据(3.3.22), 我们易知(3.3.19)成立。证毕。□

因为我们主要关心的是 $\Gamma_\psi^{(2)}(d)$ 全局最小值的正负性, 所以我们可以把恒正部分 $\|\lambda u + v\|_2^6$ 省去。于是通过计算 $\Gamma_\psi^{(2)}(u)$ 和求解一维六次多项式极小化问题的全局极小值

$$\min_{\lambda} \Gamma_\psi^{(2)}(\lambda u + v), \quad (3.3.23)$$

可以判断 x_0 是否存在位于二维子空间 $\text{span}\{u, v\}$ 中的全局下降方向。

定义 3.3.3. 任意给定一点 \bar{x} , 如果存在 $0 \neq \bar{d} \in \Upsilon$ 满足

$$f(\bar{x} + \bar{d}) < f(\bar{x}), \quad (3.3.24)$$

则我们称子空间 Υ 为 $f(x)$ 在点 \bar{x} 处的一个全局下降子空间。

既然二维子空间中全局下降方向是可计算的, 那么我们就可以通过计算某二维子空间是否存在下降方向, 来判定该二维子空间是否满足定义(3.3.3)。于是我们有如下的二维全局下降子空间的判定定理:

定理 3.3.2. 假设 x_0 是 $f(x)$ 的一个局部极小点。如果存在线性无关的单位向量 u 和 v 满足:

$$\Gamma_\psi^{(2)}(u) < 0; \quad (3.3.25)$$

或者

$$\min_{\lambda} \Gamma_\psi^{(2)}(\lambda u + v) < 0. \quad (3.3.26)$$

则 $\text{span}\{u, v\}$ 是 $f(x)$ 在 x_0 处的一个全局下降子空间; 且以

$$x^{in}(d) = x_0 - \frac{3R(d) + \text{sign}(R(d))\sqrt{9R(d)^2 - 8P(d)S(d)}}{4P(d)} \quad (3.3.27)$$

为初始点, 采用某种单调的局部优化算法求解(3.0.1), 可以得到 $f(x)$ 的一个局部极小点 x^* , 该局部极小点必满足 $f(x^*) < f(x_0)$; 这里:

$$d = \begin{cases} u, & \Gamma_\psi^{(2)}(u) < 0; \\ \lambda^{min} u + v, & \text{其它情形}(\lambda^{min} \text{是}(3.3.27)\text{的最优解}). \end{cases} \quad (3.3.28)$$

如果对任意两个线性无关的单位向量 u 和 v , (3.3.25)与(3.3.26)均不成立, 则 x_0 是 $f(x)$ 的全局极小点。

§3.3.4 算法构造

根据上一子小节提出的全局下降方向的判定方法和全局下降二维子空间的判定方法，我们可以改进本节开始时介绍的两种简单随机多初始算法3.3.2和算法3.3.3。

根据定理3.3.1，我们首先提出如下的随机多初始方向算法：

算法 3.3.4. (随机多初始方向算法)

采用随机多初始算法框架3.3.1，其中随机点评判阶段采用的规则是：

$$\forall x^{sample} \in \{ x \mid \Gamma_{\psi}^{(2)}(x^{sample} - x_0) < 0 \}, \text{ 通过}; \quad (3.3.29)$$

局部极小化阶段的初始点选取规则为：

$$\overline{x^{sample}} = x^{in}(x^{sample} - x_0). \quad (3.3.30)$$

这里我们通过随机点 x^{sample} ，来构造随机方向 $d = x^{sample} - x_0$ ，既是为了操作上容易实现；也是因为我们认为在对函数值没有任何先验假设的前提下，全局极小值点在某区域出现的可能性和该区域的测度有关，这可以用区域内的均匀分布随机点的密度来显现。

我们有理由相信算法3.3.4要比算法3.3.3好，这是因为如果随机点 x^{sample} 没有通过算法3.3.4的评判规则(3.3.29)的话，这代表整条直线

$$\mathcal{L} = \{x_0 + \lambda(x^{sample} - x_0) \mid \lambda \in \mathbf{R}\} \quad (3.3.31)$$

上所有的点都不满足算法3.3.3的评判规则(3.3.3)。因此前者会比后者错过更少的好随机点。

根据定理3.3.2，我们又可以提出如下随机多初始二维子空间算法：

算法 3.3.5. (随机多初始二维子空间算法)

我们对随机多初始算法框架3.3.1中随机点评判阶段和局部极小化阶段稍做修正。

随机点评判阶段首先采用如下的规则：

$$\Gamma_{\psi}^{(2)}(d_1) < 0, \text{ 通过}; \quad (3.3.32)$$

如果通过，则局部极小化阶段的初始点选取规则为：

$$\overline{x^{sample}} = x^{in}(x^{sample} - x_0). \quad (3.3.33)$$

如果不通过，增加一次评判过程：

$$\min_{d \in \text{span}\{d_1, d_2\}, \|d\|_2=1} \Gamma_{\psi}^{(2)}(d) < 0, \text{ 通过}; \quad (3.3.34)$$

如果通过, 则局部极小化阶段的初始点选取规则为:

$$\overline{x^{sample}} = x_0 + \lambda^{\min} d_1 + d_2. \quad (3.3.35)$$

如果仍不通过, 才视为未通过随机点评判阶段。

这里

$$d_1 = (x^{sample} - x_0) / \|x^{sample} - x_0\|_2; \quad (3.3.36)$$

$$d_2 = -\nabla f(x^{sample}), \quad (3.3.37)$$

或者

$$d_2 = (y^{sample} - x_0) / \|y^{sample} - x_0\|_2. \quad (3.3.38)$$

这里 y^{sample} 是另一个随机点。

类似前面的分析, 一般情形下, 算法3.3.5要比算法3.3.4好, 因为前者会比后者错过更少的好随机点。

§3.4 数值结果

本节将汇报我们的新算法求解二次最小二乘问题的数值表现。测试例子是[132]中的介绍的无线传感器网络定位问题的典型测试例子。

被测试的算法包括我们的算法3.3.4和算法3.3.5; 前面介绍过的求解距离几何问题(3.2.15)的半定规划松弛方法([132])、整体光滑化方法3.2.1; 求解多项式优化的稀疏SOS算法([75]); 以及简单随机多初始算法3.3.2和算法3.3.3。

§3.4.1 算法参数选取

问题构造是根据如下规则:

- d_{ij}^r 与 s_{kj}^r 是实际距离值; nf 是观测噪音; 也即观测距离 d_{ij} 与 s_{ij} 由下式确定:

$$d_{ij} = d_{ij}^r (1 + nf \cdot \text{randn}(1)); \quad (3.4.1)$$

$$s_{kj} = s_{kj}^r (1 + nf \cdot \text{randn}(1)). \quad (3.4.2)$$

- 反映真实值和计算值之间误差大小的均方根误差(Root Mean Square Deviance) $RMSD$ 定义如下:

$$RMSD = \left(\frac{1}{n} \sum_{i=1}^n \|x_i - \bar{x}_i\|_2^2 \right)^{\frac{1}{2}}, \quad (3.4.3)$$

这里 \bar{x}_i 表示传感器真实的坐标值, x_i 为计算解。

算法参数的选取规则如下：

- 局部搜索都采用“PRP”共轭梯度方法 [155];
- 全局算法参数: $OpMax = 50$; $ViMax = 10$; $JudgeMax = 50$ 。

这里需要说明的是,问题的界主要是提供了随机采点的区域,全局极小绝对位于界的内部。由于考虑到我们比较的是全局优化的策略,因此对于局部极小算法,我们并不加约束来降低算法的效率;而是采取简单的截断策略,只要超出边界,自然截断,终止极小化算法。

§3.4.2 数值结果I

我们比较半定规划松弛化算法(SDP)(算法软件包由Ye于2006年提供,其中求解半正定规划的软件是SeDuMi 1.05);整体光滑化算法(GC)(算法3.2.1;根据 [74],参数 $CGNum$ 取10,其它参数选取和我们的算法完全一致);以及我们的算法3.3.5。

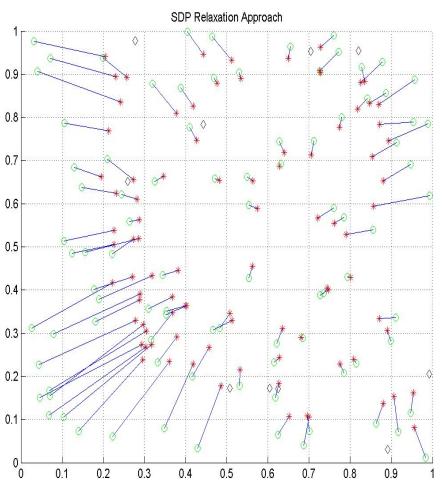
测试环境是Matlab 6.5;使用CPU为1.60GHz,内存504MB的HP Compaq nx6120型计算机。

数值结果见下表,各算法间比较的是:时间(单位:秒)/ $RMSD$ 。

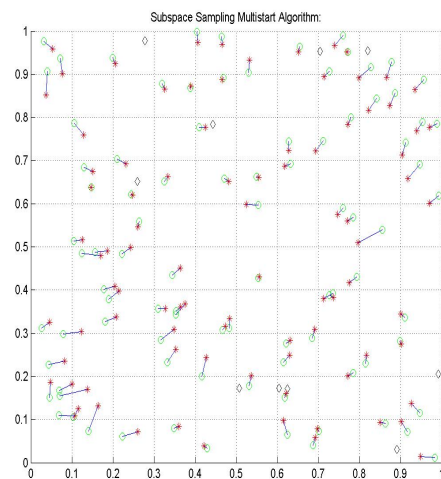
No.	n	m	nf	rd	SDP松弛算法	GC算法	算法3.3.5
1	50	3	0	0.5	3.23/2.07e-04	2.57/2.43e-07	1.28/6.01e-08
2	50	5	2%	0.4	1.99/0.0314	2.28/0.0048	0.81/ 0.0047
3	100	7	0	0.35	30.62/5.32e-04	4.47/4.49e-07	2.69/1.01e-07
4	100	10	2%	0.35	86.94/0.0147	4.79/0.0603	2.15/0.0029
5	100	10	10%	0.35	77.18/0.1111	5.38/0.0296	1.31/0.0283
6	200	10	1%	0.2	510.59/0.0128	12.46/0.0849	5.71/0.0123
7	200	10	1%	0.25	*	18.40/0.0011	8.10/0.0012
8	400	20	1%	0.2	*	33.74/7.32e-04	8.66/7.62e-04
9	500	4	0	0.3	*	176.85/8.44e-10	46.34/3.72e-09
10	1000	30	1%	0.1	*	125.62/0.0919	34.73/4.87e-04

表3.4.1 十个经典的无线传感器网络定位测试问题

观察表3.4.1,我们发现算法3.3.5总是能够在最小的时间代价内求得一个质量很好的解。相对而言,半定规划松弛化算法所需要的时间代价比较大,对于误差(nf)较大的问题数值误差比较大;整体光滑化算法在信息量($m \cdot rd$)比较少的情情况下,容易不稳定。

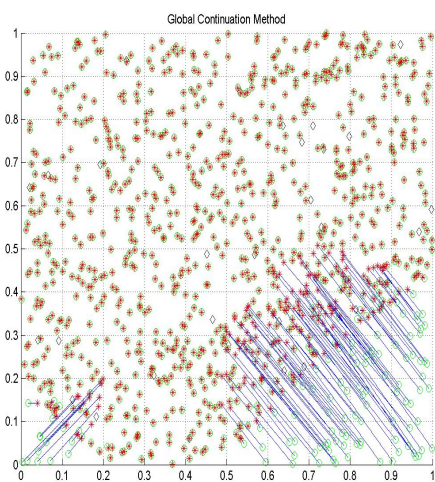


(SDP算法)

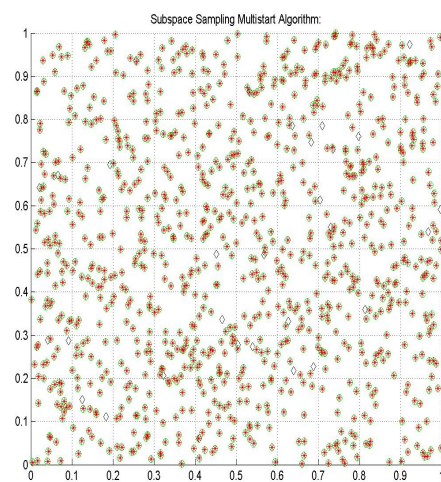


(算法3.3.5)

图 3.3



(整体光滑化算法)



(算法3.3.5)

图 3.4

下面我们给出表3.4.1中的两个最难的例子的图像结果。图3.3给出的是大误差的例子，即第5个题： $n = 100$ ， $m = 10$ ， $nf = 10\%$ ， $rd = 0.35$ 的情形。图3.4给出的是规模较大的例子，即第10题： $n = 1000$ ， $m = 30$ ， $nf = 1\%$ ， $rd = 0.1$ 的情形。

§3.4.3 数值结果II

这一节我们来关注表3.4.1中最典型的例子，即第9题(因为在各种文献中汇报得最多)： $n = 500$ ， $m = 4$ ， $nf = 0$ ， $rd = 0.3$ 的情形。参与比较的有稀疏的半正定规划松弛化算法(使用的是 [122]中的数值结果)、稀疏SOS算法 [75]、简单随机多初始算法3.3.2和算法3.3.3、我们的算法3.3.4和算法3.3.5。

方法	测试环境	时间	<i>RMSE</i>
稀疏SOS算法	0.98GB RAM, 1.46GHz CPU	85 分钟	2.9e-6
稀疏SDP松弛化算法	1.99GB RAM, 1.06GHz CPU	30 秒	1.0e-6
简单随机多初始算法I	504MB RAM, 1.60GHz CPU	78.5 秒	8.2e-9
简单随机多初始算法II	504MB RAM, 1.60GHz CPU	13.6 秒	0.1675
算法3.3.4	504MB RAM, 1.60GHz CPU	57.6 秒	4.6e-9
算法3.3.5	504MB RAM, 1.60GHz CPU	46.3 秒	3.7e-9

表3.4.2 典型测试例子9

通过表3.4.2可以看出，稀疏SOS算法虽然是求解多项式优化最有效的算法之一，但是对于特殊问题，它的效率不一定高。简单随机多初始算法II对距离几何问题而言，几乎无效。我们设计的新算法虽然理论上没法说明优于简单随机多初始算法I，但是从数值表现上来看还是具有优越性的。Ye [122]提出的稀疏的半正定规划松弛化算法利用了距离几何问题的特殊性，大大加快了半正定规划松弛化方法的计算时间。

第四章 L_1 范数极小化问题

L_1 范数极小化问题属于一类特殊的复合非光滑优化问题。虽然在很多实际的工程问题中有重要的应用价值,但是由于没有高效稳定的算法, L_1 范数极小化模型很少被采用。本章首先介绍非光滑优化问题的基本概念;然后介绍求解非线性 L_1 范数极小化问题的信赖域方法;最后给出我们对此类非光滑信赖域子问题的复杂度分析及算法。

§4.1 非光滑优化问题概述

利用 L_1 精确罚函数,可将一个带约束的非光滑优化问题在一定意义下等价地转化为一个无约束的非光滑优化问题,因此这里我们只介绍无约束的非光滑优化问题

$$\min_{x \in \mathbf{R}^n} F(x), \quad (4.1.1)$$

这里函数 $F: \mathbf{R}^n \mapsto \mathbf{R}$ 不一定可微。

根据 [11,12],我们给出如下定义:

定义 4.1.1. $F(x)$ 的次梯度集合定义如下:

$$\partial F(x) := \{d \mid F(x + \Delta x) - F(x) \geq d^T \Delta x, \forall \Delta x\}. \quad (4.1.2)$$

定义 4.1.2. $F(x)$ 关于 d 的方向导数定义如下:

$$F'(x, d) := \lim_{\alpha \rightarrow 0^+} \frac{F(x + \alpha d) - F(x)}{\alpha}. \quad (4.1.3)$$

我们也可以建立非光滑优化问题的最优性条件,可参考 [112], [157]等。

定理 4.1.1. 如果 x 是(4.1.1)的局部极小点,则:

$$F'(x, d) \geq 0, \forall d; \quad (4.1.4)$$

$$0 \in \partial F(x). \quad (4.1.5)$$

我们称(4.1.4)与(4.1.5)为(4.1.1)的一阶稳定性条件;称满足一阶稳定性条件的点为一阶稳定点。

即使 $F(x)$ 几乎处处可微,利用可微问题的方法求解来解(4.1.1),并假定在每个迭代点上 $F(x)$ 均可微,我们仍将遇到两个难点。

第一是算法的终止条件不易给出。我们知道当 x 充分靠近一连续可微函数 $f(x)$ 的极小点时, 根据一阶稳定性条件, $\|\nabla f(x)\|$ 将趋于零。所以, 光滑的无约束优化方法的终止判别条件常常是

$$\|\nabla f(x)\| \leq \varepsilon. \quad (4.1.6)$$

但对于不可微函数并没有类似的结论。例如当 $n = 1$ 时, 我们取 $F(x) = |x|$, 则对任何不是解的 x , 都有:

$$|\nabla F(x)| = 1. \quad (4.1.7)$$

[h]

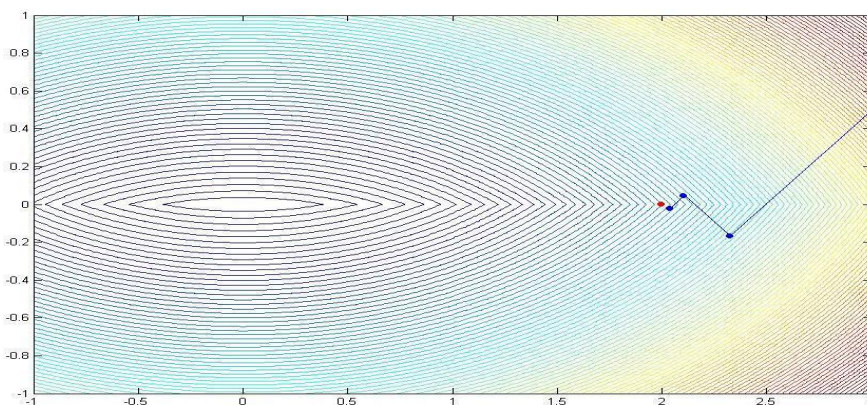


图 4.1

第二个难点是由 Wolfe(1975)指出的“折线收敛于非解”现象。当 $F(x)$ 是不可微函数时, 精确搜索下的最速下降法可能收敛于一非稳定点。例如, 设 $x = \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbf{R}^2$,

$$F(x) = \max \left\{ \frac{1}{2}u^2 + (v-1)^2, \frac{1}{2}u^2 + (v+1)^2 \right\}. \quad (4.1.8)$$

假定

$$x_k = \begin{pmatrix} 2(1 + |\varepsilon_k|) \\ \varepsilon_k \end{pmatrix}, \quad (4.1.9)$$

且 $\varepsilon_k \neq 0$, 则不难求得精确搜索下最速下降法的迭代具有如下形式:

$$x_{k+1} = \begin{pmatrix} 2(1 + |\varepsilon_k|/3) \\ -\varepsilon_k/3 \end{pmatrix}. \quad (4.1.10)$$

于是, 对于函数(4.1.8)以及给定的初始点

$$\begin{pmatrix} 2 + 2|\delta| \\ \delta \end{pmatrix}, \quad \delta \neq 0; \quad (4.1.11)$$

精确线搜索下的最速下降法将收敛于非稳定点 $x = \begin{pmatrix} 2 \\ 0 \end{pmatrix} \in \mathbf{R}^2$ 。如图4.1。

非光滑优化的一类重要方法是次梯度(subgradient)方法。这类方法最早由Shor [114]给出。假设 g_k 是(4.1.1)在迭代点 x_k 处的一个次梯度, 则次梯度法的迭代公式为:

$$x_{k+1} = x_k - \alpha_k g_k / \|g_k\|_2. \quad (4.1.12)$$

通过例子(4.1.8), 我们知道根据精确线搜索会导致(4.1.12)收敛到非稳定点。Shor [114]证明了, 对于凸函数 $F(x)$, 如果步长满足:

$$\lim_{k \rightarrow +\infty} \alpha_k = 0, \quad \sum_{k=1}^{\infty} \alpha_k = +\infty, \quad (4.1.13)$$

则可保证对 $F(x)$ 有:

$$\lim_{k \rightarrow \infty} F(x_k) = \min_{x \in \mathbf{R}^n} F(x). \quad (4.1.14)$$

即任何聚点都是极小点。但是这时, 算法的收敛速度比 R 线性收敛还要慢。关于次梯度方法的其它工作, 可参阅文献Eremin [31], Polyak [82, 83, 84], Shor [115, 116]等。

Wolfe [130]提出了共轭次梯度方法。其基本思想是将共轭梯度法推广到非光滑问题。在第 k 次迭代中, 定义一个集合

$$I_k \subset \{1, 2, \dots, k\}. \quad (4.1.15)$$

求解

$$\min_{\lambda \in \mathbf{R}^{|I_k|}} \left\| \sum_{i \in I_k} \lambda_i g_i \right\|_2^2, \quad (4.1.16)$$

$$\text{s. t.} \quad \sum_{i \in I_k} \lambda_i = 1, \quad \lambda_i \geq 0, \quad (4.1.17)$$

设 $\lambda_i^{(k)} (i \in I_k)$ 是(4.1.16)-(4.1.17)的解, 取

$$d_k = - \sum_{i \in I_k} \lambda_i^{(k)} g_i \quad (4.1.18)$$

为第 k 步迭代的线搜索方向。设 $I_{k+1} = I_k \cup \{k+1\}$, 则当 $F(x)$ 是凸的二次函数时, 精确线搜索下的共轭次梯度法实质上就是共轭梯度法。

此外, 常见的非光滑优化算法还有: 推广共轭次梯度方法得到的捆集(bundle)法(Lemaréchal [49,50,51]; Zowe [151]); 求解凸规划的割平面法(Kelley [41]; Cheney和Goldstein [17]); 本章重点介绍的信赖域方法等。

§4.2 非线性 L_1 范数极小化问题的信赖域算法

本节开始我们进入正题, 讨论如下的 L_1 范数极小化问题:

$$\min_{d \in \mathbf{R}^n} F(x) = \|f(x)\|_1 = \sum_{i=1}^m |f_i(x)|, \quad (4.2.1)$$

其中 $f: \mathbf{R}^n \mapsto \mathbf{R}^m$, $f_i: \mathbf{R}^n \mapsto \mathbf{R}$, $i = 1, \dots, m$ 是 f 的分量。这是一类特殊的复合NDO问题。

Fletcher ([33])1981年给出了一个求解复合NDO的信赖域方法。

算法 4.2.1. (求解非线性 L_1 范数极小化问题的信赖域方法)

步一 给定初始点 x_1 ; 初始Lagrange乘子 λ_0 ; 初始信赖域半径 $\Delta_1 > 0$; $\varepsilon \geq 0$;
置 $k := 1$ 。

步二 计算

$$B_k = \sum_{i=1}^m (\lambda_{k-1})_i \nabla^2 f_i(x_k). \quad (4.2.2)$$

求解如下信赖域子问题:

$$\min_{d \in \mathbf{R}^n} m_k(d) = \sum_{i=1}^m |f_i(x_k) + \nabla f_i(x_k)^T d| + \frac{1}{2} d^T B_k d; \quad (4.2.3)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta_k, \quad (4.2.4)$$

得到 d_k 。

如果 $\|d_k\| \leq \varepsilon$, 则算法停止。

步三 计算实际下降量与预估下降量之比:

$$r_k = \frac{F(x_k) - F(x_k + d_k)}{m_k(0) - m_k(d)}. \quad (4.2.5)$$

如果 $r_k < 0.25$, 则令 $\Delta_{k+1} := \|d_k\|/4$;

如果 $r_k > 0.75$, 且 $\|d_k\| = \Delta_k$, 则令 $\Delta_{k+1} = 2\Delta_k$;

如果 Δ_{k+1} 还未定义, 则令 $\Delta_{k+1} = \Delta_k$ 。

步四 如果 $r_k > 0$, 则转步五;

否则 $x_{k+1} = x_k$, $\lambda_k := \lambda_{k-1}$; 转步六。

步五 更新 $x_{k+1} = x_k + d_k$; $\lambda_k \in \partial \|f(x_k) + (\nabla f(x_k)^T)^T d_k\|_1$ 满足:

$$\nabla f(x_k)^T \lambda_k + B_k d_k + \bar{\mu}_k \mu_k = 0; \quad (4.2.6)$$

这里 $\mu_k \in \partial \|d_k\|$, $\bar{\mu}_k \geq 0$, 且

$$\bar{\mu}_k [\Delta_k - \|d_k\|] = 0. \quad (4.2.7)$$

步六 更新 $k := k + 1$, 并转步二。

上个世纪八十年代一些国际知名的优化学者建立了一系列关于用信赖域方法求解复合NDO的重要收敛性结果(如: Fletcher [34], Powell [98]及Yuan [136, 137, 138, 139, 140])。

Hald [44, 45]提出的两阶段方法也是求解复合NDO(4.2.1)的重要方法。这种方法的主要思想是切换使用求解线性近似的信赖域子问题与用牛顿法求解问题的KKT条件。前者用来近似确定解附近的各项 $f_i(x)$ 的符号, 后者的目的是在确定符号后快速收敛。

我们知道信赖域算法4.2.1在每一个迭代步都需要求解信赖域子问题(4.2.3)-(4.2.4)。据此我们认为信赖域子问题的快速算法对求解(4.2.1)的效率至关重要。

§4.3 非光滑信赖域子问题分析

本节的前三节, 我们将专门分析和讨论非线性 L_1 范数极小化问题的信赖域子问题(4.2.2)-(4.2.3)。为此我们作一些记号上的简化:

$$\min_{d \in \mathbf{R}^n} \quad \phi(d) = \|Bd + c\|_1 + \frac{1}{2} d^T A d + g^T d; \quad (4.3.1)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta, \quad (4.3.2)$$

其中 $B \in \mathbf{R}^{m \times n}$, $c \in \mathbf{R}^m$, $A \in \mathbf{R}^{n \times n}$ 是实对称矩阵, $g \in \mathbf{R}^n$, $\Delta \in \mathbf{R}^+$ 。我们记 $B = (\beta_1, \beta_2, \dots, \beta_m)^T$, 这里 $\beta_i \in \mathbf{R}^n (i = 1, 2, \dots, m)$ 是 B 矩阵的第 i 行的转置。

我们知道对一般的信赖域子问题:

$$\min_{d \in \mathbf{R}^n} \quad \bar{\phi}(d) = \frac{1}{2} d^T A d + g^T d; \quad (4.3.3)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta. \quad (4.3.4)$$

More和Sorensen(参阅 [71, 72])在1983年提出了多项式时间算法。这个算法是迄今为止精确求解信赖域子问题最好的算法。

我们发现(4.3.1)与(4.3.3)两个目标函数的差 $\phi(d) - \bar{\phi}(d) = \|Bd + c\|_1$ 是一个凸函数,也就是说以这个目标函数的差为目标函数的问题也是具有多项式时间复杂度算法的。然而两个具有多项式时间算法的问题的和问题并不一定具有多项式时间算法。下面我们就来证明非光滑的信赖域子问题(4.3.1)-(4.3.2)是NP-难的。

下述定理是我们验证一个问题是NP-难的关键依据,相关概念及其证明详见([14], [36])。

定理 4.3.1. 一个问题 H 是NP-难的当且仅当存在一个NP-完全问题可以多项式时间地归约(Turing-reduce)为问题 H (记为: $L \leq_T H$)。

我们熟知如下的划分问题是NP-完全的([36]):

定义 4.3.1. 给定一个实数集合 $\Omega = \{a_1, a_2, \dots, a_n\}$, 是否存在一个子集 I 满足

$$\sum_{i \in I} a_i = \frac{1}{2} \sum_{j \in \Omega} a_j? \quad (4.3.5)$$

找到集合 I 或者给出否定的回答, 我们称之为划分问题。

定理 4.3.2. 划分问题是NP-难的。

有了这些准备知识, 我们就清楚地知道如果能够构建划分问题与我们的非光滑信赖域子问题(4.3.1)-(4.3.2)之间的多项式归约, 我们就可以证明该子问题是NP-难的。

引理 4.3.1. 函数

$$\rho(x) = \left| \frac{\sqrt{n}}{2n} x \right| + \left| x + \frac{\sqrt{n}}{n} \right| + \left| x - \frac{\sqrt{n}}{n} \right| - x^2 + \frac{1}{2n} - \frac{2\sqrt{n}}{n} \geq 0, \quad (4.3.6)$$

对任意 $\forall x \in [-1, 1]$; 等号成立当且仅当 $x = x_1 := -\frac{\sqrt{n}}{n}$ 或者 $x = x_2 := \frac{\sqrt{n}}{n}$ 。(请参考图4.2)。

证明: 由于函数 $\rho(x)$ 关于 $x = 0$ 对称, 所以不失一般性, 我们下面仅讨论 $x \in [0, 1]$ 的情形。

首先考虑 $x \in [0, \frac{\sqrt{n}}{n}]$ 。在这个区间内,

$$\rho(x) = -x^2 + \frac{\sqrt{n}}{2n} + \frac{1}{2n}. \quad (4.3.7)$$

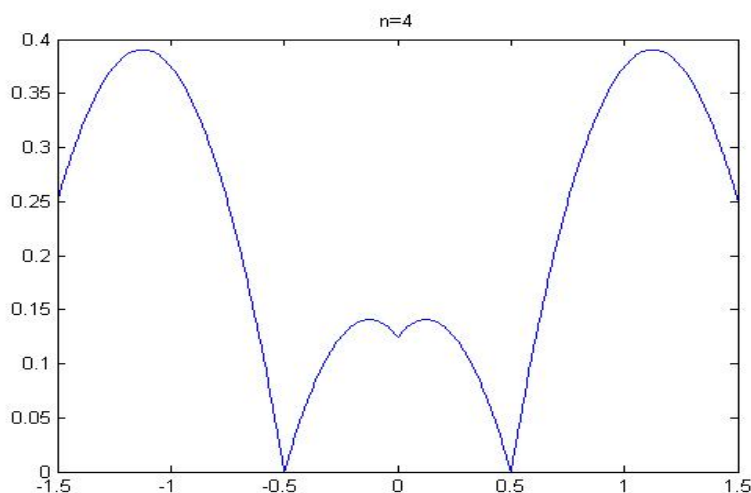


图 4.2

这时显然有 $x = x_2$ 是唯一的全局极小点, 并且 $\rho(x_2) = 0$ 。

所以当 $n = 1$ 时, 结论成立。

下面我们讨论 $n \geq 2$ 的情况, 考虑 x 在区间 $[\frac{\sqrt{n}}{n}, 1]$ 中,

$$\rho(x) = -x^2 + \left(2 + \frac{\sqrt{n}}{2n}\right) + \frac{1}{2n} - \frac{2\sqrt{n}}{n}. \quad (4.3.8)$$

我们知道此时全局最小点只可能为 $x = x_2$ 或者 $x = 1$ 。然而,

$$\rho(1) = 1 - \frac{3\sqrt{n}}{2n} + \frac{1}{2n} = \left(1 - \frac{\sqrt{n}}{2n}\right) \left(1 - \frac{\sqrt{n}}{2n}\right) > 0 = \rho(x_2). \quad (4.3.9)$$

所以 $x = x_2$ 是这个区间内唯一的全局极小点, 并且 $\rho(x_2) = 0$ 。

综上所述, 对任何 $n \leq 1$, $x = x_2$ 是区间 $[0, 1]$ 中唯一的全局极小点。因此 $x = x_1$ 和 $x = x_2$ 是函数 $\rho(x)$ 在 $[-1, 1]$ 中仅有的两个全局极小点, 并且最小值为 0。定理得证。□

定理 4.3.3. 非光滑的信赖域子问题(4.3.1)-(4.3.2)是 NP-难的。

证明: 首先给出非光滑的信赖域子问题(4.3.1)-(4.3.2)的一个特例,

$$\min_{d \in \mathbf{R}^n} \quad m(d) = \sum_{i=1}^n \rho(d_i) + \left(\sum_{i=1}^n a_i d_i\right)^2; \quad (4.3.10)$$

$$\text{s.t.} \quad \|d\|_2 \leq 1. \quad (4.3.11)$$

由于 $\|d\|_2 \leq 1$, 对每个分量我们有 $|d_i| \leq 1$ 。根据引理4.3.1, $\rho(d_i) \geq 0, \forall i = 1, 2, \dots, n$ 。因此, 我们知道 $m(d) \geq 0$ 对任何满足 $\|d\|_2 \leq 1$ 的 d 成立, 也就是说0是 $m(d)$ 的一个下界。

假设 v^* 是(4.3.10)-(4.3.11)的全局极小。因此 $v^* = 0$ 等价于

$$\begin{cases} \rho(d_i) = 0, \forall i = 1, 2, \dots, n; \\ \sum_{i=1}^n a_i d_i = 0; \\ \sum_{i=1}^n d_i^2 \leq 1. \end{cases} \quad (4.3.12)$$

根据引理4.3.1, (4.3.12)等价于

$$\begin{cases} d_i = \pm \frac{\sqrt{n}}{n}, \forall i = 1, 2, \dots, n; \\ \sum_{i=1}^n a_i d_i = 0. \end{cases} \quad (4.3.13)$$

显然(4.3.13)等价于 $\Omega = \{a_1, a_2, \dots, a_n\}$ 时由定义4.3.1定义的划分存在。因此找到(4.3.10)-(4.3.11)的全局极小, 就相当于解决了对应的划分问题。

由于(4.3.10)-(4.3.11)是非光滑的信赖域子问题(4.3.1)-(4.3.2)的一个例子, 并且 $m = 3n$ 。因此如果存在求解非光滑的信赖域子问题(4.3.1)-(4.3.2)的多项式时间算法, 那么我们就可以多项式时间求解划分问题, 这和定理4.3.1矛盾。证毕。 \square

§4.4 低维非光滑信赖域子问题的直接算法

上一节中我们证明了非光滑的信赖域子问题(4.3.1)-(4.3.2)是NP-难的。这一节中我们来讨论非光滑的信赖域子问题(4.3.1)-(4.3.2)在一维、二维这两个特殊情形下的快速多项式时间算法。本节介绍的算法将是后文介绍的求解 n 维非光滑的信赖域子问题快速算法的基础。

§4.4.1 假设和定义

为了描述清楚, 在介绍我们的算法之前, 先给出两个合理的假设条件。

假设 4.4.1. 非光滑的信赖域子问题(4.3.1)-(4.3.2)中不含有平凡项。也就是说, 对 $\forall i \in \{1, 2, \dots, m\}$, 我们有 $\beta_i \neq 0$ 。

假设 4.4.2. 非光滑的信赖域子问题(4.3.1)-(4.3.2)中不含有同类项。也就是说, 对 $\forall 1 \leq i < j \leq m$, 如果有 $\beta_i = \lambda \beta_j$, 则 $c_i \neq \lambda c_j$, 这里 λ 是非零实数。

这两个假设是合理的，因为平凡项对(4.3.1)-(4.3.2)的贡献仅仅是将函数上下平移一个常数，不会影响最优点的位置。所以即使存在平凡项，我们也完全可以不考虑那一项。对于同类项，如果存在，则只需用它们的和代替它们而不改变函数值。

上述假设是为了使算法的提出过程、描述和理论分析更明晰，在没有假设的情况下，我们的算法只需增加对于平凡项和同类项的简单处理即可。另一方面，我们也可以通过 $O(mn)$ 和 $O(m^2n)$ 的计算量将所有的平凡项和同类项侦测出来(本节中 $n = 1$ ，或者2)；而对平凡项和同类项的侦测过程完全可以融入我们本节设计的算法而不增加额外的计算量。

为了后文叙述方便，下面我们给出一些定义在 n 维空间中的量。

定义 4.4.1.

1) 称 $\Omega(d) = \{i | \beta_i^T d + c_i = 0, \quad \forall i = 1, 2, \dots, m\}$ 为 d 点处的积极非光滑指标集；
 $\beta_i^T d + c_i = 0 (\forall i \in \Omega(d))$ 为 d 点处的一个积极非光滑面；

2) $\xi(d) = \sum_{j \notin \Omega(d)} \text{sign}(\tau_j(d)) \cdot \beta_j + Ad + g$;

3) $\mathcal{S}(d) := \text{span}\{\beta_i, \forall i \in \Omega(d)\}$;

4) $\mathcal{Z}(d) := \{s | \beta_i^T s = 0, \exists i \in \Omega(d)\}$.

5) $\mathcal{Q}(d) = \bigcap_{i \in \Omega(d)} \{s | \beta_i^T s + c_i = 0\}$ 是 d 的所有积极非光滑面的交集。

定义 4.4.2. $s^{(1)}(d) = -\mathcal{P}_{\mathcal{S}(d)}^\perp \xi(d)$.

定义 4.4.3. $s^{(2)}(d) = -\sum_{i \in \Omega(d)} \eta_i^* \beta_i - \xi(d)$ ；这里 $\eta_i^* (i \in \Omega(d))$ 是如下带界约束的线性最小二乘问题

$$\min_{\eta_i \in \mathbf{R}, i \in \Omega(d)} \left\| \sum_{i \in \Omega(d)} \eta_i \beta_i + \xi(d) \right\|_2^2, \quad (4.4.1)$$

$$\text{s.t.} \quad -1 \leq \eta_i \leq 1, \quad i \in \Omega(d) \quad (4.4.2)$$

的最优解。

定义 4.4.4. 记 $\delta(d) = s^{(2)}(d) - s^{(1)}(d)$ 。

定义 4.4.5. $s^{(3)}(d) = -\zeta_0^* d - \sum_{i \in \Omega(d)} \zeta_i^* \beta_i - \xi(d)$; 这里 $\zeta_i^*(i \in \Omega(d) \cup \{0\})$ 是如下带界约束的线性最小二乘问题

$$\min_{\zeta_i \in \mathbf{R}, i \in \Omega(d) \cup \{0\}} \|\zeta_0 d + \sum_{i \in \Omega(d)} \zeta_i \beta_i + \xi(d)\|_2^2, \quad (4.4.3)$$

$$\text{s.t.} \quad \zeta_0 \geq 0, \quad (4.4.4)$$

$$-1 \leq \zeta_i \leq 1, \quad i \in \Omega(d) \quad (4.4.5)$$

的最优解。

根据上面的定义, 我们有如下的定理。

定理 4.4.1. 对任何约束(4.3.2)的内可行点 \bar{d} , 也即: $\|\bar{d}\| < \Delta$,

- 1) $s^{(1)}(\bar{d}) = 0$, 当且仅当 $\xi(\bar{d}) \in \mathcal{S}(\bar{d})$;
- 2) 如果 $s^{(1)}(\bar{d}) \neq 0$, 那么 $s^{(1)}(\bar{d})$ 是(4.3.1)中目标函数 $\phi(\bar{d})$ 在 \bar{d} 点的下降方向;
- 3) $s^{(2)}(\bar{d}) = 0$, 当且仅当 \bar{d} 是(4.3.1)的一阶稳定点;
- 4) 如果 $s^{(2)}(\bar{d}) \neq 0$, 那么 $s^{(2)}(\bar{d})$ 是(4.3.1)中目标函数 $\phi(\bar{d})$ 在 \bar{d} 点的下降方向。

对任何约束(4.3.2)的边界可行点 \bar{d} , 也即: $\|\bar{d}\| = \Delta$,

- 5) $s^{(3)}(\bar{d}) = 0$, 当且仅当 \bar{d} 是(4.3.1)-(4.3.2)的一阶稳定点;
- 6) 如果 $s^{(3)}(\bar{d}) \neq 0$, 那么 $s^{(3)}(\bar{d})$ 是(4.3.1)中目标函数 $\phi(\bar{d})$ 在 \bar{d} 点的下降方向;
- 7) $s^{(3)}(\bar{d})^T \bar{d} \leq 0$.

证明: 这里子命题1)和子命题2)显然成立, 证明从略。

对于子命题3), 我们知道(4.3.1)中目标函数 $\phi(\bar{d})$ 的次梯度 $\partial\phi(\bar{d}) = \{ \sum_{i \in \Omega(\bar{d})} \eta_i \beta_i + \xi(\bar{d}) | \eta_i \in \mathbf{R}, \forall i \in \Omega(\bar{d}) \}$ 。因此 $s^{(2)}(\bar{d}) = 0$ 当且仅当 $0 \in \partial\phi(\bar{d})$, 这等价于 \bar{d} 是(4.3.1)的一阶稳定点。

考虑子命题5), 非光滑的信赖域子问题(4.3.1)-(4.3.2)的KKT系统是

$$\begin{cases} \Omega(d) = \{i | \beta_i^T d + c_i = 0, \quad \forall i = 1, 2, \dots, m\}; & (i) \\ \zeta_0 d + \sum_{i \in \Omega(d)} \zeta_i \beta_i + \sum_{j \notin \Omega(d)} \text{sign}(\beta_j^T d + c_j) \cdot \beta_j + Ad + g = 0; & (ii) \\ \zeta_0 \cdot (\|d\|_2 - \Delta) = 0; & (iii) \\ \|d\|_2 \leq \Delta; & (iv) \\ \zeta_0 \geq 0; & (v) \\ -1 \leq \zeta_i \leq 1, \quad \forall i \in \Omega(d). & (vi) \end{cases} \quad (4.4.6)$$

当约束(4.3.2)是积极的, 也即 $\|\bar{d}\|_2 - \Delta = 0$ 时, \bar{d} 是(4.3.1)-(4.3.2)的一阶稳定点当且仅当 $s^{(3)}(\bar{d}) = 0$ 。至此子命题5)得证。

关于子命题4)和子命题6), 由于我们将在下一节中给出更强的结论, 这里暂时不给出证明。

写出带界约束的线性最小二乘问题(4.4.3)-(4.4.5)的KKT系统后, 我们知道定义4.4.5中的 ζ_0^* 需要满足如下两个方程:

$$\begin{cases} (\zeta_0^* \bar{d} + \sum_{i \in \Omega(\bar{d})} \zeta_i^* \beta_i + \sum_{j \notin \Omega(\bar{d})} \text{sign}(\beta_j^T \bar{d} + c_j) \cdot \beta_j + A\bar{d} + g)^T \bar{d} = \sigma_0^*; & (i) \\ \sigma_0^* \geq 0. & (ii) \end{cases} \quad (4.4.7)$$

这里 σ_0^* 是约束(4.4.4)的Lagrange乘子。容易发现联立不等式组(4.4.7) 等价于子命题7)中的结论。证毕。 \square

§4.4.2 一维情形

首先我们来讨论非光滑信赖域子问题(4.3.1)-(4.3.2)的一维情形:

$$\min_{d \in \mathbf{R}} \phi_{(1)}(d) = \sum_{i=1}^m |b_i d + c_i| + \frac{1}{2} \alpha d^2 + \beta d; \quad (4.4.8)$$

$$\text{s.t.} \quad -\delta \leq d \leq \delta. \quad (4.4.9)$$

由于我们的目标是找到(4.4.8)-(4.4.9)的全局极小点, 我们当然需要把所有可能的候选点找到。最直观的一种方法是先找出所有的局部极小点, 再通过比较函数值大小, 得到全局极小。

任何局部极小点都满足(4.4.8)-(4.4.9)的KKT系统:

$$\begin{cases} \Omega(d) = \{i | \tau_i(d) := b_i d + c_i = 0, \quad \forall i = 1, 2, \dots, m\}; & (i) \\ \sum_{i \in \Omega(d)} \lambda_i \cdot b_i + \sum_{i \notin \Omega(d)} \text{sign}(\tau_i(d)) \cdot b_i + \alpha d + \beta = \sigma_l - \sigma_u; & (ii) \\ -1 \leq \lambda_i \leq 1, \quad \forall i \in \Omega(d); & (iii) \\ \sigma_l \cdot (d + \delta) = 0; & (iv) \\ \sigma_u \cdot (\delta - d) = 0; & (v) \\ -\delta \leq d \leq \delta; & (vi) \\ \sigma_l \geq 0, \quad \sigma_u \geq 0. & (vii) \end{cases} \quad (4.4.10)$$

我们可以把所有满足约束条件(4.4.9)的可行点分为如下三种类型:

(一) 约束(4.4.9)的两个端点;

(二) 位于约束(4.4.9)内部, 目标函数 $\phi_{(1)}(d)$ 的不可微点;

(三) 位于约束(4.4.9)内部, 目标函数 $\phi_{(1)}(d)$ 的可微点。

在第一类点中, 局部极小点最多只可能有两个(我们称之为第一类局部极小点); 而在第二类点中, 局部极小点最多只可能有 m 个(我们称之为第二类局部极小点)。

在假设4.4.1的前提下, 我们给出下面的定义:

定义 4.4.6.

- 集合 $\Upsilon = \{s | s = -\frac{c_i}{b_i}, \forall i = 1, 2, \dots, m; \text{ 且 } -\delta \leq s \leq \delta\}$;
- 序列 $\{s_1, s_2, \dots, s_p\} := \Upsilon \cup \{-\delta, \delta\}$, 并且所有元素 $\{s_k\}$ 按升序排列, 也即 $-\delta = s_1 < s_2 < \dots < s_p = \delta$;
- 集合 $\Gamma(s) = \{i | 1 \leq i \leq m, \tau_i(s) \neq 0\}$ 。

很显然定义中的 p 满足: $2 \leq p \leq m + 2$ 。也就是说第一类和第二类局部极小点都被囊括在序列 $\{s_k\}$ 内了。

另一方面, 由于第三类点中的局部极小点(我们称之为第三类局部极小点)的个数不能直观地被估计出来, 于是我们给出下述命题, 以助于我们了解第三类局部极小点存在的条件及个数上限。

命题 4.4.1. 如果(4.4.8)-(4.4.9)存在第三类局部极小点, 那么 $\alpha \geq 0$; 进一步地, 如果存在孤立的第三类局部极小点, 则 $\alpha > 0$ 。

证明: 我们假设 d^* 是(4.4.8)-(4.4.9)的一个第三类局部极小点。因此,

$$\Omega(d^*) = \emptyset, \quad (4.4.11)$$

并且存在一个 d^* 的邻域

$$U(d^*, \delta) := \{d | |d - d^*| \leq \delta\}, \quad (4.4.12)$$

使得目标函数 $\phi_{(1)}(d)$ 在该邻域中的点上都存在导数。也即

$$m'_{(1)}(d) = \sum_{i=1}^m \text{sign}(\tau_i(d^*)) \cdot b_i + \alpha d + \beta. \quad (4.4.13)$$

因此我们可知:

$$m''_{(1)}(d^*) = \alpha, \quad (4.4.14)$$

于是结论成立。证毕。 \square

通过命题4.4.1我们知道, 当 $a < 0$ 时, (4.4.8)-(4.4.9)的所有全局极小点, 都是第一或第二类局部极小点; 当 $a = 0$ 时, 也必定存在属于第一或第二类局部极小点的全局极小点; 当 $a > 0$ 时, 我们的问题(4.4.8)-(4.4.9)是严格凸的, 因此只存在唯一的局部极小点, 也即唯一的全局极小点。

剩下的问题就是如果 $a > 0$, 我们如何找到这个唯一的极小点。根据KKT系统(4.4.10), 我们知道此时, 最关键的任务是确定每个 $\tau_i(d^*)$ 的符号。我们首先引入一个引理, 然后给出本节最关键的定理来回答这个问题。

命题 4.4.2. 假设 $d^{(1)}$ 和 $d^{(2)}$ 是任意两个满足约束条件(4.4.9)的点, 且 $d^{(1)} < d^{(2)}$ 。如果存在 $j \in \{1, 2, \dots, m\}$ 使得

$$\text{sign}(\tau_j(d^{(1)})) \cdot \text{sign}(\tau_j(d^{(2)})) = -1, \quad (4.4.15)$$

那么必定存在位于区间 $(d^{(1)}, d^{(2)})$ 中的第二类点。

证明: 由(4.4.15)可知, $\tau_j(d^{(1)})$ 与 $\tau_j(d^{(2)})$ 必定一个大于零, 一个小于零。由于 $\tau_j(d)$ 在 $[d^{(1)}, d^{(2)}]$ 中连续, 因此必定存在 $\bar{d} \in (d^{(1)}, d^{(2)})$ 满足 $\tau_j(\bar{d}) = 0$, 也即 \bar{d} 是 $\phi_{(1)}(d)$ 的不可微点。所以 \bar{d} 就是一个位于 $(d^{(1)}, d^{(2)})$ 的第二类点。 \square

定理 4.4.2. 假设序列 $\{s_k\}$ 如定义4.4.6定义, 我们记

$$k^* = \arg \min_{k=1,2,\dots,p} \phi_{(1)}(s_k). \quad (4.4.16)$$

假设(4.4.8)-(4.4.9)存在孤立的第三类局部极小点, 设为 d^* 。则必有

$$\text{sign}(\tau_i(d^*)) = \text{sign}(\tau_i(s_{k^*})), \quad \forall i \in \Gamma(s_{k^*}). \quad (4.4.17)$$

证明: 我们用反证法, 假设定理不成立, 那么存在 $j \in \Gamma(s_{k^*})$ 满足 $\text{sign}(\tau_j(d^*)) \neq \text{sign}(\tau_j(s_{k^*}))$ 。

由于 d^* 是第三类点, 我们知道 $\text{sign}(\tau_j(d^*)) \neq 0$ 。由于 $j \in \Gamma(s_{k^*})$, 我们有 $s_{k^*} \neq s_j$, 因而 $\text{sign}(\tau_j(s_{k^*})) \neq 0$ 。综上, 我们知道 $\text{sign}(\tau_j(d^*))$ 与 $\text{sign}(\tau_j(s_{k^*}))$ 必分别为1和-1。根据命题4.4.2, 必存在位于区间 $(\min\{d^*, s_{k^*}\}, \max\{d^*, s_{k^*}\})$ 中的第二类点 \hat{d} 。

由 k^* 的定义(s_{k^*} 是第一类和第二类点中函数值最小的点), 我们知道

$$\phi_{(1)}(\hat{d}) \geq \phi_{(1)}(s_{k^*}). \quad (4.4.18)$$

另一方面, 根据命题4.4.1, 由于 d^* 是(4.4.8)-(4.4.9)孤立的第三类局部极小点, 因此 $a > 0$; 进而(4.4.8)-(4.4.9)是凸优化, 它唯一的局部极小点 d^* 也是它唯一的全局极小点。因此

$$\phi_{(1)}(\hat{d}) \geq \phi_{(1)}(d^*). \quad (4.4.19)$$

由于目标函数 $\phi_{(1)}(d)$ 是凸的,且 \hat{d} 位于区间 $(\min\{d^*, s_{k^*}\}, \max\{d^*, s_{k^*}\})$ 中,再结合式(4.4.18)与式(4.4.19),我们可以推出

$$\phi_{(1)}(s_{k^*}) = \phi_{(1)}(\hat{d}) = \phi_{(1)}(d^*). \quad (4.4.20)$$

而这与 d^* 是(4.4.8)-(4.4.9)唯一的全局极小点矛盾。假设不成立,证毕。□

根据定理4.4.2,我们可以设计如下的算法求解一维非光滑信赖域子问题(4.4.8)-(4.4.9)的全局极小点。

算法 4.4.1. (一维非光滑信赖域子问题的直接算法)

步一 如果 $a \leq 0$ 转步二;否则转步三。

步二 根据定义4.4.6计算集合 Υ ;通过枚举法求解如下离散优化问题:

$$\min_{d \in \Upsilon \cup \{-\delta, \delta\}} \phi_{(1)}(d), \quad (4.4.21)$$

得 d^{record} ,并转步六。

步三 根据定义4.4.6排序,得到序列 $\{s_k\}$;通过枚举法求解如下离散优化问题:

$$\min_{k=1,2,\dots,p} \phi_{(1)}(s_k), \quad (4.4.22)$$

得 k^* ,记 $d^{record} = s_{k^*}$,并转步四。

步四 如果 $k^* < p$,求解如下一维连续优化问题:

$$\begin{aligned} \min_{s_{k^*} \leq d \leq s_{k^*+1}} & \sum_{i=1}^m \text{sign}(b_i \cdot \frac{s_{k^*} + s_{k^*+1}}{2} + c_i) \cdot (b_i d + c_i) \\ & + \frac{1}{2} \alpha d^2 + bd, \end{aligned} \quad (4.4.23)$$

得 \bar{d} ,如果 $\phi_{(1)}(\bar{d}) < \phi_{(1)}(d^{record})$,更新 $d^{record} = \bar{d}$,并转步五。

步五 如果 $k^* > 1$,求解如下一维连续优化问题:

$$\begin{aligned} \min_{s_{k^*-1} \leq d \leq s_{k^*}} & \sum_{i=1}^m \text{sign}(b_i \cdot \frac{s_{k^*-1} + s_{k^*}}{2} + c_i) \cdot (b_i d + c_i) \\ & + \frac{1}{2} \alpha d^2 + bd, \end{aligned} \quad (4.4.24)$$

得 \bar{d} ,如果 $\phi_{(1)}(\bar{d}) < \phi_{(1)}(d^{record})$,更新 $d^{record} = \bar{d}$,并转步六。

步六 返回 d^{record} ,并停止。

步三中排序算法的计算复杂度是 $O(m \log m)$;步三、步四中枚举法计算 d^{record} 和 k^* 的计算复杂度都是 $O(m^2)$;算法其它部分的复杂度均不超过 $O(m)$ 。因此算法4.4.1的计算复杂度为 $O(m^2)$ 。

§4.4.3 二维情形

本小节中，我们将考虑非光滑的信赖域子问题(4.3.1)-(4.3.2)的二维情形：

$$\min_{d \in \mathbf{R}^2} \quad \phi_{(2)}(d) = \sum_{i=1}^m |a_i d_1 + b_i d_2 + c_i| + \frac{1}{2} d^T A d + g^T d; \quad (4.4.25)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta. \quad (4.4.26)$$

首先我们给出(4.4.25)-(4.4.26)的KKT条件：

$$\left\{ \begin{array}{l} \Omega(d) = \{i | \tau_i(d) := a_i d_1 + b_i d_2 + c_i = 0, i = 1, 2, \dots, m\}; \quad (i) \\ \sum_{i \notin \Omega(d)} \text{sign}(\tau_i(d)) \begin{pmatrix} a_i \\ b_i \end{pmatrix} + \sum_{j \in \Omega(d)} \delta_j \begin{pmatrix} a_j \\ b_j \end{pmatrix} \\ \quad + (A + \lambda I) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + g = 0; \quad (ii) \\ \delta_j \in [-1, 1], \forall j \in \Omega(d); \quad (iii) \\ \lambda(d_1^2 + d_2^2 - \Delta^2) = 0; \quad (iv) \\ d_1^2 + d_2^2 \leq \Delta^2; \quad (v) \\ \lambda \geq 0. \quad (vi) \end{array} \right. \quad (4.4.27)$$

在给出我们算法的构造过程之前，和前一小节中一样，我们先要将所有满足约束条件(4.4.26)的可行点进行分类。

定义 4.4.7. 我们将 $\|d\| \leq \Delta$ 中的点分为如下五类：

- 如果点 d 位于约束边界，即： $\|d\| = \Delta$ ；并且是目标函数 $\phi_{(2)}(d)$ 的可微点，即： $|\Omega(d)| = 0$ 。我们称这样的点为 $B1$ 类点；
- 如果点 d 位于约束边界，即： $\|d\| = \Delta$ ；并且是目标函数 $\phi_{(2)}(d)$ 的不可微点，即： $|\Omega(d)| \geq 1$ 。我们称这样的点为 $B2$ 类点；
- 如果点 d 位于约束内部，即： $\|d\| < \Delta$ ；并且是目标函数 $\phi_{(2)}(d)$ 的可微点，即： $|\Omega(d)| = 0$ 。我们称这样的点为 $I1$ 类点；
- 如果点 d 位于约束内部，即： $\|d\| < \Delta$ ；该点还是目标函数 $\phi_{(2)}(d)$ 的不可微点，且满足 $|\Omega(d)| = 1$ 。我们称这样的点为 $I2$ 类点；
- 如果点 d 位于约束内部，即： $\|d\| < \Delta$ ；该点还是目标函数 $\phi_{(2)}(d)$ 的不可微点，且满足 $|\Omega(d)| \geq 2$ 。我们称这样的点为 $I3$ 类点。

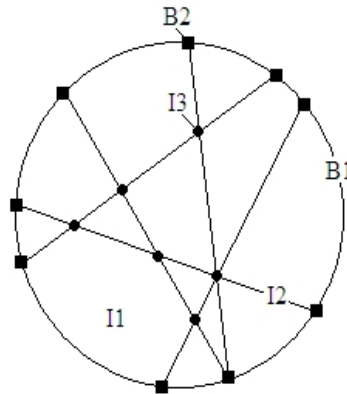


图 4.3

我们分别从这五类点中找出函数值最小的局部极小值点的难度是不一样的。比如说我们考虑I3类,则枚举法即可,因为I3类点的总个数不超过 C_m^2 个(因为 $|I3| \leq C_{|\Omega(d)|}^2 \leq C_m^2$)。而另一方面,要找属于I1类的局部极小点则相对较难,因为所有 $\tau_i(d) (\forall i = 1, 2, \dots, m)$ 的符号难以事先确定;而所有的可能性有 2^m 种,但我们不希望设计指数复杂度的算法。

比较直观地,我们可以先寻找属于B2、I2、I3类集合中函数值最小的点。这时可以把这个问题描述为如下优化问题(如图4.3所示):

$$\min_{d \in \mathbf{R}^2} \quad \phi_{(2)}(d) = \sum_{i=1}^m |a_i d_1 + b_i d_2 + c_i| + \frac{1}{2} d^T A d + g^T d; \quad (4.4.28)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta; \quad (4.4.29)$$

$$\prod_{i=1}^m \tau_i(d) = 0. \quad (4.4.30)$$

而我们发现极小化(4.4.28)-(4.4.30)的问题可以转化为具有如下形式的 m 个子问题,其中 $j = 1, 2, \dots, m$:

$$\min_{d \in \mathbf{R}^2} \quad \phi_{(2)}(d) = \sum_{i=1}^m |a_i d_1 + b_i d_2 + c_i| + \frac{1}{2} d^T A d + g^T d; \quad (4.4.31)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta; \quad (4.4.32)$$

$$\tau_j(d) = 0, \quad (4.4.33)$$

这又等价于如下优化问题

$$\min_{t \in \mathbf{R}} \quad m_{(1)}(t) = \sum_{i=1}^m |\beta_i^T w_j \cdot t + \beta_i^T v_j + c_i| + \frac{1}{2} w_j^T A w_j \cdot t^2 + (v_j^T A w_j + g^T w_j) \cdot t; \quad (4.4.34)$$

$$\text{s.t.} \quad -\sqrt{\Delta^2 - v_j^T v_j} \leq t \leq \sqrt{\Delta^2 - v_j^T v_j}. \quad (4.4.35)$$

这里 $\beta_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}$, $v_j = \begin{pmatrix} -\frac{a_j c_j}{2(a_j^2 + b_j^2)} \\ -\frac{b_j c_j}{2(a_j^2 + b_j^2)} \end{pmatrix}$, $w_j = \begin{pmatrix} \frac{b_j}{\sqrt{a_j^2 + b_j^2}} \\ -\frac{a_j}{\sqrt{a_j^2 + b_j^2}} \end{pmatrix}$.

分别求得(4.4.34)-(4.4.35)对于 $j = 1, 2, \dots, m$ 的全局极小值点 t_j^* , 代入 $x_j^* = w_j \cdot t_j^* + v_j$, 得到(4.4.31)-(4.4.33) 的全局极小值点 x_j^* 。通过求解比较目标函数 $\phi_{(2)}(d)$ 在 $x_j^* (j = 1, 2, \dots, m)$ 处的函数值, 我们可以得到(4.4.28)-(4.4.30) 的全局极小点, 我们记之为 d^{gn} (这里上标 gn 取自 *global minimizer of nondifferentiable points*)。下面的问题就在于:

- 什么时候 d^{gn} 是(4.4.25)-(4.4.26)的一个全局极小值点;
- 当 d^{gn} 不是(4.4.25)-(4.4.26)的一个全局极小值点时, 我们怎么通过 d^{gn} 找到(4.4.25)-(4.4.26)的一个全局极小值点。

下面的定理帮助我们回答了第一个问题:

定理 4.4.3. 我们假设 d^{gn} 是(4.4.28)-(4.4.30)的一个全局极小值点。

- 如果矩阵 A 半正定, 并且 d^{gn} 满足 *KKT* 系统(4.4.27), 那么 d^{gn} 是(4.4.25)-(4.4.26)的一个全局极小值点;
- 如果矩阵 A 半正定, 并且 d^{gn} 不满足 *KKT* 系统(4.4.27), 那么所有全局极小值点都属于 *B1* 或 *I1* 类。如果我们假设 d^* 是一个全局极小值点, 那么必有下式成立:

$$\{d | d = \lambda d^{gn} + (1 - \lambda) d^*, \lambda \in (0, 1)\} \cap \{d | \exists 1 \leq j \leq m, \text{s.t. } \tau_j(d) = 0\} = \emptyset. \quad (4.4.36)$$

- 如果矩阵 A 不是半正定的, 那么不存在属于 *I1* 类的局部极小值点。

证明: 首先我们考虑矩阵 A 半正定的情形, 此时(4.4.25)-(4.4.26) 是一个凸规划。因此满足 *KKT* 系统(4.4.27)的点只能是局部极小值点, 并且局部极小值点处的函

数值都相同。因此如果 d^{gn} 满足KKT系统(4.4.27),就意味着 d^{gn} 是一个全局极小值点。

而在矩阵 A 半正定的情形下,当 d^{gn} 不满足KKT条件,也就是说 d^{gn} 不是局部极小值点时,必有函数值比 d^{gn} 低的点。那么根据 d^{gn} 的定义,我们知道所有函数值比之小的点都在B1或I1类内。也即不存在属于B2、I2或I3类内的全局极小点。

因此 d^* 属于B1或I1类。下面我用反证法来证明式(4.4.36)。假设(4.4.36)不成立,那么必存在 d^{gn} 与 d^* 的凸组合

$$p = \lambda d^{gn} + (1 - \lambda)d^* \quad (4.4.37)$$

是目标函数 $\phi_{(2)}(d)$ 的一个不可微点,这里 $\lambda \in (0, 1)$ 。根据目标函数 $\phi_{(2)}(d)$ 的凸性,我们知道

$$\phi_{(2)}(p) \leq \lambda \phi_{(2)}(d^{gn}) + (1 - \lambda)\phi_{(2)}(d^*) < \phi_{(2)}(d^{gn}). \quad (4.4.38)$$

由于 p 是不可微点,因此 p 也满足约束(4.4.29)-(4.4.30),而这与 d^{gn} 是(4.4.28)-(4.4.30)的全局极小值点矛盾。因此式(4.4.36)成立。

最后,我们考虑矩阵 A 不是半正定的情形。还是用反证法,假设存在属于I1类的局部极小点,不妨设之为 d^{li} (这里上标 li 取自local minimizer from interior differentiable points)。由于矩阵 A 不是正定的, A 至少有一个负特征值,不妨设 $\sigma < 0$ 是 A 的一个负特征值,并记 u 为其对应的特征向量。我们把(4.4.25)-(4.4.26)限制到直线

$$\mathcal{L}(d^{li}, u) = \{d | d = d^{li} + \lambda u, \forall \lambda \in \mathbb{R}\} \quad (4.4.39)$$

上,得到一个一维的问题。由于 d^{li} 是内点,因此直线 $\mathcal{L}(d^{li}, u)$ 与信赖域约束 $\{d | \|d\|_2 \leq \Delta\}$ 的交的测度大于零,由此可知 d^{li} 是新的一维问题的局部极小值点。根据命题4.4.1,我们知道 $\sigma \geq 0$,这与 $\sigma < 0$ 矛盾。因此假设不成立,此时(4.4.25)-(4.4.26)不存在属于I1类的局部极小点。

综上所述,命题得证。□

上面定理告诉我们只有在矩阵 A 是半正定,且(4.4.28)-(4.4.30)的全局极小值点 d^{gn} 满足KKT系统(4.4.27)的条件下, d^{gn} 同时还是(4.4.25)-(4.4.26)的一个全局极小值点。在其它情况下,我们还需要回答第二个问题,怎么通过 d^{gn} 快速地找到(4.4.25)-(4.4.26)的一个全局极小值点。下面我们分两个情形来分析。

情形一:矩阵 A 半正定,且(4.4.28)-(4.4.30)的全局极小值点 d^{gn} 不满足KKT系统(4.4.27)。

设 d^* 是(4.4.25)-(4.4.26)的一个全局极小值点。根据定理4.4.3, 我们知道 d^{gn} 和 d^* 的连线上的所有点都是目标函数 $\phi_{(2)}(d)$ 的可微点。换句话说 d^* 位于一片和 d^{gn} 邻接的可微区域内。下面的工作就是要确定所有 $\tau_i(d^*)$ 的符号($i = 1, 2, \dots, m$)。

首先, 我们给出一个和命题4.4.2相似的命题:

命题 4.4.3. 任意给定两个约束(4.4.26)的可行点 $p^{(1)}$ 和 $p^{(2)}$, 如果存在 $j \in \{1, 2, \dots, m\}$ 使得

$$\text{sign}(\tau_j(p^{(1)})) \cdot \text{sign}(\tau_j(p^{(2)})) = -1, \quad (4.4.40)$$

那么必定存在 $p^{(1)}$ 和 $p^{(2)}$ 的凸组合

$$tp^{(1)} + (1-t)p^{(2)}, \quad \exists t \in (0, 1) \quad (4.4.41)$$

属于I2或I3类点。

证明: (4.4.40)意味着 $\tau_j(p^{(1)})$ 与 $\tau_j(p^{(2)})$ 一正一负。由于 $\tau_j(tp^{(1)} + (1-t)p^{(2)})$ 在区间(0,1)内连续, 因此, 必定存在 $t^* \in (0, 1)$ 使得

$$\tau_j(t^*p^{(1)} + (1-t^*)p^{(2)}) = 0. \quad (4.4.42)$$

也即 $t^*p^{(1)} + (1-t^*)p^{(2)}$ 是目标函数的一个不可微点。由于 $t^*p^{(1)} + (1-t^*)p^{(2)}$ 也位于约束内部, 它属于I2或者I3类。□

下面我们给出一个引理:

引理 4.4.1. 假设矩阵 A 是半正定的, (4.4.28)-(4.4.30)的全局极小值点 d^{gn} 不满足KKT系统(4.4.27), 那么

- 如果 d^{gn} 属于I2或者I3类, 则我们可以找到I1类点 d^{pd} ;
- 如果 d^{gn} 属于B2类, 则我们可以找到I1或者B1类点 d^{pd} ;

满足 $\text{sign}(\tau_i(d^{pd})) = \text{sign}(\tau_i(d^{gn}))$, $\forall i \notin \Omega(d^{gn})$, 且 $\phi_{(2)}(d^{pd}) < \phi_{(2)}(d^{gn})$ (这里上标 pd 取自point which is differentiable)。

证明: 由于 d^{gn} 不满足KKT系统(4.4.27), 所以它不是(4.4.25)-(4.4.26)的一个局部极小点。根据定理4.4.1, 当 d^{gn} 属于I2或者I3类时, $s^{(2)}(d^{gn})$ 是目标函数 $\phi_{(2)}(d)$ 在 d^{gn} 处的下降方向; 当 d^{gn} 属于B2类时, $s^{(3)}(d^{gn})$ 是目标函数 $\phi_{(2)}(d)$ 在 d^{gn} 处的下降方向。

先考虑 d^{gn} 属于I2或者I3类时的情形。我们有

$$\{d | d = d^{gn} + \lambda s^{(2)}(d^{gn}), \lambda > 0\} \cap \{d | \|d\| \leq \Delta\} \neq \emptyset. \quad (4.4.43)$$

由于 d^{gn} 是(4.4.28)-(4.4.30), 我们还可推断出

$$s^{(2)}(d^{gn}) \notin \mathcal{Z}(d^{gn}). \quad (4.4.44)$$

因此必存在 $\epsilon > 0$, 使得对 $\forall t \in (0, \epsilon)$, 有

$$d^{new(2)}(t) = d^{gn} + ts^{(2)}(d^{gn}) \quad (4.4.45)$$

属于I1类。由于 $s^{(2)}(d^{gn})$ 是下降方向, 我们必可以找到一个充分小的正数 $\bar{t} > 0$ 满足 $\phi_{(2)}(d^{new(2)}(\bar{t})) < \phi_{(2)}(d^{gn})$ 。不妨记之为 $d^{pd} = d^{new(2)}(\bar{t})$ 。

下面我们考虑 d^{gn} 属于B2类的情形, 且

$$s^{(3)}(d^{gn})^T d^{gn} < 0. \quad (4.4.46)$$

此时,

$$\{d | d = d^{gn} + \lambda s^{(3)}(d^{gn}), \lambda > 0\} \cap \{d | \|d\| \leq \Delta\} \neq \emptyset \quad (4.4.47)$$

也是成立的, 因此我们可以类似地找到属于I1类的点 d^{pd} 满足

$$\phi_{(2)}(d^{pd}) < \phi_{(2)}(d^{gn}). \quad (4.4.48)$$

现在我们考虑 d^{gn} 属于B2类, 但是不满足 $s^{(3)}(d^{gn})^T d^{gn} < 0$ 的情形。根据定理4.4.1, 可知必有

$$s^{(3)}(d^{gn})^T d^{gn} = 0. \quad (4.4.49)$$

因此必存在 $\epsilon > 0$, 使得对 $\forall t \in (0, \epsilon)$, 都有

$$\begin{aligned} d^{new(3)}(t) &= \frac{\Delta}{\sqrt{\Delta^2 + t^2 \|s^{(3)}(d^{gn})\|_2^2}} \cdot d^{gn} \\ &\quad + \frac{t\Delta}{\sqrt{\Delta^2 + t^2 \|s^{(3)}(d^{gn})\|_2^2}} \cdot s^{(3)}(d^{gn}) \end{aligned} \quad (4.4.50)$$

属于B1类。由于 $s^{(3)}(d^{gn})$ 是一个下降方向, 我们可以找到充分小的正数 $\bar{t} > 0$ 满足

$$\phi_{(2)}(d^{new(3)}(\bar{t})) < \phi_{(2)}(d^{gn}). \quad (4.4.51)$$

我们记之为

$$d^{pd} = d^{new(3)}(\bar{t}). \quad (4.4.52)$$

无论 d^{gn} 使得上面哪种情形成立, 我们都已找到了 $\phi_{(2)}(d^{pd}) < \phi_{(2)}(d^{gn})$ 。假设结论

$$\text{sign}(\tau_i(d^{pd})) = \text{sign}(\tau_i(d^{gn})), \quad \forall i \notin \Omega(d^{gn}) \quad (4.4.53)$$

不成立, 不妨设 $\exists j \notin \Omega(d^{gn})$, 使得

$$\text{sign}(\tau_j(d^{pd})) \neq \text{sign}(\tau_j(d^{gn})). \quad (4.4.54)$$

那么根据命题4.4.3, 必存在 d^{pd} 和 d^{gn} 的凸组合 $\tilde{t}d^{pd} + (1 - \tilde{t})d^{gn}$ 属于I2或I3类点。由矩阵 A 半正定时, 函数 $\phi_{(2)}(d)$ 的凸性, 我们有

$$\phi_{(2)}(\tilde{t}d^{pd} + (1 - \tilde{t})d^{gn}) \leq \tilde{t}\phi_{(2)}(d^{pd}) + (1 - \tilde{t})\phi_{(2)}(d^{gn}) < \phi_{(2)}(d^{gn}). \quad (4.4.55)$$

这与 d^{gn} 是(4.4.28)-(4.4.30)的全局极小值点矛盾。因此假设不成立, 引理证毕。□

根据引理4.4.1, 我们可以得到如下定理:

定理 4.4.4. 假设矩阵 A 是半正定的, (4.4.28)-(4.4.30)的全局极小值点 d^{gn} 不满足KKT系统(4.4.27)。如果存在属于I1或B1类的点 d^{pd} 满足

$$\phi_{(2)}(d^{pd}) < \phi_{(2)}(d^{gn}), \quad (4.4.56)$$

则任意(4.4.25)-(4.4.26)的全局极小值点 d^* 满足

$$\text{sign}(\tau_i(d^*)) = \text{sign}(\tau_i(d^{pd})), \quad \forall i = 1, 2, \dots, m. \quad (4.4.57)$$

证明: 根据定理4.4.3, 我们知道 d^* 必属于B1或者I1类。根据命题4.4.3, 如果存在 $j \in \{1, 2, \dots, m\}$ 使得

$$\tau_j(d^*) \neq \tau_j(d^{pd}), \quad (4.4.58)$$

则必存在 d^{d^*} 和 d^{pd} 的凸组合

$$d^{pn} = t^*d^{d^*} + (1 - t^*)d^{pd} \quad (4.4.59)$$

属于I2或I3类。由矩阵 A 半正定时, 函数 $\phi_{(2)}(d)$ 的凸性, 我们有

$$\phi_{(2)}(d^{pn}) \leq t^*\phi_{(2)}(d^{d^*}) + (1 - t^*)\phi_{(2)}(d^{pd}) < \phi_{(2)}(d^{gn}), \quad (4.4.60)$$

这和 d^{gn} 是(4.4.28)-(4.4.30)的全局极小值点矛盾。因此

$$\text{sign}(\tau_i(d^*)) = \text{sign}(\tau_i(d^{pd})), \quad \forall i = 1, 2, \dots, m. \quad (4.4.61)$$

证毕。 □

情形二：矩阵 A 不半正定。

根据定理4.4.3，此时没有(4.4.25)-(4.4.26)的全局极小值点属于I1类。那么对于B1类有没有类似的结果呢？答案是否定的。尽管矩阵 A 不半正定意味着(4.4.25)没有可微的局部极小点，但是约束(4.4.26)会在边界上产生可微点。

由于 d^{gn} 是目标函数 $\phi_{(2)}(d)$ 在B2、I2和I3类内函数值最小的点。因此要么 d^{gn} 是(4.4.25)-(4.4.26)的一个全局极小值点，要么所有的全局极小值点都属于B1类。

所以下面我们考虑，如何找到所有(4.4.25)-(4.4.26)的属于B1类的全局极小点。由于此时定理4.4.3中的(4.4.36)不一定再成立，所以确定每个 τ_i 项的符号变得很困难。

于是我们引入如下的变换：

$$d_1 = r\Delta \sin \theta, \quad d_2 = r\Delta \cos \theta. \quad (4.4.62)$$

这时(4.4.25)-(4.4.26)等价于：

$$\begin{aligned} \min_{\substack{\theta \in [0, 2\pi] \\ r \in [0, 1]}} p(\theta, r) &= \sum_{i=1}^m |r\Delta(a_i \sin \theta + b_i \cos \theta) + c_i| + \frac{1}{2}r^2\Delta^2 \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}^T A \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix} \\ &\quad + r\Delta g^T \begin{pmatrix} \sin \theta \\ \cos \theta \end{pmatrix}; \end{aligned} \quad (4.4.63)$$

由于我们想找到所有B1类中的局部极小点，所以我们只考虑(4.4.63)在 $r = 1$ 时的情形，也即只考虑极小化 $p(\theta, 1)$ 。我们记

$$r_i(\theta) = \Delta(a_i \sin \theta + b_i \cos \theta) + c_i = \Delta\sqrt{a_i^2 + b_i^2} \sin(\theta + \arctan \frac{b_i}{a_i}) + c_i. \quad (4.4.64)$$

下面给出我们的定理。

命题 4.4.4. 假设 α 和 β 是(4.4.63)的目标函数 $p(\theta, 1)$ 的两个可微点，并且 $0 \leq \alpha < \beta \leq 2\pi$ 。如果存在 $j \in \{1, 2, \dots, m\}$ 满足

$$\text{sign}(r_j(\alpha)) \cdot \text{sign}(r_j(\beta)) = -1, \quad (4.4.65)$$

则必存在位于区间 (α, β) 中的 $p(\theta, 1)$ 的不可微点。

证明: 由于 α 和 β 可微, 且

$$\text{sign}(r_j(\alpha)) \cdot \text{sign}(r_j(\beta)) = -1, \quad (4.4.66)$$

所以我们必有 $r_j(\alpha)$ 和 $r_j(\beta)$ 一正一负。又因为 $r_j(\theta)$ 在区间 $[\alpha, \beta]$ 中连续, 根据中值定理, 我们知道必存在 $\gamma \in (\alpha, \beta)$ 满足 $r_j(\gamma) = 0$ 。因此, γ 是 $p(\theta, 1)$ 位于区间 (α, β) 内的不可微点。证毕。 \square

我们知道每一个B1和B2类的点都和 $\theta \in [0, 2\pi)$ 一一对应, 并且相应的 $\phi_{(2)}(d)$ 和 $p(\theta, 1)$ 函数值也对应相等。我们的目标是找到(4.4.25)-(4.4.26)属于B1类的一个全局极小值点。实现方法是先找到 $p(\theta, 1)$ 的一个全局极小值点, 并找到对应的原空间的点, 如果该点是属于B1类的, 则我们任务完成。最后只要拿它和 d^{gn} 比较大, 即可找到(4.4.25)-(4.4.26)的一个全局极小值点; 如果该点是属于B2类的, 那我们就没有必要找属于B2类的全局极小值了, 因为此时 d^{gn} 已经是(4.4.25)-(4.4.26)的一个全局极小值点了。

下面我们就来找 $p(\theta, 1)$ 的一个全局极小值点。根据命题4.4.4, 所有位于 $p(\theta, 1)$ 的两个邻接不可微点之间的可微点都具有相同符号的 $r_i (i = 1, 2, \dots, m)$ 。由于区间 $[0, 2\pi)$ 的不可微点总数不超过 $2m$ 个(和原目标函数 $\phi_{(2)}(d)$ 的B2类点一一对应), 于是加上两边端点, 区间 $[0, 2\pi)$ 内最多只被划分为 $2m + 1$ 段区间, 在每段子区间内, 所有可微点的 $r_i (\forall i = 1, 2, \dots, m)$ 的符号都不变, 因此可计算任意一个内点的 $r_i (i = 1, 2, \dots, m)$ 符号来确定区间内所有可微点的符号。这样KKT系统(4.4.27)就等价于一个四次方程。

下面给出我们的整个算法。

算法 4.4.2. (二维非光滑信赖域子问题的直接算法)

步一 (B2、I2、I3类点)

- i) 判断矩阵 A 是否半正定(根据 A_{11} 、 A_{22} 和 $\det(A)$ 的正负号);
- ii) 对 $j = 1, 2, \dots, m$, 根据算法4.4.1求解(4.4.34)-(4.4.35), 分别得到最优解 $d^{n(j)}$;
- iii) 枚举法求得

$$k = \arg \min_{j=1,2,\dots,m} f(d^{n(j)}), \quad (4.4.67)$$

记 $d^{gn} := d^{n(k)}$;

- iv) 判断 d^{gn} 是否满足KKT条件(4.4.27);

- v) 根据定理4.4.3, 如果矩阵 A 半正定, 且 d^{gn} 满足 KKT 系统, 那么记 $d^* := d^{gn}$ 转步四; 如果矩阵 A 半正定, 但 d^{gn} 不满足 KKT 系统, 则转步二; 否则(矩阵 A 不半正定), 则转步三。

步二 (I1类点)

- i) 根据定义4.4.7, 判定 d^{gn} 属于哪一类点; 如果 d^{gn} 属于I2或I3类点, 则转步iii); 否则(d^{gn} 属于B2类点), 则转步iii);
- ii) 根据定义4.4.5, 计算 $s := s^{(3)}(d^{gn})$; 求解

$$\bar{t} = \min_{j \notin \Omega(d^{gn})} \frac{|\beta_j^T d^{gn} + c_j|}{2\beta_j^T s}; \quad (4.4.68)$$

如果 $s^T d^{gn} < 0$, 则转步iv); 否则转步v);

- iii) 根据定义4.4.3, 计算 $s := s^{(2)}(d^{gn})$; 求解

$$\bar{t} = \min_{j \notin \Omega(d^{gn})} \frac{|\beta_j^T d^{gn} + c_j|}{2\beta_j^T s}; \quad (4.4.69)$$

转步v);

- iv) 计算

$$d^{new}(\bar{t}) := d^{gn} + \bar{t}s; \quad (4.4.70)$$

如果 $\phi_{(2)}(d^{new}(\bar{t})) < \phi_{(2)}(d^{gn})$, 则置 $d^{pd} := d^{new}(\bar{t})$, 转步vi), 否则更新 $\bar{t} := \frac{1}{2}\bar{t}$, 转步iv);

- v) 根据(4.4.50)计算 $d^{new}(\bar{t})$; 如果 $\phi_{(2)}(d^{new}(\bar{t})) < \phi_{(2)}(d^{gn})$, 则置 $d^{pd} := d^{new}(\bar{t})$, 转步vi), 否则更新 $\bar{t} := \frac{1}{2}\bar{t}$, 转步iv)。

- vi) 求解下述二维光滑信赖域子问题

$$\min_{d \in \mathbf{R}^2} \sum_{i=1}^m \frac{1}{2} d^T A_i d + (g + \sum_{i=1}^m \text{sign}(d^{pd}) \cdot \beta_i)^T d; \quad (4.4.71)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta. \quad (4.4.72)$$

得到最优解 d^* , 转步四。

步三 (B1类点)

- i) 置 $d^* := d^{gn}$, $f^* := \phi_{(2)}(d^{gn})$; 令 $i := 1$, $j := 0$, $k := 0$; 令 $SEQ := \emptyset$;

- ii) 若 $i > m$, 则转于步 v); 否则在 $[0, 2\pi]$ 内求解 $r_i(\theta) = 0$; 若无解, 令 $i := i + 1$, 转于步 ii); 如果只有一个解, 记 $k := 1$; 否则记 $k := 2$;
- iii) 将所有解置入升序列 SEQ ; 更新 $j := |SEQ|$, $i := i + 1$, 转回于步 ii);
- iv) 若 $\theta = 0$ 或 $\theta = 2\pi$ 不在序列 SEQ 内, 则将其置入升序列 SEQ ; 更新 $j := |SEQ|$;
- v) 若 $j = 1$, 则转步四; 否则置 $\theta^c := 0.5 * (SEQ[j] + SEQ[j - 1])$;
- vi) 求解如下系统:

$$\begin{cases} (A + \lambda I) \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} + g + \sum_{i=1}^m \text{sign}(r_i(\theta^c)) \begin{pmatrix} a_i \\ b_i \end{pmatrix} = 0; & (i) \\ d_1^2 + d_2^2 = \Delta^2. & (ii) \end{cases} \quad (4.4.73)$$

得到 k 个解 $d^{pb(l)}$, $l = 1, \dots, k$ (这里上标 pb 取自 *point located on the bound*) 都满足

$$\text{sign}(\tau_i(d^{pb(l)})) = \text{sign}(r_i(\theta^c)), \quad \forall i = 1, 2, \dots, m; \quad (4.4.74)$$

- 若 $k > 0$, 则转于步 vii); 否则令 $j := j - 1$, 转于步 v);
- vii) 若 $\phi_{(2)}(d^{pb(k)}) < f^*$, 则记 $d^* := d^{pb(k)}$, $f^* := \phi_{(2)}(d^{pb(k)})$; 令 $k := k - 1$, 若 $k > 0$, 则转于步 vii); 否则令 $j := j - 1$, 转于步 v).

步四 返回 d^* , 并停止。

我们容易分析出, 算法中步一的子步 ii) 的计算复杂度是 $O(m^3)$; 算法其它部分的计算复杂度都不超过 $O(m^2)$ 。因此算法 4.4.2 的计算复杂度是 $O(m^3)$ 。

§4.5 序列二维子空间轮换搜索算法及收敛性分析

§4.5.1 方法概述

NEWUOA 是 Powell [108] 最近提出的求解无导数优化问题 (Derivative Free Optimization, 也叫直接优化问题) 的一种非常有效的算法。它的迭代过程中, 需要不断地求解信赖域子问题, 因此需要一种能够快速求得信赖域子问题 (4.3.3)-(4.3.4) 较为精确的解的算法。(4.3.3)-(4.3.4) 存在多项式时间复杂度的精确算法, 但计算量比较大, 因此一些比较高效的不精确算法被提出。我们在第二章中提到的截断共轭梯度法对于凸二次模型具有很好的性质 ([143]); 对于一般的二次模型, 这些好的性质不再成立。当共轭梯度法截断时, Powell 在 [108] 中提出了如下方法作为后处理技巧。

算法 4.5.1. (Powell的序列二维子空间技巧)

步一 使用截断共轭梯度法 [13]求解(4.3.3)-(4.3.4), 得到 $d^{(0)}$; 置 $k := 0$, 转步二。

步二 如果 d_k 与 $s_k = -(Ad_k + g)$ 的夹角接近零, 则算法停止; 否则转步三。

步三 求解(4.3.3)-(4.3.4)限制在二维子空间 $\text{span}\{d_k, s_k\}$ 中的全局极小点(等价于求解一个四次方程), 得到 d_{k+1} , 转步四。

步四 置 $k := k + 1$, 转步二。

尽管大量的数值实验显示, 经过算法4.5.1后处理的截断共轭梯度法总是能收敛到信赖域子问题的全局较小, 但是 [108]没有给出算法4.5.1的收敛性分析。

定理4.3.1告诉我们(4.3.1)-(4.3.2) 不存在多项式时间算法, 除非 $P=NP$ 。通常我们可以通过引入一些新的变量和不等式, 来转化那些 L_1 项如下式:

$$\min_{d \in \mathbf{R}^n, z \in \mathbf{R}^m} \quad \tilde{\phi}(d, z) = \frac{1}{2}d^T A d + g^T d + \sum_{i=1}^m z_i; \quad (4.5.1)$$

$$\text{s.t.} \quad c_i + \beta_i^T d \leq z_i, \quad \forall i = 1, 2, \dots, m; \quad (4.5.2)$$

$$c_i + \beta_i^T d \geq -z_i, \quad \forall i = 1, 2, \dots, m; \quad (4.5.3)$$

$$\|d\|_2 \leq \Delta_k. \quad (4.5.4)$$

但是, 仅仅是二次规划问题(4.5.1)-(4.5.3)就很难求解了(求解局部极小也是NP-难问题)。而且这种转化引入了过多的变量和约束, 特别是对于 $m \gg n$ 的情况。

所以我们希望借鉴Powell的算法4.5.1中的序列二维子空间极小技巧, 来设计求解非光滑信赖域子问题(4.3.1)-(4.3.2)的算法。

在我们的算法中, 当信赖域约束(4.3.2)不积极时, 我们选取二维子空间:

$$\text{span}\{s^{(1)}(d), s^{(2)}(d)\}; \quad (4.5.5)$$

当信赖域约束(4.3.2)积极时, 我们选取二维子空间:

$$\text{span}\{d, s^{(3)}(d)\}. \quad (4.5.6)$$

为了加快算法收敛, 我们还将引入二维子空间极小与一维非精确搜索交替轮换的技巧。这里的一维非精确搜索是指仅在非光滑点集上做搜索。

于是我们给出序列二维子空间轮换搜索算法如下:

算法 4.5.2. (序列二维子空间轮换搜索算法)

步一 初始点 $d^{(0)} = 0$; 置 $k := 0$; 算法终止参数 $0 < \varepsilon \ll 1$; 转步二。

步二 计算

$$s_k = \begin{cases} s^{(2)}(d_k), & \text{若 } \|d_k\| < \Delta; \\ s^{(3)}(d_k), & \text{若 } \|d_k\| = \Delta. \end{cases} \quad (4.5.7)$$

转步三。

步三 若 $\|s_k\| < \varepsilon$, 则算法停止; 否则, 转步四。

步四 若 $\|d_k\| < \Delta$, 则转步五; 否则, 转步六。

步五 计算

$$\bar{s}_k = s^{(1)}(d_k). \quad (4.5.8)$$

更新 $\mathcal{Y}_k = \text{span}\{\bar{s}_k, s_k\}$, 转步八。

步六 取随机数 r_k 服从 $[0, 1]$ 内的均匀分布, 如果 $r_k \leq 0.5$, 则更新 $\mathcal{Y}_k = \text{span}\{d_k, s_k\}$, 转步十(二维子空间极小); 否则, 转步七。

步七 根据(4.5.8)计算 \bar{s}_k , 若 $1 - \frac{\bar{s}_k^T d_k}{\|\bar{s}_k\| \|d_k\|} < \varepsilon$, 更新 $\mathcal{Y}_k = \text{span}\{d_k, s_k\}$, 转步十, 否则, $\mathcal{Y}_k = \text{span}\{d_k, \bar{s}_k\}$, 转步十一(一维非精确搜索)。

步八 若 $\text{rank}(\mathcal{Y}_k) = 2$, 则转步九; 否则, 随机选取 s_k 的零空间中的一个单位向量 z_k ; 更新 $\mathcal{Y}_k = \text{span}\{s_k, z_k\}$, 若 $\|\bar{s}_k\| < \varepsilon$, 转步十(二维子空间极小), 否则转步九。

步九 取随机数 r_k 服从 $[0, 1]$ 内的均匀分布, 如果 $r_k \leq 0.5$, 转步十(二维子空间极小); 否则, 转步十一(一维非精确搜索)。

步十 通过算法4.4.2求解:

$$\min \quad \phi(d) = \|Bd + c\|_1 + \frac{1}{2} d^T A d + g^T d; \quad (4.5.9)$$

$$\text{s.t.} \quad \|d\|_2 \leq \Delta, \quad (4.5.10)$$

$$d \in \mathcal{Y}_k. \quad (4.5.11)$$

得到全局极小点 d^* , 记 $d_{k+1} := d^*$, 转步十二。

步十一 求

$$\min \quad \bar{\phi}(t) = \|B(d_k + t\bar{s}_k) + c\|_1 + \frac{1}{2}(d_k + t\bar{s}_k)^T A(d_k + t\bar{s}_k) + g^T(d_k + t\bar{s}_k); \quad (4.5.12)$$

$$\text{s.t.} \quad t \in \bar{T}_k. \quad (4.5.13)$$

得极小点 t_k , 其中:

$$T_k = \{t \mid t = \frac{|\beta_j^T d_k + c_j|}{\beta_j^T \bar{s}_k}, \forall j \notin \Omega(d_k)\}; \quad (4.5.14)$$

$$\bar{T}_k = T_k \cap \left(0, \frac{\Delta}{\|d_k\|_2}\right] \cup \left\{\frac{\Delta}{\|d_k\|_2}\right\}. \quad (4.5.15)$$

如果可行集 $\bar{T}_k = \emptyset$, 或者得到的解 $\bar{d} = d_k + t_k \bar{s}_k$ 使得 $f(\bar{d}) \geq f(d_k)$, 转步十进行二维子空间极小化; 否则, 记 $d_{k+1} := \bar{d}$, 转步十二。

步十二 置 $k := k + 1$, 转步二。

§4.5.2 收敛性分析

本小节分析算法4.5.2在凸性假设下的收敛性。即在本小节中, 我们引入如下假设:

假设 4.5.1. (4.3.1)的目标函数中的矩阵 A 半正定。

定义 4.5.1. 本小节中我们用 Λ_A 表示矩阵 A 的最大特征值。

由于假设4.5.1成立, 我们可以设没有约束的情况下, (4.3.1)的解为 d^* 。我们认为算法4.5.2的收敛性分析主要包括如下两部分:

第一部分是信赖域约束(4.3.2)不积极的情形, 也即 $\|d^*\| \leq \Delta$ 。此时我们需要回答的问题是对于无约束的非光滑优化问题(4.3.1), 我们给出的序列二维子空间轮换搜索算法4.5.2能否得到充分的下降量, 并由此建立全局收敛性。

第二部分是信赖域约束(4.3.2)积极的情况, 也即 $\|d^*\| > \Delta$ 。如果第一部分得到了肯定的回答, 剩下的问题就是算法4.5.1的收敛性问题了。

因此本节我们将只对第一部分作分析。关于算法4.5.1的收敛性和收敛速度分析, 我们将在下一章的第三节详细介绍。事实上, 有了本节和下一章第三节的结果, 我们不难完成对第二部分的收敛性分析。

下面我们考虑第一部分的收敛性。这时子空间的构造主要是基于 $s^{(1)}(d_k)$ 和 $s^{(2)}(d_k)$ 。我们首先分别给出这两个方向 $s^{(i)}(d_k)$ ($i = 1, 2$)的下降性估计。为了方便分析, 我们记 $s_k^{(i)} = s^{(i)}(d_k)$ ($i = 1, 2$), 并给出如下的定义:

定义 4.5.2.

$$\bar{t}(d_k, s_k) = \begin{cases} \min_{i \notin \Omega(d_k)} \max\{-\frac{\beta_i^T d_k}{\beta_i^T s_k}, 0\}, & \text{如果 } \exists i \notin \Omega(d_k) \text{ 满足 } -\frac{\beta_i^T d_k}{\beta_i^T s_k} > 0; \\ +\infty, & \text{否则.} \end{cases} \quad (4.5.16)$$

我们记 $\bar{t}_k^{(i)} = \bar{t}(d_k, s^{(i)}(d_k)) (i = 1, 2)$ 。

引理 4.5.1. 假设4.5.1成立, 则当 $s^{(1)}(d_k) \neq 0$ 时, 必有:

$$f(d_k) - f(d_k + t_k^* s_k^{(1)}) \geq \varpi_1(d_k) \|s_k^{(1)}\|_2^2. \quad (4.5.17)$$

这里,

$$\varpi_1(d_k) = \begin{cases} \frac{1}{2\Lambda_A}, & \text{如果 } \frac{1}{\Lambda_A} \leq \bar{t}_k^{(1)}; \\ (\bar{t}_k^{(1)} - \frac{\Lambda_A}{2} (\bar{t}_k^{(1)})^2), & \text{否则.} \end{cases} \quad (4.5.18)$$

$$t_k^* = \arg \min_{t \geq 0} f(d_k + t s_k^{(1)}). \quad (4.5.19)$$

当 $s^{(1)}(d_k) = 0$ 时, d_k 必是

$$\min_{d \in \mathbf{R}^n} \phi(d) = \sum_{i \notin \Omega(d_k)} |\beta_i^T d + c_i| + \frac{1}{2} d^T A d + g^T d; \quad (4.5.20)$$

$$\text{s.t. } \beta_i^T d + c_i = 0, \quad i \in \Omega(d_k) \quad (4.5.21)$$

唯一的局部极小点。

证明: 为了使证明过程简洁, 我们暂时简化记号:

$$s_k = s^{(1)}(d_k); \quad \bar{t}_k = \bar{t}(d_k, s^{(1)}(d_k)). \quad (4.5.22)$$

我们首先考虑 $s^{(1)}(d_k) \neq 0$ 的情形。我们知道对任意 $t \in (0, \bar{t}(d_k, s_k)]$, 我们有:

$$\begin{aligned} f(d_k) - f(d_k + t s_k) &= - \sum_{i \in \Omega(d_k)} |t \beta_i^T s_k| + \sum_{j \notin \Omega(d_k)} (|\beta_j^T d_k + c_j| - |\beta_j^T d_k + c_j + t \beta_j^T s_k|) \\ &\quad + (\frac{1}{2} s_k^T A s_k + g^T d_k - \frac{1}{2} (d_k + t s_k)^T A (d_k + t s_k) - g^T (d_k + t s_k)) \\ &= - \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j^T s_k \cdot t \\ &\quad - (A d_k + g)^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\ &= -\xi(d_k)^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\ &= s_k^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\ &\geq \|s_k\|_2^2 \cdot t - \frac{1}{2} \Lambda_A \|s_k\|_2^2 \cdot t^2. \end{aligned} \quad (4.5.23)$$

如果

$$\frac{1}{\Lambda_A} \leq \bar{t}_k, \quad (4.5.24)$$

则当 $t = \frac{1}{\Lambda_A}$ 时,

$$\|s_k\|_2^2 \cdot t - \frac{1}{2}\Lambda_A\|s_k\|_2^2 \cdot t^2 = \frac{1}{2\Lambda_A}\|s_k\|_2^2. \quad (4.5.25)$$

如果

$$\frac{1}{\Lambda_A} > \bar{t}_k, \quad (4.5.26)$$

则当 $t = \bar{t}_k$ 时,

$$\|s_k\|_2^2 \cdot t - \frac{1}{2}\Lambda_A\|s_k\|_2^2 \cdot t^2 = (\bar{t}_k - \frac{\Lambda_A}{2}\bar{t}_k^2)\|s_k\|_2^2. \quad (4.5.27)$$

根据(4.5.23)、(4.5.25)及(4.5.27), 我们可得(4.5.17)。

下面考虑 $s^{(1)}(d_k) = 0$ 的情况, 我们可以写出(4.5.20)-(4.5.21)的KKT条件:

$$\xi(d) = \sum_{i \in \Omega(d_k)} \sigma_i \beta_i; \quad (4.5.28)$$

$$\beta_i^T d + c_i = 0, \quad i \in \Omega(d_k). \quad (4.5.29)$$

由于 $s^{(1)}(d_k) = 0$, 我们发现 d_k 满足(4.5.28)-(4.5.29)。又由假设4.5.1成立, 所以 d_k 是(4.5.20)-(4.5.21)唯一的局部极小值点。证毕。 \square

引理 4.5.2. 假设4.5.1成立, 则当 $s^{(2)}(d_k) \neq 0$ 时, 必有:

$$f(d_k) - f(d_k + t_k^* s_k^{(2)}) \geq \varpi_2(d_k) \|s_k^{(2)}\|_2^2. \quad (4.5.30)$$

这里

$$\varpi_2(d_k) = \begin{cases} \frac{1}{2\Lambda_A}, & \text{如果 } \frac{1}{\Lambda_A} \leq \bar{t}_k^{(2)}; \\ (\bar{t}_k^{(2)} - \frac{\Lambda_A}{2}(\bar{t}_k^{(2)})^2), & \text{否则.} \end{cases} \quad (4.5.31)$$

$$t_k^* = \arg \min_{t \geq 0} f(d_k + t s_k^{(2)}). \quad (4.5.32)$$

当 $s^{(2)}(d_k) = 0$ 时, 则 d_k 必是

$$\min_{d \in \mathbf{R}^n} \phi(d) = \sum_{i \notin \Omega(d_k)} |\beta_i^T d + c_i| + \frac{1}{2} d^T A d + g^T d \quad (4.5.33)$$

唯一的局部极小点。

证明： 为了使证明过程简洁，我们暂时简化记号：

$$s_k = s^{(2)}(d_k); \quad \bar{t}_k = \bar{t}(d_k, s^{(2)}(d_k)). \quad (4.5.34)$$

下面我们首先考虑 $s^{(2)}(d_k) \neq 0$ 的情形。

设 $s_k = - \sum_{i \in \Omega(d_k)} \eta_i^* \beta_i - \xi(d_k)$ 。写出当 $d = d_k$ 时带界约束的线性最小二乘问题(4.4.1)-(4.4.2)的KKT系统，也即 $\eta_i^*(i \in \Omega(d_k))$ 需要满足的方程：

$$\begin{cases} \left(\sum_{i \in \Omega(d_k)} \eta_i^* \beta_i + \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j + A d_k + g \right)^T \beta_i = \lambda_i^* - \sigma_i^*; \\ \lambda_i^* (\eta_i^* + 1) = 0; \\ \sigma_i^* (1 - \eta_i^*) = 0; \\ -1 \leq \eta_i^* \leq 1; \\ \lambda_i^* \geq 0, \sigma_i^* \geq 0. \end{cases} \quad (4.5.35)$$

这里 λ_i^* 与 σ_i^* 是约束(4.4.2)的Lagrange乘子。根据(4.5.35)我们易知有下列两式成立：

$$\beta_i^T s_k < 0 \Rightarrow \eta_i^* = -1, \quad \forall i \in \Omega(d); \quad (4.5.36)$$

$$\beta_i^T s_k > 0 \Rightarrow \eta_i^* = 1, \quad \forall i \in \Omega(d). \quad (4.5.37)$$

于是对任意 $t \in (0, \bar{t}(d_k, s_k)]$, 我们有:

$$\begin{aligned}
f(d_k) - f(d_k + ts_k) &= - \sum_{i \in \Omega(d_k)} |t\beta_i^T s_k| + \sum_{j \notin \Omega(d_k)} (|\beta_j^T d_k + c_j| - |\beta_j^T d_k + c_j + t\beta_j^T s_k|) \\
&\quad + \left(\frac{1}{2} s_k^T A s_k + g^T d_k - \frac{1}{2} (d_k + ts_k)^T A (d_k + ts_k) - g^T (d_k + ts_k) \right) \\
&= - \sum_{i \in \Omega(d_k)} |\beta_i^T s_k \cdot t| - \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j^T s_k \cdot t \\
&\quad - (A d_k + g)^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\
&= - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \beta_i^T s_k \cdot t + \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \beta_i^T s_k \cdot t \\
&\quad - \xi(d_k)^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\
&= \left[\sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \beta_i - \xi(d_k) \right]^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\
&= \left[- \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \eta_i \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k = 0}} \eta_i \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \eta_i \beta_i \right. \\
&\quad \left. - \xi(d_k) \right]^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\
&= s_k^T s_k \cdot t - \frac{1}{2} s_k^T A s_k \cdot t^2 \\
&\geq \|s_k\|_2^2 \cdot t - \frac{1}{2} \Lambda_A \|s_k\|_2^2 \cdot t^2. \tag{4.5.38}
\end{aligned}$$

如果

$$\frac{1}{\Lambda_A} \leq \bar{t}_k, \tag{4.5.39}$$

则当 $t = \frac{1}{\Lambda_A}$ 时,

$$\|s_k\|_2^2 \cdot t - \frac{1}{2} \Lambda_A \|s_k\|_2^2 \cdot t^2 = \frac{1}{2\Lambda_A} \|s_k\|_2^2. \tag{4.5.40}$$

如果

$$\frac{1}{\Lambda_A} > \bar{t}_k, \tag{4.5.41}$$

则当 $t = \bar{t}_k$ 时,

$$\|s_k\|_2^2 \cdot t - \frac{1}{2} \Lambda_A \|s_k\|_2^2 \cdot t^2 = (\bar{t}_k - \frac{\Lambda_A}{2} \bar{t}_k^2) \|s_k\|_2^2. \tag{4.5.42}$$

根据(4.5.38)、(4.5.40)及(4.5.42)，我们可得(4.5.30)。

当 $s^{(2)}(d_k) = 0$ 时，根据假设4.5.1和定理4.4.1，显然有 d_k 是(4.5.33)唯一的局部极小值点。证毕。 \square

根据上面引理4.5.1和引理4.5.2的证明，我们可以解释引入定义4.5.2的原因了。对于非光滑问题(4.3.1)，给定一个下降方向 s_k ，我们并不能估计沿此方向的最优下降步长的下降量。因为当步长因子 $t' = \bar{t}(d_k, s_k)$ 时，迭代 $d' = d_k + t's_k$ 使得积极非光滑面的下标集合 Ω 发生了改变，也即 $\Omega(d') \neq \Omega(d_k)$ 。所以当步长因子 $t' > \bar{t}(d_k, s_k)$ 时， s_k 就可能不是下降方向了。所以我们只能估计步长属于 $(0, \bar{t}(d_k, s_k)]$ 内的下降量。

对于非光滑优化来说，给定一个下降方向，精确线搜索算法可能不收敛的原因就是随着迭代趋于无穷，有可能是 $\bar{t}(d_k, s_k)$ 趋于零，而不是下降方向 s_k 的范数趋于零。这也是建立非光滑优化搜索算法全局收敛性的瓶颈所在。

认识到了这个本质的困难，我们构造算法时才考虑同时引入两个下降方向： $s^{(1)}(d_k)$ 和 $s^{(2)}(d_k)$ 。通过上述引理，我们可以估计二维子空间精确搜索的下降量：

推论 4.5.1. 假设4.5.1成立，且 $s^{(2)}(d_k) \neq 0$ 。当 $r_k \leq 0.5$ 时，必有：

$$f(d_k) - f(d_{k+1}) \geq \max\{\varpi_1(d_k)\|s^{(1)}(d_k)\|_2^2, \varpi_2(d_k)\|s^{(2)}(d_k)\|_2^2\}. \quad (4.5.43)$$

这里 $\varpi_1(d_k)$ 和 $\varpi_2(d_k)$ 分别由(4.5.18)和(4.5.31)定义。

证明： 由算法4.5.2及引理4.5.1和引理4.5.2立得。证毕。 \square

为了防止算法在一个非积极非光滑面 $\beta_j^T d + c_j = 0 (j \notin \Omega(d_k))$ 附近出现zigzag现象，我们特别引入了一维非精确搜索的轮换。非精确搜索只是为了算法能够更快地更新积极非光滑指标集，并不需保证充分下降，于是我们有如下下降量估计：

推论 4.5.2. 假设4.5.1成立，且 $s^{(2)}(d_k) \neq 0$ 。当 $r_k > 0.5$ 时，如果 d_{k+1} 由算法4.5.2步十一计算，则必有：

$$f(d_k) - f(d_{k+1}) > 0, \quad (4.5.44)$$

如果 d_{k+1} 由算法4.5.2步十计算，则必有：

$$f(d_k) - f(d_{k+1}) \geq \frac{1}{2\Lambda_A}\|s^{(1)}(d_k)\|_2^2. \quad (4.5.45)$$

证明: 根据算法4.5.2, 当 $r_k > 0.5$ 时, 如果 d_{k+1} 由算法4.5.2步十一计算, 显然有(4.5.44)成立。如果 d_{k+1} 由算法4.5.2步十计算, 无论是可行域 \bar{T}_k 无界, 还是 $f(\bar{d}) \geq f(d_k)$, 都可推出目标函数 f 沿下降方向 $s^{(1)}(d_k)$ 的精确搜索, 不会受到非积极非光滑面的约束, 可以走一个完整的下降步。也即:

$$\{d \mid d = d_k + ts^{(1)}(d_k), \forall t \in [0, t_k^*]\} \cap \{d \mid \beta_i^T d + c_i = 0, \exists i \notin \Omega(d_k)\} = \emptyset. \quad (4.5.46)$$

再由引理4.5.1立得(4.5.45)。证毕。 \square

推论4.5.1和推论4.5.2, 是构建算法4.5.2全局收敛性的重要基础。如果我们能够保证算法4.5.2的下降量在无穷多个迭代步达到

$$f(d_k) - f(d_{k+1}) > \bar{\omega} \|s^{(2)}(d_k)\|^2, \quad (4.5.47)$$

其中下降因子 $\bar{\omega}$ 是一个正常数, 那么自然就可保证算法4.5.2的全局收敛性。于是下一个待解决的问题就是找到 $\bar{\omega}_2(d_k)$ 在某些情形下的常数下界。

为此, 我们再引入如下几个定义:

定义 4.5.3.

$$V(d) = \begin{cases} \min_{i \notin \mathcal{S}(d)} \frac{|\beta_i^T d + c_i|}{\|\beta_i\|_2}, & \text{若 } \{i \mid i \notin \mathcal{S}(d)\} \neq \emptyset; \\ +\infty, & \text{否则.} \end{cases} \quad (4.5.48)$$

定义 4.5.4.

$$\mathcal{G} = \{\Omega \mid \{d \mid \Omega(d) = \Omega, \forall d \in \mathbf{R}^n\} \neq \emptyset, \forall \Omega \subset \{1, 2, \dots, m\}\}. \quad (4.5.49)$$

定义 4.5.5. 我们定义 d_Ω^* ($\Omega \in \mathcal{G}$)为如下优化问题的局部极小点:

$$\min_{d \in \mathbf{R}^n} \phi(d) = \sum_{i \notin \Omega} |\beta_i^T d + c_i| + \frac{1}{2} d^T A d + g^T d; \quad (4.5.50)$$

$$\text{s.t. } \beta_i^T d + c_i = 0, \quad i \in \Omega. \quad (4.5.51)$$

下面证明一个重要的引理。

引理 4.5.3. 对某个 $\Omega \in \mathcal{G}$, 如果对应的 d_Ω^* 不是(4.3.1)的局部极小点, 则有:

$$\bar{t}(d_\Omega^*, s^{(2)}(d_\Omega^*)) \geq \begin{cases} \frac{V(d_\Omega^*)}{\|s^{(2)}(d_\Omega^*)\|_2}, & \text{如果 } V(d_\Omega^*) = +\infty; \\ \frac{1}{\Lambda_A}, & \text{否则.} \end{cases} \quad (4.5.52)$$

证明: 由于 d_Ω^* 不是(4.3.1)的局部极小点, 根据定理4.4.1, $s^{(2)}(d_\Omega^*)$ 是 $\phi(d)$ 在 d_Ω^* 点的一个下降方向。如果此时存在 $i \notin \Omega(d_k)$ 满足

$$-\frac{\beta_i^T d_\Omega^*}{\beta_i^T s^{(2)}(d_\Omega^*)} > 0, \quad (4.5.53)$$

则由问题(4.3.1)确定的常数(与迭代点无关)满足:

$$V(d_\Omega^*) < +\infty. \quad (4.5.54)$$

且由定义4.5.2和定义4.5.3, 有下式成立:

$$\bar{t}(d_\Omega^*, s^{(2)}(d_\Omega^*)) \geq \frac{V(d_\Omega^*)}{\|s^{(2)}(d_\Omega^*)\|_2}. \quad (4.5.55)$$

如果不存在 $i \notin \Omega(d_k)$ 满足(4.5.53), 则

$$\bar{t}(d_\Omega^*, s^{(2)}(d_\Omega^*)) = +\infty > \frac{1}{\Lambda_A}. \quad (4.5.56)$$

综上所述, (4.5.52)成立。证毕。 \square

推论 4.5.3. 对某个 $\Omega \in \mathcal{G}$, 如果对应的 d_Ω^* 不是(4.3.1)的局部极小点, 则有

$$\varpi_2(d_\Omega^*) \geq (q - \frac{\Lambda_A}{2}q^2); \quad (4.5.57)$$

这里,

$$q = \min\{V(d_\Omega^*), \frac{1}{\Lambda_A}\}. \quad (4.5.58)$$

证明: 由引理4.5.2和引理4.5.3立得。证毕。 \square

推论4.5.3给了 $\varpi_2(d_\Omega^*)$ 的常数下界。

如果迭代序列 $\{d_k\}$ 存在一个具有共同的积极非光滑指标集的收敛子列, 并趋近于一个仍具有此积极非光滑指标集的点时, 我们可以通过推论4.5.3和下面即将给出的引理来构建这种情形下, 算法4.5.2的收敛性。

引理 4.5.4. 对于给定的 d , 我们记:

$$\mathcal{U}(d) = \mathcal{Q}(d) \cap \{d' \mid \|d' - d\|_2 \leq \frac{1}{2}V(d)\}. \quad (4.5.59)$$

则必存在一个小正数 $\delta > 0$, 及常数 $\rho_1 > 0$ 和 $\rho_2 > 0$ 使得

$$\|s^{(i)}(d') - s^{(i)}(d)\|_2 \leq \rho_i \|d' - d\|_2, i = 1, 2 \quad (4.5.60)$$

对任意 $d' \in N(d, \delta) \cap \mathcal{U}(d)$ 成立。

证明: 由(4.5.59), 我们知道:

$$\Omega(d') = \Omega(d); \quad \forall d' \in N(d, \varepsilon) \cap \mathcal{U}(d). \quad (4.5.61)$$

于是, $s^{(i)}(d)$ 和 $s^{(i)}(d')$ 所对应的 η_i^* ($i \in \Omega(d)$, 参见定义4.4.3)是同一个凸二次规划问题的解, 差别就是系数 $\xi(d)$ 的一个扰动。我们知道凸二次规划问题的KKT系统是一个线性互补问题(参阅Han, Xiu和Qi [159]), 根据线性互补问题的稳定性理论(可参阅Cottle, Pang和Stone [15]), 必有式(4.5.60)成立。证毕。 \square

如果迭代序列 $\{d_k\}$ 存在一个具有共同的积极非光滑指标集的收敛子列, 并趋近于一个具有新的积极非光滑面的点时, 我们可以使用推论4.5.2和如下引理来构建算法4.5.2的收敛性。

引理 4.5.5. 若 $s^{(2)}(d_k) \neq 0$, 令:

$$\mathcal{Y}_k = \text{span}\{s^{(1)}(d_k), s^{(2)}(d_k)\}. \quad (4.5.62)$$

则必有:

$$\delta(d_k) \perp \mathcal{Q}(d_k), \quad (4.5.63)$$

成立。进而如果有 $\text{rank}(\mathcal{Y}_k) = 1$, 则如下两条结论有且仅有一条成立:

$$s^{(1)}(d_k) = s^{(2)}(d_k); \quad (4.5.64)$$

$$s^{(1)}(d_k) = 0, \quad \text{且} \quad s^{(2)}(d_k) \perp \mathcal{Q}(d_k). \quad (4.5.65)$$

证明: 首先由定义4.4.1及定义4.4.2知:

$$\begin{aligned} s^{(1)}(d_k) &= -\mathcal{P}_{\mathcal{S}(d_k)}^\perp \xi(d_k) \\ &= -(I - J_k(J_k^T J_k)^{-1} J_k^T) \xi(d_k) \\ &= J_k(J_k^T J_k)^{-1} J_k^T \xi(d_k) - \xi(d_k). \end{aligned} \quad (4.5.66)$$

这里 J_k 是由 $\mathcal{S}(d_k)$ 的一组正交基为列向量组成的矩阵。于是我们有:

$$\begin{aligned} \delta(d_k) &= s^{(2)}(d_k) - s^{(1)}(d_k) \\ &= - \sum_{i \in \Omega(d_k)} \eta_i^* \beta_i - J_k(J_k^T J_k)^{-1} J_k^T \xi(d_k). \end{aligned} \quad (4.5.67)$$

根据(4.5.67), 显然有

$$\delta(d_k) \in \mathcal{S}(d_k). \quad (4.5.68)$$

因此, 我们立得:

$$\delta(d_k) \perp \mathcal{Q}(d_k). \quad (4.5.69)$$

当 $\text{rank}(\mathcal{Y}_k) = 1$ 时, 如果 $s^{(1)}(d_k) \neq 0$, 有

$$d_k + s^{(1)}(d_k) \in \mathcal{Q}(d_k). \quad (4.5.70)$$

于是

$$d_k + s^{(2)}(d_k) \in \mathcal{Q}(d_k). \quad (4.5.71)$$

因此我们可得

$$\delta(d_k) \in \mathcal{Q}(d_k). \quad (4.5.72)$$

由(4.5.69)与(4.5.72), 则必有

$$\delta(d_k) = 0. \quad (4.5.73)$$

也即(4.5.64)成立。如果 $s^{(1)}(d_k) = 0$, 则由(4.5.69)直接可推得(4.5.65)。证毕。 \square

引理 4.5.6. 对某个 $\Omega \in \mathcal{G}$, 对任意 $\epsilon > 0$, 都存在 $\delta_\epsilon > 0$ 使得:

$$\|s^{(2)}(d) - s^{(2)}(d_\Omega^*)\| \leq \epsilon, \quad \forall d \in N(d_\Omega^*, \delta_\epsilon) \cap \{d' \mid \Omega(d') = \Omega, \forall d' \in \mathbf{R}^n\}. \quad (4.5.74)$$

证明: 如果

$$\Omega(d_\Omega^*) = \Omega. \quad (4.5.75)$$

则根据引理4.5.4, (4.5.74)显然成立。下证(4.5.75)不成立的情形。此时必有,

$$\Omega \subsetneq \Omega(d_\Omega^*). \quad (4.5.76)$$

于是记:

$$\omega = \Omega(d_\Omega^*)/\Omega, \quad (4.5.77)$$

并记:

$$s^{(2)}(d_\Omega^*) = -\xi(d_\Omega^*) - \sum_{i \in \omega} \eta_i^* \beta_i - \sum_{i \in \Omega} \eta_i^* \beta_i, \quad (4.5.78)$$

这里 $\eta_i^*(i \in \Omega(d_\Omega^*))$ 是相应的(4.4.1)-(4.4.2)的最优解; 对充分小的 ε , 及任意的

$$d \in N(d_\Omega^*, \varepsilon \cap \{d' \mid \Omega(d') = \Omega, \forall d' \in \mathbf{R}^n\}), \quad (4.5.79)$$

有:

$$s^{(2)}(d) = -\xi(d) - \sum_{i \in \Omega} \bar{\eta}_i^* \beta_i, \quad (4.5.80)$$

这里 $\bar{e}t a_i^*(i \in \Omega)$ 是相应的(4.4.1)-(4.4.2)的最优解; 根据定义4.5.5及引理4.5.5, 我们知道:

$$s^{(2)}(d_\Omega^*) \in \mathcal{S}(d); \quad (4.5.81)$$

$$s^{(2)}(d) - s^{(1)}(d) \in \mathcal{S}(d). \quad (4.5.82)$$

进而由(4.5.78)和(4.5.80)我们可得:

$$[A(d_\Omega^* - d) - s^{(1)}(d) + \sum_{i \in \omega} (\eta_i^* - \text{sign}(\beta_i^T d + c_i)) \beta_i] \in \mathcal{S}(d). \quad (4.5.83)$$

假设存在 $j \in \omega$ 使得:

$$\eta_j^* \neq \text{sign}(\beta_j^T d + c_j), \quad (4.5.84)$$

且

$$\beta_j \notin \mathcal{S}(d), \quad (4.5.85)$$

则

$$(\eta_j^* - \text{sign}(\beta_j^T d + c_j)) \mathcal{P}_{\mathcal{S}(d)}^\perp \beta_j \neq 0, \quad (4.5.86)$$

且与 ε 无关。另一方面,

$$\lim_{\varepsilon \rightarrow 0^+} \|A(d_\Omega^* - d)\| = 0, \quad (4.5.87)$$

$$\lim_{\varepsilon \rightarrow 0^+} \|s^{(1)}(d)\| = 0, \quad (4.5.88)$$

于是导致(4.5.86)与(4.5.83)的矛盾。若(4.5.89)不成立, 即:

$$\beta_j \in \mathcal{S}(d), \quad (4.5.89)$$

则根据:

$$\beta_j d_\Omega^* + c_j = 0, \quad (4.5.90)$$

及

$$(d_{\Omega}^* - d) \perp \mathcal{S}(d), \quad (4.5.91)$$

可得:

$$(d_{\Omega}^* - d) \perp \beta_j, \quad (4.5.92)$$

于是

$$\beta_j d + c_j = \beta_j d_{\Omega}^* - \beta_j (d_{\Omega}^* - d) + c_j = 0. \quad (4.5.93)$$

这和(4.5.79)矛盾。综上所述可知:

$$\eta_j^* = \text{sign}(\beta_j^T d + c_j), \quad \forall j \in \omega. \quad (4.5.94)$$

因此(4.5.78)可改写为:

$$s^{(2)}(d_{\Omega}^*) = -\xi(d) + A(d - d_{\Omega}^*) - \sum_{i \in \Omega} \eta_i^* \beta_i. \quad (4.5.95)$$

若令:

$$\delta_{\epsilon} = \min\left\{\delta, \frac{\epsilon}{\rho_1}\right\}, \quad (4.5.96)$$

这里 δ 与 ρ_1 均由引理4.5.4定义。由(4.5.95), (4.5.80), 并根据引理4.5.4的证明过程, 可知(4.5.74)成立。证毕。 \square

最后给出的引理是关于集合论的简单结果:

引理 4.5.7. 设集合 $\Sigma = \{1, 2, \dots, n\}$, 由 Σ 子集组成的集合 $\Pi = \{\mathcal{A} \mid \mathcal{A} \subset \Sigma\}$ 。存在一个映射:

$$\mathcal{F} : \{x_k\} \mapsto \Pi, \quad (4.5.97)$$

这里 $\{x_k\}$ 是一个给定的序列, 则必存在集合 $\mathcal{A}^* \in \Pi$, 满足:

(1) 存在无限个正整数 k 满足:

$$\mathcal{F}(x_k) = \mathcal{A}^*; \quad (4.5.98)$$

(2) 不存在无限个正整数 k 满足:

$$\mathcal{F}(x_k) = \mathcal{A}, \quad (4.5.99)$$

对任意的 $\mathcal{A}^* \subsetneq \mathcal{A} \in \Pi$.

证明: 首先由于映射是将一个无限集映射到一个有限集, 所以必存在一个 Π 中的元素满足条件(1)。下面假设 \mathcal{A}^* 是所有满足条件(1)的 Π 的元素中长度 $|\mathcal{A}^*|$ 最大的, 则易证 \mathcal{A}^* 满足第二个条件。证毕。 \square

下面给出我们的全局收敛性定理:

定理 4.5.1. 假设4.5.1成立, 且最优解处信赖域约束(4.3.2)不积极。则算法4.5.2有限终止, 或者算法产生的点列 $\{d_k\}$ 满足:

$$\liminf_{k \rightarrow +\infty} \|s^{(2)}(d_k)\| = 0. \quad (4.5.100)$$

证明: 我们考虑算法4.5.2不有限终止的情形, 即算法将无限迭代。

由于集合 \mathcal{G} 的元素个数有限, 根据引理4.5.7, 必存在 $\bar{\Omega} \in \mathcal{G}$, 使得有无限个 $d_k (k \in \Theta \subset \{1, 2, \dots\})$ 满足:

$$\Omega(d_k) = \bar{\Omega}, \quad \forall k \in \Theta. \quad (4.5.101)$$

且对任何包含 $\bar{\Omega}$ 的 $\tilde{\Omega}$, 仅有有限个 d_k 满足:

$$\Omega(d_k) = \tilde{\Omega}. \quad (4.5.102)$$

由于问题是凸的, 且 $\phi(d_k)$ 单调下降有下界, 因此迭代点列 $\{d_k\}$ 有界, 于是这无限个 $d_k (k \in \Theta)$ 必有收敛子列, 不妨设为 $\{d_{k_j}, (j = 1, 2, \dots)\}$, 并记

$$\bar{d}_{\bar{\Omega}} := \lim_{j \rightarrow +\infty} d_{k_j}. \quad (4.5.103)$$

下面我们将分两种情况来分析。

首先考虑:

$$\Omega(\bar{d}_{\bar{\Omega}}) = \bar{\Omega} \quad (4.5.104)$$

成立的情况。

此时必有, $\bar{d}_{\bar{\Omega}}$ 是问题(4.3.1)限制在 $\mathcal{Q}(\bar{\Omega})$ 上的局部极小点, 即:

$$\bar{d}_{\bar{\Omega}} = d_{\bar{\Omega}}^*. \quad (4.5.105)$$

也即子列 d_{k_j} 收敛到问题限制于如下约束的解:

$$\beta_i^T d + c_i = 0, \quad i \in \bar{\Omega}. \quad (4.5.106)$$

假设(4.5.105)不成立, 则根据定理4.4.1 有 $s^{(1)}(\bar{d}_{\bar{\Omega}}) > 0$ 且与 j 无关. 由引理4.5.4, 对任意充分小的正数 ε , 必存在足够大的 N_ε , 使得对 $\forall j > N_\varepsilon$, 都有:

$$\|s^{(1)}(d_{k_j}) - s^{(1)}(\bar{d}_{\bar{\Omega}})\|_2 < \varepsilon. \quad (4.5.107)$$

于是存在常数 $C_1 > 0$, 满足:

$$f(d_{k_j}) - f(d_{k_{j+1}}) \geq (q - \frac{\Lambda_A}{2} q^2) \|s^{(1)}(\bar{d}_{\bar{\Omega}})\|_2^2 - C_1 \varepsilon; \quad (4.5.108)$$

这里

$$q = \min\{V(\bar{d}_{\bar{\Omega}}), \frac{1}{\Lambda_A}\}. \quad (4.5.109)$$

而且 q 也是与 j 无关的常量.

另一方面, 对任意充分小的正数 ϵ , 都存在足够大的 N_ϵ 使得对 $\forall j > N_\epsilon$, 都有:

$$\|d_{k_j} - \bar{d}_{\bar{\Omega}}\|_2 < \epsilon. \quad (4.5.110)$$

于是必存在 $C_2 > 0$, 使得:

$$f(d_{k_j}) - f(d_{k_{j+1}}) < C_2 \epsilon. \quad (4.5.111)$$

根据(4.5.108)和(4.5.111)可得:

$$\begin{aligned} (q - \frac{\Lambda_A}{2} q^2) \|s^{(1)}(\bar{d}_{\bar{\Omega}})\|_2^2 - C_1 \varepsilon &\leq \\ f(d_{k_j}) - f(d_{k_{j+1}}) &< f(d_{k_j}) - f(d_{k_{j+1}}) < C_2 \epsilon. \end{aligned} \quad (4.5.112)$$

从而得到矛盾, 因此(4.5.105)成立。

此时利用引理4.5.3和引理4.5.4, 可完全类似地证明 $d_{\bar{\Omega}}^*$ 是(4.3.1) 的局部极小点, 也即:

$$\lim_{j \rightarrow +\infty} s^{(2)}(d_{k_j}) = 0. \quad (4.5.113)$$

于是(4.5.100)成立. 值得注意的是, 这里我们没有考虑二维子空间退化的情形. 因为如果二维子空间无穷次退化, 根据引理4.5.5, 只可能是结论(4.5.64)的情形. 此时, 上面已经证明了的命题(4.5.112)对 $s^{(2)}(\bar{d}_{\bar{\Omega}})$ 也成立, 于是 $\bar{d}_{\bar{\Omega}}$ 是(4.3.1)局部极小点, 即有(4.5.100)成立. 我们易知有穷次退化不影响结果.

下面考虑(4.5.104)不成立的情况, 我们知道此时必有:

$$\bar{\Omega} \subsetneq \Omega(\bar{d}_{\bar{\Omega}}). \quad (4.5.114)$$

由(4.5.103), 我们可得:

$$\lim_{j \rightarrow +\infty} f(d_{k_j}) = f(\bar{d}_{\bar{\Omega}}). \quad (4.5.115)$$

由于算法4.5.2步十一会被调用无穷次, 如果其中有无穷次是通过转回步十计算 d_{k+1} 的, 则根据推论4.5.2, 同样可得 $\bar{d}_{\bar{\Omega}}$ 是问题(4.3.1)限制在 $Q(\bar{\Omega})$ 上的局部极小点, 也即(4.5.105)成立。

对任意充分小的正数 ι , 都存在足够大的 N_ι 使得对 $\forall j > N_\iota$, 都有:

$$\|d_{k_j} - \bar{d}_{\bar{\Omega}}\| < \iota. \quad (4.5.116)$$

且存在常数 $C_3 > 0$ 使得:

$$f(d_{k_j}) - f(d_{k_{j+1}}) < f(d_{k_j}) - f(\bar{d}_{\bar{\Omega}}) < C_3 \iota. \quad (4.5.117)$$

另一方面, 根据(4.5.3), 如果此时 $\bar{d}_{\bar{\Omega}}$ 不是(4.3.1)的局部极小点, 则必存在 \tilde{d} 满足:

$$f(\bar{d}_{\bar{\Omega}}) - f(\tilde{d}) \geq (q - \frac{\Lambda_A}{2} q^2) \|s^{(2)}(\bar{d}_{\bar{\Omega}})\|_2^2, \quad (4.5.118)$$

这里 q 如(4.5.109)定义是个与 j 无关的常数。如果设:

$$\tilde{d} = \bar{d}_{\bar{\Omega}} + q^* s^{(2)}(\bar{d}_{\bar{\Omega}}), \quad (4.5.119)$$

并记:

$$\tilde{d}_{k_j} = d_{k_j} + q^* s^{(2)}(d_{k_j}). \quad (4.5.120)$$

于是根据引理4.5.6, 我们可知:

$$\|\tilde{d} - \tilde{d}_{k_j}\| \leq \|\bar{d}_{\bar{\Omega}} - d_{k_j}\| + q^* \|s^{(2)}(\bar{d}_{\bar{\Omega}}) - s^{(2)}(d_{k_j})\| \leq C_3 \iota + q^* \epsilon. \quad (4.5.121)$$

于是必存在常数 $C_4 > 0$ 使得:

$$|f(\tilde{d}) - f(\tilde{d}_{k_j})| \leq C_4 \iota. \quad (4.5.122)$$

于是根据(4.5.116), (4.5.117), (4.5.118)及(4.5.122), 我们有

$$\begin{aligned}
\left(q - \frac{\Lambda_A}{2}q^2\right) \|s^{(2)}(\bar{d}_{\bar{\Omega}})\|_2^2 &\leq |f(\bar{d}_{\bar{\Omega}}) - f(\tilde{d})| \\
&= |f(\bar{d}_{\bar{\Omega}}) - f(\tilde{d}_{k_j}) + f(\tilde{d}_{k_j}) - f(\tilde{d}_{k_j}) + f(\tilde{d}_{k_j}) - f(\tilde{d})| \\
&\leq |f(\bar{d}_{\bar{\Omega}}) - f(\tilde{d}_{k_j})| + |f(\tilde{d}_{k_j}) - f(\tilde{d}_{k_j})| + |f(\tilde{d}_{k_j}) - f(\tilde{d})| \\
&\leq C_3\iota + |f(\tilde{d}_{k_j}) - f(d_{k_{j+1}})| + C_4\iota \leq (2C_3 + C_4)\iota.
\end{aligned} \tag{4.5.123}$$

矛盾。因此 $\bar{d}_{\bar{\Omega}}$ 是(4.3.1)局部极小点, 即有(4.5.100)成立。

这里我们同样没有考虑二维子空间退化的情形。因为如果二维子空间无穷次退化, 根据引理4.5.5, 只可能是结论(4.5.64)的情形, 于是可知推论4.5.2的结论同样适用于 $s^{(2)}(d_{k_j})$, 这样直接可证明 $\bar{d}_{\bar{\Omega}}$ 是(4.3.1)局部极小点, 即有(4.5.100)成立。有穷次退化同样不影响结果。

最后考虑情况, 如果有无穷次是根据步十一计算 d_{k+1} 的, 则根据推论4.5.2, 必存在一个充分大的正数 M , 使得对 $\forall j > M$, 必存在 $k_j < l_{j-M} < k_{j+1}$, 满足:

$$f(d_{k_j}) > f(d_{l_{j-M}}) > f(k_{j+1}), \tag{4.5.124}$$

及

$$\bar{\Omega} \subsetneq \Omega(d_{l_{j-M}}). \tag{4.5.125}$$

这里(4.5.125)是因为对充分大的正数 M , 我们总可以保证对 $\forall j > M$, 必有 d_{k_j} 到 $\bar{d}_{\bar{\Omega}}$ 间的距离充分小于 $\bar{d}_{\bar{\Omega}}$ 到任何 $\beta_i^T d + c_i = 0 (i \notin \Omega(\bar{d}_{\bar{\Omega}}))$ 的距离。

这样我们就找到一个新的无限序列 $\{d_{l_j}\}$, 于是必存在一个 $\hat{\Omega}$, 使得序列 $\{d_{l_j}\}$ 中的某个子列 $\{d_{l_i}\}$ 满足:

$$\Omega(d_{l_i}) = \hat{\Omega}, \quad \forall l_i. \tag{4.5.126}$$

由(4.5.125), 我们知道

$$\bar{\Omega} \subset \hat{\Omega}, \tag{4.5.127}$$

于是存在无限个序列 $\{d_k\}$ 中的点, 使得(4.5.101)成立, 这与(4.5.102)矛盾。所以当(4.5.104)不成立时, d_{k+1} 最多只有有限次是由算法4.5.2中步十一计算的, 其它都转回由步十计算。因此也即必有结论(4.5.100)成立。

综上所述, 定理证毕。 \square

由定理4.5.1, 我们就得到了当信赖域约束(4.3.2)不积极的情形时算法4.5.2的收敛性。关于当信赖域约束(4.3.2)是积极的情形时的收敛性, 可由定理4.5.1和下一章介绍的关于算法4.5.1的收敛性分析过程类似地完成。在此不作赘述。下面仅给出方向 $s^{(3)}(d_k)$ 的下降性分析。

引理 4.5.8. 假设4.5.1成立, 则当 $s^{(3)}(d_k) \neq 0$ 时, 必有:

$$f(d_k) - f(d_k + t_k^* s_k^{(3)}) \geq \varpi_3(d_k) \|s_k^{(3)}\|_2^2. \quad (4.5.128)$$

这里:

$$\varpi_3(d_k) = \begin{cases} \frac{1}{2\Lambda_A}, & \text{如果 } \frac{1}{\Lambda_A} \leq \bar{t}_k^{(3)}; \\ (\bar{t}_k^{(3)} - \frac{\Lambda_A}{2}(\bar{t}_k^{(3)})^2), & \text{否则.} \end{cases} \quad (4.5.129)$$

$$t_k^* = \arg \min_{t \geq 0} f(d_k + t s_k^{(3)}). \quad (4.5.130)$$

证明: 为了使证明过程简洁, 我们暂时简化记号:

$$s_k = s^{(3)}(d_k); \quad \bar{t}_k = \bar{t}(d_k, s^{(3)}(d_k)). \quad (4.5.131)$$

由于 $s^{(3)}(d_k) \neq 0$, 设 $s_k = -\zeta_0^* d_k - \sum_{i \in \Omega(d_k)} \zeta_i^* \beta_i - \xi(d_k)$ 。写出当 $d = d_k$ 时带界约束的线性最小二乘问题(4.4.3)-(4.4.5)的KKT系统, 也即 $\zeta_i^* (i \in \Omega(d_k))$ 需要满足的方程:

$$\begin{cases} (\zeta_0^* d_k + \sum_{i \in \Omega(d_k)} \zeta_i^* \beta_i + \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j + A d_k + g)^T d_k = \lambda_0^*; \\ (\zeta_0^* d_k + \sum_{i \in \Omega(d_k)} \zeta_i^* \beta_i + \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j + A d_k + g)^T \beta_i = \lambda_i^* - \sigma_i^*; \\ \lambda_i^* \cdot \zeta_0^* = 0; \\ \lambda_i^* (\zeta_i^* + 1) = 0; \\ \sigma_i^* (1 - \zeta_i^*) = 0; \\ -1 \leq \zeta_i^* \leq 1; \\ \lambda_0^* > 0, \lambda_i^* \geq 0, \sigma_i^* \geq 0. \end{cases} \quad (4.5.132)$$

这里 λ_0^* , λ_i^* 与 σ_i^* 是约束(4.4.2)的Lagrange乘子。根据(4.5.132)我们易知有下列三式成立:

$$\beta_i^T s_k < 0 \Rightarrow \zeta_i^* = -1, \quad \forall i \in \Omega(d); \quad (4.5.133)$$

$$\beta_i^T s_k > 0 \Rightarrow \zeta_i^* = 1, \quad \forall i \in \Omega(d); \quad (4.5.134)$$

$$\zeta_0^* d_k^T s_k = 0. \quad (4.5.135)$$

于是对任意 $t \in (0, \bar{t}(d_k, s_k)]$, 我们有:

$$\begin{aligned}
f(d_k) - f(d_k + ts_k) &= - \sum_{i \in \Omega(d_k)} |t\beta_i^T s_k| + \sum_{j \notin \Omega(d_k)} (|\beta_j^T d_k + c_j| - |\beta_j^T d_k + c_j + t\beta_j^T s_k|) \\
&\quad + \left(\frac{1}{2}s_k^T A s_k + g^T d_k - \frac{1}{2}(d_k + ts_k)^T A(d_k + ts_k) - g^T(d_k + ts_k)\right) \\
&= - \sum_{i \in \Omega(d_k)} |\beta_i^T s_k \cdot t| - \sum_{j \notin \Omega(d_k)} \text{sign}(\beta_j^T d_k + c_j) \cdot \beta_j^T s_k \cdot t \\
&\quad - (Ad_k + g)^T s_k \cdot t - \frac{1}{2}s_k^T A s_k \cdot t^2 \\
&= - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \beta_i^T s_k \cdot t + \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \beta_i^T s_k \cdot t \\
&\quad - \xi(d_k)^T s_k \cdot t - \frac{1}{2}s_k^T A s_k \cdot t^2 \\
&= \left[\sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \beta_i - \xi(d_k) \right]^T s_k \cdot t - \frac{1}{2}s_k^T A s_k \cdot t^2 \\
&= \left[- \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k < 0}} \zeta_i \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k = 0}} \zeta_i \beta_i - \sum_{\substack{i \in \Omega(d_k), \\ \beta_i^T s_k > 0}} \zeta_i \beta_i \right. \\
&\quad \left. \zeta_0 d_k - \xi(d) \right]^T s_k \cdot t - \frac{1}{2}s_k^T A s_k \cdot t^2 \\
&= s_k^T s_k \cdot t - \frac{1}{2}s_k^T A s_k \cdot t^2 \\
&\geq \|s_k\|_2^2 \cdot t - \frac{1}{2}\Lambda_A \|s_k\|_2^2 \cdot t^2. \tag{4.5.136}
\end{aligned}$$

我们知道, 由(4.5.34)定义的 (s_k, \bar{t}_k) 满足的(4.5.39)-(4.5.42)式在 (s_k, \bar{t}_k) 由(4.5.131)定义时照样成立。

于是根据(4.5.136)、(4.5.40)及(4.5.42), 我们可得(4.5.128)成立。证毕。 \square

第五章 其它相关工作

§5.1 最小二乘参数估计与观测次数的关系

§5.1.1 最小二乘估计

我们考虑线性模型:

$$y = X\beta + \varepsilon, \quad E(\varepsilon) = 0, \quad \text{Cov}(\varepsilon) = \sigma^2 I; \quad (5.1.1)$$

的参数 β 的估计。其中 $y \in \mathbf{R}^m$ 是观测向量(每个分量是一次观测的结果); $X \in \mathbf{R}^{m \times n}$ 是设计矩阵(每一个行向量是一次观测的数据); $\beta \in \mathbf{R}^n$ 是未知参数向量; $\varepsilon \in \mathbf{R}^m$ 是随机误差向量(每个分量是一次观测的观测误差); $\sigma^2 > 0$ 为误差方差。

我们考虑模型中 $m > n$ 、 $\text{rank}(X) = n$ (也即满秩)的情况, 这通过选择适合的设计矩阵可以办到。这时我们容易知道最小二乘解是对参数 β 的最好估计(无偏估计, 可参阅 [154]), 也即

$$\hat{\beta} = (X^T X)^{-1} X^T y. \quad (5.1.2)$$

由(5.1.1)、(5.1.2)式, 我们容易知道

$$\hat{\beta} - \beta = (X^T X)^{-1} X^T \varepsilon. \quad (5.1.3)$$

我们知道对任意的观测次数都有

$$E(\hat{\beta} - \beta) = 0. \quad (5.1.4)$$

这解释了为什么说(5.1.2)是对模型的最好估计。

本文将考察观测次数 m 与模型估计的关系。我们将验证 $\text{Var}(\hat{\beta}_i - \beta_i)$ 是关于 m 单调递减的, $\forall i = 1, 2, \dots, n$ 。

定理 5.1.1. 假设 β 是满足(5.1.1)式的实际参数, $\hat{\beta}$ 是根据(5.1.2)式得到的最小二乘估计, 那么必定有

$$\text{Var}(\hat{\beta} - \beta) = \frac{1}{X^T X} \sigma^2, \quad \text{若 } n = 1; \quad (5.1.5)$$

$$\text{Var}(\hat{\beta}_i - \beta_i) = \frac{\det(X^T X)_{ii}}{\det(X^T X)} \sigma^2, \quad \text{若 } n > 1, \quad \forall i = 1, 2, \dots, n. \quad (5.1.6)$$

其中, $(X^T X)_{ii}$ 是 $(X^T X)$ 的 i 行 i 列元对应的伴随矩阵。

证明: 对于 $n = 1$ 的情况, 显然有

$$\text{Var}(\hat{\beta} - \beta) = \text{Var}\left(\frac{X^T \varepsilon}{X^T X}\right) = \sum_{i=1}^m \frac{X_i^2 \sigma^2}{(X^T X)^2} = \frac{1}{X^T X} \sigma^2. \quad (5.1.7)$$

也即(5.1.5)式成立。对于 $n > 1$ 的情况, 我们有

$$\begin{aligned} \text{Var}(\hat{\beta} - \beta) &= \text{Var}((X^T X)^{-1} X^T \varepsilon) = (X^T X)^{-1} X^T \text{Var}(\varepsilon) (X^T X)^{-1} X^T \\ &= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} I = \sigma^2 (X^T X)^{-1}. \end{aligned} \quad (5.1.8)$$

于是, 我们容易知道有

$$\text{Var}(\hat{\beta}_i - \beta_i) = \frac{\det(X^T X)_{ii}}{\det(X^T X)} \sigma^2. \quad (5.1.9)$$

成立。证毕。□

§5.1.2 几个线性代数引理

在我们建立估计值与实际值偏差的方差 $\text{Var}(\hat{\beta}_i - \beta_i)$ 与观测次数 m 的关系定理之前, 我们有必要给出几个线性代数的引理, 它们的证明可参见Wang, Wu和Jia [162]。

定义 5.1.1.

- $A \succ (\succeq) 0$ 表示 A 是一个(半)正定矩阵;
- $A \succ (\succeq) B$ 表示 $A - B$ 是一个(半)正定矩阵。

引理 5.1.1. 假设 $A \succeq 0$ 和 $B \succeq 0$ 是两个实对称矩阵, 则

$$\det(A) + \det(B) \leq \det(A + B). \quad (5.1.10)$$

引理 5.1.2. 假设 $A \succeq 0$ 和 $B \succeq 0$ 是两个实对称矩阵, 如果 $A \succeq B$, 则

$$\det(A) \geq \det(B). \quad (5.1.11)$$

定义 5.1.2. 如果方阵 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 的子阵 $A_{11} \succ 0$, 则定义 A_{11} 的Shur补如下:

$$(A/A_{11}) = A_{22} - A_{21} A_{11}^{-1} A_{12}. \quad (5.1.12)$$

引理 5.1.3. 如果方阵 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 的子阵 $A_{11} \succ 0$, (A/A_{11}) 是 A_{11} 的 *Shur* 补, 那么

$$\det(A/A_{11}) = \frac{\det A}{\det A_{11}}. \quad (5.1.13)$$

引理 5.1.4. 如果同阶实对称矩阵 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 和 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 的同阶子阵 $A_{11} \succ 0$, $B_{11} \succ 0$, (A/A_{11}) , (B/B_{11}) 与 $((A+B)/(A_{11}+B_{11}))$ 分别是 A_{11} , B_{11} 与 $(A_{11}+B_{11})$ 的 *Shur* 补, 那么

$$((A+B)/(A_{11}+B_{11})) \succeq (A/A_{11}) + (B/B_{11}). \quad (5.1.14)$$

引理 5.1.5. 如果同阶实对称半正定矩阵 $A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$ 和 $B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}$ 的同阶子阵 $A_{11} \succ 0$, $B_{11} \succ 0$, 那么

$$\frac{\det(A+B)}{\det(A_{11}+B_{11})} \geq \frac{\det A}{\det A_{11}} + \frac{\det B}{\det B_{11}}. \quad (5.1.15)$$

§5.1.3 估计与观测次数的关系定理

本节中我们将建立及证明有关估计值与实际值偏差的方差 $\text{Var}(\hat{\beta}_i - \beta_i)$ ($\forall i = 1, 2, \dots, n$) 随观测次数 m 单调递减的定理。

定理 5.1.2. 如果矩阵 $X \in \mathbf{R}^{m \times n}$ ($m > n$) 是满秩的, 矩阵 $\bar{X} = \begin{pmatrix} X \\ \delta^T \end{pmatrix} \in \mathbf{R}^{(m+1) \times n}$, 其中 $\delta \in \mathbf{R}^n$, 则对 $\forall i = 1, 2, \dots, n$ 有如下结论成立

$$\frac{\det(\bar{X}^T \bar{X})_{ii}}{\det(\bar{X}^T \bar{X})} \leq \frac{\det(X^T X)_{ii}}{\det(X^T X)}. \quad (5.1.16)$$

证明: 由 X 满秩, 可得 $X^T X > 0$ 及 \bar{X} 满秩, 于是又可以推得 $\bar{X}^T \bar{X} > 0$; 由 $X^T X > 0$ 及 $\bar{X}^T \bar{X} > 0$, 又可推得 $(X^T X)_{ii} > 0$ 及 $(\bar{X}^T \bar{X})_{ii} > 0$. 于是我们可以通过简单的分析得到我们要证明的结论(5.1.16)等价于

$$\frac{\det(\bar{X}^T \bar{X})}{\det(\bar{X}^T \bar{X})_{ii}} \geq \frac{\det(X^T X)}{\det(X^T X)_{ii}}. \quad (5.1.17)$$

又等价于

$$\frac{\det(X^T X + \delta \delta^T)}{\det(X^T X + \delta \delta^T)_{ii}} \geq \frac{\det(X^T X)}{\det(X^T X)_{ii}} \quad (5.1.18)$$

下面我们将通过使用引理(5.1.5)来证明(5.1.18)式。对 $\forall t \in (0, 1)$ ，我们知道矩阵

$$A(t) = (1-t)X^T X > 0 \quad (5.1.19)$$

及 $A_{ii}(t) > 0$ 是显然成立的；另一方面矩阵

$$B(t) = tX^T X + \delta\delta^T = (X(t)^T X(t)), \quad (5.1.20)$$

其中矩阵

$$X(t) = \begin{pmatrix} \sqrt{t}X \\ \delta^T \end{pmatrix}, \quad (5.1.21)$$

由此可知 $B(t) > 0$ 及 $B_{ii}(t) > 0$ ，这样我们对矩阵 $A(t)$ 及 $B(t)$ 使用引理(5.1.5)可得

$$\frac{\det(A(t) + B(t))}{\det(A_{ii}(t) + B_{ii}(t))} \geq \frac{\det A(t)}{\det A_{ii}(t)} + \frac{\det B(t)}{\det B_{ii}(t)} > \frac{\det A(t)}{\det A_{ii}(t)}. \quad (5.1.22)$$

(5.1.22)式对 $\forall t \in (0, 1)$ 成立。我们再来构造如下函数

$$f(t) = \frac{\det(A(t) + B(t))}{\det(A_{ii}(t) + B_{ii}(t))} - \frac{\det A(t)}{\det A_{ii}(t)}, \quad (5.1.23)$$

化简后得

$$f(t) = \frac{\det(X^T X + \delta\delta^T)}{\det(X^T X + \delta\delta^T)_{ii}} - (1-t) \frac{\det(X^T X)}{\det(X^T X)_{ii}}. \quad (5.1.24)$$

我们容易从(5.1.24)式看出 $f(t)$ 是 \mathbb{R} 上的连续函数，于是

$$f(0) = \lim_{t \rightarrow 0^+} f(t), \quad (5.1.25)$$

根据(5.1.22)式，我们知道

$$f(t) > 0, \quad \forall t \in (0, 1), \quad (5.1.26)$$

所以

$$f(0) = \lim_{t \rightarrow 0^+} f(t) \geq 0. \quad (5.1.27)$$

而

$$0 \leq f(0) = \frac{\det(X^T X + \delta\delta^T)}{\det(X^T X + \delta\delta^T)_{ii}} - \frac{\det(X^T X)}{\det(X^T X)_{ii}} \quad (5.1.28)$$

即有(5.1.18)式成立，也即(5.1.16)成立。证毕。 \square

定理 5.1.3. 假设 β 是满足(5.1.1)式的实际参数, $\hat{\beta}$ 是根据(5.1.2)式得到的最小二乘估计, 那么必有估计值与实际值偏差的方差 $\text{Var}(\hat{\beta}_i - \beta_i)$ ($\forall i = 1, 2, \dots, n$)随观测次数 m 单调递减。

证明: 对于 $n = 1$ 的情况, 由定理5.1.1知,

$$\text{Var}(\hat{\beta} - \beta) = \frac{1}{X^T X} \sigma^2, \quad (5.1.29)$$

当增加一次观测后, X 变为 $\bar{X} = \begin{pmatrix} X \\ d \end{pmatrix}$, 其中 $d \in \mathbf{R}$, 显然有

$$\frac{1}{X^T X} \sigma^2 \geq \frac{1}{X^T X + d^2} \sigma^2 = \frac{1}{\bar{X}^T \bar{X}} \sigma^2. \quad (5.1.30)$$

对于 $n > 1$ 的情况, 由定理5.1.1, 我们有

$$\text{Var}(\hat{\beta}_i - \beta_i) = \frac{\det(X^T X)_{ii}}{\det(X^T X)} \sigma^2. \quad (5.1.31)$$

当增加一次观测后, X 变为 $\bar{X} = \begin{pmatrix} X \\ \delta^T \end{pmatrix}$, 其中 $\delta \in \mathbf{R}^n$, 由定理5.1.2知

$$\frac{\det(X^T X)_{ii}}{\det(X^T X)} \sigma^2 \geq \frac{\det(\bar{X}^T \bar{X})_{ii}}{\det(\bar{X}^T \bar{X})} \sigma^2. \quad (5.1.32)$$

证毕。 □

§5.2 凸二次规划的不可行原始对偶积极集方法

本节考虑一种凸二次规划的不可行原始对偶积极集方法。本文第四章的算法4.4.2和算法4.5.1中都调用了本节介绍的算法来求方向 $s^{(i)}(d)(i = 2, 3)$ 。

§5.2.1 问题描述

我们知道所有的二次规划问题都能化成如下只具有等式约束和界约束的形式:

$$\min_{x \in \mathbf{R}^n} \quad \frac{1}{2} x^T A x + g^T x, \quad (5.2.1)$$

$$\text{s.t.} \quad P^T x = q; \quad (5.2.2)$$

$$l \leq x \leq u. \quad (5.2.3)$$

这里 $A \in \mathbf{R}^{n \times n}$ 是正定矩阵; $g \in \mathbf{R}^n$; $P \in \mathbf{R}^{n \times m}$; $q \in \mathbf{R}^m$; $l, u \in \mathbf{R}^n$ 。(5.2.3)表示 $l_i \leq x_i \leq u_i (\forall i = 1, 2, \dots, n)$ 。

下面将介绍我们提出的求解(5.2.1)-(5.2.3)的不可行原始对偶积极集方法。我们注意到Kunisch和Rendl [43]及Voglis和Lagaris [127]都提出了类似的算法思想, 不过前者是针对单向界约束, 后者是针对双向界约束。而我们独立提出的算法适用范围更广。此外 [43]对单向界约束的问题所做的收敛性分析并不实用, 因为文中的分析是基于矩阵 A 满足两个极其苛刻的条件。

§5.2.2 算法描述

我们首先考虑(5.2.1)-(5.2.3)的KKT系统,

$$\left\{ \begin{array}{ll} Ax + g = \lambda - \sigma - Pr; & (i) \\ P^T x = q; & (ii) \\ (x - l) \circ \lambda = 0; & (iii) \\ (u - x) \circ \sigma = 0; & (iv) \\ l \leq x \leq u; & (v) \\ \lambda \geq 0, \sigma \geq 0. & (vi) \end{array} \right. \quad (5.2.4)$$

在给出我们的具体算法步骤之前, 我们先引入如下概念。

定义 5.2.1. 我们称一组 (x, λ, σ) 是互补的, 当且仅当(5.2.4)中的(iii)式和(iv)式同时被满足。

定义 5.2.2. 我们称 $(\mathcal{D}, \mathcal{L}, \mathcal{U})$ 是全集 $\Omega = \{1, 2, \dots, n\}$ 的一组划分, 当且仅当如下两个条件被满足:

$$\mathcal{D} \cup \mathcal{L} \cup \mathcal{U} = \Omega, \quad (5.2.5)$$

$$(\mathcal{D} \cap \mathcal{L}) \cup (\mathcal{D} \cap \mathcal{U}) \cup (\mathcal{L} \cap \mathcal{U}) = \emptyset. \quad (5.2.6)$$

定义 5.2.3. 我们称一组划分 $(\mathcal{D}, \mathcal{L}, \mathcal{U})$ 是一组互补的 (x, λ, σ) 对应的广义积极集, 当

且仅当:

$$\begin{cases} x_{\mathcal{D}} = z^*; \\ x_{\mathcal{L}} = l_{\mathcal{L}}; \\ x_{\mathcal{U}} = u_{\mathcal{U}}; \\ \lambda_{\mathcal{D}} = 0; \\ \lambda_{\mathcal{L}} = (Ax + g + Pr^*)_{\mathcal{L}}; \\ \lambda_{\mathcal{U}} = 0; \\ \sigma_{\mathcal{D}} = 0; \\ \sigma_{\mathcal{L}} = 0; \\ \sigma_{\mathcal{U}} = -(Ax + g + Pr^*)_{\mathcal{U}}; \end{cases} \quad (5.2.7)$$

这里 z^* 是如下等式约束凸二次规划问题的解:

$$\min_{z \in \mathbf{R}^{|\mathcal{D}|}} \frac{1}{2} z^T A_{\mathcal{D}, \mathcal{D}} z + (g_{\mathcal{D}} + A_{\mathcal{D}, \mathcal{L}} l_{\mathcal{L}} + A_{\mathcal{D}, \mathcal{U}} u_{\mathcal{U}})^T z, \quad (5.2.8)$$

$$\text{s.t.} \quad P_{\mathcal{D}}^T z = q_{\mathcal{D}} - P_{\mathcal{L}}^T l_{\mathcal{L}} - P_{\mathcal{U}}^T u_{\mathcal{U}}. \quad (5.2.9)$$

r^* 是对应的Lagrange乘子; n 维向量或 $n \times m$ 矩阵加下标集表示行序为下标集的子向量或子阵; $n \times n$ 矩阵加双下标集表示行序、列序分别为对应下标集的子阵。

等式约束凸二次规划问题(5.2.8)-(5.2.9)等价于求解如下的线性方程组:

$$\begin{pmatrix} A_{\mathcal{D}, \mathcal{D}} & P_{\mathcal{D}} \\ P_{\mathcal{D}}^T & 0 \end{pmatrix} \begin{pmatrix} z \\ r \end{pmatrix} = \begin{pmatrix} -g_{\mathcal{D}} - A_{\mathcal{D}, \mathcal{L}} l_{\mathcal{L}} - A_{\mathcal{D}, \mathcal{U}} u_{\mathcal{U}} \\ q_{\mathcal{D}} - P_{\mathcal{L}}^T l_{\mathcal{L}} - P_{\mathcal{U}}^T u_{\mathcal{U}} \end{pmatrix}. \quad (5.2.10)$$

(5.2.10)称为鞍点问题。关于鞍点问题的比较成熟的算法可参阅Elman和Golub [30]; Bramble [10]; Yuan [158]; Golub [6]等文献。然而当问题(5.2.1)-(5.2.3)中等式约束(5.2.2)不存在时, (5.2.10)就成了一般的线性方程组求解问题了。求解线性方程组的算法可参见: [163], [153], [158], [164]等。

下面给出我们设计的算法:

算法 5.2.1. (不可行原始对偶积极集方法)

步一 设置初始广义积极集 $(\mathcal{D}^{(1)}, \mathcal{L}^{(1)}, \mathcal{U}^{(1)})$ 为 $\{\Omega, \emptyset, \emptyset\}$ 置 $k := 1$ 。

步二 求解鞍点问题(5.2.10), 得到对应的 $(x^{(k)}, \lambda^{(k)}, \sigma^{(k)})$ 。

步三 更新广义积极集如下:

$$\begin{aligned} \mathcal{D}^{(k+1)}(x^{(k)}, \lambda^{(k)}, \sigma^{(k)}) &= \{i \mid i \in \mathcal{D}^{(k)}, \text{ 且 } l_i \leq x_i^{(k)} \leq u_i\} \\ &\cup \{i \mid i \in \mathcal{L}^{(k)}, \text{ 且 } \lambda_i^{(k)} < 0\} \\ &\cup \{i \mid i \in \mathcal{U}^{(k)}, \text{ 且 } \sigma_i^{(k)} < 0\}; \end{aligned} \quad (5.2.11)$$

$$\begin{aligned} \mathcal{L}^{(k+1)}(x^{(k)}, \lambda^{(k)}, \sigma^{(k)}) &= \{i \mid i \in \mathcal{D}^{(k)}, \text{ 且 } x_i^{(k)} < l_i\} \\ &\cup \{i \mid i \in \mathcal{L}^{(k)}, \text{ 且 } \lambda_i^{(k)} \geq 0\}; \end{aligned} \quad (5.2.12)$$

$$\begin{aligned} \mathcal{U}^{(k+1)}(x^{(k)}, \lambda^{(k)}, \sigma^{(k)}) &= \{i \mid i \in \mathcal{D}^{(k)}, \text{ 且 } x_i^{(k)} > u_i\} \\ &\cup \{i \mid i \in \mathcal{U}^{(k)}, \text{ 且 } \sigma_i^{(k)} \geq 0\}. \end{aligned} \quad (5.2.13)$$

步四 如果

$$\mathcal{D}^{(k+1)} = \mathcal{D}^{(k)}. \quad (5.2.14)$$

(也即KKT系统(5.2.4)被满足), 则算法终止;
否则置 $k := k + 1$, 并转步二。

§5.2.3 数值实验

本小节报告用我们的算法在求解凸二次规划问题中的数值表现。

§5.2.3.1 算法和问题参数的关系

我们使用Zhang [148]中关于非负最小二乘问题的数值实例, 即:

$$\min_{x \in \mathbf{R}^n} \|Ax - b\|_2^2, \quad (5.2.15)$$

$$s.t. \quad x \geq 0. \quad (5.2.16)$$

其中 $A \in \mathbf{R}^{m \times n}$ (一般来说 $m > n$), $b \in \mathbf{R}^m$ 。这类测试问题不含等式约束。

实验参数在matlab中的实现代码为:

```
rand('state', 0); density = 0.05; condA = 1.e + 2;
```

```
A = sprand(m, n, density, 1/condA)。
```

由于本实验目的是考察算法和问题参数的关系, 因此我们对不同的初始广义积极集的取法进行了三组测试。它们的参数选取方式分别为:

- 测试一: 初始广义积极集为 (Ω, \emptyset) , $b = rand(m, 1)$;
- 测试二: 初始广义积极集为 (Ω, \emptyset) , $b = 2 * rand(m, 1) - 1$;

- 测试三：初始广义积极集为 (\emptyset, Ω) , $b = 2 * rand(m, 1) - 1$ 。

所有数据都为计算10次随机生成问题的平均迭代步数。测试环境是Matlab 7.0; 使用CPU为2.66GHz, 内存1.96GB的Dell Optiplex755型计算机。

No.	m	n	测试一	测试二	测试三
1	200	100	2.4	3.1	2.7
2	400	200	3.0	3.8	3.6
3	500	300	3.9	4.3	4.2
4	800	400	3.9	4.7	4.7
5	2000	1000	4.9	5.8	5.4
6	5000	2000	6.0	7.0	6.6

表5.2.1 问题规模及初始广义积极集的比较

No.	稠密度	条件数	m	n	测试一
1	0.2	1.e+3	1000	800	8.8
2	0.5	1.e+3	1000	800	9.2
3	1	1.e+3	1000	800	10.2
4	1	1.e+2	1000	800	7.9
5	0.2	1.e+4	1000	800	11.7
6	0.2	1.e+3	1000	1000	9.8
6	0.2	1.e+3	3000	1000	8.8

表5.2.2 稠密度、条件数及问题形状的比较

根据数值试验结果表5.2.1和表5.2.2的观察, 我们发现算法有很好性质: 就是对于问题的规模、初始广义积极集和问题稠密度都不是很敏感; 也不会因为条件数或问题的形状(退化性)的改变而产生很大的影响。此外, 我们还发现, 除了前两次迭代, 后面的每一次迭代中广义积极集的变化都不是很大, 也就是说每相邻两个迭代步的步二中的线性最小二乘问题(这里求解的线性方程组问题等价于一个线性最小二乘问题), 有很强的继承性。因此我们可以结合某种热启动的技巧设计节省计算量的算法来求解线性最小二乘问题, 以提高算法效率。另外我们算法的精度即是求解线性最小二乘问题的精度。

§5.2.3.2 和投影BB方法的比较

本节比较算法5.2.1与Dai和Fletcher [23]提出的非单调的投影BB方法的比较结

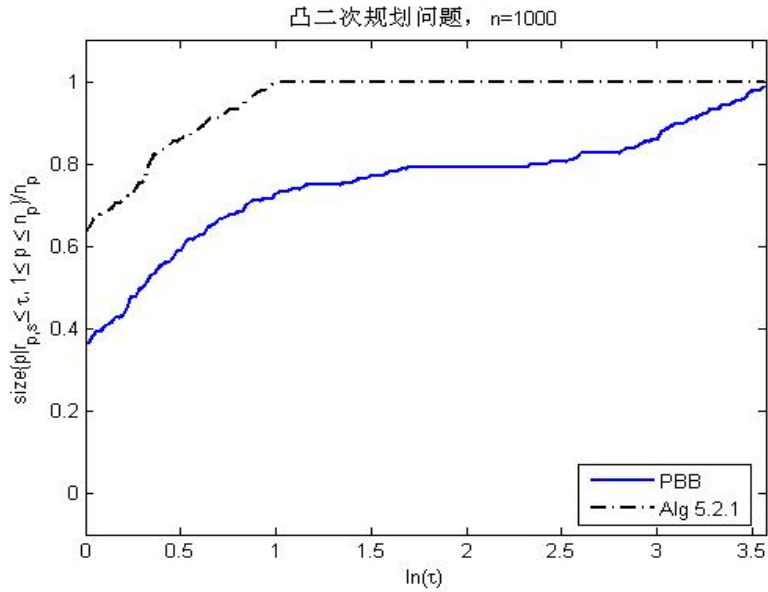


图 5.1

果。我们的测试问题选用“随机SPD测试问题”，详细请参见 [73]或 [22]。问题是具有界约束和一个等式约束的凸二次规划问题。问题构造参数有四个，分别是：

- (1) $ncond$: 问题的条件数, 取值范围 $\{4, 5, 6, 7\}$;
- (2) $ndeg$: 问题的退化程度, 取值范围 $\{1, 3, 5, 7, 9\}$;
- (3) $rais$: 最优解满足界约束的分量所占的比例, 取值范围 $\{10\%, 50\%, 90\%\}$;
- (4) $raii$: 初始点满足界约束的分量所占的比例, 取值范围 $\{10\%, 50\%, 90\%\}$ 。

这里分量(4)只对Dai和Fletcher的方法适用, 因为我们的方法统一取初始广义积极集为: $(\Omega, \emptyset, \emptyset)$ 。

Dolan和Moré [29]在2002年提出了一种算法性能表现(Performance Profiles)方法, 成为当前最流行的算法比较方法。它可以比较公平有效地比较各个算法在某一方面性能, 比如运行时间、迭代次数、函数值计算次数、梯度值计算次数等。这里以运行时间为例, 简要介绍一下算法性能表现的原理。

我们假设我们有 n_s 个算法 \mathcal{S} , n_p 个算例 \mathcal{P} 。我们记

$$t_{p,s} = \text{使用 } s \text{ 算法求解 } p \text{ 问题所需的时间。} \quad (5.2.17)$$

然后我们将其和求解 p 问题最快的算法比较, 并由此得到算法性能比:

$$r_{p,s} = \frac{t_{p,s}}{\min\{t_{p,s} : s \in \mathcal{S}\}}. \quad (5.2.18)$$

这样我们可以定义:

$$\rho_s(\tau) = \frac{1}{n_p} \text{size}\{p \in \mathcal{P} : r_{p,s} \leq \tau\}, \quad (5.2.19)$$

问题库 \mathcal{P} 中有多少百分比的问题, 使用算法 s 求解时, 所花时间至多是算法集 \mathcal{S} 中最快算法计算时间的 τ 倍。而算法性能表现评估图就是 ρ_s 关于 τ 的变化曲线。关于这套算法性能表现方法的性质可以参阅 [29]。

图5.1是 $n = 1000$ 时, 算法关于时间(由于两种算法机理不同, 无法比较算法迭代步数) 的性能表现图。我们发现算法5.2.1明显优于非单调的投影BB方法。下面抽取这180个例子中的24个, 来比较两种算法受问题参数取值变化的影响:

测试问题				投影BB方法		算法5.2.1	
n_{cond}	n_{deg}	r_{ais}	r_{aii}	时间(秒)	迭代步数	时间(秒)	迭代步数
4	3	10%	10%	0.34375	88	0.828	3
7	3	10%	10%	1.8594	462	1.063	4
4	9	10%	10%	0.35938	89	0.797	3
7	9	10%	10%	13.688	2000	1.094	4
4	3	50%	10%	0.3125	76	0.421	3
7	3	50%	10%	13.422	2000	0.609	5
4	9	50%	10%	0.3125	75	0.61	4
7	9	50%	10%	13.391	2000	0.813	5
4	3	90%	10%	0.32813	77	0.422	4
7	3	90%	10%	14.141	2000	0.437	5
4	9	90%	10%	0.34375	80	0.453	4
7	9	90%	10%	13.766	2000	0.485	4

表5.2.3(i) 随机SPD测试问题

测试问题				投影BB方法		算法5.2.1	
<i>ncond</i>	<i>ndeg</i>	<i>rais</i>	<i>raii</i>	时间(秒)	迭代步数	时间(秒)	迭代步数
4	3	10%	50%	0.35938	88	0.766	3
7	3	10%	50%	1.7656	437	0.797	3
4	9	10%	50%	0.32813	80	0.828	3
7	9	10%	50%	3.9531	751	0.844	3
4	3	50%	50%	0.29688	73	0.531	4
7	3	50%	50%	13.422	2000	0.641	4
4	9	50%	50%	0.34375	81	0.562	3
7	9	50%	50%	13.016	2000	0.797	5
4	3	90%	50%	0.375	89	0.391	4
7	3	90%	50%	14.047	2000	0.438	4
4	9	90%	50%	0.35938	84	0.438	4
7	9	90%	50%	1.8906	450	0.516	4

表5.2.3(ii) 随机SPD测试问题(续)

我们发现投影BB方法的表现主要受条件数的制约,也会受到最优解和初始值满足界约束的分量所占比例的影响。而我们的算法5.2.1不受任何参数影响。

§5.3 序列二维子空间方法的收敛性分析

我们在前一章已经介绍了Powell [108]提出的序列二维子空间技巧4.5.1。作为用截断共轭梯度法([13]; [143])求解信赖域子问题(4.3.3)-(4.3.4)的后处理技巧,算法有很好的数值表现,然而却没有很好的收敛性分析。

Dai [20]提出了求椭球外一点到椭球距离的序列二维子空间算法,该算法有很好的数值表现。[20]分析了算法收敛性和收敛速度。椭球外一点到椭球距离的问题在某种意义上可以看成是信赖域子问题的特例。

对于信赖域子问题(4.3.3)-(4.3.4),我们设它的解为 d^* 。如果 $\nabla m_k(d^*) = 0$,我们知道此时 $m_k(d)$ 是凸函数,且约束(4.3.4)并不积极;由于我们的二维子空间算法包含最速下降方向,此时序列二维子空间算法总是不输于最速下降法的收敛性和收敛速度。

因此在本节的分析中,我们将考虑 $\nabla m_k(d^*) \neq 0$,即信赖域约束积极的情形,

此时球约束等价于球面约束。更一般地, 我们考虑如下球面约束问题:

$$\min_{d \in \mathbf{R}^n} f(d) \quad (5.3.1)$$

$$\text{s.t.} \quad \|d\|_2 = \Delta. \quad (5.3.2)$$

而我们考虑的序列二维子空间算法也不仅仅是作为截断共轭梯度法的后处理技巧, 而是直接的序列二维子空间算法:

算法 5.3.1. (序列二维子空间算法)

步一 给定初始点 d_0 满足约束 (5.3.2), 如可取:

$$d^{(0)} = -\frac{\Delta}{\|\nabla f(d_0)\|_2} \cdot \nabla f(d_0); \quad (5.3.3)$$

$0 < \varepsilon \ll 1$; 置 $k := 0$, 转步二。

步二 如果 $\cos(d_k, -\nabla f(d_k)) > 1 - \varepsilon$, 则算法停止; 否则转步三。

步三 求解 (5.3.1)-(5.3.2) 限制在二维子空间 $\text{span}\{d_k, -\nabla f(d_k)\}$ 上的全局极小点, 得到 d_{k+1} , 转步四。

步四 置 $k := k + 1$, 转步二。

所以算法 4.5.1 是我们考虑的算法 5.3.1 的一个特例。

我们知道在等式约束的情形下, 如果 f 在信赖域 $\|d\|_2 \leq \Delta$ 内不凸, 我们可以在目标函数上加上正则项:

$$\tilde{f}(d) = f(d) + \lambda \|d\|_2^2, \quad \lambda > 0; \quad (5.3.4)$$

使得目标函数在 $\|d\|_2 \leq \Delta$ 内变凸而不改变极小点的位置。

根据上面的分析, 在下文的分析中我们将有如下假设:

假设 5.3.1. 问题 (5.3.1)-(5.3.2) 中的目标函数 f 是 $\{d \mid \|d\|_2 \leq \Delta\}$ 内的二次连续可微的凸函数。

假设 5.3.2. 若 d^* 是 (5.3.1)-(5.3.2) 的一个局部极小, 则 $\nabla f(d^*) \neq 0$ 。

于是我们可以定义:

定义 5.3.1.

- $M = \sup_{\|d\|_2 \leq \Delta} \|\nabla^2 f(d)\|_2$;
- $N = \sup_{\|d\|_2 \leq \Delta} \|\nabla f(d)\|_2$.

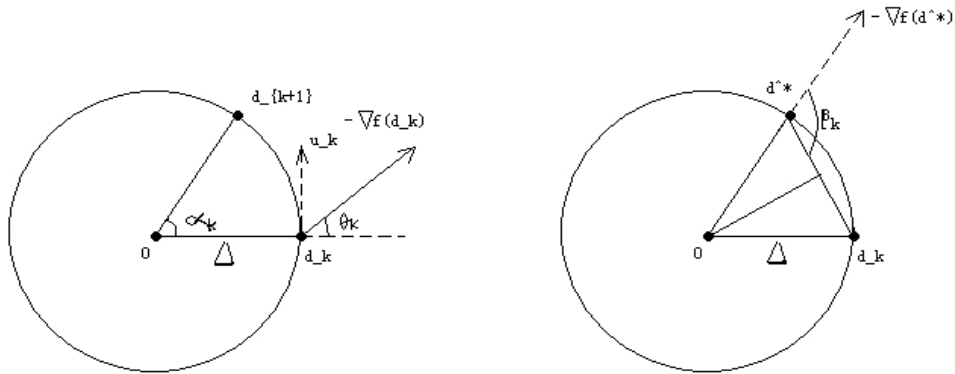


图 5.2

§5.3.1 算法的收敛性分析

首先给出本节中的常用记号的定义:

定义 5.3.2. 定义 $u(d)$ 为 d 点处负梯度方向在 d 的零空间内投影:

$$u(d) = -\nabla f(d) + \frac{\nabla f(d)^T d}{\Delta^2} \cdot d. \quad (5.3.5)$$

我们有如下命题成立:

命题 5.3.1. d 为(5.3.1)-(5.3.2)的局部极小值点, 当且仅当:

$$u(d) = 0. \quad (5.3.6)$$

证明: 根据(5.3.1)-(5.3.2)的KKT条件:

$$\begin{cases} \nabla f(x) = \lambda \frac{d}{\|d\|_2}; \\ \|d\|_2 = \Delta. \end{cases} \quad (5.3.7)$$

如果 d 为(5.3.1)-(5.3.2)的局部极小值点, 则由(5.3.7):

$$\begin{aligned} u(d) &= -\nabla f(d) + \frac{\nabla f(d)^T d}{\Delta^2} \cdot d \\ &= \frac{\lambda}{\Delta} \cdot d + \frac{\lambda}{\Delta} \cdot \frac{d^T d}{\Delta^2} \cdot d = 0. \end{aligned} \quad (5.3.8)$$

反之, 如果 $u(d) = 0$, 则取

$$\lambda = \frac{\nabla f(d)^T d}{\Delta}. \quad (5.3.9)$$

则(5.3.7)成立, 即 d 为(5.3.1)-(5.3.2)的局部极小值点, 证毕。□

为了简化记号, 我们记:

$$u_k = u(d_k); \quad u^* = u(d^*). \quad (5.3.10)$$

引理 5.3.1. 设 $\{d_k\}$ 为算法5.3.1产生的迭代点列, 则:

$$f(d_k) - f(d_{k+1}) \geq \frac{\Delta}{2M\Delta + \sqrt{5}N} \cdot \|u_k\|_2^2; \quad (5.3.11)$$

证明: 如图5.2左所示, 我们记 θ_k 为 d_k 和负梯度方向 $-\nabla f(d_k)$ 的夹角。则球面 $\|d\|_2 = \Delta$ 与二维子空间 $\text{span}\{d_k, -\nabla f(d_k)\}$ 的交上的任何一点都可以表示为:

$$d(\alpha) = d_k - (1 - \cos \alpha) \cdot d_k + \frac{\Delta}{\|u_k\|_2} \sin \alpha \cdot u_k. \quad (5.3.12)$$

我们记

$$c_1 = -(1 - \cos \alpha) = -2 \sin^2 \frac{\alpha}{2}; \quad c_2 = \frac{\Delta}{\|u_k\|_2} \sin \alpha. \quad (5.3.13)$$

并令

$$\alpha_k = \arg \min_{\alpha \in [0, \pi]} f(d(\alpha)). \quad (5.3.14)$$

因此我们有 $d_{k+1} = d(\alpha_k)$ 。

根据定义5.3.1, $\|\nabla^2 f(d)\|_2 \leq M, \quad \forall \|d\|_2 \leq \Delta$ 。于是我们有:

$$f(d_k + s) \leq f(d_k) + s^T \nabla f(d_k) + \frac{1}{2} M \|s\|_2^2. \quad (5.3.15)$$

因此, 我们有

$$\begin{aligned} f(d_k) - f(d(\alpha)) &\geq -\frac{1}{2} M \|c_1 d_k + c_2 u_k\|_2^2 - (c_1 d_k + c_2 u_k)^T \nabla f(d_k) \\ &= -\frac{1}{2} M c_1^2 \|d_k\|_2^2 - \frac{1}{2} M c_2^2 \|u_k\|_2^2 \\ &\quad - (c_1 d_k + c_2 u_k)^T \left(-u_k + \frac{\nabla f(d_k)^T d_k}{\Delta^2} \cdot d_k \right) \\ &= -\frac{1}{2} M c_1^2 \|d_k\|_2^2 - \frac{1}{2} M c_2^2 \|u_k\|_2^2 - c_1 \nabla f(d_k)^T d_k + c_2 \|u_k\|_2^2 \\ &= -2M\Delta^2 \sin^4 \frac{\alpha}{2} - \frac{M}{2} \Delta^2 \sin^2 \alpha - 2\Delta \|\nabla f(d_k)\|_2 \cos \theta \sin^2 \frac{\alpha}{2} \\ &\quad + \Delta \|u_k\|_2 \sin \alpha \\ &= -2M\Delta^2 \sin^2 \frac{\alpha}{2} - 2\Delta \|\nabla f(d_k)\|_2 \cos \theta \sin^2 \frac{\alpha}{2} + \Delta \|u_k\|_2 \sin \alpha \\ &= -(M\Delta^2 + \Delta \|\nabla f(d_k)\|_2 \cos \theta) (1 - \cos \alpha) + \Delta \|u_k\|_2 \sin \alpha \\ &= -T + \sqrt{T^2 + \Delta^2 \|u_k\|_2^2} \sin(\alpha + \varphi), \end{aligned} \quad (5.3.16)$$

这里:

$$T = M\Delta^2 + \Delta\|\nabla f(d_k)\|_2 \cos \theta; \quad \varphi = \arctg \frac{T}{\Delta\|u_k\|_2}. \quad (5.3.17)$$

令 $\bar{\alpha} := \frac{\pi}{2} - \varphi$, 并根据(5.3.16), 我们有

$$\begin{aligned} f(d_k) - f(d_{k+1}) &\geq f(d_k) - f(d(\bar{\alpha})) \\ &= -T + \sqrt{T^2 + \Delta^2\|u_k\|_2^2} \\ &= \frac{\Delta^2\|u_k\|_2^2}{T + \sqrt{T^2 + \Delta^2\|u_k\|_2^2}} \\ &\geq \frac{\Delta^2\|u_k\|_2^2}{2T + \Delta\|u_k\|_2} \\ &= \frac{\Delta^2\|u_k\|_2^2}{2(M\Delta^2 + \Delta\|\nabla f(d_k)\|_2 \cos \theta) + \Delta\|u_k\|_2} \\ &= \frac{\Delta^2\|u_k\|_2^2}{2M\Delta^2 + 2\Delta\|\nabla f(d_k)\|_2 \cos \theta + \Delta\|\nabla f(d_k)\|_2 \sin \theta} \\ &\geq \frac{\Delta^2\|u_k\|_2^2}{2M\Delta^2 + \sqrt{5}\Delta\|\nabla f(d_k)\|_2} \\ &\geq \frac{\Delta}{2M\Delta + \sqrt{5}N} \cdot \|u_k\|_2^2. \end{aligned} \quad (5.3.18)$$

证毕。 □

引理 5.3.2. 设 $\{d_k\}$ 为算法5.3.1产生的迭代点列, d^* 为(5.3.1)-(5.3.2)的局部极小点, 则:

$$\lim_{k \rightarrow +\infty} \frac{\|u_k\|_2^2}{f(d_k) - f(d^*)} \geq \frac{2\|\nabla f(d^*)\|_2}{\Delta}. \quad (5.3.19)$$

证明: 如图5.2右所示, 我们记 β_k 为 $d_k - d^*$ 和 $-\nabla f(d^*)$ 的夹角, 于是我们有:

$$\cos \beta_k = -\frac{\|d_k - d^*\|_2}{2\Delta}. \quad (5.3.20)$$

我们记:

$$p = \frac{\|\nabla f(d^*)\|_2}{\Delta}, \quad (5.3.21)$$

于是有：

$$\begin{aligned}
f(d_k) - f(d^*) &= (d_k - d^*)^T \nabla f(d^*) + \frac{1}{2}(d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) + o(\|d_k - d^*\|_2^2) \\
&= -\|d_k - d^*\|_2 \|\nabla f(d^*)\|_2 \cos \beta_k + \frac{1}{2}(d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) \\
&\quad + o(\|d_k - d^*\|_2^2) \\
&= \frac{1}{2}p\|d_k - d^*\|_2^2 + \frac{1}{2}(d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) \\
&\quad + o(\|d_k - d^*\|_2^2). \tag{5.3.22}
\end{aligned}$$

根据命题5.3.1, 我们知道 $u^* = 0$ 。因此：

$$\begin{aligned}
\|u_k\|_2^2 &= \|u_k - u^*\|_2^2 \\
&= \left\| -\nabla f(d_k) + \frac{\nabla f(d_k)^T d_k}{\Delta^2} \cdot d_k + \nabla f(d^*) - \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot d^* \right\|_2^2 \\
&= \left\| -(\nabla f(d_k) - \nabla f(d^*)) + \frac{\nabla f(d_k)^T d_k}{\Delta^2} \cdot d_k - \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot d^* \right\|_2^2 \\
&= \left\| -\nabla^2 f(d^*)(d_k - d^*) + \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot (d_k - d^*) + \frac{\nabla f(d^*)^T (d_k - d^*)}{\Delta^2} \cdot d^* \right. \\
&\quad \left. + \frac{(\nabla^2 f(d^*)(d_k - d^*))^T d^*}{\Delta^2} \cdot d^* + o(\|d_k - d^*\|_2) \right\|_2^2; \tag{5.3.23}
\end{aligned}$$

再根据(5.3.20), 我们有：

$$\nabla f(d^*)^T (d_k - d^*) = o(\|d_k - d^*\|_2), \quad (d_k - d^*)^T d^* = o(\|d_k - d^*\|_2). \tag{5.3.24}$$

进而有：

$$\begin{aligned}
\|u_k\|_2^2 &= \left\| -\nabla^2 f(d^*)(d_k - d^*) + \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot (d_k - d^*) + \frac{(d_k - d^*)^T \nabla^2 f(d^*) d^*}{\Delta^2} \cdot d^* \right. \\
&\quad \left. + o(\|d_k - d^*\|_2) \right\|_2^2; \\
&= \|\nabla^2 f(d^*)(d_k - d^*)\|_2^2 + \left\| \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot (d_k - d^*) \right\|_2^2 \\
&\quad + \left\| \frac{(d_k - d^*)^T \nabla^2 f(d^*) d^*}{\Delta^2} \cdot d^* \right\|_2^2 - 2 \frac{\nabla f(d^*)^T d^*}{\Delta^2} \cdot (d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) \\
&\quad - 2 \frac{((d_k - d^*)^T \nabla^2 f(d^*) d^*)^2}{\Delta^2} + o(\|d_k - d^*\|_2^2) \\
&= \|\nabla^2 f(d^*)(d_k - d^*)\|_2^2 + p^2 \|d_k - d^*\|_2^2 - \left\| \frac{(d_k - d^*)^T \nabla^2 f(d^*) d^*}{\Delta^2} \cdot d^* \right\|_2^2 \\
&\quad + 2p \cdot (d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) + o(\|d_k - d^*\|_2^2) \\
&= \|\nabla^2 f(d^*)(d_k - d^*)\|_2^2 - \left\| (\nabla^2 f(d^*)(d_k - d^*))^T \left(\frac{1}{\Delta} d^* \right) \right\|_2^2 \\
&\quad + p^2 \|d_k - d^*\|_2^2 + 2p \cdot (d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) + o(\|d_k - d^*\|_2^2) \\
&\geq p^2 \|d_k - d^*\|_2^2 + 2p \cdot (d_k - d^*)^T \nabla^2 f(d^*)(d_k - d^*) + o(\|d_k - d^*\|_2^2). \tag{5.3.25}
\end{aligned}$$

再由(5.3.21)、(5.3.22), 我们就有:

$$\begin{aligned} \lim_{k \rightarrow +\infty} \frac{\|u_k\|_2^2}{f(d_k) - f(d^*)} &\geq \lim_{k \rightarrow +\infty} \frac{p^2 \|d_k - d^*\|_2^2 + 2p \cdot (d_k - d^*)^T \nabla^2 f(d^*) (d_k - d^*)}{\frac{1}{2} p \|d_k - d^*\|_2^2 + \frac{1}{2} (d_k - d^*)^T \nabla^2 f(d^*) (d_k - d^*)} \\ &\geq 2p = \frac{2\|\nabla f(d^*)\|_2}{\Delta}. \end{aligned} \quad (5.3.26)$$

证毕。□

定理 5.3.1. 假设问题(5.3.1)-(5.3.2)满足假设5.3.1。则算法5.3.1或者有限终止, 或者产生的迭代点列 $\{d_k\}$ 满足:

$$\lim_{k \rightarrow +\infty} \|u_k\|_2 = 0. \quad (5.3.27)$$

证明: 如果算法5.3.1不有限终止, 则根据引理5.3.1, 对任意 $k > 0$, 有

$$\begin{aligned} f(d_0) - f(d_{k+1}) &= \sum_{i=0}^k [f(d_i) - f(d_{i+1})] \\ &\geq \frac{\Delta}{2M\Delta + \sqrt{5}N} \sum_{i=1}^k \|u_k\|_2^2. \end{aligned} \quad (5.3.28)$$

由于 $f(d)$ 在 $\|d\| = \Delta$ 上是下方有界的, 因此 $\exists V > 0$ 满足 $f(d_0) - f(d_{k+1}) < V, \forall k > 0$ 。所以(5.3.27)成立, 证毕。□

定理 5.3.2. 假设问题(5.3.1)-(5.3.2)满足假设5.3.1和假设5.3.2。则迭代点列 $\{d_k\}$ 满足:

$$\lim_{k \rightarrow +\infty} \sup \frac{f(d_{k+1}) - f(d^*)}{f(d_k) - f(d^*)} \leq \frac{2M\Delta + \sqrt{5}N - 2\|\nabla f(d^*)\|_2}{2M\Delta + \sqrt{5}N} < 1. \quad (5.3.29)$$

证明: 根据引理(5.3.1)及引理(5.3.2), 我们有:

$$\begin{aligned} \lim_{k \rightarrow +\infty} \sup \frac{f(d_{k+1}) - f(d^*)}{f(d_k) - f(d^*)} &= 1 - \lim_{k \rightarrow +\infty} \inf \frac{f(d_{k+1}) - f(d_k)}{f(d_k) - f(d^*)} \\ &\leq 1 - \frac{2\|\nabla f(d^*)\|_2}{2M\Delta + \sqrt{5}N} \\ &= \frac{2M\Delta + \sqrt{5}N - 2\|\nabla f(d^*)\|_2}{2M\Delta + \sqrt{5}N} < 1. \end{aligned} \quad (5.3.30)$$

最后一个不等号是根据假设5.3.2得到的。证毕。□

§5.3.2 拉格朗日牛顿型子空间取法

上一小节我们分析了二维子空间算法的收敛性和收敛速度。我们发现在球面极小化中, 我们的子空间选取方法 $\text{span}\{d_k, -\nabla f(d_k)\} = \text{span}\{d_k, u_k\}$, 就好比无约束优化中的最速下降法。于是我们就考虑是否存在其它的子空间选取方法, 使得该方法在球面上有高效的收敛速度。

很自然地, 我们首先想到的是构造和牛顿法相对应的方法。于是我们考虑在 d_k 的零空间中找拉格朗日牛顿方向:

$$v_k = B_k^{-1}u_k - \frac{d_k^T B_k^{-1}u_k}{d_k^T B_k^{-1}d_k} \cdot B_k^{-1}d_k. \quad (5.3.31)$$

来代替 u_k 。这里:

$$B_k = \nabla^2 f(d_k) - \lambda_k \cdot I; \quad (5.3.32)$$

根据命题5.3.1的证明, 我们知道:

$$\lambda_k = \frac{d_k^T \nabla f(d_k)}{\Delta^2} \quad (5.3.33)$$

是一个比较合理的拉格朗日因子的估计(为了记号简洁, 在定义(5.3.32)和(5.3.33)时, 我们引入了和(5.3.2)等价的约束 $\frac{1}{2}\|d\|_2^2 = \frac{1}{2}\Delta^2$)。

下面的命题将保证用 v_k 代替 u_k 可行性的第一步:

命题 5.3.2. d_k 为 (5.3.1)-(5.3.2) 的局部极小值点, 当且仅当:

$$v_k = 0. \quad (5.3.34)$$

证明: 首先如果 $u_k = 0$, 则根据5.3.31, 显然有 $v_k = 0$ 。反之, 如果 $v_k = 0$, 根据5.3.31, 我们有:

$$0 = B_k v_k = u_k - \frac{d_k^T B_k^{-1}u_k}{d_k^T B_k^{-1}d_k} \cdot d_k; \quad (5.3.35)$$

两端左乘 u_k^T 可得:

$$0 = u_k^T u_k - \frac{d_k^T B_k^{-1}u_k}{d_k^T B_k^{-1}d_k} \cdot u_k^T d_k; \quad (5.3.36)$$

根据定义5.3.2, $u_k^T d_k = 0$ 。由(5.3.36), $\|u_k\|_2^2 = u_k^T u_k = 0$, 也即 $u_k = 0$ 。

综上所述, 我们证明了 $v_k = 0$ 当且仅当 $u_k = 0$ 。由命题5.3.1, $u_k = 0$ 等价于 d_k 为 (5.3.1)-(5.3.2) 的局部极小值点。证毕。□

于是我们给出拉格朗日牛顿型的序列二维子空间算法:

算法 5.3.2. (拉格朗日牛顿型的序列二维子空间算法)

基于算法5.3.1, 仅改变步三为:

步三: 求解(5.3.1)-(5.3.2)限制在二维子空间 $\text{span}\{d_k, v_k\}$ 上的全局极小点(等价于求解四次方程), 得到 d_{k+1} , 转步四。

根据假设5.3.2, 我们可以证明只要初始点在解的邻域内, 我们总保证算法中 $\lambda_k \leq 0 (k \geq 1)$, 这样 B_k 总是正定的。下面给出的命题对 $f(x)$ 是二次函数的情形, 初始点的可选区域作了一个粗糙的估计:

命题 5.3.3. 若 $f(x)$ 是二次函数, 则对任何 d_k 满足:

$$\|d_k - d^*\|_2 \leq \sqrt{\frac{2p}{p+M}} \Delta. \quad (5.3.37)$$

有 $\lambda_k \leq 0$ 成立。这里 M 、 p 及 λ_k 分别由定义5.3.1、(5.3.21)及(5.3.33)定义。

证明: 因为 $f(x)$ 是二次函数, 因此我们有:

$$\nabla f(d_k) = \nabla f(d^*) + \nabla^2 f(d^*)(d_k - d^*). \quad (5.3.38)$$

两边左乘 d_k^T

$$\begin{aligned} d_k^T \nabla f(d_k) &= d_k^T \nabla f(d^*) + d_k^T \nabla^2 f(d^*)(d_k - d^*) \\ &= -p d_k^T d^* + d_k^T \nabla^2 f(d^*)(d_k - d^*) \\ &= -p d_k^T (d_k + d^* - d_k) + d_k^T \nabla^2 f(d^*)(d_k - d^*) \\ &\leq -p \Delta^2 + (p+M) d_k^T (d_k - d^*) \\ &= -p \Delta^2 + (p+M) \Delta \|d_k - d^*\|_2 \cdot \frac{\|d_k - d^*\|_2}{2\Delta} \\ &= -p \Delta^2 + \frac{p+M}{2} \|d_k - d^*\|_2^2 \leq 0. \end{aligned} \quad (5.3.39)$$

最后一个不等号成立是因为(5.3.37)式。证毕。□

下面我们首先给出算法5.3.2的局部收敛速度的分析。为此我们还需要追加一个假设:

假设 5.3.3. $\nabla^2 f(d^*)$ 的最小特征值是正的。

定义 5.3.3. $m = \inf_{\|d\|_2 \leq \Delta} \|\nabla^2 f(d)\|_2$.

引理 5.3.3. 在(5.3.1)-(5.3.2)的局部极小点 d^* 的小邻域内的任意两点 $d^{(1)}$ 和 $d^{(2)}$, 必定满足:

$$f(d^{(1)}) - f(d^{(2)}) = O(\|d^{(1)} - d^{(2)}\|_2^2). \quad (5.3.40)$$

证明: (5.3.39)可由(5.3.22)式直接得到。□

引理 5.3.4. 对任意正定矩阵 A , 任意充分小的向量 $x \in \mathbf{R}^n$, 和任意向量 $y \in \mathbf{R}^n$, 都有:

$$1 - \frac{(x^T y)^2}{(x^T x)(y^T y)} = O(x^2) \Leftrightarrow 1 - \frac{(x^T A y)^2}{(x^T A x)(y^T A y)} = O(x^2). \quad (5.3.41)$$

证明: 我们先证明“ \Rightarrow ”。如果

$$1 - \frac{(x^T y)^2}{(x^T x)(y^T y)} = O(x^2), \quad (5.3.42)$$

成立, 则

$$\sin \theta = O(x). \quad (5.3.43)$$

这里 θ 是向量 x 与 y 的夹角。由于左右两边分式的分子分母都是关于 y 同次的齐次多项式, 因而 y 的长度不影响我们的结果, 所以我们不妨假设 $\|y\|_2 = \|x\|_2$ 。如图5.3-(i)所示, 我们记

$$s = x - y, \quad y' = \frac{x^T A y}{y^T A y} \cdot y, \quad s' = x - y'. \quad (5.3.44)$$

根据(5.3.43), 我们有:

$$\|s\|_2 = 2\|x\|_2 \sin \frac{\theta}{2} \leq 2\|x\|_2 \sin \theta = O(x^2). \quad (5.3.45)$$

于是:

$$y' = \frac{x^T A y}{y^T A y} \cdot y = \frac{(y + s)^T A y}{y^T A y} \cdot y = \left(1 + \frac{y^T A s}{y^T A y}\right) \cdot y. \quad (5.3.46)$$

我们令 $\sigma_{\max}(A)$ 和 $\sigma_{\min}(A)$ 分别表示 A 的最大和最小特征值。

$$\left| \frac{y^T A s}{y^T A y} \right| \leq \frac{\sigma_{\max}(A)\|s\|_2}{\sigma_{\min}(A)\|y\|_2} = O(x). \quad (5.3.47)$$

因此有 $y' = O(x)$, 进而

$$\begin{aligned} \|s'\|_2 &= \|x - y'\|_2 \leq \|x - y\|_2 + \|y - y'\|_2 \\ &= \|s\|_2 + \left\| \frac{y^T A s}{y^T A y} \cdot y \right\|_2 = O(x^2). \end{aligned} \quad (5.3.48)$$

另一方面我们有:

$$\begin{aligned}
 1 - \frac{(x^T Ay)^2}{(x^T Ax)(y^T Ay)} &= 1 - \frac{(x^T Ay)^2}{(x^T Ax)(y'^T Ay')} \\
 &= \frac{((y' + s')^T A(y' + s'))(y'^T Ay') - ((y' + s')^T Ay')^2}{((y' + s')^T A(y' + s'))(y'^T Ay')} \\
 &= \frac{s'^T As'}{y'^T Ay' + s'^T As'} = O(x^2). \tag{5.3.49}
 \end{aligned}$$

下面证明“ \Leftarrow ”。由于 A 是对称正定的, 我们可推得存在可逆矩阵 L 使得 $A = L^T L$, 及 A^{-1} 也是对称正定的。根据上面左推右的证明, 我们知道有:

$$1 - \frac{(x'^T y')^2}{(x'^T x')(y'^T y')} = O(x'^2) \Rightarrow 1 - \frac{(x'^T A^{-1} y')^2}{(x'^T A^{-1} x')(y'^T A^{-1} y')} = O(x'^2). \tag{5.3.50}$$

于是我们令 $x' = Lx$, $y' = Ly$, 即可得右推左。证毕。 \square

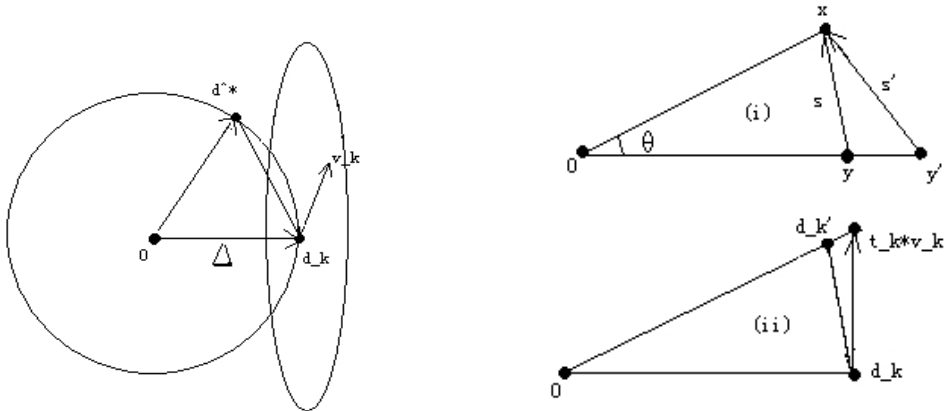


图 5.3

引理 5.3.5. 记 $s_k = d^* - d_k$, 则有:

$$1 - \frac{(s_k^T B_k v_k)^2}{(s_k^T B_k s_k)(v_k^T B_k v_k)} = O(s_k^2). \tag{5.3.51}$$

这里 v_k 及 B_k 分别如 (5.3.31) 式和 (5.3.32) 式定义。

证明: 根据 (5.3.23) 式, 我们可以直接得到下述结论:

$$\begin{aligned}
 u_k &= \nabla^2 f(d_k) s_k - \lambda_k s_k + \eta_k d_k + O(s_k^2) \\
 &= B_k s_k + \eta_k d_k + O(s_k^2), \tag{5.3.52}
 \end{aligned}$$

这里 $\eta_k = O(s_k)$ 。我们记：

$$w_k = u_k - B_k s_k - \eta_k d_k. \quad (5.3.53)$$

于是：

$$\|w_k\|_2 = O(s_k^2). \quad (5.3.54)$$

因此根据(5.3.52)式及(5.3.20)式，有：

$$\begin{aligned} v_k^T B_k v_k &= u_k^T B_k^{-1} u_k - \frac{(d_k^T B_k^{-1} u_k)^2}{d_k^T B_k^{-1} d_k} \\ &= (B_k s_k + \eta_k d_k + w_k)^T B_k^{-1} (B_k s_k + \eta_k d_k + w_k) \\ &\quad - \frac{(d_k^T B_k^{-1} (B_k s_k + \eta_k d_k + w_k))^2}{d_k^T B_k^{-1} d_k} \\ &= s_k^T B_k s_k + \eta_k^2 d_k^T B_k^{-1} d_k + 2\eta_k s_k^T d_k + 2w_k^T s_k + 2\eta_k w_k^T B_k^{-1} d_k \\ &\quad - \eta_k^2 d_k^T B_k^{-1} d_k - 2\eta_k s_k^T d_k - 2\eta_k w_k^T B_k^{-1} d_k + O(s_k^4) \\ &= s_k^T B_k s_k + 2w_k^T s_k + O(s_k^4). \end{aligned} \quad (5.3.55)$$

于是我们有：

$$v_k^T B_k v_k = O(s_k^2). \quad (5.3.56)$$

而另一方面

$$\begin{aligned} v_k^T B_k s_k &= u_k^T s_k - \frac{d_k^T B_k^{-1} u_k}{d_k^T B_k^{-1} d_k} \cdot d_k^T s_k \\ &= (B_k s_k + \eta_k d_k + w_k)^T s_k - \frac{d_k^T B_k^{-1} (B_k s_k + \eta_k d_k + w_k)}{d_k^T B_k^{-1} d_k} \cdot d_k^T s_k \\ &= s_k^T B_k s_k + \eta_k d_k^T s_k + w_k^T s_k - \eta_k d_k^T s_k + O(s_k^4) \\ &= s_k^T B_k s_k + w_k^T s_k + O(s_k^4). \end{aligned} \quad (5.3.57)$$

根据(5.3.55)式和(5.3.57)式，我们可以推出下式：

$$\begin{aligned} &v_k^T B_k v_k \cdot s_k^T B_k s_k - (v_k^T B_k s_k)^2 \\ &= (s_k^T B_k s_k + 2w_k^T s_k + O(s_k^4)) \cdot s_k^T B_k s_k - (s_k^T B_k s_k + w_k^T s_k + O(s_k^4))^2 \\ &= -(w_k^T s_k)^2 + O(s_k^6) = O(s_k^6). \end{aligned} \quad (5.3.58)$$

由(5.3.56)式及(5.3.58)式，我们可以推得(5.3.51)成立。证毕。 \square

引理 5.3.6. 如图 5.3-(ii)所示, 我们记:

$$d'_k = \frac{\Delta}{\|d_k + t_k \cdot v_k\|_2} \cdot (d_k + t_k \cdot v_k); \quad (5.3.59)$$

这里,

$$t_k = \frac{\|s_k\|_2}{\|v_k\|_2}. \quad (5.3.60)$$

则有:

$$\|d^* - d'_k\|_2 = O(s_k^2). \quad (5.3.61)$$

证明: 根据引理5.3.5及引理5.3.4, 我们可得:

$$\tau_k = 1 - \frac{(s_k^T v_k)^2}{(s_k^T s_k)(v_k^T v_k)} = O(s_k^2). \quad (5.3.62)$$

于是:

$$1 - \frac{|s_k^T v_k|}{\|s_k\|_2 \cdot \|v_k\|_2} = 1 - \sqrt{1 - \tau_k} = \frac{\tau_k}{1 + \sqrt{1 - \tau_k}} = O(s_k^2). \quad (5.3.63)$$

(5.3.63)表示当 $k \rightarrow +\infty$ 时, s_k 与 v_k 的夹角接近于0或者 π , 显然如果后者成立说明 $\exists \lambda > 0$ 使得:

$$v_k = -\lambda s_k + O(s_k^2); \quad (5.3.64)$$

代入(5.3.57)即得矛盾. 于是 s_k 与 v_k 的夹角接近于0, 也即 $s_k^T v_k > 0$:

$$1 - \frac{s_k^T v_k}{\|s_k\|_2 \cdot \|v_k\|_2} = 1 - \sqrt{1 - \tau_k} = \frac{\tau_k}{1 + \sqrt{1 - \tau_k}} = O(s_k^2). \quad (5.3.65)$$

根据(5.3.60)和(5.3.65)我们可得:

$$\begin{aligned} \|d^* - (d_k + t_k \cdot v_k)\|_2^2 &= \|s_k - t_k \cdot v_k\|_2^2 \\ &= \|s_k\|_2^2 + t_k^2 \|v_k\|_2^2 - 2t_k \|s_k\|_2 \cdot \|v_k\|_2 \frac{s_k^T v_k}{\|s_k\|_2 \cdot \|v_k\|_2} \\ &= 2\|s_k\|_2^2 \left(1 - \frac{s_k^T v_k}{\|s_k\|_2 \cdot \|v_k\|_2}\right) \\ &= O(s_k^4). \end{aligned} \quad (5.3.66)$$

根据(5.3.66), 我们有

$$\|d^* - (d_k + t_k \cdot v_k)\|_2 = O(s_k^2). \quad (5.3.67)$$

由于 d^* 和 d'_k 都在球面上, 因此我们容易知道:

$$\|d^* - (d_k + t_k \cdot v_k)\|_2 > \|d^* - d'_k\|_2. \quad (5.3.68)$$

由(5.3.67)和(5.3.68)可得(5.3.61)。证毕。□

定理 5.3.3. 假设问题(5.3.1)-(5.3.2)满足假设5.3.1、5.3.2和5.3.3。则迭代点列 $\{d_k\}$ 满足:

$$\|d^* - d_{k+1}\|_2 = O(\|d^* - d_k\|_2^2). \quad (5.3.69)$$

证明: 根据 d'_k 的定义, 我们知道

$$f(d^*) \leq f(d_{k+1}) \leq f(d'_k). \quad (5.3.70)$$

并根据(5.3.22), 引理5.3.6, 我们可以得到(5.3.69)。证毕。□

下面考虑算法5.3.2的全局收敛性:

引理 5.3.7. 设 $\{d_k\}$ 为算法5.3.2产生的迭代点列, 则:

$$f(d_k) - f(d_{k+1}) \geq \frac{\Delta m^2 \sqrt{m}}{2\sqrt{m}M\Delta + \sqrt{M} + 4mN} \cdot \|v_k\|_2^2; \quad (5.3.71)$$

这里 M 、 N 和 m 由定义5.3.1和定义5.3.3定义。

证明: 参照引理5.3.1, 球面 $\{d \mid \|d\|_2 = \Delta\}$ 与二维子空间 $\text{span}\{d_k, -\nabla f(d_k)\}$ 的交上的任何一点都可以表示为:

$$d(\alpha) = d_k - (1 - \cos \alpha) \cdot d_k + \frac{\Delta}{\|v_k\|_2} \sin \alpha \cdot v_k. \quad (5.3.72)$$

我们记:

$$c_1 = -(1 - \cos \alpha) = -2 \sin^2 \frac{\alpha}{2}, \quad c_2 = \frac{\Delta}{\|v_k\|_2} \sin \alpha. \quad (5.3.73)$$

并令:

$$\alpha_k = \arg \min_{\alpha \in [0, \pi]} f(d(\alpha)), \quad (5.3.74)$$

因此我们有: $d_{k+1} = d(\alpha_k)$ 。

根据定义5.3.1, 我们有 $\|\nabla^2 f(d)\|_2 < M, \forall \|d\|_2 \leq \Delta$, 因此:

$$f(d_k + s) \leq f(d_k) + s^T \nabla f(d_k) + \frac{1}{2} M \|s\|_2^2, \quad (5.3.75)$$

于是:

$$\begin{aligned}
f(d_k) - f(d(\alpha)) &\geq -\frac{1}{2}M\|c_1d_k + c_2v_k\|_2^2 - (c_1d_k + c_2v_k)^T \nabla f(d_k) \\
&= -\frac{1}{2}Mc_1^2\|d_k\|_2^2 - \frac{1}{2}Mc_2^2\|v_k\|_2^2 \\
&\quad - (c_1d_k + c_2v_k)^T \left(-u_k + \frac{\nabla f(d_k)^T d_k}{\Delta^2} \cdot d_k\right) \\
&= -\frac{1}{2}Mc_1^2\|d_k\|_2^2 - \frac{1}{2}Mc_2^2\|v_k\|_2^2 \\
&\quad - c_1 \nabla f(d_k)^T d_k + c_2(v_k^T u_k). \tag{5.3.76}
\end{aligned}$$

又由于:

$$\begin{aligned}
v_k^T u_k &= (B_k^{-1}u_k - \frac{d_k^T B_k^{-1}u_k}{d_k^T B_k^{-1}d_k} \cdot B_k^{-1}d_k)^T u_k \\
&= u_k^T B_k^{-1}u_k - \frac{(d_k^T B_k^{-1}u_k)^2}{d_k^T B_k^{-1}d_k} \\
&= v_k^T B_k v_k. \tag{5.3.77}
\end{aligned}$$

由(5.3.76)和(5.3.77), 我们可得:

$$\begin{aligned}
f(d_k) - f(d(\alpha)) &\geq -\frac{1}{2}Mc_1^2\|d_k\|_2^2 - \frac{1}{2}Mc_2^2\|v_k\|_2^2 - c_1 \nabla f(d_k)^T d_k + c_2(v_k^T B_k v_k) \\
&= -2M\Delta^2 \sin^4 \frac{\alpha}{2} - \frac{M}{2}\Delta^2 \sin^2 \alpha \\
&\quad - 2\Delta\|\nabla f(d_k)\|_2 \cos \theta \sin^2 \frac{\alpha}{2} + \Delta \frac{v_k^T B_k v_k}{\|v_k\|_2} \cdot \sin \alpha \\
&= -2M\Delta^2 \sin^2 \frac{\alpha}{2} - 2\Delta\|\nabla f(d_k)\|_2 \cos \theta \sin^2 \frac{\alpha}{2} + \Delta \frac{v_k^T B_k v_k}{\|v_k\|_2} \cdot \sin \alpha \\
&= -(M\Delta^2 + \Delta\|\nabla f(d_k)\|_2 \cos \theta)(1 - \cos \alpha) + \Delta \frac{v_k^T B_k v_k}{\|v_k\|_2} \cdot \sin \alpha \\
&= -T + \sqrt{T^2 + \Delta^2 \frac{(v_k^T B_k v_k)^2}{\|v_k\|_2^2}} \sin(\alpha + \varphi) \tag{5.3.78}
\end{aligned}$$

这里

$$T = M\Delta^2 + \Delta\|\nabla f(d_k)\|_2 \cos \theta, \quad \varphi = \arctg \frac{T\|v_k\|_2}{\Delta v_k^T B_k v_k}. \tag{5.3.79}$$

我们记 $\bar{\alpha} := \frac{\pi}{2} - \varphi$, 由(5.3.78)可得:

$$\begin{aligned}
f(d_k) - f(d_{k+1}) &\geq f(d_k) - f(d(\bar{\alpha})) \\
&= -T + \sqrt{T^2 + \Delta^2 \frac{(v_k^T B_k v_k)^2}{\|v_k\|_2^2}}. \tag{5.3.80}
\end{aligned}$$

另一方面我们有:

$$\frac{v_k^T B_k v_k}{\|v_k\|_2} \geq \sigma_{\min}(B_k) \|v_k\|_2 \geq \sigma_{\min}(\nabla^2 f(d_k)) \|v_k\|_2 \geq m \|v_k\|_2; \quad (5.3.81)$$

及:

$$\begin{aligned} \frac{v_k^T B_k v_k}{\|v_k\|_2} &\leq \frac{1}{\sqrt{\sigma_{\min}(B_k)}} \sqrt{v_k^T B_k v_k} \leq \frac{1}{\sqrt{m}} \sqrt{u_k^T B_k u_k} \\ &\leq \frac{\sqrt{M}}{\sqrt{m}} \|u_k\|_2 = \frac{\sqrt{M}}{\sqrt{m}} \cdot \|\nabla f(d_k)\|_2 \sin \theta. \end{aligned} \quad (5.3.82)$$

因此, 我们有:

$$\begin{aligned} f(d_k) - f(d_{k+1}) &\geq -T + \sqrt{T^2 + \Delta^2 \frac{(v_k^T B_k v_k)^2}{\|v_k\|_2^2}} = \frac{\Delta^2 \frac{(v_k^T B_k v_k)^2}{(\|v_k\|_2)^2}}{T + \sqrt{T^2 + \Delta^2 \frac{(v_k^T B_k v_k)^2}{(\|v_k\|_2)^2}}} \\ &\geq \frac{m^2 \Delta^2 \|v_k\|_2^2}{2T + \frac{\sqrt{M}}{\sqrt{m}} \cdot \Delta \|\nabla f(d_k)\|_2 \sin \theta} \\ &= \frac{\Delta^2 m^2 \sqrt{m} \|v_k\|_2^2}{2\sqrt{m}(M\Delta^2 + \Delta \|\nabla f(d_k)\|_2 \cos \theta) + \sqrt{M}\Delta \|\nabla f(d_k)\|_2 \sin \theta} \\ &= \frac{\Delta^2 m^2 \sqrt{m} \|v_k\|_2^2}{2\sqrt{m}M\Delta^2 + 2\sqrt{m}\Delta \|\nabla f(d_k)\|_2 \cos \theta + \sqrt{M}\Delta \|\nabla f(d_k)\|_2 \sin \theta} \\ &\geq \frac{\Delta^2 m^2 \sqrt{m} \|v_k\|_2^2}{2\sqrt{m}M\Delta^2 + \sqrt{M} + 4m\Delta \|\nabla f(d_k)\|_2} \\ &\geq \frac{\Delta m^2 \sqrt{m}}{2\sqrt{m}M\Delta + \sqrt{M} + 4mN} \cdot \|v_k\|_2^2. \end{aligned} \quad (5.3.83)$$

证毕。□

定理 5.3.4. 假设问题(5.3.1)-(5.3.2)满足假设5.3.1和假设5.3.3。则算法5.3.1或者有限终止, 或者产生的迭代点列 $\{d_k\}$ 满足:

$$\lim_{k \rightarrow +\infty} \|v_k\|_2 = 0. \quad (5.3.84)$$

证明: 如果算法5.3.1不有限终止。则根据引理5.3.7, 对任意 $k > 0$, 有

$$\begin{aligned} f(d_0) - f(d_{k+1}) &= \sum_{i=0}^k [f(d_i) - f(d_{i+1})] \\ &\geq \frac{\Delta m^2 \sqrt{m}}{2\sqrt{m}M\Delta + \sqrt{M} + 4mN} \sum_{i=1}^k \|v_i\|_2^2. \end{aligned} \quad (5.3.85)$$

由于 $f(d)$ 在 $\|d\|_2 = \Delta$ 上是下方有界的, 因此 $\exists V > 0$ 满足 $f(d_0) - f(d_{k+1}) < V, \forall k > 0$ 。所以(5.3.84)成立, 证毕。□

§5.3.3 小结

这里我们给出了直接的序列二维子空间算法5.3.1的全局收敛性和局部收敛速度的分析。在分析收敛速度时,我们做了假设5.3.2,这个假设在我们的证明中是无法消除的,因为在式(5.3.25)中,我们在最后一个不等式中丢弃了:

$$\begin{aligned} & \|\nabla^2 f(d^*)(d_k - d^*)\|_2^2 - \|(\nabla^2 f(d^*)(d_k - d^*))^T (\frac{1}{\Delta} d^*)\|_2^2 \\ &= \|\nabla^2 f(d^*)(d_k - d^*)\|_2^2 \sin^2 \gamma_k. \end{aligned} \quad (5.3.86)$$

这里 $\nabla^2 f(d^*)(d_k - d^*)$ 与 d^* 的夹角 γ_k 是无法估计出来的。事实上,虽然 $(d_k - d^*)$ 与 d^* 的夹角会趋近于 $\frac{\pi}{2}$,但不排除 γ_k 会接近于0。

Dai [20]中介绍的椭球外一点到椭球的距离,如果转化为信赖域子问题后,总是满足假设5.3.2的。因此可以说我们使用了完全不同的分析技巧(更为简洁),得到了比 [20]更一般的收敛性结论。

关于拉格朗日牛顿型的子空间选取方法,这只是我们对序列二维子空间算法的研究过程的一个中间步骤。我们的主要目的是以拉格朗日牛顿型的选取方法为基础,以期找到和它近似的子空间取法,使得新的取法具有较好的收敛性质(超线性收敛)并易于计算(不使用二阶信息)。而对基于拉格朗日牛顿型的子空间的算法5.3.2的局部收敛速度的分析,提出了球面上二维子空间二次收敛的分析方法。

§5.4 无约束优化的随机性算法

本节考虑求解如下无约束凸二次优化问题的随机性算法:

$$\min_{x \in \mathbf{R}^n} f(x) = \frac{1}{2} x^T A x + g^T x; \quad (5.4.1)$$

这里 A 是 $n \times n$ 的实对称正定矩阵。

发现和分析(5.4.1)的高效、低计算量的算法,是所有无约束优化算法设计的基础。

我们知道在求解(5.4.1)的梯度型算法中,采用精确线搜索的最速下降法,下面给出算法:

算法 5.4.1. (最速下降方法)

步一 给出初始点 $x_0 \in \mathbf{R}^n$; $0 < \varepsilon \ll 1$; 置 $k := 0$ 。

步二 计算 $d_k = -\nabla f(x_k)$; 如果 $\|d_k\|_2 \leq \varepsilon$, 则算法终止; 否则, 转步三。

步三 求精确搜索步长:

$$\alpha_k = \arg \min_{\alpha} f(x_k + \alpha d_k). \quad (5.4.2)$$

步四 更新 $x_{k+1} = x_k + \alpha_k d_k$; 置 $k := k + 1$; 转步二。

因为算法5.4.1会产生Zigzag效应 [124, 166], 效率并不高; 而Barzilai和Borwein [5]发现的BB步长虽然不能保证每步迭代单调下降, 但是数值表现非常不错, 在一些条件下BB方法具有超线性收敛性 [21]; Yuan和Dai [145, 144] 提出了其它一些单调的步长取法, 也具有不错的数值表现。最近又有学者提出了一种步长随机选取的方法:

算法 5.4.2. (随机步长算法)

基于算法5.4.1, 并在其步三与步四间插入一步:

步NEW: 随机取 λ , 服从 $[0.5, 1.5]$ 中的均匀分布; 置 $\alpha_k := \lambda \alpha_k$ 。

另外一个比较直观的随机算法是方向取为随机:

算法 5.4.3. (随机方向算法)

基于算法5.4.1, 并在其步二与步三间插入一步:

步NEW: 随机取 r_k , 服从单位球面 $\{z \mid \|z\|_2 = 1\}$ 上的均匀分布;

若 $r_k^T d_k \geq 0$, 置 $d_k := r_k$; 否则置 $d_k := -r_k$ 。

与负梯度方向加权, 我们可以改进随机方向为:

算法 5.4.4. (改进的随机方向算法)

基于算法5.4.1, 并在其步二与步三间插入一步:

步NEW: 随机取 r_k , 服从单位球面 $\{z \mid \|z\|_2 = 1\}$ 上的均匀分布;

若 $r_k^T d_k \geq 0$, 置 $d_k := \frac{1}{2}(r_k + d_k / \|d_k\|_2)$; 否则置 $d_k := \frac{1}{2}(-r_k + d_k / \|d_k\|_2)$ 。

此外还能构造随机水平集方法:

算法 5.4.5. (随机水平集方法)

步一 给出初始点 $x_0 \in \mathbb{R}^n$; $0 < \varepsilon \ll 1$; 置 $k := 0$ 。

步二 计算 $d_k = -\nabla f(x_k)$; 如果 $\|d_k\|_2 \leq \varepsilon$, 则算法终止; 否则, 转步三。

步三 随机选取 x_{k+1} , 服从椭球 $\{z \mid f(z) \leq f(x_k)\}$ 内的均匀分布;

置 $k := k + 1$; 转步二。

最后回到本文第三章、第四章和本章第三节中多次提到的序列二维子空间极小方法, 我们知道二维子空间的选取将对算法的效率有重要的影响。假设我们用序列二维子空间极小化方法求解(5.4.1), 假设当前迭代为 x_k , 令

$$\mathcal{Y} = \text{span}\{\nabla f(x_k), s_k\}; \quad (5.4.3)$$

这里 $\nabla f(x_k) = Ax_k + g$, $s_k = x_k - x_{k-1}$ 。我们只要每步迭代都根据(5.4.2)选取二维子空间 \mathcal{Y} , 就将产生和共轭梯度法 [124, 166]相同的点列。

为了保证算法的下降, 一般来说我们总是会选取 $\nabla f(x_k)$ 作为子空间的一个方向。下面给出另一个方向采用随机选取的序列随机二维子空间方法:

算法 5.4.6. (序列随机二维子空间方法)

步一 给出初始点 $x_0 \in \mathbb{R}^n$; $0 < \varepsilon \ll 1$; 置 $k := 0$ 。

步二 计算 $d_k = -\nabla f(x_k)$; 如果 $\|d_k\|_2 \leq \varepsilon$, 则算法终止; 否则, 转步三。

步三 随机取 r_k , 服从单位球面 $\{z \mid \|z\|_2 = 1\}$ 上的均匀分布; 构建子空间 $\mathcal{Y} = \text{span}\{d_k, r_k\}$ 。

步四 求解二维子问题的精确全局极小:

$$s_k = \arg \min_{s \in \mathcal{Y}} f(x_k + s). \quad (5.4.4)$$

步五 更新 $x_{k+1} = x_k + s_k$; 置 $k := k + 1$; 转步二。

§5.4.1 数值比较

本节我们就来分析一下各种随机算法的数值表现。我们利用Stafford [119]提出的近似方法来产生球内服从均匀分布的随机变量。其基本思想是利用如下近似公式:

$$\frac{x}{\|x\|_2} \left(\frac{1}{\Gamma(\frac{n}{2})} \int_0^{\frac{\|x\|_2}{2}} t^{\frac{n}{2}-1} e^{-t} dt \right)^{\frac{1}{n}}. \quad (5.4.5)$$

这里随机向量 x 是满足协方差矩阵为 I 的正态分布。这里:

$$\Gamma\left(\frac{n}{2}\right) = \begin{cases} \frac{n!}{2}, & n \text{ 是偶数;} \\ \frac{(n-2)! \sqrt{\pi}}{2^{n-2}}, & n \text{ 是奇数.} \end{cases} \quad (5.4.6)$$

关于 Γ 函数的一般定义见下一节。

我们的试验问题为：

$$\min f(x) = \frac{1}{2}x^T Ax. \quad (5.4.7)$$

这里对角矩阵 A 满足：

$$A_{ii} = \text{power}(10, \frac{i-1}{n-1} \cdot \text{ncond}). \quad (5.4.8)$$

初始点为：

$$x_0 = \text{rand}(n, 1). \quad (5.4.9)$$

终止准则为：

$$\|g\| < 1.e - 6; \text{ 或者 } ITMAX = 10000. \quad (5.4.10)$$

下表为计算结果，数据为十个不同初始值的平均迭代步数。

No.	算法5.4.1	算法5.4.2	算法5.4.3
$n = 2$ ncond = 3	5.7	428.5	218.5
$n = 2$ ncond = 5	4.2	3776.6	1850.9
$n = 2$ ncond = 10	9.0	ITMAX	ITMAX
$n = 5$ ncond = 3	4344.6	464.1	3506.6
$n = 5$ ncond = 5	ITMAX	5576.5	ITMAX
$n = 5$ ncond = 10	ITMAX	ITMAX	ITMAX
$n = 10$ ncond = 3	5861.9	532.5	ITMAX
$n = 10$ ncond = 5	ITMAX	5334.0	ITMAX
$n = 10$ ncond = 10	ITMAX	ITMAX	ITMAX
$n = 100$ ncond = 3	6570.5	583.1	ITMAX
$n = 100$ ncond = 5	ITMAX	6625.1	ITMAX
$n = 1000$ ncond = 3	7173.2	617.1	ITMAX
$n = 1000$ ncond = 5	ITMAX	6386.8	ITMAX

表5.4.1(i) 随机算法数值表现比较

No.	算法5.4.4	算法5.4.5	算法5.4.6
$n = 2$ ncond = 3	227.6	40.2	1
$n = 2$ ncond = 5	2208.7	45.3	1
$n = 2$ ncond = 10	ITMAX	70.1	1
$n = 5$ ncond = 3	1674.1	98.7	445.7
$n = 5$ ncond = 5	ITMAX	111.4	6136.6
$n = 5$ ncond = 10	170.6	ITMAX	ITMAX
$n = 10$ ncond = 3	3517.9	196.6	2026.2
$n = 10$ ncond = 5	ITMAX	229.2	ITMAX
$n = 10$ ncond = 10	ITMAX	339.8	ITMAX
$n = 100$ ncond = 3	6310.2	ITMAX	6093.8
$n = 100$ ncond = 5	ITMAX	ITMAX	ITMAX
$n = 1000$ ncond = 3	7405.9	ITMAX	7178.9
$n = 1000$ ncond = 5	ITMAX	ITMAX	ITMAX

表5.4.1(ii) 随机算法数值表现比较(续)

通过表5.4.1, 我们知道随机步长算法5.4.2对于大规模、病态的问题表现最好。序列随机二维子空间方法并没有比最速下降法好多少, 因此, 我们可以认为序列二维子空间方法中方向的选取还是特别重要的。随机水平集方法对问题的条件数并不敏感, 但对于维数的变化很敏感。

§5.4.2 随机水平集算法的收敛性

不失一般性, 我们不妨对(5.4.1)中的目标函数做如下假设:

$$f(x) = x^T A x. \quad (5.4.11)$$

且 $A = \text{diag}\{\frac{1}{a_1}, \frac{1}{a_2}, \dots, \frac{1}{a_n}\}$ 。于是我们容易知道最优解是 $x = 0$ 。

假设当前迭代点是 x_k , 下一个迭代点是 x_{k+1} (随机点, 服从 $\{z \mid f(z) \leq f(x_k)\}$ 内的均匀分布, 见算法5.4.5步三)。下面我们考虑当 x_k 确定时, $f(x_{k+1})/f(x_k)$ 的期望:

$$\mathbf{E}(f(x_{k+1})/f(x_k)). \quad (5.4.12)$$

我们可以用(5.4.12)来衡量算法关于函数值的收敛速度。

首先我们考虑二维的情形, 来梳理(5.4.12)的计算步骤。

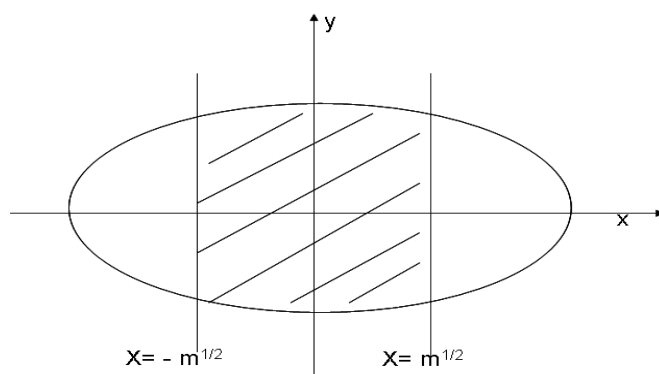


图 5.4

引理 5.4.1. 已知

$$f(z) = \frac{x^2}{a^2} + \frac{y^2}{b^2}, \quad (5.4.13)$$

随机变量: $Z := \{X, Y\}$ 服从椭球 $\Omega = \{z | z^T A z \leq 1\}$ 内的均匀分布, 则:

$$\mathbf{E}(Z^T A Z) = \frac{1}{2}. \quad (5.4.14)$$

证明: 我们知道:

$$\begin{aligned} \mathbf{E}(Z^T A Z) &= \mathbf{E}\left(\frac{X^2}{a^2} + \frac{Y^2}{b^2}\right) \\ &= \frac{\mathbf{E}X^2}{a^2} + \frac{\mathbf{E}Y^2}{b^2}. \end{aligned} \quad (5.4.15)$$

下面我们先来考虑随机变量 X^2 的期望 $\mathbf{E}X^2$ 。我们首先需要计算 X^2 的分布函数:

$$\begin{aligned} \Pr(X^2 \leq m) &= \Pr(-\sqrt{m} \leq X \leq \sqrt{m}) \\ &= \frac{S(\text{阴影部分面积})}{S(\text{整个椭圆面积})} \quad (\text{如图5.4}) \\ &= \frac{4 \int_0^{\sqrt{m}} b \sqrt{1 - \frac{x^2}{a^2}} dx}{\pi ab} \\ &= \frac{4 \int_0^{\arcsin \frac{\sqrt{m}}{a}} ab \cos^2 \theta d\theta}{\pi ab} \\ &= \frac{ab(\sin 2\theta + 2\theta)_0^{\arcsin \frac{\sqrt{m}}{a}}}{\pi ab} \\ &= \frac{2}{\pi} \left(\frac{\sqrt{m}}{a} \times \sqrt{1 - \frac{m}{a^2}} + \arcsin \frac{\sqrt{m}}{a} \right). \end{aligned} \quad (5.4.16)$$

根据(5.4.15), 我们可以得到随机变量 X^2 的密度函数:

$$p(m) = \frac{2}{\pi a} m^{-\frac{1}{2}} \left(1 - \frac{m}{a^2}\right)^{\frac{1}{2}}. \quad (5.4.17)$$

于是我们就可以计算期望了:

$$\begin{aligned} \mathbf{E}(X^2) &= \int_0^{a^2} \frac{2}{\pi a} m^{\frac{1}{2}} \left(1 - \frac{m}{a^2}\right)^{\frac{1}{2}} dm \\ &= \int_0^{\frac{\pi}{2}} \frac{4}{\pi} a^2 \sin^2 \theta \cos^2 \theta d\theta \\ &= \frac{a^2}{\pi} \left(\frac{\theta}{2} - \frac{\sin 4\theta}{8}\right) \Big|_0^{\frac{\pi}{2}} \\ &= \frac{a^2}{4}. \end{aligned} \quad (5.4.18)$$

根据对称性, 我们有:

$$\mathbf{E}(Y^2) = \frac{b^2}{4}. \quad (5.4.19)$$

由(5.4.15)式、(5.4.18)式和(5.4.19)式, 我们得到:

$$\begin{aligned} \mathbf{E}(Z^T AZ) &= \frac{\mathbf{E}X^2}{a^2} + \frac{\mathbf{E}Y^2}{b^2} \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2}. \end{aligned} \quad (5.4.20)$$

证毕。 □

在考虑 n 维情形之前, 我们先复习一下B函数的相关定义和性质([150], [3])。

定义 5.4.1. 我们称 $B(x, y)$ 为一个B函数(或者第一类欧拉积分), 如果:

$$B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt. \quad (5.4.21)$$

对任何 $x > 0, y > 0$ 。

定义 5.4.2. 我们称 $\Gamma(x, y)$ 为一个伽玛函数, 如果:

$$\Gamma(z) = \int_0^{+\infty} e^{-t} t^{z-1} dt. \quad (5.4.22)$$

对任何 $z > 0$ 。

关于B函数和伽玛函数, 我们有如下命题成立:

命题 5.4.1.

- 1) $B(x, y) = \frac{\Gamma(x+y)}{\Gamma(x)\Gamma(y)}$;
- 2) $B(x, y) = 2 \int_0^{\frac{\pi}{2}} \sin^{2x-1} \theta \cos^{2y-1} \theta d\theta$;
- 3) $\Gamma(x+1) = x\Gamma(x)$.

下面我们开始证明本节最关键的引理:

引理 5.4.2. 已知 $f(x)$ 如 (5.4.11) 定义, 随机变量 $X := \{X_1, X_2, \dots, X_n\} (n \geq 2)$ 服从椭球 $\Omega = \{x | x^T A x \leq 1\}$ 内的均匀分布, 则:

$$\mathbf{E}(X^T A X) = \frac{n}{n+2}. \quad (5.4.23)$$

证明: 我们采用引理 5.4.1 的证明技巧:

$$\begin{aligned}
Pr(X_1^2 \leq m_1) &= Pr(-\sqrt{m} \leq X_1 \leq \sqrt{m}) \\
&= \frac{\int_{-\sqrt{m}}^{\sqrt{m}} \int_{\frac{x_2^2}{a_2^2} + \dots + \frac{x_n^2}{a_n^2} \leq 1 - \frac{x_1^2}{a_1^2}} dx_n \dots dx_2 dx_1}{\int_{\frac{x_1^2}{a_1^2} + \dots + \frac{x_n^2}{a_n^2} \leq 1} dx_n \dots dx_2 dx_1} \\
&= \frac{\int_{-\sqrt{m}}^{\sqrt{m}} \int_{\frac{x_2^2}{a_2^2} + \dots + \frac{x_{n-1}^2}{a_{n-1}^2} \leq 1 - \frac{x_1^2}{a_1^2}} 2a_n \Delta_{n-1}^{\frac{1}{2}} dx_{n-1} \dots dx_2 dx_1}{\int_{\frac{x_1^2}{a_1^2} + \dots + \frac{x_{n-1}^2}{a_{n-1}^2} \leq 1} 2a_n \Delta_{n-1}^{\frac{1}{2}} dx_{n-1} \dots dx_2 dx_1} \\
&= \frac{\int_{\frac{x_2^2}{a_2^2} + \dots + \frac{x_{n-2}^2}{a_{n-2}^2} \leq 1 - \frac{x_1^2}{a_1^2}} 2a_{n-1} a_n \Delta_{n-2}^{\frac{2}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta dx_{n-2} \dots dx_2 dx_1}{\int_{\frac{x_1^2}{a_1^2} + \dots + \frac{x_{n-2}^2}{a_{n-2}^2} \leq 1} 2a_{n-1} a_n \Delta_{n-2}^{\frac{2}{2}} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta dx_{n-2} \dots dx_2 dx_1} \\
&= \dots \\
&= \frac{2a_2 a_3 \dots a_n \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta \dots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2} \theta d\theta \int_{-\arcsin \frac{\sqrt{m}}{a_1}}^{\arcsin \frac{\sqrt{m}}{a_1}} \cos^n \theta d\theta}{2a_2 a_3 \dots a_n \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta d\theta \dots \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^{n-2} \theta d\theta \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n \theta d\theta} \\
&= \frac{\int_0^{\arcsin \frac{\sqrt{m}}{a_1}} \cos^n \theta d\theta}{\int_0^{\frac{\pi}{2}} \cos^n \theta d\theta}, \quad (5.4.24)
\end{aligned}$$

这里:

$$\Delta_n = \left(1 - \frac{x_1^2}{a_1^2} - \dots - \frac{x_n^2}{a_n^2}\right). \quad (5.4.25)$$

因此密度函数为:

$$p(m) = \frac{\left(1 - \frac{m}{a_1^2}\right)^{\frac{n-1}{2}} m^{-\frac{1}{2}}}{2a_1 \int_0^{\frac{\pi}{2}} \cos^n \theta d\theta}. \quad (5.4.26)$$

所以:

$$\begin{aligned} \mathbf{E}(X_1^2) &= \frac{\left(1 - \frac{m}{a_1^2}\right)^{\frac{n-1}{2}} m^{\frac{1}{2}}}{2a_1 \int_0^{\frac{\pi}{2}} \cos^n \theta d\theta} \\ &= \frac{a_1^2 \int_0^{\frac{\pi}{2}} \cos^n \theta \sin^2 \theta d\theta}{\int_0^{\frac{\pi}{2}} \cos^n \theta d\theta}. \end{aligned} \quad (5.4.27)$$

于是我们可以得到 n 维情形的期望:

$$\begin{aligned} \mathbf{E}(X^T A X) &= \sum_{i=1}^n \frac{\mathbf{E}X_i^2}{a_i^2} \\ &= n \cdot \frac{\int_0^{\frac{\pi}{2}} \cos^n \theta \sin^2 \theta d\theta}{\int_0^{\frac{\pi}{2}} \cos^n \theta d\theta}. \end{aligned} \quad (5.4.28)$$

根据定义5.4.1、定义5.4.2及命题5.4.1, 我们有

$$\begin{aligned} \mathbf{E}(X^T A X) &= n \cdot \frac{B\left(\frac{3}{2}, \frac{n+1}{2}\right)}{B\left(\frac{1}{2}, \frac{n+1}{2}\right)} \\ &= n \cdot \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+4}{2}\right)} / \frac{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n+2}{2}\right)} \\ &= n \cdot \frac{\Gamma\left(\frac{3}{2}\right)\Gamma\left(\frac{n+2}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{n+4}{2}\right)} \\ &= n \cdot \frac{1}{2} \cdot \frac{2}{n+2} \\ &= \frac{n}{n+2}. \end{aligned} \quad (5.4.29)$$

证毕。□

由引理5.4.2, 我们就可以得到下面的收敛性定理:

定理 5.4.1. 假设用算法5.4.5求解(5.4.1)产生的点列是 $\{x_k\}$, 记 x^* 为(5.4.1)的最优解。则:

$$\mathbf{E}((f(x_{k+1}) - f(x^*)) / (f(x_k) - f(x^*))) = \frac{n}{n+2}. \quad (5.4.30)$$

证明： 首先由于我们知道算法5.4.5具有平移不变性、旋转不变性、数乘不变性，因此我们可以不妨假设 $f(x)$ 具有(5.4.11)的形式且 $f(x_k) = 1$ ，于是有 $x^* = 0$ ， $f(x^*) = 0$ 。根据引理5.4.2，(5.4.30)成立。证毕。 \square

根据(5.4.1)，随机水平集算法5.4.5的平均收敛速度仅仅和问题的维数相关，而与问题的条件数无关。这个结论和实验结果非常吻合。

第六章 结论与展望

我们讨论了可分非线性最小二乘问题的求解，提出了一种结构型变量不分离的算法框架。基于该框架，我们提出了三种算法格式。数值实验和算法收敛性分析都显示出我们提出的算法具有和经典的变量投影方法一样的收敛性质。同时我们的算法可以很容易地推广到约束优化问题，这是变量投影方法不具备的优点。

对可分非线性最小二乘问题，还有一些问题有待考虑。比如对于约束具有和目标函数不同的可分离格式的问题，我们该如何设计算法使得约束的特殊形式也能同时被利用。最近我们发现在求解三角形或者四面体元上的数值积分点问题 [147]，最后也归结为可分非线性最小二乘问题，不过问题的难点在于要求问题的全局极小点。我们下一步工作将考虑利用问题的变量可分离性质来设计这类问题的全局极小算法。

我们讨论了二次最小二乘问题的全局极小算法。我们提出全局下降二维子空间方向，并以之设计了随机二维子空间算法。我们的方法在求解来源于无线传感器网络定位问题的距离几何模型时，有很好的数值表现。

我们未来的工作，主要集中在试图找到我们的随机二维子空间方法比一般的随机多初始方法好的理论依据，以及考虑算法中参数选取对解的影响。在求解来源于蛋白质空间构型的距离几何模型时，我们的算法表现不尽如人意，下一步可以考虑的是结合几何构造方法 [131]来构造专门适用于蛋白质结构计算的算法。

我们分析了求解 L_1 范数极小化问题的信赖域子问题的计算复杂度；设计了低维非光滑信赖域子问题的求解方法；并提出了一种序列二维子空间方法来求解非光滑的信赖域子问题。我们发现使用两个下降方向杂交的线搜索方法可以保证充分下降量，并以之分析了算法的收敛性。

下一步需要考虑的问题是，如何结合我们设计的信赖域子问题与信赖域框架算法来计算非线性 L_1 范数极小化问题。我们还计划研究非光滑信赖域子问题的截断共轭梯度方法。此外，考虑到非线性 L_1 范数极小化问题的非光滑点的测度为零，也就是说绝大部分的点都是可微的，我们考虑研究非线性 L_1 范数极小化问题的光滑化技巧，使得在不可微点及其附近能够定义一个连续的梯度函数，这样就可以结合一般的信赖域算法来求解非线性 L_1 范数极小化问题。

我们构建了线性最小二乘参数估计与观测次数的关系，证明了线性最小二乘估计的误差随着观测次数的增大单调下降。下一步可以考虑的是如何在非线性最

小二乘估计中，得到类似的结果。

我们提出了凸二次规划的不可行原始对偶积极集方法，并发现该算法有很好的数值表现。未来的工作是分析这个算法的收敛性及解释它的数值表现优异的原因。我们还可以结合求解鞍点问题的算法来提高算法的效率，使之能够适应实际工程计算。

我们给出了序列二维子空间极小方法的收敛性和收敛速度的分析。提出并分析了拉格朗日牛顿型子空间取法，下一步工作的计划是构造拟拉格朗日型子空间取法，使之能够真正适应实际计算。

我们比较了各种无约束问题的随机性算法；并且证明了随机水平集方法的收敛性。未来的工作计划是据此分析随机步长的梯度型方法的收敛性。

参考文献

- [1] Bernardetta Addis, Marco Locatelli, Fabio Schoen, Local Optima Smoothing for Global Optimization, *Technical report DSI*, 5-2003, 2003.
- [2] Bernardetta Addis, Sven Leyffer, A Trust-Region Algorithm for Global Optimization, *ANL/MCS-P1190-0804*, 2004.
- [3] R. A. Askey and R. Roy, Beta function, *Digital Library of Mathematical Functions*, *N.I.S.T.*, R. F. Boisvert, C. F. Ronald, C. W. Clark, and D. M. Lozier et al., 2008.
- [4] M. C. Bartholomew-Biggs, The Estimation of the Hessian Matrix in Nonlinear Least Squares Problems with Non-zero Residuals, *Math. Prog.*, 12(1977):67–80.
- [5] J. Barzilai, J. M. Borwein, Two-point Step Size Gradient Methods, *IMA J. Numer. Anal.*, 8(1988): 141–148.
- [6] M. Benzi, G. H. Golub, and J. Liesen, Numerical solution of saddle point problems, *Acta Numerica*, 14(2005): 1–137.
- [7] Å. Björck, Numerical Methods for Least Squares Problems, *SIAM*, 1996.
- [8] J. Bochnak, M. Coste, and M. F. Roy, Real Algebraic Geometry, *Springer*, 1998.
- [9] C. Böckmann, A modification of the trust-region Gauss-Newton method to separable nonlinear least squares problems, *J. Math. Systems, Estimation and Control*, 5(1995), 1–16.
- [10] J. H. Bramble and J. E. Pasciak, Analysis of the Inexact Uzawa Algorithm for Saddle Point Problems, *SIAM J. Numer. Anal.*, 34(1997): 1072–1092.
- [11] F. H. Clark, Generalized Gradients and Applications, *Trans. of Amer. Math. Soc.*, 205(1975): 247–262.
- [12] F.H. Clark, Optimization and Nonsmooth Analysis, *New York: John Wiley and Sons*, 1983.
- [13] A. R. Conn, N. I. M. Gould and Ph. L. Toint, Trust-Region Methods, *MPS SIAM Series on Optimization (Philadelphia)*, 2000.
- [14] S. A. Cook, The Complexity of theorem Proving Procedures, *New York: ACM Proceedings, Third Annual ACM Symposium on the Theory of Computing*, 1971, 151–158.
- [15] R. Cottle, J. S. Pang and R. E. Stone, The Linear Complementarity Problem, *Academic Press*, 1992.

- [16] D. A. Cox, J. B. Little, and D. O'Shea, Ideals, varieties, and algorithms: an introduction to computational algebraic geometry and commutative algebra. *Springer*, 1997.
- [17] E. W. Cheney and A. A. Goldstein, Newton's Method for Convex Programming and Chebyshev Approximation. *Numerische Mathematik*, 1(1959): 253–268.
- [18] David A. Cox, John Little, Donal O' Shea, Algebraic Geometry, *GTM 185*, Springer-Verlag, New York 2005.
- [19] Y. H. Dai, Fast Algorithms on Projection on an Ellipsoid, *University of Dundee*, Report NA/220(2004).
- [20] Y. H. Dai, Fast Algorithms on Projection on an Ellipsoid, *SIAM Journal on Optimization*, 16(2006): 986–1006.
- [21] Y. H. Dai and R. Fletcher, On the Asymptotic Behaviour of Some New Gradient Methods, *Mathematical Programming (Series A)*, 13(2003): 541–559
- [22] Y. H. Dai and R. Fletcher, Projected Barzilai-Borwein Methods for Large-Scale Box-Constrained Quadratic Programming, *Numerische Mathematik*, 100(2003): 21–47.
- [23] Y. H. Dai and R. Fletcher, New Algorithms for Singly Linearly Constrained Quadratic Programs Subject to Lower and Upper Bounds, *Math. Prog.*, 106(2006): 403–421.
- [24] Y. H. Dai and Y. Yuan, Analyses of Monotone Gradient Methods, *Journal of Industry and Management Optimization*, 1(2005): 181-192.
- [25] J. E. Dennis, Some Computational Techniques for the Nonlinear Least Squares Problems, *Numerical Solution of System of Nonlinear Algebraic Equations*, Academic Press, New York, 1973: 157-183
- [26] J. E. Dennis, D. M. Gay, and R. E. Welsch, Algorithm 573 NL2SOL - An Adaptive Nonlinear Least Squares Algorithm, *ACM Transaction on Math. Software*, 7(1981): 369–383.
- [27] J. E. Dennis, H. J. Martinez, R. A. Tapia, Convergence Theory for the Structured BFGS Secant Method with an Application to Nonlinear Least Squares, *J. Optim. Theory Appl.*, 61(1989): 161–178.
- [28] J. E. Dennis, Jr. and R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations (Classics in Applied Mathematics 16), *SIAM*, Philadelphia, 1996.
- [29] E. D. Dolan and J. J. Moré, Benchmarking optimization software with performance profiles, *Mathematical Programming*, 91(2002): 201-213.

-
- [30] H. C. Elman and G. H. Golub, Inexact and Preconditioned Uzawa Algorithms for Saddle Point Problems, *SIAM Journal on Numerical Analysis*, 31(1994): 1645–1661.
- [31] I. I. Eremin, A Generalization of the Motzkin-Agmon Relaxation Method, *Soviet Math. Doklady*, 6(1965): 219–221.
- [32] J.Y. Fan and Y.X. Yuan, On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption, *Computing*, 74(2005) 23–39.
- [33] R. Fletcher, Practical Methods of Optimization, Vol. 2, *Chechester: John Wiley and Sons*, 1981.
- [34] R. Fletcher, A Model Algorithm for Composite NDO Optimization, *Math. Prog., Study*, 17: 67–76.
- [35] R. Fletcher and C. Xu, Hybrid Methods for Nonlinear Least Squares, *IMA J. Numer. Anal.*, 7(1987): 371–389.
- [36] M. R. Garey and D. S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness *New York: W.H. Freeman*, 1979.
- [37] G.H. Golub and V. Pereyra. The differentiation of pseudo-inverses and nonlinear least squares whose variables separable. *SIAM J. Numer. Anal.*, 10(1973), 413–432.
- [38] G.H. Golub and V. Pereyra. Separable nonlinear least squares: the variable projection method and applications, *Inverse Problems*, 19(2002), 1-26.
- [39] L. Kaufman, A variable projection method for solving separable nonlinear least squares problems, *BIT*, 15(1975), 49–57.
- [40] L. Kaufman and V. Pereyra, The separable nonlinear least squares problems with separate nonlinear equality constraints, *SIAM J. Numer. Anal.*, 15(1978), 12–20.
- [41] J. E. Kelley, The Cutting Plane Method for Solving Convex Programs. *J. of SIAM*, 8(1960): 703–712.
- [42] F.T. Krogh, Efficient implementation of a variable projection algorithm for nonlinear least squares problems, *Comm. ACM*, 17(1974), 167–169.
- [43] K. Kunisch and F. Rendl, An Infeasible Active Set Method for Quadratic Problems with Simple Bounds, *SIAM Journal on Optimization*, 14(2003): 35-52.
- [44] J. Hald, A 2-stage Algorithm for Nonlinear L_1 Optimization, *University of Denmark*, Report No. NI-81-03, Inst. for Num. Analysis, Tech.(1981).

- [45] J. Hald and K. Madsen, Combined LP and Quasi-Newton Methods for Nonlinear L_1 Optimization, *SIAM J. Numerical Analy.*, 22(1985): 68–80.
- [46] E. T. Hale, W. Yin, and Y. Zhang, A Fixed-Point Continuation Method for L_1 -Regularized Minimization with Applications to Compressed Sensing, *Rice University, CAAM Technical Report TR07-07*(2007).
- [47] R. Horst and P. M. Pardalos, *Handbook of Global Optimization*, Kluwer Academy Publishers, 1995.
- [48] J. B. Lasserre, Global Optimization with Polynomials and the Problem of Moments, *SIAM Journal of Optimiziation*, 11(2001): 796–817.
- [49] C. Lemaréchal, Bundle Methods in Nonsmooth Optimization, *Nonsmooth Optimization, Pergamon, Oxford*, C. Lemaréchal, and R. Mifflin eds., (1978): 79–102.
- [50] C. Lemaréchal C. Nonsmooth Optimization and Descent Methods, *RR-78-4, IIASA report*, 1978.
- [51] C. Lemaréchal, J. J. Strodiot, and A. Bihain, On a Bundle Algorithm for Nonsmooth Optimization *Nonlinear Programming 4. New York: Academic Press*, O. L. Mangasarian, P. R. Meyer, and S. M. Robinson ed., (1981): 245–282.
- [52] K. Levenberg, A Method for the Solution of Certain Nonlinear Problems in Least Squares. *Qart. Appl. Math.* 2(1944): 164-166.
- [53] T. Y. Li, Numerical Solution of Multivariate Polynomial Systems by Homotopy Continuation Methods, *Acta Numerical*, 6(1997): 399-436.
- [54] T. Li and F. S. Bai, Minimizing Multi-homogeneous Bézout Numbers, *Mathematics of Computing*, 70(2000): 767–787.
- [55] T. Li, Z. J. Lin and F. S. Bai, Heuristic Methods for Computing the Minimal Multi-homogeneous Bézout Number, *Applied Mathematics and Computation*, 146(2003): 237–256.
- [56] X. Liu, An Efficient Unseparated Scheme for Separable Nonlinear Least Squares Problem, Proceedings of the Eighth National Conference of Operation Research Society of China(Global-Link, Hong Kong), (2006): 132–137.
- [57] X. Liu and Ya-xiang Yuan, On the Separable Nonlinear Least Squares Problems, *Journal of Computational Mathematics*, 26(2008): 390–403.

-
- [58] X. Liu, Global Minimization of Quadratic Least-Squares Problems, Proceedings of the Ninth National Conference of Operation Research Society of China(Global-Link, Hong Kong), (2008): 188–193.
- [59] X. Liu and Y. X. Yuan, On the Trust Region Subproblem for Nonlinear L1 Norm Minimization Problem, finished.
- [60] X. Liu, S. McKee, J. Y. Yuan and Y. X. Yuan, Uniform Bound on the 1-norm of the Inverse of Lower Triangular Toeplitz Matrices, finished.
- [61] X. Liu, Efficient Unseparated Algorithms for Separable Nonlinear Least Squares Problems, presented on the International Conference on Numerical Analysis and Optimization dedicated to M. J. D. Powell's 70th Birthday, Beijing, China, 2006.
- [62] X. Liu, Global Nonlinear Least Squares in Data Analysis, post presentation on the First Joint China Operations Research Days Workshop, IBM Research Center, Beijing, China, 2007.
- [63] X. Liu, Global Minimization on Quadratic Least Squares Problems, presented on the 6th International Congress on Industrial and Applied Mathematics, Zurich, Switzerland, 2007.
- [64] X. Liu, Global Minimization on Quadratic Least Squares Problems, presented on the 6th International Conference on Numerical Linear Algebra and Optimization, Urumqi, China, 2007.
- [65] X. Liu, New Efficient Sampling Methods for Distance Geometry Problems, presented at Siemens Corporate Technology, Munich, Germany, 2007.
- [66] X. Liu, New Sampling Techniques for Distance Geometry Problems, presented at Technische Universitat Darmstadt, Darmstadt, Germany, 2007.
- [67] X. Liu, New Methods for Distance Geometry Problems, presented at Zuse Institute Berlin, Berlin, Germany, 2007.
- [68] X. Liu, L_1 Norm Minimization for Quadratic System, presented on the VII Brazilian Workshop on Continuous Optimization, Campinas, Brazil, 2008.
- [69] D. W. Marquardt, An Algorithm for Least-squares Estimation of Nonlinear Inequalities. *SIAM J. Appl. Math.* 11(1963): 431-441.
- [70] J. J. Moré, The Levenberg-Marquardt Algorithm: Implementation and Theory, *Lecture Notes in Mathematics 630: Numerical Analysis*, Springer-Verlag, Berlin 1978: 105–116.

- [71] J. J. Moré, Recent Developments in Algorithms and Software for Trust Region Methods, in: A. Bachem, M. Groetschel and B. Korte, eds., *Mathematical Programming: The State of the Art*, Springer, Berlin, 1983: 258–287.
- [72] J. J. Moré and D. C. Sorensen, Computing a Trust Region Step. *SIAM J. Sci. Statist. Comput.*, 4(1983): 553–572.
- [73] J. J. Moré and G. Toraldo, Algorithms for Bound Constrained Quadratic Programming Problems, *Numer. Math.*, 55(1989), 377–400.
- [74] J. J. Moré and Z. J. Wu, Global Continuation for Distance Geometry Problems. *SIAM Journal on Optimization*, Volume 7, Issue 4 (1997): 814–836.
- [75] J. W. Nie, Sum of Squares Methods for Sensor Network Localization. *Computational Optimization and Applications*, to appear.
- [76] J. Nocedal and S.J. Wright, Numerical Optimization, *Springer-Verlag, New York, Inc.*, 1999.
- [77] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, *SIAM, Philadelphia*, 2000.
- [78] P. M. Pardalos and H. Edwin, Handbook of Global Optimization (II), *Kluwer Academy Publishers*, 2002.
- [79] P. A. Parrilo, Structured Semidefinite Programs and Semialgebraic Geometry Methods in Robustness and Optimization, *Ph.D. thesis, California Institute of Technology*, 2000.
- [80] P. A. Parrilo, Semidefinite programming relaxations for semialgebraic problems, *Mathematical Programming*, 96(2003): 293–320.
- [81] P. A. Parrilo, B. Sturmfels, Minimizing Polynomial Functions, *Algorithmic and quantitative real algebraic geometry, DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 60(2003): 83–99.
- [82] B. T. Polyak, A General Method of Solving Extremal Problems, *Soviet Math. Doklady*, 8(1967): 14–29.
- [83] B. T. Polyak, The Conjugate Gradient Method in Extremum Problems. *USSR Comp. Math. and Math. Phys.*, 9(1969): 94–112.
- [84] B. T. Polyak, Subgradient Methods: A Survey of Soviet Research, *Nonsmooth Optimization*, Pergamon, Oxford, Lemaréchal C, Mifflin R. ed.. (1978): 5–30.

-
- [85] M. J. D. Powell, An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives. *The Computer J.*, 7(1964): 155–162.
- [86] M. J. D. Powell, On the Calculation of Orthogonal Vectors. *Computer J.*, 11(1968): 302–304.
- [87] M. J. D. Powell, A Theory on Rank One Modifications to a Matrix and Its Inverse. *The Computer J.*, 12(1969): 288–290.
- [88] M. J. D. Powell, A New Algorithm for Unconstrained Optimization, *Nonlinear Programming*, New York: Academic Press, J. B. Rosen, O. L. Mangasarian, K. Ritter eds.. (1970): 31–66.
- [89] M. J. D. Powell, On the Convergence of the Variable Metric Algorithm. *J. Inst. Maths. Appl.*, 7(1971): 21–36.
- [90] M. J. D. Powell, Quadratic Termination Properties of Minimization Algorithms, Part I and Part II. *J. Inst. Maths. Appl.*, 10(1972): 332–357.
- [91] M. J. D. Powell, Convergence Properties of a Class of Minimization Algorithms *Nonlinear Programming 2*, New York: Academic Press, O. L. Mangasarian, R. R. Meyer, S. M. Robinson eds., 1–27.
- [92] M. J. D. Powell, Some Global Convergence Properties of a Variable Metric Algorithm for Minimization Without Exact Line Searches *Nonlinear Programming. SIAM-AMS Proceedings vol. IX. SIAM publications, Philadelphia*, R. W. Cottle, C. E. Lemke eds., (1976): 53–72.
- [93] M. J. D. Powell, Some Convergence Properties of the Conjugate Gradient Method. *Math. Prog.*, 11(1976b): 42–49.
- [94] M. J. D. Powell, Restart Procedure for the Conjugate Gradient Method. *Math. Prog.*, 12(1977): 241–254.
- [95] M. J. D. Powell, A Fast Algorithm for Nonlinearly Constrained Optimization Calculations *Numerical Analysis Berlin: Springer-Verlag*, (1978): 144–157.
- [96] M. J. D. Powell, Nonconvex Minimization Calculations and the Conjugate Gradient Method *Numerical Analysis, Lecture Notes in Mathematics 1066. Berlin: Springer-Verlag*, Griffiths D F. ed.. (1984): 122–141.
- [97] M. J. D. Powell, On the Rate of Convergence of Variable Metric Algorithms for Unconstrained Optimization *Proceeding of the International Congress of Mathematicians New York: Elsevier*, Z. Ciesielki, C. Olech eds., (1984): 1525–1539.

- [98] M. J. D. Powell, General Algorithms for Discrete Nonlinear Approximation Calculations, *Approximation Theory IV. New York: Academic Press*, L. L. Schumaker ed., (1984): 187–218.
- [99] M. J. D. Powell, On the Global Convergence of Trust Region Algorithms for Unconstrained Optimization. *Math. Prog.*, 29(1984): 297–303.
- [100] M. J. D. Powell, On the Quadratic Programming Algorithm of Goldfarb and Indani, *Math. Prog. Study*, 25(1985): 46–61.
- [101] M. J. D. Powell, How Bad Are the BFGS and DFP Methods When the Objective Function is Quadratic. *Math. Prog.*, 34(1986): 34–47.
- [102] M. J. D. Powell, Updating Conjugate Directions by the BFGS Formular. *Math. Prog.*, 38(1987): 29–46.
- [103] M. J. D. Powell, A Tolerant Algorithm for Linearly Constrained Optimization Calculations, *Math. Prog.*, 45(1989): 547–566.
- [104] M. J. D. Powell, UOBYQA: Unconstrained Optimization by Quadratic Approximation. *Math. Programm.*, 92(2002): 555–582.
- [105] M. J. D. Powell and Y. Y. Yuan, Conditions for Superlinear Convergence in L_1 and L_∞ Solutions of Overdetermined Nonlinear Equations. *IMA J. Numer. Anal.*, 4(1984): 241–251.
- [106] M. J. D. Powell and Y. Y. Yuan, A Recursive Quadratic Programming Algorithm That Use Differentiable Exact Penalty Function. *Math. Prog.*, 35(1986): 265–278.
- [107] M. J. D. Powell and Y. Y. Yuan, A Trust Region Algorithm for Equality Constrained Optimization. *Math. Prog.*, 49(1989): 189–211.
- [108] M.J.D. Powell, The NEWUOA Software for Unconstrained Optimization Without Derivatives, *University of Cambridge DAMTP 2004/NA05*.
- [109] L. Q. Qi, Z. Wan, and Y. F. Yang, Global Minimization of Normal Quartic Polynomials Based on Global Descent Directions. *SIAM Journal on Optimization*, Vol 15, No. 1: pp. 275–302.
- [110] B. Reznick, Uniform denominators in Hilbert’s seventeenth problem. *Math Z.*, 220(1995): 75–97.
- [111] B. Reznick, Some concrete aspects of Hilbert’s 17th problem, *Contemporary Mathematics, American Mathematical Society*, 253(2000): 251–272.

-
- [112] R. T. Rockafellar, The Theory of Subgradient and Its Application to Problems of Optimization: Convex and Not Convex Functions, *West Berlin: Heldermann Verlag*, 1981
- [113] A. Ruhe and P.Å. Wedin, Algorithms for separable nonlinear least squares problems, *SIAM Review*, 22(1980), 318–337.
- [114] N. Z. Shor, Application of the Gradient Method for Solution of Network Transportation Problems(in Russian), *Notes Scientific Seminar on Theory and Application of Cybernetics and Operations Research, Acad. of Sciences*, 1962
- [115] N. Z. Shor, Utilization of the Operation of Space Dilatation in the Minimization of Convex Function(in Russian), *Kibernetika*, 6(1970): 6–12.
- [116] N. Z. Shor, Quadratic optimization problems, *Soviet J. Comput. Systems Sci.*, 25 (1983): 1-11.
- [117] N.Z. Shor, Nondifferentiable Optimization and Polynomial Problems, *Kluwer, Dordrecht*, 1998.
- [118] T. P. Shuai and X. D. Hu, Connected Set Cover Problem and Its Applications, *Algorithmic Aspects in Information and Management, Lecture Notes in Computer Science*, Springer Berlin/Heidelberg, 4041(2006): 243–254.
- [119] R. Stafford, Random Points in an n-Dimensional Hypersphere, *MATLAB CENTRAL: <http://www.mathworks.com/matlabcentral/fileexchange/9443>*, 2005.
- [120] T. Steihaug, The Conjugate Gradient Method and Trust Regions in Large Scale Optimization, *SIAM J. Numer. Anal.* 20(1983): 626–637.
- [121] B. Sturmfels, Gröbner Bases and Convex Polytopes, *University Lectures, American Mathematical Society*, 8(1995).
- [122] A. M. So and Y. Y. Ye, Theory of semidefinite programming for Sensor Network Localization. *Mathematical Programming*, Volume 109, Numbers 2-3, 2007, pp. 367–384.
- [123] W. Y. Sun, R. J. B. Sampaio and J. Y. Yuan, Quasi-Newton Trust Region Algorithm for Nonsmooth Least Squares Problems, *Appl. Math. Comput.*, 105(1999): 183–194.
- [124] W. Y. Sun and Y. X. Yuan, Optimization Theory and Methods, *Springer Press, New York*, 2006.
- [125] Ph.L. Toint, Towards an efficient sparsity exploiting Newton method for minimization, in: I. Duff, ed., *Sparse Matrices and Their Uses*, Academic Press, Berlin, 1981, pp. 57–88.

- [126] Aimo Torn, Antanas Zilinskas, *Lectual Notes in Global Optimization*, Springer-Verlag, 1989.
- [127] C. Voglis and I.E. Lagaris, BOXCQP: An Algorithm for Bound Constrained Convex Quadratic Problems, *1st International Conference "From Scientific Computing to Computational Engineering"*, Athens, 2004.
- [128] C. Y. Wang and N. H. Xiu, Convergence of Projected Gradient Methods for Generalized Convex Minimization, *Comput. Optim. Appl.*, 16(2000): 111–120.
- [129] J. Wolberg, *Data Analysis Using the Method of Least Squares: Extracting the Most Information from Experiments*, Springer, 2005.
- [130] P. Wolfe, A method of conjugate subgradients for minimizing nondifferentiable function. *Math. Prog. Study*, 3(1975):145–173.
- [131] D. Wu, Z. J. Wu, Y. X. Yuan, Rigid versus unique determination of protein structures with geometric buildup, *Optimization Letters* 2(2008): 319–331.
- [132] Z. Wang, S. Zheng, S. Boyd, and Y. Y. Ye, Further Relaxations of the SDP Approach to Sensor Network Localization. *Technical report, Stanford University*, 2006.
- [133] N. H. Xiu, C. Y. Wang and L. C. Kong, A Note on the Gradient Projection Method with Exact Step-size Rule, *J. Comput. Math.*, 25(2007): 221–230.
- [134] H. Yabe and T. Takahashi, Factorized quasi-Newton methods for nonlinear least squares problems, *Mathematical Programming*, 51(1991), 75–100.
- [135] N. Yamashita and M. Fukushima, On the Rate of Convergence of the Levenberg-Marquardt Method, *Report: the first Sino-Japan Optimization Conference*, 2000.
- [136] Y. X. Yuan, Some Properties of Trust Region Algorithms for Nonsmooth Optimization, *University of Cambridge DAMTP 1983/NA4*.
- [137] Y. X. Yuan, Global Convergence of Trust Region Algorithms for Nonsmooth Optimization, *University of Cambridge DAMTP 1983/NA13*.
- [138] Y. X. Yuan, An Example of Only Linearly Convergence of Trust Region Algorithms for Nonsmooth Optimization. *IMA J. Numer. Anal.* 4(1984): 327–335.
- [139] Y. X. Yuan, Conditions for Convergence of Trust Region Algorithms for Nonsmooth Optimization. *Math. Prog.* 31(1985): 220–228.
- [140] Y. X. Yuan, On the Superlinear Convergence of a Trust Region Algorithm for Nonsmooth Optimization. *Math. Prog.* 31(1985): 269–285.

- [141] Y. X. Yuan, On a Subproblem of Trust Region Algorithms for Constrained Optimization. *Math. Prog.* 47(1990): 53–63.
- [142] Y. X. Yuan, Trust Region Algorithms for Nonlinear Equations, *Information*, 1(1998): 7–20.
- [143] Y. X. Yuan, On the Truncated Conjugate Gradient Method. *Math. Programm.* 87(2000): 561–571.
- [144] Y. X. Yuan, Step-sizes for the gradient method, *the Third International Congress of Chinese Mathematicians, AMS/IP Studies in Advanced Mathematics*, K.S. Liu, Z.P. Xin and S.T. Yau, eds., (2008), pp 785–796.
- [145] Y. X. Yuan and Y. H. Dai, Alternate Minimization Gradient Method, *IMA J. Num. Anal.* 23(2003): 377–393.
- [146] Y. X. Yuan and F. Hickernell, A simple multi-start algorithm for global optimization, *Operational Research Transaction, China*: 1(1997) no.2. 1–12.
- [147] L. B. Zhang, T. Cui and H. Liu, A Set of Symmetric Quadrature Rules on Triangles and Tetrahedra, *Journal of Computational Mathematics*, 27(2009): 89–96.
- [148] Y. Zhang, Interior-Point Gradient Methods with Diagonal-Scalings for Simple-Bound Constrained Optimization, *CAAM technical report TR04-06, Rice University*, 2004.
- [149] B. Zhou, L. Gao and Y. H. Dai, Monotone Projected Gradient Methods for Large-scale Box-constrained Quadratic Programming, *Science in China Series A: Mathematics*, 49(2006): 688–702.
- [150] M. Zelen and N. C. Severo, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. New York: Dover, M. Abramowitz and I. A. Stegun, eds., 1972.
- [151] J. Zowe, Nondifferentiable Optimization a Motivation and a Short Introduction in to the Subgradient and the Bundle Concept, *Computational Mathematical Programming, Berlin: Springer-Verlag*, K. Schittkowski ed., (1985): 321–356.
- [152] 白峰杉, 数值分析引论, 高等教育出版社, 2004.
- [153] 蔡大用, 白峰杉, 高等数值分析, 清华大学出版社, 1997.
- [154] 陈家鼎, 孙山泽, 李东风, 数理统计学讲义, 高等教育出版社, 2003.
- [155] 戴戡虹, 袁亚湘, 非线性共轭梯度法, 上海科技出版社, 2000.

- [156] 范金燕, 非线性方程组和无约束优化的数值方法, 博士学位论文, 中科院数学与系统科学研究院, 2002.
- [157] 高岩, 非光滑优化, 科学出版社, 2008.
- [158] G. H. 戈卢布, C. F. 范洛恩, 袁亚湘译, 矩阵计算, 科学出版社, 2001.
- [159] 韩继业, 修乃华, 戚后铎, 非线性互补理论与算法, 上海科技出版社, 2006.
- [160] 李庆扬, 莫孜中, 祁力群, 非线性方程组的数值解法, 科学出版社, 1999.
- [161] 王宜举, 修乃华, 非线性规划理论与算法, 陕西科技出版社, 2004.
- [162] 王松桂, 吴密霞, 贾忠贞, 矩阵不等式(第二版), 科学出版社, 2006.
- [163] 徐树方, 矩阵计算的理论与方法, 北京大学出版社, 1995.
- [164] 徐树方, 高立, 张平文, 数值线性代数, 北京大学出版社, 2002.
- [165] 袁亚湘, 孙文瑜, 最优化理论与方法, 科学出版社, 1997.
- [166] 袁亚湘, 非线性优化计算方法, 科学出版社, 2008.

发表文章目录

- [1] X. Liu, An Efficient Unseparated Scheme for Separable Nonlinear Least Squares Problem, Proceedings of the Eighth National Conference of Operation Research Society of China(Global-Link, Hong Kong), (2006): 132–137.
- [2] X. Liu and Y. X. Yuan, On the Separable Nonlinear Least Squares Problems, Journal of Computational Mathematics, 26(2008): 390–403.
- [3] X. Liu, Global Minimization of Quadratic Least Squares Problems, Proceedings of the Ninth National Conference of Operation Research Society of China(Global-Link, Hong Kong), (2008): 188–193.

待发表文章

- [4] X. Liu and Y. X. Yuan, On the Trust Region Subproblem for Nonlinear L1 Norm Minimization Problem, finished.
- [5] X. Liu, S. McKee, J. Y. Yuan and Y. X. Yuan, Uniform Bound on the 1-norm of the Inverse of Lower Triangular Toeplitz Matrices, finished.

国际学术报告

- [6] X. Liu, Efficient Unseparated Algorithms for Separable Nonlinear Least Squares Problems, presented on the International Conference on Numerical Analysis and Optimization dedicated to M. J. D. Powell's 70th Birthday, Beijing, China, 2006.
- [7] X. Liu, Global Nonlinear Least Squares in Data Analysis, post presentation on the First Joint China Operations Research Days Workshop, IBM Research Center, Beijing, China, 2007.
- [8] X. Liu, Global Minimization on Quadratic Least Squares Problems, presented on the 6th International Congress on Industrial and Applied Mathematics, Zurich, Switzerland, 2007.
- [9] X. Liu, Global Minimization on Quadratic Least Squares Problems, presented on the 6th International Conference on Numerical Linear Algebra and Optimization, Urumqi, China, 2007.
- [10] X. Liu, New Efficient Sampling Methods for Distance Geometry Problems, presented at Siemens Corporate Technology, Munich, Germany, 2007.

-
- [11] X. Liu, New Sampling Techniques for Distance Geometry Problems, presented at Technische Universitat Darmstadt, Darmstadt, Germany, 2007.
 - [12] X. Liu, New Methods for Distance Geometry Problems, presented at Zuse Institute Berlin, Berlin, Germany, 2007.
 - [13] X. Liu, L_1 Norm Minimization for Quadratic System, presented on the VII Brazilian Workshop on Continuous Optimization, Campinas, Brazil, 2008.

致 谢

谨在学位论文完成之际，把我最深挚的谢意献给我的恩师袁亚湘研究员。恩师是我遇见过的最完美的英雄，他知识渊博、思维敏锐、为人正直、待人亲切；更难能可贵的是年近半百的他有着一颗朝气蓬勃的心和雄健的体魄，他犹如一台永动机，把永不枯竭的精力全部奉献给了他所热爱的科研事业和学生培养。能够受业于恩师，我感到无比的光荣与幸运。

感谢恩师，首先要感谢他的悉心指导。无论是聆听他所授的课程、或是所作的报告，还是和他交流，都能使我感受到豁然开朗的美感，使我体会到融会贯通的重要性。在恩师的殷殷教导下，我已慢慢地扣启了科研的大门。感谢恩师，我还要感谢他的言传身教。恩师有着先贤般的睿智和洞察力；有着无比坚定的共产主义信念和爱国情怀。在感佩之余，我的人生观、价值观也潜移默化地受到涤染。感谢恩师，我更要感谢他的信任。因为这份信任，我才能保持着难能可贵的自信；因为这份信任，我得到无数的机遇和挑战，对我的成长有着无可替代的帮助。

德高为师、身正为范，恩师正是师中的典范。他给我一生取之不尽、用之不竭的精神财富。他的谆谆教导，我将永镌于心。

特别感谢戴彧虹研究员、北京大学高立教授长期以来对我的关心和指导。他们勤奋刻苦的钻研精神、勇于探索的创新精神和严谨务实的治学态度，非常让我崇敬。他们渊博的学识和高尚的品格，永远是我学习的榜样。

非常感谢同济大学殷俊锋博士、北京航空航天大学夏勇博士和美国哥伦比亚大学博士研究生马士谦，他们不仅仅是我的师兄，更是我的良师益友。他们传授给我的宝贵经验使我受益匪浅。非常感谢这五年来相识的好朋友Thorsten Koch博士、崔涛、戴小英、丁晓东、程明厚、杨熙，以及我的老朋友们，和他们交往带给我的快乐和感动，是我永远无法忘怀的。我将永远珍惜这份友谊。

攻读博士学位期间，还有不计其数的老师和同学们给予我关怀、帮助和支持。在此，对他们一并表示衷心的感谢和深深的祝福。尤其是Martin Grötschel教授、周爱辉研究员、袁富宇博士、吴继萍、白英、邵欣、尹永华、丁晓蕾、胡洁、韦晓鹏、孙燕新等老师；王彦飞、范金燕、黑龙、王周宏、罗新龙、徐玲玲、严涛、李多全、郑真真、胡毅庆、袁春明、王增琦、赵延龙、李春、徐姿、吴新明、王辛、谢和虎等师兄师姐们；阴小波、胡捷、高兴誉、陈星玓、吴涛涛、张强劲、李在禾、牛凌峰、牛新艳、翟方曼、赵礼、孙波等同学们；付云姗、唐明筠、聂宁明、

胡娟、田霞、陈景润、杜锐、滕飞、陈华杰、张娅、翟淑星、刘伟、张静静、刘辉、张磊、寇彩霞、费存林、宫鲁津、杨章、陈黄鑫、郝春林、刘亚锋、张在坤、吴乐秦、王晓、姜波、孙聪等可爱的师弟师妹们。

最后，我要对我的爸爸、妈妈和佩佩深深地道一声谢谢。是他们无私的爱和鼓励，使我能够在自己选择的道路上不畏挫折，坚定地走下去；是他们无限的爱和宽容，使我在失落和悲伤的时候，永远有一片疗伤的港湾；是他们无尽的爱和希望，使我永远能够满怀热情地面对未来。我将竭尽所能去报答和爱护他们，虽然我知道他们给予我的爱，是我穷尽一生也无以回报其万一的。

谨以此文献给我至爱的爸爸、妈妈、恩师和佩佩。