摘 要

约束非线性规划在经济金融、工程控制、技术物理、物流配送、计算机科学及生物工程等各个领域有着广泛的应用。近年来，随着理论研究的深入和计算机技术的普及和发展，人们开始尝试把一些计算复杂度小、稳定性好、收敛性强的算法拓展到求解非线性规划问题。其中，内点算法的研究尤为引人注目。内点算法的基本思想是从问题可行域中的某一点出发，沿着中心路径进行搜索，直达问题的最优解。不过，内点算法在非线性规划中的实际研究、证明和测试中还是遇到了许多的障碍。

首先，对于具有大规模约束的问题，如何寻找一个初始的可行点是内点法中研究的课题。在线性规划问题中，我们可以采用一些非可行内点算法的技巧，例如在某一步迭代过程中选取全牛顿步长等等。研究表明，在锥优化模型中，可以通过引入自对偶嵌入模型来克服初始点选取的困难。但这些技巧不适用于一般的非线性规划问题。其次，在路径跟踪内点算法中，由于正交性条件的不满足，如何证明算法的收敛性，也是要进行探索的问题。

本文主要的工作是结合内外罚函数给出了求解约束非线性规划的内点方法。针对上的问题，我们作了以下二方面的工作，一、通过引入辅助变量来构造原问题的等价问题，从而克服了初始点选取的难题。然后，给出相应的KKT条件和罚内点算法，并采用Wolf条件设计了一个可调的内嵌算法，进一步证明了算法的收敛性，数值试验也说明了新给出的算法是可行的、有效的。二、在前工作的基础上，构造修正的KKT条件，给出了大步长路径跟踪内点算法，通过添加关系

不等式条件，给出并证明路径跟踪算法的收敛性定理，相应的数值算例也说明了新给出的算法是可行的、有效的。

本文结构共分为四章，在第一及第二章，我们简单介绍了内点算法基本概念、发展历史及分类，并对对数障碍函数法和原对偶-路径跟踪法的思想作了较详细的介绍。第三章,给出了线搜索下的罚内点算法，并证明了算法的收敛性，第四章，给出了大步长路径跟踪内点算法，并证明了算法的全局收敛性。相应的数值算例也说明了新给出的算法是可行的、有效的。

关键词：约束非线性规划；罚函数方法；内点法；牛顿方程；KKT条件

Abstract

Constrained nonlinear programming has been widely used in economy & finance, engineering control, technical physics, logistics & transshipment, computer science, biochemical construction and many other fields. Recently, with development of theory research and computer technology, some effective algorithms with less time complexity, good stability and convergence have been tried to solve nonlinear programming. Among them, interior point method won most attentions. The basic idea of interior point method is to select an initial point in feasible region and then iterate along the central path till reaching the optimal point. However, many difficulties occur while researching, testing, convergent proofing the algorithm for solving nonlinear programming.

Firstly, for problem with many constraints, how to find an initial feasible point is a research topic. In linear programming, researchers present several strategies to overcome this difficulty, see infeasible-interior-point method, such as, take a full Newton step at any iteration. Within the framework of conic optimization, a possible remedy of the problem is to embed the primal and dual formulations of the problem into a single self-dual model, which then has an easy available initial feasible solution. However these strategies are not effective for nonlinear programming. Secondly, due to unsatisfactory of orthogonality condition, how to proof convergence of path-following algorithm is still a topic.

The major work is to integrate interior point method with inner-outer penalty functions for solving constrained nonlinear programming. In light of difficulties mentioned above, our work includes two parts. Firstly, incorporate an auxiliary variable which is driven to zero by penalization and then construct a new problem equal to original one. Due to this auxiliary variable, the problem of initialization is circumvented. The corresponding KKT conditions as well as penalized interior point algorithm are given subsequently. We design an inner algorithm based on Wolf conditions and proof convergence of outer and inner algorithms respectively. Results of the numerical experiment are reported to show the algorithm is practical and effective. Secondly, based on the work done before, we propose modified KKT conditions as well as long-step path following algorithm. Formula assumption is described and then the algorithm is proved to be convergent. Results of the numerical experiment are reported to show the algorithm is practical and effective.

The paper consists of four chapters. In chapter one and two, we describe the basic concept and development history of interior point method. Some classic interior point methods are presented there, where logarithm barrier function method and primal-dual path following method are introduced in details. In chapter three, we design a penalized interior point algorithm based on line search conditions and present convergent theorems. A penalized interior point algorithm based on path following method is given in chapter four while several convergent theorems are proved there. Numerical results are reported to show algorithms are practical and effective respectively.

Key Words. Constrained Nonlinear Programming; Penalty Function Method; Interior Point Method; Newton Equation; KKT Conditions.

原创性声明

本人声明：所呈交的论文是本人在导师指导下进行的研究工作。除了文中特别加以标注和致谢的地方外,论文中不包含其他人已发表和撰写过的研究成果。参与同一工作的其他同志对本研究所做的任何贡献均已在论文中作了说明并表示了谢意。

签名：陆文婷 日期：2009.6.10

本论文使用授权说明

本人完全了解上海大学有关保留、使用学位论文的规定，即：学校有权保留论文及送交论文复印件，允许论文被查阅和借阅；学校可以公布论文的全部或部分内容。

签名：陆文婷 导师签名：[illegible] 日期：2009.6.10

第一章 绪论

非线性规划是20世纪50年代才开始形成的一门新兴学科，现已成为运筹学的一个重要分支，它在工程、管理、经济、科研、军事等方面都有广泛的应用，为最优设计提供了有力的工具。1951年H.W.库恩和A.W.塔克发表的关于最优性条件(后来称为库恩－塔克条件)的论文是非线性规划正式诞生的一个重要标志。在20世纪50年代还得出了可分离规划和二次规划的多种解法，它们大都是以G.B.丹齐克提出的解线性规划的单纯形法为基础的。50年代末到60年代末出现了许多解非线性规划问题的有效的算法，70年代、80年代又得到进一步的发展。

随着研究的深入，人们开始尝试把一些成熟的、有效的线性规划算法拓展到求解非线性规划问题。其中内点算法的研究最引人注目。内点算法的提出最初是为了克服单纯性算法的缺陷。单纯性算法的时间复杂度是指数形式的，在运用计算机编程调试时，每一步迭代必然导致内存需要存贮大量的数据，因此在求解一些大规模的问题时就显得效率不高。这个缺陷促使人们迫切寻找到一种新的、更行之有效的算法，对此人们作了很多的尝试也得到了一些结果。

1979年，苏联学者哈奇扬提出了第一个多项式算法-椭球算法[1]，并证明了计算复杂性是$O(n^4L)$，这引起了人们极大的热情，对算法复杂度理论产生了巨大的影响。但是该法在实际上并没有如期所希望的那样在计算速度上超过单纯形法。原因主要有二：一是迭代次数仍然很多；二是不便应用稀疏矩阵技术，每次迭代的计算比单纯形法慢很多。

1984年，N.Karmarkar提出了线性规划的一种新的多项式算法[2]，Karm arkar算法不仅比椭球算法具有更优越的计算复杂度，而且在实际计算中也可以与单纯形法相媲美，尤其对大规模问题更显其高效性。与单纯形算法沿着可行区域的边界寻优不同，Karmarkar算法是建立在单纯形结构之上的，它是从初始内点出发，沿着最速下降方向，从可行区域内部逐渐走向最优解，因此Karmarkar算法又被称为内点算法。

自从Karmarkar划时代的论文发表以来，内点算法一直是数学规划领域一个非常活跃的研究方向。很多研究者在Karmarkar法的基础上对算法作了各种修正或改进。首先的改进包括各种问题形式的列出；起始可行解的计算；近似最速下降投影方向的计算以及的带稀疏性矩阵的方程组的解算；每步跌代的目标函数下界的计算；以及每步跌代时为加快收敛的参数α的选择等。这方面已发表的文章很多，典型的可举出[5, 6]等等。

另一方面，从理论上的不同方面发展了作为Karmarkar法的变种的其他内点法。在这些方法中，值得提出的有[7]和[8]等人发展的原仿射比例调节法和Adler[9, 10]等人提出的对偶仿射比例调节法等。此外，1986年Gill[11]等人第一次把原来用于非线性规划的对数障碍函数法应用于线性规划，并且证明了对数障碍函数法和Karmarkar投影法是等价的，以后的研究进一步表明了Karmarkar法实际上是广义对数障碍函数法的一个特殊情形。此后Megiddo, Kojima, Lustig和Mehrotra等人又提出了原-对偶路径跟踪对数障碍函数法，以及"极限可行方向"原-对偶路径跟踪法，全面收敛不可行内点算法和Mehrotra预计改正等改进方法[12]-[15]，这些方法都是很有效的，计算复杂性达到$O(n^3L)$。类似的方法还有Gonzaga和Ye等提出势函数下降法[4, 16, 17]。他们方法优点是采用大步长，并且也证明了计算复杂性也达到$O(n^3L)$。

在Karmarkar法提出后，在对内点法的实际编程并与单纯形法的

优秀商用软件对比方面，开始直接按Karmarkar法编程的测试并未表明他的明显优越性。1986年Adler等人应用Karmarkar法变种的所谓对偶仿射法编程，并对DavidGay所收集的约50个问题的NETLIB问题集与单纯形法商用软件Minos4.0进行了对比测试。对比表明，对较大型问题来说，内点法要明显的优于单纯形法[9, 11]。1992年Lustig[18]等人用原-对偶路径跟踪法并采用一系列的改进技术进行了编程，并与比MINOS系统快2-10倍的最新单纯形法商用软件IBM OSLRelease2[19]进行了对比测试，对比的问题是比NETLIB问题大得多的8个（几千几万行和几万几十万列）的大问题。结果表明，除了一个问题特殊之外，对其他问题内点法比单纯形法快2.5-20倍。更为壮观的是，用原-对偶有效地解决了大至99533（行）×117117（列）和270796×30396的两个大问题，而这种大问题单纯形法是从未涉足过的。

内点算法区别其他算法的一个特性是由算法产生的迭代点都是向最优点"迈进了一大步，从不走任何弯路"。它不是围绕着可行区域的边界寻找最优点，而是在可行域内部迭代，因此当他到达可行域的边界时就说明已经到达了最优点。

内点算法的诸多优点都吸引着人们把它拓展到求解其他的问题，包括非线性规划问题。1994年，Nesterov和Nemirovski提出了self-concordant障碍理论[37]，它是把内点算法思想拓展到求解一般凸最优化问题的理论基础。在此理论基础上，内点算法被应用于有效地解决一类的凸优化问题。现在，内点算法的思想已被广泛的应用于研究和求解非线性规划问题[48]-[56]，不过，内点算法在非线性规划中的实际研究、证明和测试中还是遇到了许多的障碍。

由于内点算法的特性保证由一个初始可行点产生的一系列迭代点都是严格可行的，因此选择一个初始的可行点就显得尤为重要。但对于

一般的大规模问题，一个可行的初始点并不是容易得到的。为此人们也作了大量的研究工作，希望来克服这一难题。1994年，Ye,Y., Todd, Mizuno提出了求解线性规划的自对偶嵌入方法[33]，解决了寻找初始可行点的困难且保证嵌入问题有最优解，并可利用任何内点算法来求解，从而也得到了原问题的最优解。对一般的线性规划问题也可采用一些非可行内点算法的技巧，例如在某一步迭代过程中选取全牛顿步长等等。到了2000年，Luo, Z.Q, Sturm, J.F.和Zhang, S.提出了把自对偶嵌入技术推广到SDP及更一般锥优化问题[36]，从而使初始化难题得到解决。2004年，Zhang,S.提出了解决凸规划的一个新自对偶嵌入技术[39]。该方法的思想是先把凸规划转化为等价的锥优化问题，然后运用自对偶嵌入技术来解决由此产生的锥优化模型。其优点是初始化难题得到了克服。不过这些方法并不适用于一般的非线性规划问题，因此，关于求解非线性规划问题的初始化难题也成为了一个重要的研究课题。

作为内点算法独有的特征，中心路径在算法中扮演着重要的角色，因而路经跟踪算法也就成为了内点算法中一个重要的组成部分。路经跟踪算法的思想是使每一步产生的迭代点均落在中心路径周围的一个带状区域，直至问题的最优解。在该算法的收敛性证明中，要求迭代方向相互正交。这一条件在一般的线性规划问题中自然成立。但对于非线性的问题，就不一定满足了。这也成为了该算法能普遍有效地应用于求解非线性规划问题的一大障碍。

本文主要是结合内外罚函数给出了求解约束非线性规划的内点方法。针对上述提到的两个问题，作了以下二方面的工作，一、通过引入辅助变量，构造原问题的等价问题，并在罚内点算法的设计中，使辅助变量逐步被"罚"为零，从而克服了初始点选取的困难。然

后，给出了相应的KKT条件、在牛顿法的基础上，设计了一个罚内点算法，并证明了所给算法是收敛的。同时结合线搜索在Wolf条件下设计了可调的内嵌算法，来修正牛顿步。进一步证明了该算法的全局收敛性，数值试验说明了所给算法是可行的、有效的。二、在前工作的基础上，构造修正的KKT条件，然后通过调节迭代参数，控制每一步的迭代都落在中心路径的一个较宽泛的带状区域内，从而设计出大步长路径跟踪内点算法。通过添加关系不等式条件，证明了算法的收敛性。相应的数值算例也说明了新给出的算法是可行的、有效的。

第二章　几类经典的内点算法

在本章中，我们将主要介绍几类经典的内点算法，它们分别是

- Karmarkar算法
- 仿射比例调节法
- 对数障碍函数法
- 原-对偶路径跟踪法
- 势函数下降法

这些算法都是用以求解线性规划问题。通过对这些算法的研究和分析，掌握各类算法的思想和优缺点，从而为后面两章中运用内点算法求解非线性规划问题的研究打下基础。

§2.1　内点算法简介

在非线规划问题中，对数障碍函数法和原-对偶路径跟踪法应用得最为广泛[41]-[47]，因此，会后面的几节中作详细的介绍，这里就对剩下的三个算法作一个简单的介绍。

Karmarkar算法：1984年，在美国贝尔实验室工作的数学家Karmarkar提出了一个多项式算法——Karmarkar算法，他的时间复杂度是$O(n^{3.5}L)$，而且声称他比单纯刑法更为有效。在当时，Karmarkar的名字和Karmarkar的算法被刊登在纽约时代杂志的头版，尽管当时他的主张受到同行许多专家学者的质疑。今天看来，Karmarkar显然开创了线性规划的一个

新领域——内点算法。在此之后，Karmarkar算法得到不断地改进。线性规划、凸二次规划的内点算法相继问世。已经证实，如果通过好的程序来实现内点算法，包括最初的Karmarkar算法，的确可以得到比单纯形法更好的效果。特别是对于含几千个以上变量的大规模问题，他的收敛性态完全优越于单纯形法。更令人吃惊的是，在实际计算中发现内点算法的迭代次数几乎与问题的规模无关，这一事实也是内点算法在理论尤其是实践中引人注目的重要因素。

仿射比例调节法：是Karmarkar法中变换方法。它是用一个仿射变换$D_k^{-1}x$代替投影变换，以投影目标函数替代势函数，把坐标系的正挂限（而不是单纯形）映射到自身，把每次跌代的变量值x^k变换成与各坐标面等距的e。

早在1967年苏联学者Dikin[7, 20]就首先提出了这种方法，并于1974年给出了收敛性的证明。但他的这些工作直到Barnes等人再次研究该法后被人们所发现。1985年，Barnes [8]和Vanderrei[21]等人详细研究称为原仿射法，到了1987年Adler[9]等提出另一种仿射比例调节法，由于它实际上是从原问题的对偶问题出发，因而被称之为对偶放射比例调节法。

势函数下降法：是由Gonzaga[4, 28]和Ye与Todd[27, 29]等提出的基于势函数下降的原势函数下降法和原-对偶势函数下降法，其迭代次数为$O(\sqrt{n}L)$。在这方面作出贡献的还有Tanabe[30]、Güler[31]和Kojima等[32]。从历史的角度，最早的Karmarkar原算法就是用势函数进行推导的[2]。这种早期的方法对随后发展的路径跟踪方法提供了有用的理论上的透视，因此很自然的激起人们继续研究这类方法。另一方面，就路径跟踪法来说，实际所用的方法已消除了很多对路径跟踪的限制，因此导致了理论与实践之间的变样。而势函数法正好能把理论与实践更

紧密地联系起来，因为诸如搜索方向的选择，步长的确定以及它们的分析都可依据一个确定的势函数（或优势函数）来进行。这个势函数可作为测度来衡量一个点的质量以及指示怎样对他进行改进。采用一个适当的算法在每次迭代都使该势函数尽可能减少，因而导致相应算法的计算复杂性。

§2.2 对数障碍函数法

对数障碍函数法首先是引进解非线性规划问题，即

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \geq 0, \quad i = 1, 2, \ldots, m \end{aligned} \tag{2.2.1}$$

式中,$x = (x_1, x_2, \ldots, x_n)^T$。引用对数障碍函数把上面问题转换为如下形式的非约束条件问题：

$$\min f(x) - \mu_k \sum_{i=1}^{m} \ln g_i(x) \tag{2.2.2}$$

式中，μ_k是一个障碍参数，且有$\mu_k > 0$和

$$\lim_{k \to \infty} \mu_k = 0$$

算法是从一严格可行点出发，以一迭代形式选取μ_k和使(2.2.2)式为最小的x^k，产生的一系列可行点x^k收敛于(2.2.2)的解，而

$$\lim_{k \to \infty} \frac{\mu_k}{g_i(x^k)} = \lambda_i \tag{2.2.3}$$

这里λ_i是关于$g_i(x^k)$的最优拉格朗日乘数。

对数障碍函数法应用于线性规划问题，可从原问题形式或对偶问题形式出发。前者称为原对数障碍函数法，后者称为对偶对数障碍函数法。下面分别论述。

Gill等[11] (1986) 把原对数障碍函数法应用于线性规划，且考虑如下标准形式的线性规划问题：

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned} \tag{2.2.4}$$

式中，$A \in R^{m\times n}$ $(m \leq n)$。

把上面问题转换成

$$\begin{aligned} \min \quad & f(x) = c^T x - \mu \sum_{j=1}^{n} \ln x_j \\ \text{s.t.} \quad & Ax = b \end{aligned} \tag{2.2.5}$$

这里是对非负约束引用了障碍参数μ，而对等式约束，由于不能用障碍变换处理，因此仍以约束条件形式直接处理。对这些等式约束条件引进相应的拉格朗日乘数列向量π，则(2.2.5)的拉格朗日式为

$$L(x,\pi,\mu) = c^T x - \mu \sum_{j=1}^{n} \ln x_j - \pi^T (Ax - b) \tag{2.2.6}$$

求其最小值，条件为其对x和π的偏导数为零，即

$$\begin{aligned} \nabla_x L \quad & = c - \mu D^{-1} e - A^T \pi = 0 \\ \nabla_\pi L \quad & = -Ax + b = 0 \end{aligned} \tag{2.2.7}$$

注意，式中用D代表其对角线元素分别为x_j $(j = 1, 2, \ldots, n)$的对角方阵，e为n个元素都为1的n维列向量。

非线性规划中的Newton法，是把函数$F(x)$按如下的泰勒级数展开：

$$F(x^{k+1}) = F(x^k) + \nabla^T F(x^k) \triangle x^k + \frac{1}{2} (\triangle x^k)^T \nabla^2 F(x^k) \triangle x^k$$

其中略去高于二次的近似，增量为$\triangle x^k = x^{k+1} - x^k$。求该函数在$\triangle x^k$方向上的最小值，是该函数$\triangle x^k$的每个分量求导数后并令其为零，因而得出下列关系：

$$-\nabla^2 F(x^k)\triangle x^k = \nabla^T F(x^k) \tag{2.2.8}$$

注意，这里取$\triangle x^k$为Newton方向（最速下降方向）。

对问题(2.2.5)应用Newton法（即可行点最速下降法）。如当前的迭代点x^k满足$Ax^k = b$的条件，下一次估计的最优解为

$$x^{k+1} = x^k + \alpha p_B \tag{2.2.9}$$

即处于Newton搜索方向上，这里α是某个适当的大于零的步长参数，且p_B和α的计算必须保证$Ax^{k+1} = b$和$F(x^{k+1}) < F(x^k)$。

对应上述拉格朗日式的最速下降方向是向量$(p_b, \triangle\pi)^T$，因此按(2.2.8)列出的方程组为

$$\begin{bmatrix} -\mu D^{-2} & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} p_B \\ \triangle\pi \end{bmatrix} = \begin{bmatrix} c - \mu D^{-1}e - A^T\pi \\ 0 \end{bmatrix}$$

即

$$\begin{aligned} -\mu D^{-2}p_B + A^T\pi_B &= c - \mu D^{-1}e \\ Ap_B &= 0 \end{aligned} \tag{2.2.10}$$

注意，这里$\pi_B = \pi + \triangle\pi_B$是相应Newton搜索方向$p_B$的拉格朗日乘数变量。以后单称Newton搜索方向为Newton方向。右边第二项是可行性条件所致。

由上面第一式可知

$$p_B = \left(\frac{D}{\mu}\right)[DA^T\pi_B - (Dc - \mu e)]$$

考虑第二式，得出拉格朗日乘数变量π_B和Newton搜索方向p_B为

$$\pi_B = (AD^2A^T)^{-1}AD(Dc - \mu e) \tag{2.2.11}$$

$$p_B = -\left(\frac{D}{\mu}\right) r_B \tag{2.2.12}$$

其中

$$r_B = [I - DA^T(AD^2A^T)^{-1}AD](Dc - \mu e) \tag{2.2.13}$$

为下列最小二乘法问题的最优残差向量

$$\min_{\pi} \|Dc - \mu e - DA^T\pi\| \tag{2.2.14}$$

注意，这里$\|v\|$表示$(v^Tv)^{\frac{1}{2}}$之意。

关于原对数障碍函数法的算法可总结如下：

给出参数μ_0和满足$Ax^0 = b$和$x^0 > 0$的起始可行解

begin

$k := 0$

$x := x^0$

$\mu := \mu_0$

while $\|r\| > \varepsilon$ or $\mu > \varepsilon_\mu$

do

$\mu := \rho\mu$

$D := diag(x_1, x_2, \ldots, x_n)$

$g := Dc - \mu e$

$\pi := (AD^2A^T)^{-1}ADg$

$\eta := c - A^T\pi$

$r := D\eta - \mu e$

$p := -\frac{1}{\mu} Dr$

$\alpha := \gamma \min_{1 \le j \le n} \{ \frac{-x_j^{k-1}}{p_j} \mid p_j < 0 \}$

$x^{k+1} := x^k + \alpha p$

$k := k+1$

end do

end

关于算法，还有几点应加以说明。

1) 起始可行解

算法是假定已有起始可行点x^0，且$x^0 > 0$和$Ax^0 = b$。这个假定可通过在原问题(2.2.4)中引进人工变量列，并用“两阶段”法或“大M”法来处理。如果用后者方法，则问题变为

$$\begin{aligned} (\mathrm{P}) \qquad \min \quad & c^T x + M x_{n+1} \\ \text{s.t.} \quad & Ax + (b - Ax^0) x_{n+1} = b \\ & x \ge 0, \quad x_{n+1} \ge 0 \end{aligned} \tag{2.2.15}$$

式中，M是一大数。该问题的起始可行解为$x = x^0$，$x_{n+1} = 1$，其中x^0为任意值，由于M足够大，因此假定问题具有可行解，则必有$x_{n+1} = 0$，即问题(2.2.15)的最优解就是(2.2.4)的最优解。

2) 步长参数α的选择

为加快收敛，可以在保证不破坏可行性的条件下适当加大每步迭代的步长。这可在每步迭代时按下式找到α:

$$\alpha = \gamma \min_{1 \le j \le n} \left\{ \frac{-x_j^{k-1}}{p_j} \middle| p_j < 0 \right\} \tag{2.2.16}$$

式中，$\gamma < 1$（经验表明0.995是可取的）。

3)结束准则

残差向量$\|r\| \leq \varepsilon$和$\mu_k \rightarrow 0$可作为结束准则。但μ_k变化到一最小值$\mu_{\min}$时不再减小。

下面再论述对偶对数障碍函数法。

引进松弛变量$z(z \in R^n)$，对原问题(2.2.4)的对偶问题可写为

$$\begin{aligned} \text{(P)} \qquad \max \quad & b^T y \\ \text{s.t.} \quad & A^T y + z = c \\ & z \geq 0 \end{aligned} \tag{2.2.17}$$

引用对数障碍函数，上问题可写为

$$\max F(y) = b^T y + \mu \sum_{j=1}^{n} \ln(c_j - a_j^T y) \tag{2.2.18}$$

式中，a_j是矩阵A的第j列。上式最优的一阶条件式为

$$b - \mu A G^{-1} e = 0 \tag{2.2.19}$$

式中，G是$n \times n$对角矩阵，其元素是$z_j = c_j - a_j^T y$。

同样应用Newton法（参见(2.2.8)式），可知

$$\nabla F(y) = b - \mu A G^{-1} e$$

$$\nabla^2 F(y) = -\mu A G^{-2} A^T$$

若直接用Δy表示Newton方向，则有

$$\mu A G^{-2} A^T \Delta y = b - \mu A G^{-1} e$$

因此

$$\Delta y = \frac{1}{\mu}(A G^{-2} A^T)^{-1} b - (A G^{-2} A^T)^{-1} A G^{-1} e \tag{2.2.20}$$

(2.2.20)中第二项表示找出对偶空间的解析中心，而第一项则是求最优解的。

关于对偶对数障碍函数的起始可行点，也可对问题(2.2.17)引进人工变量列，使问题为

$$
\begin{aligned}
(\mathrm{P}) \qquad & \max \quad b^T y + M y_{m+1} \\
& \text{s.t.} \quad A^T y + (A^T y^0 + z^0 - c) y_{m+1} = c \\
& \qquad\quad z \geq 0, \quad y_{m+1} \leq 0
\end{aligned} \tag{2.2.21}
$$

这里问题的起始可行解为$y = y^0$，$y_{m+1} = -1$，M是足够大的正数。

显然引进大数M和人工变量，会给计算带来很大的不便。为了避免它们的引进，Lustig提出了$(\Delta x, \Delta y, \Delta z)$极限方向的方法[14]，此外有各种不可行内点法的引进[13]，将会在后面的章节中再作介绍。

§2.3 原-对偶路径跟踪法

最早分别提出路径跟踪法的有Megiddo[12]和Renegar[22]等人。所谓原-对偶路径跟踪法(primal-dual path following method)，按Megiddo(1986)的论述，实际上就是原-对偶障碍函数法(primal-dual log barrier method)，它包含了对数障碍函数的问题(2.2.5)，且在障碍参数$\mu > 0$时有惟一的最优解，这个惟一的最优解所构成的曲线$\{x(\mu) | \mu > 0\}$称为一条路径或中心轨迹(central trajectory)，当$\mu \to 0$时$x(\mu)$的极限即为原问题的最优解。

Kojima等最早(1987)[13, 23]提出收敛的算法。之后Monteiro和Adler[24, 25]以及其他研究者[14, 26]对算法都作了进一步的改进，使计算复杂性达到$O(n^{\frac{1}{2}}L)$ 迭代和全部$O(n^3L)$次算术运算的水平。

为讨论其算法，考虑如下标准形式的原问题和对偶问题对:

$$\begin{aligned}(\mathrm{P})\qquad &\min\ \ c^T x\\ &\mathrm{s.t.}\ \ Ax = b\\ &\qquad\ x \geq 0\end{aligned}\tag{2.3.22}$$

$$\begin{aligned}(\mathrm{D})\qquad &\max\ \ b^T y\\ &\mathrm{s.t.}\ \ A^T y + z = c\\ &\qquad\ z \geq 0\end{aligned}\tag{2.3.23}$$

这里A是$m \times n$矩阵，b和c分别是为m和n维向量，z是对偶问题中加入的松弛变量(n维向量)。

对对偶问题(D)引进对数障碍函数，则问题转换为

$$\begin{aligned}()\qquad &\max\ \ b^T y + \mu \sum_{j=1}^{n} \ln z_j\\ &\mathrm{s.t.}\ \ A^T y + z = c\end{aligned}\tag{2.3.24}$$

式中，μ为障碍参数($\mu > 0$)。显然x为(2.3.24)的约束条件的拉格朗日乘数，因此相应的拉格朗日函数为

$$L(x, y, z, \mu) = b^T y + \mu \sum_{j=1}^{n} \ln z_j - x^T (A^T y + z - c)\tag{2.3.25}$$

第一阶最优性条件，即其分别对z、y和x的一阶导数为零，导致下列方程组:

$$\begin{aligned} DGe - \mu e &= 0\\ Ax - b &= 0\\ A^T y + z - c &= 0\end{aligned}\tag{2.3.26}$$

式中，D和G分别为对角元为x_j和z_j的三角矩阵，因此下面的计算式中DG^{-1}与$G^{-1}D$的结果一致。注意上式的后两式，分别是通常的原问

题和对偶问题的可行性条件，而第一式即为$\mu \to 0$的极限状态时地互余松弛条件。

以$(\Delta x, \Delta y, \Delta z)$表示Newton搜索方向，根据(2.2.8)式，从(2.3.26)式可得出如下方程组：

$$\begin{bmatrix} -G & 0 & -D \\ -A & 0 & 0 \\ 0 & -A^T & -I \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} DGe - \mu e \\ 0 \\ 0 \end{bmatrix}$$

即

$$\begin{aligned} G\Delta x + D\Delta z &= \mu e - DGe \\ A\Delta x &= 0 \\ A^T\Delta y + \Delta z &= 0 \end{aligned} \tag{2.3.27}$$

解该方程组得出按如下计算步骤地式子:

$$\begin{aligned} \Delta y &= -(ADG^{-1}A^T)^{-1}AG^{-1}v(\mu) \\ \Delta z &= -A^T\Delta y \\ \Delta x &= G^{-1}v(\mu) - DG^{-1}\Delta z \end{aligned} \tag{2.3.28}$$

式中，$v(\mu) = \mu e - DGe$

(2.3.28)的全显示表示式则为

$$\begin{aligned} \Delta x &= [G^{-1} - DG^{-1}A^T(ADG^{-1}A^T)^{-1}AG^{-1}][\mu e - DGe] \\ \Delta y &= -[(ADG^{-1}A^T)^{-1}AG^{-1}][\mu e - DGe] \\ \Delta z &= [A^T(ADG^{-1}A^T)^{-1}AG^{-1}][\mu e - DGe] \end{aligned} \tag{2.3.29}$$

式中，$v(\mu) = \mu e - DGe$

根据计算的方向，考虑适当的步长参数α，则可写出迭代公式。在Monteiro和Adler[24]提出的算法中每步取步长参数$\alpha = 1$，且由于从一

步到下步迭代时μ的减少缓慢，因此这种算法收敛慢。在Mcshane等[26]提出分别对原问题空间和对偶问题空间采用不同的步长参数，即

$$\begin{aligned} x^{k+1} &= x^k + \alpha_p^k \Delta x^k \\ y^{k+1} &= y^k + \alpha_D^k \Delta y^k \\ z^{k+1} &= z^k + \alpha_D^k \Delta z^k \end{aligned} \tag{2.3.30}$$

这里处理方法被证实很有效。

该算法有两个直接的优点，一是对步长参数的选择，可以首先找出迭代时可能使某个x_j或某个z_j分别变成负值（不可行）的最大步长，分别为$\bar{\alpha}_p$和$\bar{\alpha}_D$，即

$$\bar{\alpha}_p = \min_{1 \le j \le n} \{\frac{-x_j}{\Delta x_j} | \Delta x_j < 0\} \tag{2.3.31a}$$

$$\bar{\alpha}_D = \min_{1 \le j \le n} \{\frac{-z_j}{\Delta z_j} | \Delta z_j < 0\} \tag{2.3.31b}$$

进而考虑步长α_p和α_D为用一定常数$\gamma(< 1)$乘这个最大步长，即

$$\alpha_p = \gamma \bar{\alpha}_p \tag{2.3.32a}$$

$$\alpha_D = \gamma \bar{\alpha}_D \tag{2.3.32b}$$

对原仿射法何对偶仿射法来说，可取$\gamma = 0.95$等；而对(2.3.28)的原-对偶法的计算，由于含有参数μ，经验表明，在$\alpha_p \le 1$和$\alpha_D \le 1$的条件下可取$\gamma = 0.9995$。另一个优点是，对当前原问题可行解x和对偶问题可行解y与z来说，可知精确的对偶间隙值，即把(2.3.26)的第三式乘以x^T和第二式转置后再乘以y，可直接导出对偶间隙为

$$c^T x - b^T y = x^T z \tag{2.3.33}$$

这说明对可行点(x, y, z)来说，每次迭代都可给出它们与最优解的接近程度。

为说明迭代过程中对偶间隙的变化。用d^k表示第k次迭代的对偶间隙，由(2.3.30)和(2.3.29)，并注意到$c^T=(y^k)^TA+(z^k)^T$，因此

$$\begin{aligned} d^{k+1} &= c^Tx^{k+1}-b^Ty^{k+1} \\ &= (c^Tx^k-b^Ty^k)+\alpha_p^k[(y^k)^TA+(z^k)^T]\Delta x^k-\alpha_D^k b^T\Delta y^k \end{aligned}$$

代入(2.3.28)中的Δx^k和Δz^k，并考虑到$A\Delta x=0$，$(z^k)^TG^{-1}=e^T$，$De=x^k$，化简后得到

$$d^{k+1}=d^k+\alpha_p^k[n\mu-(z^k)^Tx^k]+(\alpha_p^k-\alpha_D^k)b^T\Delta y^k$$

由于$d^k=(z^k)^Tx^k$，因此上式为

$$d^{k+1}=d^k+\alpha_p^k[n\mu-d^k]+(\alpha_p^k-\alpha_D^k)b^T\Delta y^k \tag{2.3.34}$$

可见当$\alpha_p^k=\alpha_D^k=\alpha^k$时，有

$$d^{k+1}=d^k+\alpha^k[n\mu-d^k]$$

这时只要$\mu<\frac{d^k}{n}$时，对偶空隙总是递减的。因此一般取$\mu_k=\rho\frac{d^k}{n}(0<\rho<1)$。而当$\alpha_p^k\neq\alpha_D^k$时，对偶空隙递减的情况不一定成立。

注意，(2.3.29)中含有(μ)的项是表示找出迭代点x^k和z^k的中心，而不含(μ)的项是表示求最优解，即缩小对偶空隙。对求中心来说，允许较大步长。这表明障碍参数法要优于仿射方法。

关于起始可行解，Lustig提出把人工变量x_{n+1}和y_{m+1}分别引入原问题和对偶问题，并且相互增加相应的附加约束条件，构成如下的增广原问题：

$$\begin{aligned} \min\quad & c^Tx+M_px_{n+1} \\ \text{s.t.}\quad & Ax+d_px_{n+1}=b \\ & d_D^Tx+x_{n+2}=M_D \\ & x\geq 0,\ \ x_{n+1},\ \ x_{n+2}\geq 0 \end{aligned} \tag{2.3.35}$$

和增广对偶问题:

$$\begin{aligned}
\max \quad & b^T y + M_D y_{m+1} \\
\text{s.t.} \quad & A^T y + d_D y_{m+1} + z = c \\
& d_p^T y + z_{n+1} = M_p \\
& y_{m+1} + z_{n+2} = 0 \\
& z \geq 0, \quad z_{n+1}, \quad z_{n+2} \geq 0
\end{aligned} \tag{2.3.36}$$

式中,$d_p = b - Ax^0$, $d_D = A^T y^0 + z^0 - c$; 这里M_p和M_D是很大的正数, x^0, y^0和z^0是任意的起始点。再令$x_{n+1}^0 = 1$和$y_{m+1}^0 = -1$, 则它们构成了(2.3.35)和(2.3.36)的一组其实可行解。由于M_p和M_D是足够大的正数，故上两问题的最优解即分别为问题(2.3.22)和(2.3.23)的最优解，且有$x_{n+1} = 0$ 和$y_{m+1} = 0$.

因此可总结算法如下:

给出起始可行解x^0, y^0和z^0

begin

$k := 0$

$d := 100\varepsilon$

$x^k := x^0$, $z^k := z^0$, $y^k := y^0$

while $d > \varepsilon$

do

$\mu := \rho \times \frac{d}{n}$

$D := diag(x_1^k, x_2^k, \cdots, x_n^k)$

$G := diag(z_1^k, z_2^k, \cdots, z_n^k)$

$\Delta y := -(ADG^{-1}A^T)^{-1}AG^{-1}(\mu e - DGe)$

$\Delta z := -A^T \Delta y$

$\Delta x := G^{-1}v(\mu) - DG^{-1}\Delta z$

Find α_p, α_D

$x^{k+1} := x^k + \alpha_p^k \Delta x^k$

$y^{k+1} := y^k + \alpha_D^k \Delta y^k$

$z^{k+1} := z^k + \alpha_D^k \Delta z^k$

$d := (x^{k+1})^T z^{k+1}$

$k := k + 1$

end do

end

§2.4 不可行的内点算法

上面介绍的各种内点算法均需要有一个初始的严格可行解，而对多数实际问题来说，一开始就能提供便利的可行起始点的情况是很少有的。这时，一方面可以通过引入人工变量和人工约束的办法来解决，但计算起来会有累计误差。另一方面，直接采用不可行内点算法。这方面已有的方法基本上是一般的原-对偶路径跟踪法。在基于势函数的方面，已有用原势函数下降法的方法[3, 5, 33]，但都没有达到原-对偶路径跟踪法那样的有效性。而从另一方面，Ye、Todd和Misunol1994年发表了构建“齐次自对偶问题”的方法[33]，它把原问题及其对偶问题的一个已知的不可行起始点变为齐次自对偶问题的严格可行起始点，从该齐次自对偶问题的解又可得出原问题和其对偶问题的解。

后来许多学者进行研究，把这种自对偶嵌入的思想拓展到求解一些更一般的约束凸优化问题。主要的工作有两个方面，一方面，对凸椎优化问题，包括半正定规划，Luo, Sturm 和Zhang[36]提出了一个自对偶潜入模型。JosSturm和SeDuMi进行了相关的算法数值测试，用

这种方法有效的求解了对称椎优化问题。另一方面，对具有不等式约束的椎优化问题，Andersen和Ye[34, 35]从Xu、Hung和Ye[38]关于线性规划问题建立的一些基本模型中得到启发，提出了一些不同的自对偶嵌入模型。实际上，Andersen和Ye提出的方法是针对非线性规划的互补问题，因此更具一般性。不过由于非线性规划的非线性，使得这种方法在实际的迭代过程中产生一些不可避免的误差。

到了2004年，Zhang[39]对具有不等式约束的椎优化问题提出一个特定的椎方程式，即把原问题中的不等式约束转化为椎约束，运用自对偶嵌入技术以，结合Nesterov和Nemiroviski提出的自协调障碍函数理论[37]克服了凸规划中初始点选取的困难。并证明了如果原问题的约束均为二次凸函数，那么自对偶嵌入问题的障碍函数是自协调的。而这一类问题的时间复杂度为$O(\sqrt{r}\log\frac{1}{\epsilon})$，其中，r为约束的个数，$\epsilon > 0$是需要的算法精度。通过对大量数值算例的测试，表明了该方法是可行的、有效的[40]。

第三章 结合线搜索的罚内点算法

通过前面的介绍可以看出，对于凸规划和椎优化问题，初始可行点的难题已被较好的解决了，现在很多人尝试把自对偶嵌入技术继续拓展到求解更一般的有约束非线性规划问题，不过还未有最新的进展。本文中，我们也对这一难题作了研究，并提出了通过引入辅助变量，构造原问题的等价问题从而来克服这一难题。

§3.1 障碍KKT条件

首先，我们考虑具有如下形式的约束最优化问题

$$\begin{aligned} (\mathrm{P}) \qquad & \min \ f(x) \\ & \text{s.t.} \ \ g_i(x) \le 0, i = 1, \cdots, m. \end{aligned} \tag{3.1.1}$$

其中，函数$f, g_i : R^n \to R, i = 1, \cdots, m$是二次连续可微的。

引入变量x_0, (3.1.1)则变成

$$\begin{aligned} (\mathrm{P}) \qquad & \min \ f(x) \\ & \text{s.t.} \ \ x_0 - g_i(x) \ge 0, i = 1, \cdots, m. \\ & \qquad x_0 = 0 \end{aligned} \tag{3.1.2}$$

对(3.1.2)定义拉格朗日函数

$$L(w) = f(x) - yx_0 - \sum_{i=1}^{m} z_i(x_0 - g_i(x)) \tag{3.1.3}$$

其中$w = (\bar{x}, y, z)^T, \bar{x} = (x_0, x)^T \in R^1 \times R^n$是决定变量，$y \in R^1$和$z \in R^n$分别是等式和不等式约束的Lagrange乘子。那么如上问题的K-K-T条件就

是

$$\gamma_o(w) \equiv \begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 \\ \bar{G} Z e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.1.4}$$

$$\bar{g} = (x_0 - g_1(x), \cdots, x_0 - g_m(x))^T \geq 0, z = (z_1, \cdots, z_m)^T \geq 0 \tag{3.1.5}$$

其中

$$\begin{aligned} \nabla_{\bar{x}} L(w) &= \begin{bmatrix} \nabla_{x_0} L \\ \nabla_x L \end{bmatrix} = \begin{bmatrix} 0 \\ \nabla f(x) \end{bmatrix} - y \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \sum_{i=1}^{m} z_i \begin{bmatrix} 1 \\ -\nabla g_i(x) \end{bmatrix} \\ &= \begin{bmatrix} -y - \sum_{i-1}^{m} z_i \\ \nabla f(x) + \sum_{i=1}^{m} z_i \nabla g_i(x) \end{bmatrix} = \begin{bmatrix} -y - z^T e \\ \nabla f(x) + \nabla g(x)^T z \end{bmatrix} \end{aligned} \tag{3.1.6}$$

$$\nabla g(x) = \begin{bmatrix} \nabla^T g_1(x) \\ \vdots \\ \nabla^T g_m(x) \end{bmatrix}, Z = diag(z_1, \cdots, z_m),$$

$$\bar{G} = diag(x_0 - g_1(x), \cdots, x_0 - g_m(x))^T, e = (1, \ldots, 1)^T \in R^m$$

在运用内点法求解的过程中，我们普遍采用Newton法来求解上述KKT条件等式，即

$$\begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 \\ \bar{G} Z e - \mu e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \bar{g}(x) > 0, \quad and \quad z > 0 \tag{3.1.7}$$

其中$\mu > 0$是罚参数。在这种情况下，Newton迭代方向$d_w = (d_{\bar{x}}, d_y, d_z)^T$

具有如下形式

$$\begin{bmatrix} 0 & 0 & -1 & -e^T \\ 0 & H & 0 & \nabla g(x) \\ 1 & 0 & 0 & 0 \\ z & -\nabla^T g(x) z & 0 & \bar{G} \end{bmatrix} d_w = - \begin{bmatrix} -y - z^T e \\ \nabla f(x) + \nabla g(x) z \\ x_0 \\ \bar{G} Z e - \mu e \end{bmatrix}$$

其中矩阵H是$\nabla_{xx} L(w)$或是其近似值。

§3.2 内点算法

首先，引入障碍法函数$Q(\bar{x}, \mu) : R^{n+1} \to R^1$

$$Q(\bar{x}, \mu) = f(x) - \mu \sum_{i=1}^{m} log(x_0 - g_i(x)) + \frac{1}{\mu} x_0^2, \quad \mu \geq 0 \tag{3.2.8}$$

其中，罚参数$\mu > 0$是一给定的的常数。可以看到，当μ的取值足够小时，问题(3.2.8)便近似等价于原问题(3.1.2)。问题(3.2.8)的必要最优性条件是

$$\nabla Q(\bar{x}, \mu) = \begin{bmatrix} 0 \\ \nabla f(x) \end{bmatrix} - \mu \sum_{i=1}^{m} \begin{bmatrix} \frac{1}{x_0 - g_i(x)} \\ -\frac{\nabla g_i(x)}{x_0 - g_i(x)} \end{bmatrix} + \frac{1}{\mu} \begin{bmatrix} 2x_0 \\ 0 \end{bmatrix} = 0 \tag{3.2.9}$$

即

$$-\mu \sum_{i=1}^{m} \frac{1}{x_0 - g_i(x)} + \frac{1}{\mu} 2x_0 \quad = 0 \tag{3.2.10}$$

$$\nabla f(x) + \mu \sum_{i=1}^{m} \frac{\nabla g_i(x)}{x_0 - g_i(x)} \quad = 0 \tag{3.2.11}$$

等价于

$$\frac{2x_0}{\mu} - \mu (\bar{g}^{-1})^T e \quad = 0 \tag{3.2.12}$$

$$\nabla f(x) + \mu \nabla g(x)^T \bar{g}^{-1} \quad = 0 \tag{3.2.13}$$

其中

$$\bar{g}^{-1} = \left(\frac{1}{x_0 - g_1(x)}, \cdots, \frac{1}{x_0 - g_m(x)}\right)^T 。$$

我们引入两个辅助变量$z = \mu\bar{G}^{-1}e \in R^m$和$y = -z^Te$，上述问题可转化为

$$\gamma(w,\mu) = \begin{bmatrix} \nabla_{\bar{x}}L(w) \\ x_0 + \frac{\mu}{2}y \\ \bar{G}Ze - \mu e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{3.2.14}$$

其中$\bar{g} > 0, z > 0, -\frac{1}{\mu} \leq y \leq \frac{1}{\mu}$。

在这里，我们定义条件(3.2.14)为障碍罚KKT条件，满足这一条件的点$w(\mu) = (\bar{x}(\mu), y(\mu), z(\mu))^T \in R^{n+1} \times R^1 \times R^m$称为障碍罚KKT点。当$\mu \downarrow 0$时，$w(\mu)$满足K-K-T条件。

现在，我们定义内点算法的框架

算法IP

步骤0. 令$\varepsilon > 0, \bar{x}^0 = (x_0^0, x^0) \in R^{n+1}$ 满足$x_0^0 > \max\limits_{1\leq i\leq m} g_i(x^0), M_c > 0$，当$k :=$ 0，给定$\{\mu_k\}$，使$\mu_k \downarrow 0$。

步骤1. 如果$\|\gamma_0(w^k)\| \leq \varepsilon$，停止。

步骤2. 寻找一内点w^{k+1} 使其满足$\|\gamma(w^{k+1}, \mu_k)\| \leq M_c\mu_k$。

步骤3. 令$k := k+1$ 转向步骤1。

接下来给出算法的收敛性定理

定理 3.1 *假设w^k是由算法产生的点列，$\{\bar{x}^k\}$和$\{y^k\}$是有界的。则有$\{z^k\}$ 有界，且$\{w^k\}$的聚点满足KKT条件(3.1.4)，(3.1.5)。*

证明 假设存在i使得$(z_i^k) \to \infty$，由步骤2可知

$$\left|\frac{-y^k - z^k e}{z_i^k}\right| \leq M_c \frac{\mu_{k-1}}{(z_i^k)}$$

即有

$$\left| \frac{-y^k - \sum\limits_{j \neq i} z_j^k}{z_i^k} - 1 \right| \leq M_c \frac{\mu_{k-1}}{(z_i^k)}$$

矛盾，因此$\{z^k\}$有界。■

我们定义条件(3.2.14)是精确的最障碍罚优性条件，满足这一条件的点为精确的障碍罚最优点。

为了寻找当$\mu \geq 0$给定时的精确的障碍罚最优点，我们采用Newton方法来进行求解。令$d_w = (d_{\bar{x}}, d_y, d_z)^T$是如下问题的解：

$$N(w)d_w = -\gamma(w, \mu) \tag{3.2.15}$$

其中$d_{\bar{x}} = (d_{x_0}, d_x)$。且由(3.2.14)

$$\gamma(w, \mu) = \begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 + \frac{\mu}{2} y \\ \bar{G}Ze - \mu e \end{bmatrix} = \begin{bmatrix} -y - z^T e \\ \nabla f(x) + \nabla g(x)^T z \\ x_0 + \frac{\mu}{2} y \\ \bar{G}Ze - \mu e \end{bmatrix}$$

可得

$$N(w) = \begin{bmatrix} 0 & 0 & -1 & -e^T \\ 0 & H & 0 & \nabla g(x) \\ 1 & 0 & \frac{\mu}{2} & 0 \\ z & -\nabla g(x)^T z & 0 & \bar{G} \end{bmatrix}$$

H取$\nabla_{xx} L(w)$或其近似值。且当$H = \nabla_{xx} L(w)$时，$N(w)$ 是$\gamma(w, \mu)$关于w的Jacobian矩阵。

下面的引理给出了(3.2.15)可解得充分条件。

引理 3.2.1 如果$\frac{1}{2} z^T H z + H - \nabla g(x) \bar{G}^{-1} \nabla g(x)^T Z$是正定矩阵，那么$N(w)$非奇异。

证明 考虑等式方程

$$N(w)\begin{pmatrix} v_{x0} \\ v_x \\ v_y \\ v_z \end{pmatrix} = 0$$

其中$(v_{x0}, v_x, v_y, v_z) \in R^1 \times R^n \times R^1 \times R^m$。则可得

$$\begin{aligned} (\tfrac{1}{2}z^T Hz + H - \nabla g(x)\bar{G}^{-1}\nabla g(x)^T Z) \quad v_x &= 0 \\ v_y &= v_z \\ v_{x0} &= \tfrac{1}{2}v_z \\ v_z &= -\nabla g(x)^{-T} H v_x \end{aligned}$$

由假设可知$v_x = 0$，因此有$v_{x0} = 0$，$v_y = 0$，$v_z = 0$。得证。■

Newton法的基本迭代步骤可写成：

$$w^{k+1} = w^k + t_k d_{w^k}$$

其中t_k的取值是为了保证障碍罚函数$Q(\bar{x}, \mu)$值的下降。

§3.3 线搜索算法

当给定一个$\mu > 0$时，为了设计一个能寻找到障碍KKT点的全局收敛算法，有必要对全步长Newton迭代过程进行修正。迭代公式具有如下形式：

$$w^{k+1} = w^k + t_k d_{w^k} \tag{3.3.16}$$

其中t_k就是由下面介绍的线搜索方法来确定的。

在这里，我们用*Wolfe*条件来作为线搜索准则。对一点w^k，利用如下公式求解t_k:

$$Q(\bar{x}^k + t_k d_{\bar{x}^k}, \mu) \le Q(\bar{x}^k, \mu) + c_1 t_k \nabla Q_k^T d_{\bar{x}^k} \tag{3.3.17a}$$

$$\nabla Q(\bar{x}^k + t_k d_{\bar{x}^k}, \mu)^T d_{\bar{x}^k} \ge c_2 \nabla Q_k^T d_{\bar{x}^k} \tag{3.3.17b}$$

其中$0 < c_1 < c_2 < 1$。

现给出线搜索算法，定义为算法LS。这个算法可以看作为算法IP的一个内嵌算法（参见算法IP的第二步）。注意到这里给出的ε'是与算法IP中的$M_c\mu_k$相对应。

算法LS

步骤0. 令$\omega_0 \in R^{n+1} \times R^1 \times R^m$,$\mu > 0$，且$\varepsilon' > 0,0 < c_1 < c_2 < 1$，$k = 0$。

步骤1. 如果$\| \gamma(\omega^k, \mu_k) \| \le \varepsilon'$，停止。

步骤2. 由(3.2.15)计算迭代方向d_ω。

步骤3. 求解t_k，使之满足

$$Q(\bar{x}^k + t_k d_{\bar{x}^k}, \mu_k) \le Q(\bar{x}^k, \mu_k) + c_1 t_k \nabla Q_k^T d_{\bar{x}^k}$$

$$\nabla Q(\bar{x}^k + t_k d_{\bar{x}^k}, \mu_k)^T \ge c_2 \nabla Q_k^T d_{\bar{x}^k}$$

with $0 < c_1 < c_2 < 1$ 其中$0 < c_1 < c_2 < 1$。

步骤4.

$$w^{k+1} = w^k + t_k d_{w^k}$$

步骤5. 令$k := k + 1$，转向步骤1。

为了证明算法LS的全局收敛性，需要如下假设。

假设

1. 矩阵$\frac{1}{2}z^THz+H-\nabla g(x)\bar{G}^{-1}\nabla g(x)^TZ$是正定的。

2. ∇Q在Ω上是Lipschitz连续的，即，存在一个常数$L>0$使得$\|\nabla Q(w,\mu)-\nabla Q(\widetilde{w},\mu)\|\le L\|w-\widetilde{w}\|$，对所有的$w,\widetilde{w}\in\Omega$。

下面给出算法LS的收敛性证明。

定理 3.2 由假设，函数f, $g_i(x),i=1,\dots,m$是二次连续可微的。假定算法LS产生的无限序列$\{w^k\}$仍在紧集Ω内。那么序列$\{w^k\}$至少有一个聚点，且$\{w^k\}$的任意聚点均为障碍罚KKT点。

证明 由假设知，函数$Q(\bar{x},\mu)$在Ω内是有界的。注意到当采用拟Newton法来求解矩阵$\frac{1}{2}z^THz+H-\nabla g(x)\bar{G}^{-1}\nabla g(x)^TZ$时，则要求$f$及$g_i(x),i=1,\dots,m$的连续性。

定义

$$\cos\theta_k=\frac{-\nabla Q_k^Td_{\bar{x}^k}}{\|\nabla Q_k\|\|d_{\bar{x}^k}\|} \tag{3.3.18}$$

由(3.3.17b), (3.3.16)可得

$$(\nabla Q_{k+1}-\nabla Q_k)^Td_{\bar{x}^k}\ge(c_2-1)\nabla Q_k^Td_{\bar{x}^k}$$

又由于Lipschitz条件，则有

$$(\nabla Q_{k+1}-\nabla Q_k)^Td_{\bar{x}^k}\le t_kL\|d_{\bar{x}^k}\|^2$$

结合上面两个关系式，可得

$$t_k\ge\frac{c_2-1}{L}\frac{\nabla Q_k^Td_{\bar{x}^k}}{\|d_{\bar{x}^k}\|^2}$$

把这个不等式代入Wolfe条件的第一式(3.3.17a)中，有

$$Q_{k+1}\le Q_k-c_1\frac{1-c_2}{L}\frac{(\nabla Q_k^Td_{\bar{x}^k})^2}{\|d_{\bar{x}^k}\|^2}$$

由(3.3.18)的定义，有如下关系式

$$Q_{k+1} \le Q_k - c\cos^2\theta_k \|\nabla Q_k\|^2$$

其中，$c = c_1(1-c_2)/L$。把下标小于等于k的所有表达式相加，得

$$Q_{k+1} \le Q_0 - c\sum_{j=0}^{\infty} \cos^2\theta_j \|\nabla Q_j\|^2 \tag{3.3.19}$$

因为Q有下界的，则$Q_0 - Q_{k+1}$是某个正常数，对所有的k。
代入(3.3.19)即有

$$\sum_{k=0}^{\infty} \cos^2\theta_k \|\nabla Q_k\|^2 < \infty \tag{3.3.20}$$

则

$$\cos^2\theta_k \|\nabla Q_k\|^2 \to 0 \tag{3.3.21}$$

由$\cos\theta_k$的定义可知，存在正常数δ使得

$$\cos\theta_k \ge \delta > 0, \quad for \quad all \quad k$$

因此由(3.3.21)容易得出

$$\lim_{k\to\infty} \|\nabla Q_k\| = 0$$

■

接下来给出障碍罚参数μ_k的修正方法。假设现有一个精确的障碍罚KKT点w^{k+1}满足

$$\|\gamma(w^{k+1}, \mu_k)\| \le M_c \mu_k$$

（算法IP的第二步）。则μ_{k+1}的计算公式为

$$\mu_{k+1} = \max\left(\frac{\|\gamma(w^{k+1}, \mu_k)\|}{M_\mu}, \frac{\mu_k}{M_0}\right)$$

其中$0 < M_c < M_\mu$，$M_0 > 1$。在本文的算法程序调试中，我们取$M_\mu = 4$，$M_0 = 10^6$，$M_c = 3$。

§3.4 数值算例

在如下的部分，我们测试了一些数值算例，结果表明以上提出的算法是可行的、有效的。

例1.

$$
\begin{aligned}
\text{(P)} \qquad \min \quad & f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\
\text{s.t.} \quad & g_1(x) = -x_1 - x_2^2 \le 0 \\
& g_2(x) = -x_1^2 - x_2 \le 0 \\
& -0.5 \le x_1 \le 0.5 \\
& x_2 \le 1
\end{aligned}
$$

该问题的障碍罚函数为

$$
\begin{aligned}
& Q(\bar{x}, \mu) = f(x) - \mu \textstyle\sum_{i=1}^{m} \log(x_0 - g(x)) + \frac{1}{\mu} x_0^2 \\
& = f(x) - \mu \log(x_0 - (-x_1 - x_2^2)) - \mu \log(x_0 - (-x_1^2 - x_2)) - \mu \log(x_0 - (-0.5 - x_1)) \\
& -\mu \log(x_0 - (-0.5 + x_1)) - \mu \log(x_0 - (-1 + x_2)) + \frac{1}{\mu} x_0^2
\end{aligned}
$$

定义初始点: $x_0 = (9.9999997E-10, 0.5000000, 0.5000000), \varepsilon = 1E-05$

表1 共迭代5步

迭代步骤	$\bar{x}^k$	$Q(\bar{x}^k, \mu_k)$	$f(\bar{x}^k)$	μ_k	$\gamma(w)$
0	9.9999997E-10,0.5000000,0.5000000	6.500000	6.500000	9.9999997E-8	0.2500000
1	7.0000000E-10,0.5000000,0.2750000	0.3125001	0.3125000	2.5000004E-08	2.5000004E-02
2	1.0111600E-09,0.5000000,0.2525000	0.2506256	0.2506250	6.2500010E-09	2.4999869E-03
3	1.1389349E-08,0.5000000,0.2502500	0.2500064	0.2500063	1.5625002E-09	2.5001052E-04
4	3.9522856E-09,0.5000000,0.2500250	0.2500001	0.2500001	3.9062506E-10	2.5001265E-05
5	1.0981988E-09,0.5000000,0.2500025	0.2500000	0.2500000	9.7656265E-11	2.5023062E-06

由表1可知，当经过5步迭代后，得到点$x_5^* = (1.0981988E-09, 0.5000000, 0.2500025)$目标函数值$f(x_5^*) = 0.2500000$。注意到该问题的最优解点为$(0.5, 0.25)$，最优值为0.25。

例2.

$$\text{(P)} \qquad \min \quad f(x) = (x_1-2)^4 + (x_1-2x_2)^2$$
$$\text{s.t.} \quad g(x) = x_1^2 - x_2 \le 0$$

该问题的障碍罚函数为

$$Q(\bar{x},\mu) = f(x) - \mu\sum_{i=1}^{m}\log(x_0 - g(x)) + \frac{1}{\mu}x_0^2 = f(x) - \mu\log(x_0 - (x_1^2 - x_2) + \frac{1}{\mu}x_0^2$$

定义初始点: $x_0 = (4.9999999E-03, 1.000000, 1.000000), \varepsilon = 1E-05$

表2 共迭代4步

迭代步骤	$\bar{x}^k$	$Q(\bar{x}^k,\mu_k)$	$f(\bar{x}^k)$	μ_k	$\gamma(w)$
0	1.0000000E-02,1.000000,1.000000	2.007908	2.000000	0.2000000	0.7199850
1	1.0451545E-03,0.9900161,0.9817969	1.990646	1.982330	5.0000001E-02	5.1898486E-03
2	1.0806412E-03,0.9808956,0.9636081	1.976677	1.968040	1.2500000E-02	6.5903723E-04
3	1.1300817E-03,0.9630076,0.9285494	1.958725	1.949543	3.1250000E-03	5.1537370E-05
4	1.1586071E-03,0.9460912,0.8960579	1.952797	1.949465	7.8125000E-04	8.2657667E-06

由表2可知，当经过4步迭代后，得到点$x_4^* = (1.1586071E-03, 0.9460912, 0.8960579)$目标函数值$f(x_4^*) = 1.949465$。注意到该问题的最优解点为$(0.945, 0.894)$，最优值为1.94。

例3.

$$\text{(P)} \qquad \min \quad f(x) = (x_1-2)^2 + (x_2-1)^2$$
$$\text{s.t.} \quad g(x) = x_1 + x_2 - 2 \le 0$$

该问题的障碍罚函数为

$$Q(\bar{x},\mu) = f(x) - \mu\sum_{i=1}^{m}\log(x_0 - g(x)) + \frac{1}{\mu}x_0^2$$
$$= f(x) - \mu\log(x_0 - (x_1 + x_2 - 2)) + \frac{1}{\mu}x_0^2$$

定义初始点: $x_0 = (1.0000000E-06, 1.000000, 1.000000), \varepsilon = 1E-05$

表3 共迭代2步

迭代步骤	$\bar{x}^k$	$Q(\bar{x}^k,\mu_k)$	$f(\bar{x}^k)$	μ_k	$\gamma(w)$
0	1.0000000E-06,1.000000,1.000000	1.000014	1.000000	1.0000000E-06	0.7094432
1	5.5330622E-07,1.451487,0.5485129	0.5047206	0.5047075	2.5000000E-07	6.8861194E-02
2	5.0376644E-07,1.495310,0.5046899	0.5000446	0.5000445	6.2500000E-08	5.6593649E-07

由表3可知，当经过2步迭代后，得到点$x_2^* = (5.0376644E{-}07, 1.495310, 0.5046899)$目标函数值$f(x_2^*) = 0.5000445$。注意到该问题的最优解点为$(1.5, 0.5)$，最优值为0.5。

第四章　结合大步路径的罚内点算法

在本章中，我们将对一类具体的内点算法——路径跟踪算法——在非线性规划中的应用进行研究分析。在这一算法的收敛性证明中需要运用正交性条件，而这个条件在一般的非线性规划问题不一定得到满足。因此在本章中，我们不仅需要通过引入辅助变量来克服初始点选取的难题，还要通过构造新的关系不等式来绕开正交性难题。具体的分析如下：

§4.1　KKT条件

首先，考虑如下的约束优化问题：

$$\begin{aligned} (\mathrm{P}) \qquad & \min \ f(x) \\ & \text{s.t.} \ \ g_i(x) \le 0, i = 1, \cdots, m. \end{aligned} \tag{4.1.1}$$

其中，$f, g_i : R^n \to R, i = 1, \cdots, m$是二次连续可微的。

引入变量x_0，则(4.1.1)变成：

$$\begin{aligned} (\mathrm{P}) \qquad & \min \ f(x) \\ & \text{s.t.} \ \ x_0 - g_i(x) \ge 0, i = 1, \cdots, m. \\ & \qquad x_0 = 0 \end{aligned} \tag{4.1.2}$$

定义上述问题的拉格朗日函数：

$$L(w) = f(x) - yx_0 - \sum_{i=1}^{m} z_i(x_0 - g_i(x)) \tag{4.1.3}$$

其中，$w = (\bar{x}, y, z)^T$，$\bar{x} = (x_0, x)^T \in R^1 \times R^n$是决定变量，$y \in R^1$和$z \in R^n$分别是等式和不等式约束对应的拉格朗日乘子向量。因此上述问题

的KKT最优性条件为：

$$\begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 \\ \bar{G}Ze \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.1.4}$$

其中

$$\bar{g} = (x_0 - g_1(x), \cdots, x_0 - g_m(x))^T \geq 0, \quad z = (z_1, \cdots, z_m)^T \geq 0 \tag{4.1.5}$$

$$\begin{aligned} \nabla_{\bar{x}} L(w) &= \begin{bmatrix} \nabla_{x_0} L \\ \nabla_x L \end{bmatrix} = \begin{bmatrix} 0 \\ \nabla f(x) \end{bmatrix} - y \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \sum_{i=1}^{m} z_i \begin{bmatrix} 1 \\ -\nabla g_i(x) \end{bmatrix} \\ &= \begin{bmatrix} -y - \sum_{i-1}^{m} z_i \\ \nabla f(x) + \sum_{i=1}^{m} z_i \nabla g_i(x) \end{bmatrix} = \begin{bmatrix} -y - z^T e \\ \nabla f(x) + \nabla g(x)^T z \end{bmatrix} \end{aligned} \tag{4.1.6}$$

$$\nabla g(x) = \begin{bmatrix} \nabla^T g_1(x) \\ \vdots \\ \nabla^T g_m(x) \end{bmatrix}, Z = diag(z_1, \cdots, z_m),$$

$$\bar{G} = diag(x_0 - g_1(x), \cdots, x_0 - g_m(x)), \quad e = (1, \cdots, 1)^T \in R^m$$

现在，对KKT条件引入一个松弛变量$s = (s_1, s_2, \cdots, s_m)^T$，从而得到(4.1.4)的一个等价形式：

$$\begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 \\ \bar{g} - s \\ SZe \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.1.7}$$

其中$(s, z) \geq 0, \quad S = diag(s_1, s_2, \cdots, s_m)$。

在实际求解过程中，我们通常采用Newton法来求解上述KKT条件等式，那么Newton方程的解即为Newton迭代方向$\Delta w = (\Delta\bar{x}, \Delta s, \Delta y, \Delta z)^T$，有

$$\begin{bmatrix} 0 & 0 & 0 & -1 & -e^T \\ 0 & H & 0 & 0 & \nabla g(x) \\ 1 & 0 & 0 & 0 & 0 \\ e & -\nabla g(x) & -e & 0 & 0 \\ 0 & 0 & Z & 0 & S \end{bmatrix} \begin{bmatrix} \Delta x_0 \\ \Delta x \\ \Delta s \\ \Delta y \\ \Delta z \end{bmatrix} = - \begin{bmatrix} -y - z^T e \\ \nabla f(x) + \nabla g(x) z \\ x_0 \\ \bar{g} - s \\ SZe \end{bmatrix} \tag{4.1.8}$$

其中矩阵H等于$\nabla_{xx}L(w)$或其近似值。

§4.2 路径跟踪算法

众所周知，中心路径$\mathcal{C}$在内点算法中扮演着非常重要的角色，他是由一类严格可行点组成的弧形轨迹。若有参数$\tau > 0$，则每个点$(\bar{x}_\tau, s_\tau, y_\tau, z_\tau) \in \mathcal{C}$均可通过下列方程式进行求解：

$$r(w) \equiv \begin{bmatrix} \nabla_{\bar{x}} L(w) \\ x_0 \\ \bar{g} - s \\ SZe - \tau e \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad s > 0, \quad 以及 \quad z > 0 \tag{4.2.9}$$

本文中，我们称上述条件(4.2.9)为修正的KKT条件。与KKT条件(4.1.7)相比，唯一的差别在于最后一行中引入参数τ，其实就是把(4.1.7)中第三行中的互补条件换成了要求$s_i z_i$乘积值对任意下标i均相等的条件。由(4.2.9)可定义中心路径

$$\mathcal{C} = \{(\bar{x}_\tau, s_\tau, y_\tau, z_\tau) | \tau > 0\}$$

当τ趋进于零时，条件(4.2.9)等价于(4.1.7)。如果当$\tau\downarrow 0$, $\mathcal{C}$收敛于某一点时，那么该点必为原问题的最优点。因此，沿着中心路径搜索的方法给我们提供了一条找到最优解的途径，沿着这条途径不仅可以保证s的z各个分量都严格大于零，还可以使所有s_iz_i的乘积值几乎以相同的速率下降到零，这就使我们少走了很多"弯路"。

为了描述搜索方向的偏差，我们引入中心参数$\sigma\in[0,1)$和对偶参数μ，由$\mu=\frac{1}{n}\sum\limits_{i=1}^{n}s_iz_i=\frac{s^Tz}{n}$的定义知$\mu$衡量了$s_iz_i$乘积的均值。令$\tau=\sigma\mu$并用Newton法来求解(4.2.9)，即得

$$N(w)\Delta w=\begin{bmatrix} 0 & 0 & 0 & -1 & -e^T \\ 0 & H & 0 & 0 & \nabla g(x) \\ 1 & 0 & 0 & 0 & 0 \\ e & -\nabla g(x) & -e & 0 & 0 \\ 0 & 0 & Z & 0 & S \end{bmatrix}\begin{bmatrix} \Delta x_0 \\ \Delta x \\ \Delta s \\ \Delta y \\ \Delta z \end{bmatrix}=-\begin{bmatrix} -y-z^Te \\ \nabla f(x)+\nabla g(x)z \\ x_0 \\ \bar{g}-s \\ SZe-\sigma\mu e \end{bmatrix} \tag{4.2.10}$$

其中矩阵H等于$\nabla_{xx}L(w)$或其近似值。

当s_iz_i的各个分量均等于$\sigma\mu$时，Newton迭代方向$(\Delta\bar{x},\Delta s,\Delta y,\Delta z)$指向点$(\bar{x}_{\sigma\mu},s_{\sigma\mu},y_{\sigma\mu},z_{\sigma\mu})\in\mathcal{C}$。反之，由(4.1.8)求得的迭代方向则是直接指向满足KKT条件(4.1.7)的点。

接下来的引理给出了一个(4.2.10)可解的充分条件。

引理 4.2.1 如果矩阵$H+\nabla g(x)^TS^{-1}Z\nabla g(x)$是正定的，则矩阵$N(w)$非奇异。

证明 考虑方程式

$$\begin{bmatrix} 0 & 0 & 0 & -1 & -e^T \\ 0 & H & 0 & 0 & \nabla g(x) \\ 1 & 0 & 0 & 0 & 0 \\ e & -\nabla g(x) & -e & 0 & 0 \\ 0 & 0 & Z & 0 & S \end{bmatrix} \begin{pmatrix} v_{x0} \\ v_x \\ v_s \\ v_y \\ v_z \end{pmatrix} = 0$$

其中$(v_{x0}, v_x, v_s, v_y, v_z) \in R^1 \times R^n \times R^m \times R^1 \times R^m$。则有

$$\begin{aligned} H + \nabla g(x)^T S^{-1} Z \nabla g(x) \quad & v_x = 0 \\ & v_y = -e^T v_z \\ & v_{x0} = 0 \\ & v_s = -\nabla g(x) v_x \\ & v_z = -\nabla g(x)^{-T} H v_x \end{aligned}$$

由假设知$v_x = 0$，因此$v_{x0} = 0$，$v_s = 0$，$v_y = 0$，$v_z = 0$。由此可求得$(\Delta\bar{x}, \Delta s, \Delta y, \Delta z)^T$，得证。■

路径跟踪算法要求每步产生的迭代点均落于一个包含中心路径C的区域内，且沿着这条中心路径搜索至原问题的最优解。为了避免迭代点过于接近非负区域的边界，我们要求每次迭代中的搜索方向必须使下一步产生的点列更加接近最优点。

最优性算法的一个要点是如何在搜索空间中衡量满足条件的点，在路径跟踪算法中，对偶参数μ担当了这个重要的角色。当$k \to \infty$时，μ_k下降至零，因此迭代点$(\bar{x}^k, s^k, y^k, z^k)$逐步满足KKT条件(4.1.7)。

现定义单边∞范数区域$\mathcal{N}_{-\infty}(\gamma)$

$$\mathcal{N}_{-\infty}(\gamma) = \{(\bar{x}, s, y, z) \in \mathcal{F}^0 | s_i z_i \geq \gamma\mu \quad \text{对所有的} \quad i = 1, 2, \ldots, n\} \tag{4.2.11}$$

其中$\gamma \in (0,1]$，$\mathcal{F}^0 = \{(\bar{x}, s, y, z) | s > 0\}$。若存在一点落在$\mathcal{N}_{-\infty}(\gamma)$中，则$s_i z_i$的每个分量必定大于$\mu$与$\gamma$的乘积。其实这个约束是非常松的，当$\gamma$趋近于零时，$\mathcal{N}_{-\infty}(\gamma)$就包含可行域$\mathcal{F} = \{(\bar{x}, s, y, z) | s \geq 0\}$中的大部分点，这里，我们取$\gamma = 10^{-3}$。

下面介绍的大步长路径跟踪算法，由于它采用了一个较为广泛的区域$\mathcal{N}_{-\infty}(\gamma)$（对$\gamma$趋于零），因此具有良好的实用性。通过(4.2.10)可求得搜索方向，步长α_k则是保证迭代点落于$\mathcal{N}_{-\infty}(\gamma)$的最大取值。

给出一些定义：

$$(\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha)) = (\bar{x}^k, s^k, y^k, z^k) + \alpha(\Delta\bar{x}^k, \Delta s^k, \Delta y^k, \Delta z^k) \tag{4.2.11a}$$

$$\mu_k(\alpha) = s^k(\alpha)^T z^k(\alpha)/n \tag{4.2.11b}$$

大步长算法

步骤0. 给定$\varepsilon > 0$，$\gamma, \sigma \in (0,1)$，取x^0及$x_0^0 > \max(g_i(x^0))$，$s^0 = \bar{g}^0$，y^0和z^0；

步骤1. 如果$\|r(w^k)\| \leq \varepsilon$，停止；

步骤2. 由(4.2.10)求得搜索方向$(\Delta\bar{x}^k, \Delta s^k, \Delta y^k, \Delta z^k)$；

步骤3. 取最大值α_k使得α在区间[0,1]内，且有

$$(\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma)$$

步骤4.

$$(\bar{x}^{k+1}, s^{k+1}, y^{k+1}, z^{k+1}) = (\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha))$$

步骤5. 令$k := k+1$，转步骤1。

§4.3 收敛性分析

下面给出几个定理来说明上述算法的收敛性。首先介绍引理(4.3.1)——见[48]，并用它来证明引理(4.3.2)，给出$\Delta\bar{g}_i\Delta z_i$，$i=1,2,\ldots,n$向量乘积的边界。定理(4.1)给出了$\alpha_k$的一个下确界以及迭代过程中参数$\mu$下降量的估计值。由定理(4.1)即得全局收敛性。为了证明算法的收敛性，给出如下假设

假设

1. 矩阵$H+\nabla g(x)^T S^{-1} Z \nabla g(x)$正定

2. 不等式$0 \le \Delta s^T \Delta z \le (1-\sigma)s^T z$成立

引理 4.3.1 假设u、v是任意两个n维向量，且有$u^T v \ge 0$。那么

$$\|UVe\| \le 2^{-3/2}\|u+v\|^2$$

其中

$$U = diag(u_1, u_2, \ldots, u_n), \qquad V = diag(v_1, v_2, \ldots, v_n)$$

引理 4.3.2 如果$(\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha) \in \mathcal{N}_{-\infty}(\gamma)$，那么

$$\|\Delta S \Delta Z e\| \le 2^{-3/2}(1+1/\gamma)n\mu$$

证明 由假设易得

$$\Delta s^T \Delta z \ge 0 \tag{4.3.12}$$

(4.2.10)的最后一行两边分别乘以$(SZ)^{-1/2}$，并令$D = S^{1/2}Z^{-1/2}$，得到

$$D^{-1}\Delta s + D\Delta z = (SZ)^{-1/2}(-SZe + \sigma\mu e) \tag{4.3.13}$$

因为$(D^{-1}\Delta s)^T(D\Delta z) = \Delta s^T \Delta z \geq 0$，令$u = D^{-1}\Delta s$，$v = D\Delta z$，由引理(4.3.1)得

$$\begin{aligned}
\|\Delta S \Delta Z e\| &= \|(D^{-1}\Delta s)(D\Delta z)e\| \\
&\leq 2^{-3/2}\|D^{-1}\Delta s + D\Delta z\|^2 && \text{由引理 (4.3.1)} \\
&= 2^{-3/2}\|(SZ)^{-1/2}(-SZe + \sigma\mu e)\|^2 && \text{由 (4.3.13)}
\end{aligned}$$

利用欧拉范数和关系式$s^T z = n\mu$，$e^T e = n$，有

$$\begin{aligned}
\|\Delta S \Delta Z e\| &\leq 2^{-3/2}[s^T z - 2\sigma\mu e^T e + \sigma^2\mu^2 \sum_{i=1}^{n} \tfrac{1}{s_i z_i}] \\
&\leq 2^{-3/2}[s^T z - 2\sigma\mu e^T e + \sigma^2\mu^2 \tfrac{n}{\gamma\mu}] && \text{因为} \quad s_i z_i \geq \gamma\mu \\
&\leq 2^{-3/2}[1 - 2\sigma + \tfrac{\sigma^2}{\gamma}]n\mu \\
&\leq 2^{-3/2}(1 + 1/\gamma)n\mu
\end{aligned}$$

得证。■

定理 4.1 给定算法参数γ和σ，则存在一个常数$\delta \in (0,1)$使得

$$\mu_{k+1} \leq (1-\delta)\mu_k \tag{4.3.14}$$

对所有的$k \geq 0$

证明 首先考虑

$$(\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha)) \in \mathcal{N}_{-\infty}(\gamma) \qquad for\ all\ \alpha \in [0, 2^{3/2}\gamma\tfrac{1-\gamma}{1+\gamma}\tfrac{\sigma}{n+\gamma}] \tag{4.3.15}$$

因为α_k是[0,1]区间内α的最大值，那么α_k的下确界可以写成

$$\alpha_k \geq 2^{3/2}\gamma\tfrac{1-\gamma}{1+\gamma}\tfrac{\sigma}{n+\gamma} \tag{4.3.16}$$

由引理(4.3.2)知，对任意的$i = 1, 2, \ldots, n$，有

$$|\Delta s_i^k \Delta z_i^k| \leq \|\Delta S^k \Delta Z^k e\|_2 \leq 2^{-3/2}(1 + 1/\gamma)n\mu_k \tag{4.3.17}$$

利用(4.2.10)、(4.3.17)和关系式$s_i^k z_i^k \geq \gamma\mu_k$

$$\begin{aligned}
s_i^k(\alpha)z_i^k(\alpha) &= (s_i^k + \alpha\Delta s_i^k)(z_i^k + \alpha\Delta z_i^k) \\
&= s_i^k z_i^k + \alpha(s_i^k\Delta z_i^k + z_i^k\Delta s_i^k) + \alpha^2\Delta s_i^k\Delta z_i^k \\
&\geq s_i^k z_i^k(1-\alpha) + \alpha\sigma\mu_k - \alpha^2|\Delta s_i^k\Delta z_i^k| \\
&\geq \gamma(1-\alpha)\mu_k + \alpha\sigma\mu_k - \alpha^2 2^{-3/2}(1+1/\gamma)n\mu_k
\end{aligned}$$

累加方程$S^k\Delta z^k + Z^k\Delta s^k = -S^kZ^ke + \sigma\mu_k e$（由(4.2.10)的最后一行求得）的n个分量，利用(4.3.17)和μ_k、$\mu_k(\alpha)$)的定义，即得

$$\begin{aligned}
\mu_k(\alpha) &= (1-\alpha(1-\sigma))\mu_k + \tfrac{\alpha^2}{n}\Delta s_i^k\Delta z_i^k \\
&\leq (1-\alpha(1-\sigma))\mu_k + \tfrac{\alpha^2}{n}|\Delta s_i^k\Delta z_i^k| \\
&\leq (1-\alpha(1-\sigma))\mu_k + \alpha^2 2^{-3/2}(1+1/\gamma)\mu_k
\end{aligned}$$

为了满足邻近性条件

$$s_i^k(\alpha)z_i^k(\alpha) \geq \gamma\mu_k(\alpha)$$

则有

$$\gamma(1-\alpha)\mu_k+\alpha\sigma\mu_k-\alpha^2 2^{-3/2}(1+1/\gamma)n\mu_k \geq \gamma((1-\alpha(1-\sigma))\mu_k+\alpha^2 2^{-3/2}(1+1/\gamma)\mu_k)$$

整理上式

$$\alpha\sigma\mu_k(1-\gamma) \geq \alpha^2 2^{-3/2}(n+\gamma)\mu_k(1+1/\gamma)$$

当

$$\alpha \leq \frac{2^{3/2}}{n+\gamma}\sigma\gamma\frac{1-\gamma}{1+\gamma}$$

邻近性条件成立。我们证明了当α的取值范围为(4.3.15)时，$(\bar{x}^k(\alpha), y^k(\alpha), s^k(\alpha), z^k(\alpha))$满足$\mathcal{N}_{-\infty}(\gamma)$的邻近性条件。在本章的开头就提到，$z_i^k(\alpha)$是对应第k不迭代后第i个不等式约束的拉格朗日乘子，拉格朗日乘子的性质保证了$z_i^k(\alpha) > 0$。又因为对所有的k有$\mu_k > 0$，则容易得出对任意α在给定

的区域内，有$(\bar{x}^k(\alpha), s^k(\alpha), y^k(\alpha), z^k(\alpha)) \in \mathcal{F}^0$。因此(4.3.15)得证，继而得出(4.3.16)。

接下来估计在第k步迭代中，参数μ的下降量。由(4.3.17)、(4.3.16)、(4.2.10)的最后一行以及假设，得

$$
\begin{aligned}
\mu_{k+1} &= s^k(\alpha_k)^T z^k(\alpha_k)/n \\
&= [(s^k)^T z^k + \alpha_k((s^k)^T \Delta z^k + (z^k)^T \Delta s^k) + \alpha_k^2 (\Delta s^k)^T \Delta z^k]/n \\
&= \mu_k + \alpha_k(-(s^k)^T z^k/n + \sigma\mu_k) + \alpha_k^2 (\Delta s^k)^T \Delta z^k/n \\
&\le (1-\alpha_k(1-\sigma))\mu_k + \alpha_k^2(1-\sigma)\tfrac{(s^k)^T z^k}{n} \\
&= (1-\alpha_k(1-\sigma) + \alpha_k^2(1-\sigma))\mu_k \\
&= (1-\alpha_k(1-\alpha_k)(1-\sigma))\mu_k
\end{aligned}
\tag{4.3.18}
$$

可知$\alpha(1-\alpha)$是一个关于α的二次凹函数，所以在任意给定的区间范围内，函数的最小值在区间的端点处取到。在(4.3.18)中引入替代估计参数

$$\delta = (1-\sigma)\min\{\alpha_k(1-\alpha_k)\} \qquad 其中 \quad \alpha_k \in [\frac{2^{3/2}}{n+\gamma}\sigma\gamma\frac{1-\gamma}{1+\gamma}, 1]$$

得证。■

§4.4 数值算例

在如下的部分，我们测试了一些数值算例，结果表明以上提出的算法是可行的、有效的。

例1.

$$
\begin{aligned}
\text{(P)} \qquad \min \quad & f(x) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2 \\
\text{s.t.} \quad & g_1(x) = -x_1 - x_2^2 \le 0 \\
& g_2(x) = -x_1^2 - x_2 \le 0 \\
& -0.5 \le x_1 \le 0.5 \\
& x_2 \le 1
\end{aligned}
$$

定义初始点: $\bar{x}^0 = (1E{-}05, 0.5, 0.5)$; $s^0 = (0.7500100, 0.7500100, 9.9999997E{-}06, 1.000010,$
$0.5000100)$; $y^0 = 1.0$; $z^0 = (1.0, 1.0, 1.0, 1.0, 1.0)$; $\sigma = 0.5$; $\varepsilon = 1E - 08$.

表1 共迭代9步

迭代步骤	$\bar{x}^k$	$f(\bar{x}^k)$	$\gamma\mu^k$	$r(w^k)$
0	(9.9999997E-06,0.5000000,0.5000000)	6.500000	6.0001004E-04	3.820576
1	(1.0000002E-06,0.4955080,0.2736015)	0.3333232	6.3562758E-05	1.238662
2	(1.0000004E-07,0.4978577,0.2506172)	0.2529059	1.0329035E-05	0.2167492
3	(1.0000007E-08,0.4995730,0.2498444)	0.2504346	1.0498230E-06	2.4334069E-02
4	(1.0000010E-09,0.4999518,0.2499787)	0.2500483	1.0494578E-07	2.4822732E-03
5	(1.0000012E-10,0.4999951,0.2499978)	0.2500049	1.0493346E-08	2.4882317E-04
6	(1.0000015E-11,0.4999995,0.2499998)	0.2500005	1.0468115E-09	2.4862107E-05
7	(1.0000017E-12,0.4999999,0.2500000)	0.2500001	1.0706392E-10	2.4913159E-06
8	(1.0000019E-13,0.5000000,0.2500000)	0.2500000	9.5142817E-12	2.7764307E-07
9	(1.0000022E-14,0.5000000,0.2500000)	0.2500000	9.5142839E-13	7.7662126E-09

由表1可知，当经过9步迭代后，得到点$x_9^* = (0.5000000, 0.2500000)$目标函数值$f(x_9^*) = 0.2500000$。注意到该问题的最优点为$(0.5, 0.25)$，最优值为0.25。

例2.

$$
\begin{aligned}
\text{(P)} \qquad \min \quad & f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2 \\
\text{s.t.} \quad & g(x) = x_1^2 - x_2 \le 0
\end{aligned}
$$

定义初始点: $\bar{x}^0 = (1,0,0)$; $s^0 = 1.0$; $y^0 = 1.0$; $z^0 = (2.0)$; $\sigma = 0.5$; $\varepsilon = 1E - 06$.

表2 共迭代6步

迭代步骤	$\bar{x}^k$	$f(\bar{x}^k)$	$\gamma\mu^k$	$r(w^k)$
0	(1.000000,0.0000000E+00,0.0000000E+00)	16.00000	2.0000001E-03	2.612066
1	(0.1000000,0.5633932,0.4058072)	4.321046	2.2473413E-04	1.060262
2	(1.0000006E-03,0.8354847,0.9962838)	3.177837	1.3856229E-05	1.4947158E-02
3	(1.0000008E-04,0.9105619,0.9153758)	2.255422	3.1786121E-07	4.1167527E-03
4	(1.0000011E-05,0.9393997,0.8971907)	1.996333	5.0372124E-09	6.3856736E-04
5	(1.0000012E-06,0.9446525,0.8947638)	1.954271	8.0987284E-11	7.4648298E-05
6	(1.0000015E-07,0.9452673,0.8944804)	1.949388	3.2067162E-12	1.7406819E-07

由表2可知，当经过6部迭代后，得到点$x_6^* = (0.9452673, 0.8944804)$目标函数值$f(x_6^*) = 1.949388$。注意到该问题的最优点为$(0.945, 0.894)$，最优值为1.94。

例3.

$$
\begin{aligned}
\text{(P)} \qquad \min \quad & f(x) = (x_1-2)^2 + (x_2-1)^2 \\
\text{s.t.} \quad & g_1(x) = x_1^2 - x_2 \le 0 \\
& g_2(x) = x_1 + x_2 - 2 \le 0
\end{aligned}
$$

定义初始点: $\bar{x}^0 = (1E-05, 0.5, 0.5)$; $s^0 = (0.2500100, 1.000010)$; $y^0 = 1.0$; $z^0 = (1.0, 1.0)$; $\sigma = 0.5$; $\varepsilon = 1E-07$.

表3 共迭代10步

迭代步骤	$\bar{x}^k$	$f(\bar{x}^k)$	$\gamma\mu^k$	$r(w^k)$
0	(9.9999997E-06,0.5000000,0.5000000)	2.500000	6.2501006E-04	2.374507
1	(1.0000002E-06,0.8978315,0.8010017)	1.254376	1.1968108E-04	0.4671720
2	(1.0000004E-07,0.9822894,0.9587578)	1.037436	1.9406780E-05	7.0330746E-02
3	(1.0000007E-08,0.9969561,0.9931979)	1.006143	3.0959800E-06	1.4664013E-02
4	(1.0000010E-09,0.9996504,0.9992228)	1.000700	3.5033440E-07	1.7573243E-03
5	(1.0000012E-10,0.9999644,0.9999210)	1.000071	3.5591224E-08	1.7975713E-04
6	(1.0000015E-11,0.9999964,0.9999921)	1.000007	3.5763579E-09	1.8019182E-05
7	(1.0000017E-12,0.9999996,0.9999992)	1.000001	3.5762929E-10	1.8137798E-06
8	(1.0000019E-13,0.9999999,0.9999999)	1.000000	5.9604724E-11	1.7206375E-07
9	(1.0000022E-14,1.000000,1.000000)	1.000000	6.6613388E-18	4.8666983E-08
10	(1.0000024E-15,1.000000,1.000000)	1.000000	6.2912643E-19	4.8666983E-08

由表3可知，当经过10步迭代后，得到点$x_{10}^* = (1.000000, 1.000000)$目标函数值$f(x_{10}^*) = 1.000000$。注意到该问题的最优解为$(1.0, 1.0)$，最优值为1。

第五章 结论与展望

在这篇硕士学位论文中，我们对于用内点算法求解约束非线性规划问题进行了研究和创新。在充分吸取前人优秀思想的基础上，针对内点算法在求解非线性规划问题时遇到的两个难题，分别给出了两个新的算法思想，设计了不同算法，并对他们进行了较为丰富的数值试验。从理论分析及数值结果来看，我们所提出的这两个算法都是有效、可行的。然而，这些算法还是都有着一些不足之处:

(1)对一些大规模的实际问题，往往会出现原问题的约束条件个数增加或是减少的情况，这时，原本的迭代数据就会显得毫无价值。重新的、反复的计算会增加计算成本和时间成本。很多研究者已经发现的这一问题，"warm start"思想也在近年来被逐步提出[57]，这也将成为今后对内点法的一个研究趋势。

(2)在算法的迭代中主要运用了牛顿方程，不过对一些大规模的非线性规划问题而言，当接近最优解时，两阶导数矩阵往往会出现“病态”。其实这一现象在无约束优化算法中也出现过。因此，有提出用BFGS方法来拟合迭代过程中所需要计算的两阶导数矩阵[58]，不过这类方法只适合一类具有凸性要求的函数，并不具有普遍性。

尽管如此，内点算法仍不失为求解约束非线性规划的一类有效途径，上面所提到的各种问题可以通过其他手段来改进与发展。

参考文献

[1] Khachiyan, L.G., *A Polynomial algorithm for linear programming*, Soviet Mathematics Doklady, (1979) 20:191-194.

[2] Karmarkar, N.K., *A new polynomial-time algorithm for linear programming*, Combinatorica, (1984) 4:373-395.

[3] Anstreicher, K. M., *A combined phase I-phase II scaled potential algorithm for linear programming*, Mathematical Programming, (1991) 52:429-439.

[4] Gonzaga, C. C, *Path following methods for linear programming*, SIAM Review, (1992) 34(2):167-227.

[5] Anstreicher, K. M., *A combined phase I-phase II projective algorithm for linear programming*, Mathematical Programming, (1989) 43:209-223.

[6] Anstreicher, K. M. and R.A. Bosch, *Long step in an $O(n^3L)$ algorithm for linear programming*, Mathematical Programming, (1992) 54:251-265.

[7] Dikin, I. I., *Iteration solution of problems of linear and quadratic programming*, Soviet Mathematics Doklady, (1967) 8:674-675.

[8] Barnes, E. R., *A variation on Karmarkar's algorithm of solving linear programming problems*, Mathematical Programming, (1986) 36:174-182.

[9] Adler, I., M. G. C. Resende, G. Veiga and N. Karmarkar, *An implementation of Karmarkar's algorithm for linear programming*, Mathematical Programming, (1989) 44:297-335.

[10] Adler, I. and R.D.C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Mathematical Programming, (1991) 50:29-51.

[11] Gill, P. E., W. Murray, M.A. Saunders, J.A. Tomlin and M.H. Wright, *On projected Newton barrier method for linear programming and an equivalence to Karmarkar's projective method*, Mathematical Programming, (1986) 36:183-209.

[12] Megiddo, N., *Pathways to the optimal set in linear programming*, In: N. Megiddo ed., Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, New York, (1989):131-158.

[13] Kojima, M., S. Mizuno and A. Yoshise, *A primal-dual interior point algorithm for linear programming*, In: N. Megiddo ed., Progress in Mathematical Programming: Interior Point and Related Methods, Springer Verlag, New York, (1989):29-47.

[14] Lustig, I. J., *Feasibility issues in a primal-dual interior point methods for linear programming*, Mathematical Programming, (1991) 49:145-162.

[15] Mehrotra, S., *On the implementation of a primal-dual interior point method*, SIAM Journal on Optimization, (1992) 2(4):575-601.

[16] Gonzaga, C. C., *Interior point algorithm for linear programming with inequality constraints*, Mathematical Programming, (1991) 52:209-225.

[17] Ye. Y., *An $O(n^3L)$ potential algorithm for linear programming*, Mathematical Programming, (1991) 50:239-258.

[18] Lustig, I. J., R. E. Marsten and D. F. Shanno, *On implementing Mehrotra's predictor-corrector interior point method for linear programming*, SIAM Journal on Optimization, (1992) 2:435-449.

[19] IBM Corporation, Optimization subroutine library guide and reference. Release 2, (1991).

[20] Dikin, I. I., *On the convergence of an iterative process*, Upravlyaemye Sistemi (in Russian), (1974) 12:54-60.

[21] Vanderrei, R. J., M. S. Meketon and B. A. Freedman, *On a modification of Karmarkar's linear programming algorithm*, Algorithmica, (1986) 1(4):395-408.

[22] Renegar, J., *A polynomial-time algorithm based on Newton's method for linear programming*, Mathematical Programming, (1988) 40:59-94.

[23] Kojima, M., S. Mizuno and A. Yoshise, *A polynomial-time algorithm for a class of linear complementarity problems*, Mathematical Programming, (1989) 44:1-26.

[24] Monteiro, R. D. C., and I. Adler, *Interior path following primal-dual algorithm: Part I: linear programming*, Mathematical Programming, (1989) 44:27-41.

[25] Monteiro, R. D. C., and I. Adler, *Interior path following primal-dual algorithm: Part II: linear programming*, Mathematical Programming, (1989) 44:43-66.

[26] McShane, K. A., C. L., Monma and D. F. Shanno, *An implementation of a primal-dual interior point method for linear programming*, ORSA Journal on Computing, (1989) 1:70-83.

[27] Zhang Y, *On the convergence of a class infeasible interior-point algorithms for the horizontal linear complementarity problem*, SIAM Journal on Optimization, (1994) 4:208-227.

[28] Gonzaga, C. C. and M. J. Todd, *An $O(L)$-iteration large-step primal-dual affine algorithm for linear programming*, SIAM Journal on Optimization, (1992) 2:349-359.

[29] Todd, M. J. and Y. Ye, *A centered projective algorithm for linear programming*, Mathematics of Operations Research, (1990) 15:508-529.

[30] Tananbe, K., *Centered Newton method for mathematical programming*, In: M. Iri, and K. Yajima eds., System Modeling and Optimization, Springer-Verlag, New York, (1988):197-226.

[31] Güler, O. and Y. Ye, *Convergence behavior of interior-point algorithms*, Mathematical Programming, (1993) 60:215-228.

[32] Kojima, M., S. Mizuno and A. Yoshise, *An $O(L)$ iteration potential reduction algorithm for linear complenentarity problems*, Mathematical Programming, (1991) 50:331-342.

[33] Ye. Y., M. J. Todd and S. Mizuno, *An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm*, Mathematics of Operations Research, (1994) 19:52-67.

[34] Anderson, E. D. and Ye, Y., *A computational study of the homogenous algorithm for large-scale convex optimization*, Working Paper, (1997).

[35] Anderson, E. D. and Ye, Y., *On a homogenous algorithm for monotone complementary problem*, Mathematical Programming, (1999) 84:375-399.

[36] Luo, Z. Q., Sturm, J. F. and Zhang, S, *Conic convex programming and self-dual embedding*, Optimization Methods and Software, (2000) 14:169-218.

[37] Y. Nesterov, A. Nemirovsky. Interior point polynomial methods in convex programming, Studies in Applied Mathematics, (1994) 13.

[38] Xu, X., Hung, P. F. and Ye, Y., *A simplified homogeneous self-dual linear programming algorithm and its implementation*, Annuals of Operations Research, (1996) 62:151-171.

[39] Shuzhong Zhang. *A new self-dual embedding method for convex programming*, Journal of Global Optimization, (2004) 29(4):479-496.

[40] J.T. W. Cheng and S. Zhang, *On implementation of a self-dual embedding method for convex programming*, Optimization Methods and Software, (2005) 00(00):1-29.

[41] Bo Yua, Qing Xub, Guochen Feng, *On the complexity of a combined homotopy interior method for convex programming*, Journal of Computational and Applied Mathematics, (2007) 200(1):32-46.

[42] Georg Stadler, *Path-following and augmented Lagrangian methods for contact problems in linear elasticity*, Journal of Computational and Applied Mathematics, (2007) 203(2):533-547.

[43] Arkadi Nemirovski, Levent Tuncel, *"Cone-free" primal-dual path-following and potential-reduction polynomial time interior-point methods*, Mathematical Programming, (2005) 102(2):261-294.

[44] Y. B. Zhao, D. Li, *A New Path-Following Algorithm for Nonlinear P^* Complementarity Problems*, Computational Optimization and Applications, (2006) 34(2):183-214.

[45] Renato D. C. Monteiro, Takashi Tsuchiya, *A strong bound on the integral of the central path curvature and its relationship with the iteration-complexity of primal-dual path-following LP algorithms*, Mathematical Programming, (2008) 115(1):105-149.

[46] Ai Wenbao, Zhang Kecun, *General central path and the largest step general central path following algorithm for linear programming*, Acta Mathematicae Applicatae Sinica (English Series), (2001) 17(3):296-303.

[47] M. Argáez, R. Tapia, L. Velázquez, *Numerical Comparisons of Path-Following Strategies for a Primal-Dual Interior-Point Method for Nonlinear Programming*, Journal of Optimization Theory and Applications, (2002) 114(2):255-272.

[48] Jorge Nocedal, Stephen J.Wright, *Numerical Optimization*, Springer, (1999).

[49] QIAN Zhong-gen, BAI Yan-qin, *Primal-Dual Interior-Point Algorithms with Dynamic Step-Size Based on Kernel Functions for Linear Programming*, Journal of Shanghai University(English Edition), (2005) 9(5):391-396.

[50] WANG Guo-qiang, BAI Yan-qin, *A class of polynomial primal-dual interior-point algorithms for semidefinite optimization*, Journal of Shanghai University(English Edition), (2006) 10(3):198-207.

[51] E.D. Andersen, C. Roos, T. Terlaky, *On implementing a primal-dual interior-point method for conic quadratic optimization*, Mathematical Programming, (2003) 95(2):249-277.

[52] Y. Q. Bai, M. El Ghami, C. Roos, *A comparative study of kernel functions for primal-dual interior-point algorithms in linear optimization*, Society for Industry and Applied Mathematics Journal on Optimization, (2004) 15(1):101-128.

[53] Luís N. Vicente, Stephen J. Wright, *Local convergence of a primal-dual method for degenerate nonlinear programming*, Computational Optimization and Applications, (2002) 22(3):311-328.

[54] Nicholas I. M. Gould, Dominique Orban, Annick Sartenaer, Philippe L. Toint, *Superlinear convergence of primal-dual interior point algorithms for nonlinear programming*, Society for Industry and Applied MathematicsJournal on Optimization, (2001) 11(4):974-1002.

[55] C. Grossmann; M. Zadlo, *A general class of penalty/barrier path-following Newton methods for nonlinear programming*, Optimization, (2005), 54(2):161-190.

[56] L. Chen, D. Goldfar, *Interior-point l_2-penalty methods for nonlinear programming with strong global convergence properties*, Mathematical Programming, (2006) 108(1):1-36.

[57] Alexander Engau, Miguel F. Anjos and Anthony Vannelli, *A primal-dual slack approach to warmstarting interior-point methods for linear programming*, Operations Research and Cyber-Infrastructure, Springer US, (2009):195-217.

[58] Paul Armand, Jean Charles Gilbert and Sophie Jan-Jegou, *A feasible BFGS interior point algorithm for solving convex minimization problems*, SIAM Journal on Optimization, (2000) 11(1):199-222.

作者攻读硕士学位期间已完成的论文

1. Wenting Lu,Yirong Yao, Liansheng Zhang. *A Penalized Primal-Dual Interior Point Approach for Constrained Nonlinear Programming*, Journal of Shanghai University (English Edition), (2009) 13(3):230-237.

2. Yirong Yao, Wenting Lu, Cunjia Zhang. *A New Long-Step Path Following Approach for Constrained Nonlinear Programming*, Submitted to Journal of Computational and Applied Mathematics.

3. 陈锐、陆文婷、姚奕荣、郑权. 盐业配送优化方案设计，应用数学与计算数学学报，(2008) 22(1):1-6.

4. Y.R. Yao, L.S. Zhang and W.T. Lu. *A one Parameter Filled Function Method for Global Minimization*, Submitted to Applied Mathematics and Computation.

致 谢

值此论文完成之际，我衷心感谢我的导师姚奕荣副教授三年以来对我的严格要求与悉心指导。

另外还要感谢数学系运筹学与控制论专业的各位老师对我的关心与帮助，同时还要把真挚的谢意送给在学习和生活上曾给予我帮助的师兄妹们。

最后，感谢我的父母及所有给过我支持和鼓励的人。