广西大学

硕士学位论文

约束Minimax问题的一个带积极集识别的简单序列二次约束二次规划算法

姓名：邱丽娟

申请学位级别：硕士

专业：运寿学与控制论

指导教师：简金宝

20110527

约束 Minimax 问题的一个带积极集识别的简单

序列二次约束二次规划算法

摘 要

本学位论文在光滑约束优化问题的简单序列二次约束二次规划算法和 minimax 问题积极集识别技术基础上，提出了一个求解带非线性约束的 minimax 问题的新算法. 与先前工作不同的是，在每次迭代过程中，可行下降方向是通过解一个带简单二次约束的子问题获得. 该子问题是由一个凸的二次目标函数和不带二阶导数的简单二次不等式约束条件组成. 用于修正主方向的高阶修正方向 (避免 Maratos 效应) 是通过计算一个线性方程组产生. 该线性方程组仅包括积极识别集中的约束函数及其梯度，从而进一步减少高阶修正方向计算的规模和计算量. 该文的算法在弱 Mangasarian-Fromovitz 约束规格 (MFCQ) 条件下具有全局收敛性. 在上层严格互补等较温和的条件下具有超线性收敛性. 最后，数值试验表明，对于所测试的问题，本文的算法是有效的. 下面简述本文的主要内容.

第 1 章为绪论，除研究背景外，重点介绍和回顾与本文有关的可行方向法的思想和积极集识别技术，由此引出本文的主要思想.

第 2 章，给出本文的算法步骤和算法所具有的一些性质.

第 3 章，在弱 MFCQ 条件下证明算法具有全局收敛性.

第 4 章，在包含上层严格互补等较温和条件下证明算法具有强收敛性和超线性收敛性.

第 5 章，对算法进行一些初步的数值试验，以说明算法的有效性.

关键词：约束 minimax 问题　序列二次约束二次规划　可行方向法

积极集识别　全局和超线性收敛性

A SIMPLE SEQUENTIAL QUADRATICALLY CONSTRAINED QUADRATIC PROGRAMMING ALGORITHM WITH ACTIVE IDENTIFICATION SETS FOR CONSTRAINED MINIMAX PROBLEMS

ABSTRACT

In this thesis, the nonlinear minimax problems with inequality constraints are discussed. Based on the ideas of simple sequential quadratically constrained quadratic programming algorithm for smooth constrained optimization and active identification sets for minimax problems, we propose an algorithm for solving the discussed problems. Unlike the previous work, at each iteration, a feasible direction of descent called main direction is obtained by solving only one subprogram which is composed of a convex quadratically objective function and simple quadratic inequality constraints without the second derivatives of the constrained functions. Then a high-order correction direction (to avoid the Maratos effect) is yielded by solving a system of linear equations (SLE) which includes only the constraints and their gradients corresponding to the active identification set, thus, the scale and the conputation cost of the high-order correction direction are deceased. The proposed algorithm possesses global convergence under weak Mangasarian-Fromovitz constrained qualification (MFCQ) and superlinear convergence under mild conditions with the upper-level strict complementarity. At last, the numerical results show that the proposed algorithm is promising for the tested problems. The main contents of this paper are given as follows:

Chapter 1: the research background, the related ideas and algo-

rithms of feasible direction algorithm as well as active set identifying technique are recalled, then we introduced the idea of this paper.

Chapter 2: the details of our algorithm and its basic properties are given.

Chapter 3: global convergence of the proposed algorithm is proved under weak MFCQ.

Chapter 4: the proposed algorithm possesses superlinear convergence under mild conditions including the upper-level strict complementarity.

Chapter 5: some elementary numerical experiments are tested, and these show that the algorithm is promising.

KEY WORDS: constrained minimax problems; sequential quadratically constrained quadratic programming; methods of feasible directions; active set identification; global and superlinear convergence

第 1 章 绪论

1.1 研究背景

最优化是运筹学方面的一个非常重要的分支，是计算数学和运筹学方面乃至工程科学方面的一门交叉学科. 最优化有着非常重要的理论研究意义和实际应用意义，在工程设计，经济方面，交通运输，国防与优化控制等领域的应用越来越重要. 目前，最优化的方法已经广泛的应用于电气工程，经济，卫生，管理，科技，社会，航空航天等众多方面，为这些方面的研究作出了重要的贡献.

随着时代的发展，数学背景的加强，科学技术的进步，最优化的研究越来越成熟，越来越复杂. 尤其是对于非线性光滑约束优化问题的研究更加重要，国内外众多学者对这方面的理论研究出了众多的方法来解决. 如序列二次规划算法 (SQP), 序列二次约束二次规划算法 (SQCQP), 简单序列二次约束二次规划算法 (SSQCQP), 序列线性方程组算法 (SSLE), 原始对偶内点算法等等. 因此对最优化方法的研究有着重要的实际背景和理论意义.

1.2 国内外研究现状

近年来，越来越多的学者对非线性光滑约束优化问题进行研究，得到了众多优秀的理论成果. 本学位论文首先简单介绍一下可行方向法 (MFD) 的研究的发展过程. MFD 是求解光滑约束优化

$$\begin{aligned} \min\ & f(x) \\ \text{s.t.}\ & g_j(x) \le 0, \quad \forall\, j \in J \end{aligned} \tag{1.1}$$

的一类重要方法. 其首先由 Zoutendijk[1] 提出. Zoutendijk 的 MFD 的主要优点是其可行下降方向通过求解一个线性规划问题得到，但是 Wolfe[2] 用一个反例证明出了 Zoutendijk 的 MFD 不具有全局收敛性. Topiks 和 Veinott[3] 改进了 Zoutendijk 的 MFD, 使得算法产生的点列收敛于 Fritz-John 点. Pironneau 和 Polak[4] 通过增加二次项 $\frac{1}{2}\|d\|^2$ 进一步改进文 [3] 的算法，使得算法具有线性收敛性. 为提高 [4] 中算法

的收敛速度, Cawood, Kostreva, Chen[5, 6] 提出了一种模松弛序列二次规划 (SQP) 算法. 在文 [6] 中, 对于 (1.1) 的当前可行迭代点 x^k, 产生搜索方向的子问题为:

$$
\begin{aligned}
&\min_{(z,d)} z + \tfrac{1}{2}d^T B_k d \\
&\text{s.t. } \nabla f(x^k)^T d - \gamma_0 z \le 0, \\
&\qquad g_j(x^k) + \nabla g_j(x^k)^T d - \gamma_j z \le 0, \quad \forall\, j \in J,
\end{aligned} \tag{1.2}
$$

其中 γ_j $(j \in \{0\} \cup J)$ 都是正的常数 , B_k 是对称正定阵.

由于文 [6] 中的算法要求初始点可行, 简金宝、柯晓艳及郑海艳等 [7] 在上面子问题和强次可行方向法基础上 (强次可行方向法思想见文献 [10] 的第二章), 提出了针对 (1.1) 的模松弛 SQP 的超线性算法. 对于当前迭代点 $x^k \in R^n$, 文 [7] 产生搜索方向的子问题为:

$$
\begin{aligned}
&\min_{(z,\ d)} \gamma_0 z + \tfrac{1}{2}d^T B_k d \\
&\text{s.t. } \nabla f(x^k)^T d - \gamma_0 z - \gamma\varphi(x^k)^\sigma \le 0, \\
&\qquad g_j(x^k) + \nabla g_j(x^k)^T d - \gamma_j \eta_k z \le 0, \quad j \in I_k^-, \\
&\qquad g_j(x^k) + \nabla g_j(x^k)^T d - \gamma_j \eta_k z - \varphi(x^k) \le 0, \quad j \in I_k^+,
\end{aligned} \tag{1.3}
$$

其中 γ, σ 和 γ_j 都是正的常数, η_k 是与 x^k 有关的正参数, $I_k^- = \{j \in I:\ -\varepsilon \le g_j(x^k) \le 0\}$, $I_k^+ = \{j \in J:\ g_j(x^k) > 0,\ -\varepsilon \le g_j(x^k) - \varphi(x^k) \le 0\}$, $\varphi(x^k) = \max\{0;\ g_j(x^k),\ j \in I\}$. 此外, 为克服 Maratos 效应 [8], 通过求解线性方程组:

$$
\begin{pmatrix} B_k & N_k \\ N_k^T & -D^k \end{pmatrix} \begin{pmatrix} d \\ h \end{pmatrix} = \begin{pmatrix} 0 \\ -\max\{\|d^k\|^\tau,\ \mid \eta_k^\nu z_k \mid \|d^k\|\} e_{I_k} - \widetilde{g}^k \end{pmatrix}
$$

产生高阶修正方向. 其中矩阵 D^k 是一个合适的对角矩阵, $N_k = (\nabla f_j(x^k),\ j \in I_k = I_k^- \cup I_k^+)$.

简金宝、胡清洁及唐春明等 [9] (亦可参见 [10] 中 12.4 节) 通过引进一个简单的二次约束, 提出了一个针对 (1.1) 的简单序列二次约束二次规划可行方向法, 其产生主

搜索方向子问题为：

$$
\begin{aligned}
\min_{(z,\ d)}\ & \gamma z + \frac{1}{2}d^T B_k d \\
\text{s.t.}\ & \nabla f(x^k)^T d - \gamma z \le 0, \\
& g_j(x^k) + \nabla g_j(x^k)^T d - \gamma_j \sigma_k z + \varepsilon_k z^2 \le 0, \quad j \in I_k,
\end{aligned} \tag{1.4}
$$

其中近似积极集 I_k 是由一个转轴运算产生，参数 γ, γ_j ($j \in I_k$) 是正的常数， σ_k 和 ε_k 是非负参数且不同时为零. 克服 Maratos 效应的高阶修正方向由显式公式

$$
\tilde{d}^k = -A_k(A_k^T A_k)^{-1}(\psi_k e_k + \tilde{f}^k) \tag{1.5}
$$

计算. 其中 $A_k = (\nabla g_j(x^k),\ j \in I_k)$, 式 (1.5) 中其它量的定义可参考文 [9, 10]. 在文 [9, 10] 中，通篇需要线性无关约束规格 (LICQ), 且运用转轴运算产生近似积极集，对于某些算例，转轴运算的计算量较大且不稳定.

Minimax 问题是一类重要的有着广泛应用背景的非光滑优化问题，引起了许多学者对 minimax 问题的研究兴趣 [11]-[18]. 2007 年, 简金宝、全然及张雪露 [15] 对 minimax 问题 (1.8) 提出了带广义单调线搜索的 SQP 算法. 在文献 [15] 中， 产生主方向的子问题为：

$$
\begin{aligned}
\min_{(z,\ d)}\ & z + \frac{1}{2}d^T B_k d \\
\text{s.t.}\ & f_i(x^k) + \nabla f_i(x^k)^T d - F(x^k) - z \le 0, \quad i \in I^k_{\varepsilon_k}, \\
& g_j(x^k) + \nabla g_j(x^k)^T d - c_k z \le 0, \quad j \in J^k_{\theta_k},
\end{aligned} \tag{1.6}
$$

其中 $I^k_{\varepsilon_k} = \{i \in I: f_i(x^k) - F(x^k) \ge -\varepsilon_k\}$, $J^k_{\theta_k} = \{j \in J: g_j(x^k) \ge -\theta_k\}$, $\varepsilon_k, \theta_k, c_k > 0$, 高阶修正方向由另一个 QP 子问题产生. 该算法在弱 MFCQ 条件下具有全局收敛性，在 LICQ 以及严格互补等条件下具有超线性收敛性.

2011 年，李杰 [18] 针对 minimax 问题 (1.8) 提出了具有超线性收敛的模松弛拟强次可行方向法. 该算法通过求解一个二次规划子问题

$$
\begin{aligned}
\min_{(z,\ d)}\ & \gamma_0 z + \frac{1}{2}d^T B_k d \\
\text{s.t.}\ & f_j(x^k) - F(x^k) + g_j(x^k)^T d \le \gamma_0 z + \gamma \varphi_k^\sigma,\ j \in J_k, \\
& \bar{f}_j(x^k) + g_j(x^k)^T d \le \gamma_j c_k z, \quad j \in I_k
\end{aligned} \tag{1.7}
$$

得到主搜索方向，其中 $J_k=\{j\in J: -\varepsilon\le f_j(x)-F(x)\le 0\}$, $I_k=\{j\in I: -\varepsilon\le \bar{f}_j(x)\le 0\}$. 式 (1.7) 中的其他定义见文献 [18]. 此文献通过引入一种新的非单调的曲线搜索，使得算法总能保证有限步迭代后迭代点恒落入可行域. 该算法在弱 MFCQ 且无严格互补等较温和的假设条件下，具有全局收敛和超线性收敛性.

因为积极识别集技术可以最大限度地降低子问题的规模和计算量，所以该技术是设计约束优化算法的重要技术，吸引了众多学者的研究，例如文献 [19] – [21] 及 [10] 中的 1.5 节. 韩道兰、简金宝及李杰 [21] 建立了约束 minimax 问题积极集的精确识别技术.

1.3 论文的主要研究内容及结构安排

本学位论文研究求解如下带不等式约束的非线性 minimax 问题：

$$(\mathrm{P})\quad \begin{array}{ll} \min & F(x) \\ \text{s.t.} & g_j(x)\le 0,\quad \forall\, j\in J, \end{array} \tag{1.8}$$

其中 $F(x)=\max\{f_i(x),\ i\in I\}$, $I=\{1,2,\ \ldots,m\}$, $J=\{1,2,\ \ldots,m'\}$, 函数 $f_i\ (i\in I)$ 及 $g_j\ (j\in J)$: $R^n\rightarrow R$ 都是一阶连续可微的. 如果存在乘子向量 $\alpha_I=(\alpha_i,\ i\in I)$, $\beta_J=(\beta_j,\ j\in J)$ 满足

$$\begin{cases} \sum\limits_{i\in I}\alpha_i\nabla f_i(x)+\sum\limits_{j\in J}\beta_j\nabla g_j(x)=0;\quad \sum\limits_{i\in I}\alpha_i=1; \\ 0\le\alpha_i\perp(f_i(x)-F(x))\le 0,\ \forall\, i\in I;\quad 0\le\beta_j\perp g_j(x)\le 0,\ \forall\, j\in J, \end{cases} \tag{1.9}$$

则可行点 x 称为 minimax 问题 (1.8) 的稳定点 (亦称为 KKT 点)[22]–[24], 其中 $\omega\perp y$ 表示 $\omega^T y=0$.

本文在简单二次约束二次规划和 minimax 问题积极识别集技术的基础上，提出了一个新的解约束 minimax 问题 (1.8) 的快速算法. 该算法具有如下性质：

(1) 初始点是可行的，所以是一个可行方向法；

(2) 主方向通过求解一个带简单二次约束的子问题获得，这个子问题是由一个规模较小的凸的二次目标函数和不带二阶导数的简单二次不等式的约束条件组成；

(3) 高阶修正方向是通过解一个线性等式组产生避免 Maratos 效应获得，该方程组只包括积极识别集中的约束及其梯度；

(4) 在弱 MFCQ 条件下具有全局收敛性，在包含上层严格互补等较温和的条件下具有超线性收敛性.

根据本学位论文的研究内容，现将本论文的主要结构介绍如下：

第 1 章 绪论：主要介绍最优化理论的研究背景和国内外研究现状，并且简单叙述本学位论文的研究内容和论文的逻辑结构.

第 2 章 算法：给出了算法的主搜索方向和高阶修正方向计算子问题，对本文的算法进行描述，并给出了算法所具有的一些性质.

第 3 章 全局收敛性：证明了算法在弱 MFCQ 的条件下具有全局收敛性.

第 4 章 强收敛和超线性收敛性：证明算法在有上层严格互补的较温和的条件下具有强收敛和超线性收敛性.

第 5 章 数值试验：通过一些初步的数值试验，表明了本文的算法是有效的.

第 2 章 算 法

2.1 子问题的构造

为简化分析，记问题 (1.8) 的可行集 X, 积极集 $I(x)$ 和 $J(x)$ 分别为

$$X = \{x \in R^n \colon g_j(x) \le 0,\ \forall\, j \in J\},$$

$$I(x) = \{i \in I \colon f_i(x) = F(x)\},\ \ J(x) = \{j \in J : g_j(x) = 0\}.$$

设可行集 X 非空，且下面的两个基本假设成立.

假设 2.1 函数 f_i $(i \in I)$ 和 g_j $(j \in J)$ 在可行集 X 中至少是一阶连续可微的.

假设 2.2 假设弱 MFCQ 对任意的 $x \in X$ 成立，即存在向量 $d \in R^n$ 使得 $\nabla g_j(x)^T d < 0,\ \forall\, j \in J(x)$ 成立.

注 2.3 因为假设中不包括目标中积极函数梯度 $\nabla f_i(x)$ $(i \in I(x))$, 所以称之为弱 MFCQ.

为便于后面的分析，下面给出问题 (1.8) 稳定点的等价条件.

命题 2.4 设在可行点 x 处，弱 MFCQ 条件成立，则稳定点系统 (1.9) 等价于：

$$\begin{cases} \sum\limits_{i\in I}\alpha_i\nabla f_i(x) + \sum\limits_{j\in J}\beta_j\nabla g_j(x) = 0;\ \ \sum\limits_{i\in I}\alpha_i + \sum\limits_{j\in J}\beta_j = 1, \\ 0 \le \alpha_i \perp (f_i(x) - F(x)) \le 0,\ \forall\, i \in I;\ 0 \le \beta_j \perp g_j(x) \le 0,\ \forall\, j \in J, \end{cases} \tag{2.1}$$

进而等价于

$$\begin{cases} \sum\limits_{i\in I(x)}\alpha_i\begin{pmatrix}\nabla f_i(x)\\ 1\end{pmatrix} + \sum\limits_{j\in J(x)}\beta_j\begin{pmatrix}\nabla g_j(x)\\ 1\end{pmatrix} = \begin{pmatrix}0\\ 1\end{pmatrix}, \\ \alpha_i \ge 0,\ \forall\, i \in I(x);\ \beta_j \ge 0,\ \forall\, j \in J(x). \end{cases} \tag{2.2}$$

基于稳定点系统 (1.9), 定义稳定点识别函数:

$$\phi(x,\alpha,\beta)=\begin{pmatrix}\nabla_x L(x,\alpha,\beta)\\ \min\{F(x)e-f_I(x),\ \alpha\}\\ \min\{-g_J(x),\ \beta\}\end{pmatrix},\quad \varrho(x,\alpha,\beta)=\|\phi(x,\alpha,\beta)\|^{\delta},\tag{2.3}$$

其中 $e=(1,1,\cdots,1)^T\in R^m$, $\delta\in(0,1)$, 易知 $\varrho(x,\alpha,\beta)$ 是 (1.8) 的最优识别函数, 即 (x,α,β) 是 (1.8) 的稳定点对当且仅当 $\varrho(x,\alpha,\beta)=0$. 下面引入 [21] 的积极约束识别集:

$$\begin{cases}I(x,\alpha,\beta)=\{i\in I \mid f_i(x)-F(x)+\varrho(x,\alpha,\beta)\geq 0\},\\ J(x,\alpha,\beta)=\{j\in J \mid g_j(x)+\varrho(x,\alpha,\beta)\geq 0\}.\end{cases}\tag{2.4}$$

在文 [21] 中已证明, 当 (x^k,α^k,β^k) 充分接近于稳定点对 (x^*,α^*,β^*), 且弱 MFCQ 和强二阶充分条件在稳定点对 (x^*,α^*,β^*) 均成立时, $I(x^k,\alpha^k,\beta^k)$ 和 $J(x^k,\alpha^k,\beta^k)$ 分别是积极集 $I(x^*)$, $J(x^*)$ 的精确识别集. 对于当前迭代点 x^k, 本文不再增加新的计算成本产生与之相应的近似乘子 α^k 及 β^k, 而是采用上一次迭代的数值信息产生 α^k 及 β^k.

对于当前迭代点 $x^k\in X$, 为产生搜索方向, 考虑以下带简单二次约束的子问题:

$$\begin{aligned}\min_{(z,d)}\quad & \gamma_0 z+\tfrac{1}{2}d^T B_k d\\ \text{s.t.}\quad & f_i(x^k)+\nabla f_i(x^k)^T d-F(x^k)\leq\gamma_0 z,\quad i\in I_k\triangleq I(x^k,\alpha^k,\beta^k),\\ & g_j(x^k)+\nabla g_j(x^k)^T d\leq\gamma_j\theta_k z-\varepsilon_k z^2,\quad j\in J_k\triangleq J(x^k,\alpha^k,\beta^k),\end{aligned}\tag{2.5}$$

其中参数 $\gamma_0,\gamma_j>0$, θ_k 和 ε_k 都是非负参数, 且满足 $(\theta_k,\varepsilon_k)\neq 0$, 而对称正定矩阵 B_k 是 (1.8) 的拉格朗日 Hesse 阵的近似阵. 设 (z_k,d^k) 是 (2.5) 的最优解, 记 (2.5) 的积极约束集为:

$$L_k=\{i\in I_k\colon f_i(x^k)+\nabla f_i(x^k)^T d^k-F(x^k)=\gamma_0 z_k\},\tag{2.6}$$

$$L_k'=\{j\in J_k\colon g_j(x^k)+\nabla g_j(x^k)^T d^k=\gamma_j\theta_k z_k-\varepsilon_k z_k^2\}.\tag{2.7}$$

子问题 (2.5) 的 KKT 条件为：

$$\begin{cases} \begin{pmatrix} \gamma_0 \\ B_k d \end{pmatrix} + \sum\limits_{i\in I_k} u_i^k \begin{pmatrix} -\gamma_0 \\ \nabla f_i(x^k) \end{pmatrix} + \sum\limits_{j\in J_k} v_j^k \begin{pmatrix} 2\varepsilon_k z - \gamma_j\theta_k \\ \nabla g_j(x^k) \end{pmatrix} = 0; \\ 0 \le u_i \perp (f_i(x^k) + \nabla f_i(x^k)^T d - F(x^k) - \gamma_0 z) \le 0,\ \forall\, i \in I_k; \\ 0 \le v_j \perp (g_j(x^k) + \nabla g_j(x^k)^T d - \gamma_j\theta_k z + \varepsilon_k z^2) \le 0,\ \forall\, j \in J_k. \end{cases} \tag{2.8}$$

下面的引理给出了子问题 (2.5) 的一些重要性质.

引理 2.5 设假设 2.1, 2.2 成立，B_k 对称正定，$\gamma_j\ (j \in \{0\} \cup J_k)$ 都是正参数，则

(i) 子问题 (2.5) 有唯一最优解，记之为 (z_k, d^k)；

(ii) $\gamma_0 z_k + \frac{1}{2}(d^k)^T B_k d^k \le 0,\ z_k \le 0$，且 $z_k = 0 \Longleftrightarrow d^k = 0$；

(iii) $d^k = 0 \Longleftrightarrow x^k$ 是 (1.8) 的稳定点；

(iv) (z_k, d^k) 是 (2.5) 的最优解当且仅当它是 (2.5) 的 KKT 点；

(v) 如果 $d^k \ne 0$，则 $z_k < 0$，进而 d^k 是 (1.8) 在点 x^k 的可行下降方向.

证明：结论 (i), (ii) 的证明分别类似于文 [9] 中的引理 2.2 和 2.3(i, ii) 的证明，从略.

(iii) 假设 $d^k = 0$，由结论 (ii) 可知 $z_k = 0$. 再设 $x \triangleq x^k$ 不是 (1.8) 的稳定点，故由命题 2.4 知道系统 (2.2) 无解，进而由 Farkas 择一性定理 ([10], 定理 1.1.13) 知道，系统

$$\begin{cases} \begin{pmatrix} \nabla f_i(x^k) \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ h \end{pmatrix} \le 0,\ \forall\, i \in I(x^k),\quad \begin{pmatrix} \nabla g_j(x^k) \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ h \end{pmatrix} \le 0,\ \forall\, j \in J(x^k), \\ \begin{pmatrix} 0 \\ 1 \end{pmatrix}^T \begin{pmatrix} d \\ h \end{pmatrix} > 0 \end{cases}$$

有解，其中参数 $d \in R^n$，$h \in R^1$. 亦即系统

$$h > 0,\ \nabla f_i(x^k)^T d \le -h,\ \forall\, i \in I(x^k),\ \nabla g_j(x^k)^T d \le -h,\ \forall\, j \in J(x^k) \tag{2.9}$$

有解. 假设 (h', d') 是 (2.9) 的解, 由 (2.5) 易知, 当 $t>0$ 充分小时 , $(z, d) \triangleq (0, td')$ 是 (2.5) 的严格内点, 再结合子问题 (2.5) 的凸性和连续性, 可知 (2.5) 的 Slater 约束规格 (见文 [10] 中 1.3 节) 成立. 因此最优解 $(z_k, d^k)=(0,0)$ 是 (2.5) 的 KKT 点. 由 KKT 条件 (2.8) 有

$$\begin{cases} \sum\limits_{i\in I_k} u_i^k \nabla f_i(x^k) + \sum\limits_{j\in J_k} v_j^k \nabla g_j(x^k) = 0; \\ \sum\limits_{i\in I_k} u_i^k \gamma_0 + \sum\limits_{j\in J_k} v_j^k \gamma_j \theta_k = \gamma_0; \\ 0 \le u_i^k \perp (f_i(x^k) - F(x^k)) \le 0, \ \forall\, i \in I_k; \\ 0 \le v_j^k \perp g_j(x^k) \le 0, \ \forall\, j \in J_k. \end{cases} \tag{2.10}$$

由假设 2.2 及 (2.10) 的第一, 二式得到 $\sum\limits_{i\in I_k} u_i^k > 0$. 因此 (2.10) 可表示为

$$\begin{cases} \sum\limits_{i\in I} \bar{u}_i^k \nabla f_i(x^k) + \sum\limits_{j\in J} \bar{v}_j^k \nabla g_j(x^k) = 0; \ \sum\limits_{i\in I} \bar{u}_i^k = 1; \\ 0 \le \bar{u}_i^k \perp (f_i(x^k) - F(x^k)) \le 0, \ \forall\, i \in I; \\ 0 \le \bar{v}_j^k \perp g_j(x^k) \le 0, \ \forall\, j \in J, \end{cases} \tag{2.11}$$

其中,

$$\bar{u}_{I_k}^k = u_{I_k}^k \Big/ \sum_{i\in I_k} u_i^k, \quad \bar{v}_{J_k}^k = v_{J_k}^k \Big/ \sum_{i\in I_k} u_i^k, \quad \bar{u}_{I\setminus I_k}^k = 0, \quad \bar{v}_{J\setminus J_k}^k = 0. \tag{2.12}$$

因此 $(x^k, \bar{u}^k, \bar{v}^k)$ 是 (1.8) 的稳定点对, 与假设矛盾.

反之, 如果 $x \triangleq x^k$ 是 (1.8) 的稳定点, 由命题 2.4 和 Farkas 择一性定理可知系统 (2.9) 无解. 因为 $z_k \le 0$, 若 $z_k < 0$, 则 d^k 再加上一个合适的 s_k 可组成系统 (2.9) 的一个解, 与系统 (2.9) 无解矛盾. 所以 $z_k = 0$, 进而由 (ii) 得 $d^k = 0$.

(iv) 假如 (z_k, d^k) 是 (2.5) 的最优解, 分两种情况证明 (z_k, d^k) 是 (2.5) 的 KKT 点.

情况 1. 如果 $d^k = 0$, 由 (ii), (iii) 可知 $z_k = 0$ 且 x^k 是 (1.1) 的稳定点. 结合 $(z_k, d^k) = (0, 0)$ 和 KKT 条件 (2.8), 易知 (z_k, d^k) 是 (2.5) 的 KKT 点.

情况 2. 如果 $d^k \ne 0$, 由 (ii), (iii) 和 (2.5) 可知 $z_k < 0$, 且

$$f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) \le \gamma_0 z_k < 0, \ i \in I_k,$$

$$g_j(x^k) + \nabla g_j(x^k)^T d^k \le \gamma_j \theta_k z_k - \varepsilon_k z_k^2 < 0, \ j \in J_k.$$

因此，$(z, d) \triangleq (0, d^k)$ 是 (2.5) 的一个严格内点，结合子问题 (2.5) 的连续性和凸性可知 (2.5) 的 Slater 约束规格成立. 因此最优解 (z_k, d^k) 是 (2.5) 的 KKT 点.

反之，如果 (z_k, d^k) 是 (2.5) 的 KKT 点，由于 (2.5) 是凸规划，因此它是最优解.

(v) 如果 $d^k \neq 0$, 由 (ii) 可知 $z_k < 0$. 注意到 (2.5) 的约束条件和 $\gamma_0 > 0$, 有

$$\nabla f_i(x^k)^T d \le \gamma_0 z_k + F(x^k) - f_i(x^k) = \gamma_0 z_k < 0, \ i \in I(x^k) \subseteq I_k,$$

$$\nabla g_j(x^k)^T d \le \gamma_j \theta_k z_k - \varepsilon_k z_k^2 < 0, \ \forall\, j \in J(x^k) \subseteq J_k.$$

故 d^k 是 (1.8) 在点 x^k 的可行方向，且 $F(x)$ 在点 x^k 沿着方向 d^k 的方向导数满足

$$F'(x^k; d^k) = \lim_{t \to 0^+} \frac{F(x^k + td^k) - F(x^k)}{t} = \max\{\nabla f_i(x^k)^T d, \ i \in I(x^k)\} \le \gamma_0 z_k < 0. \tag{2.13}$$

这意味着 d^k 是 $F(x)$ 在点 x^k 的下降方向.

2.2 高阶修正方向的构造

虽然子问题 (2.5) 可以产生可行下降的搜索方向，但仍不能克服 Maratos 效应. 因此必须采取一个合适的技术修正 d^k, 产生一个高阶修正方向. 本文在积极集识别技术的基础上，引进如下线性方程组：

$$V_k \begin{pmatrix} d \\ s \end{pmatrix} = \begin{pmatrix} 0 \\ -\max\{\|d^k\|^\tau, \ -\bar{\gamma}\theta_k^\nu z_k \|d^k\| + \varepsilon_k z_k^2 \|d^k\|^\nu\}\varpi_k - \tilde{h}^k \end{pmatrix}, \tag{2.14}$$

其中

$$V_k \triangleq \begin{pmatrix} B_k & R_k \\ R_k^T & 0 \end{pmatrix}, \ R_k = (\nabla f_i(x^k) - \nabla f_{i_k}(x^k), \ i \in I_k \backslash \{i_k\}; \ \nabla g_j(x^k), \ j \in J_k), \tag{2.15}$$

$$\begin{cases} \tilde{h}^k = \begin{pmatrix} \tilde{f}^k \\ \tilde{g}^k \end{pmatrix}, \quad \tilde{f}^k = (\tilde{f}_i^{\,k}, \ i \in I_k \backslash \{i_k\}), \ \tilde{g}^k = (\tilde{g}_j^{\,k}, \ j \in J_k), \\ \tilde{f}_i^{\,k} = f_i(x^k + d^k) - f_i(x^k) - \nabla f_i(x^k)^T d^k - \Delta_k, \ i \in I_k \backslash \{i_k\}, \\ \Delta_k = f_{i_k}(x^k + d^k) - f_{i_k}(x^k) - \nabla f_{i_k}(x^k)^T d^k, \\ \tilde{g}_j^{\,k} = g_j(x^k + d^k) - g_j(x^k) - \nabla g_j(x^k)^T d^k + \gamma_j \theta_k z_k - \varepsilon_k z_k^2, \ j \in J_k, \end{cases} \tag{2.16}$$

上式中 $i_k \in I(x^k)$, $\varpi_k = (1, 1, \ldots, 1)^T \in R^{|I_k|+|J_k|-1}$, $\tau \in (2, 3)$, $\bar{\gamma} = \max\{\gamma_j, \ j \in J\}$.

2.3 算法描述

基于以上分析，下面给出具体的算法步骤.

算法 2.6

步骤 0. 选取参数 $\gamma_0, \gamma_j > 0$, $\tau \in (2, 3)$, $\alpha \in (0, \frac{1}{2})$, $\beta \in (0, 1)$, $\nu \in (0, 1)$, $\delta \in (0, 1)$. 选择初始对称正定矩阵 B_0, 初始可行点 $x^0 \in X$, 令 $(\alpha^0, \beta^0) = (1, 1, \ldots, 1)^T \in R^{m+m'}$. 令 $k: = 0$.

步骤 1. 通过 (2.3) 计算 $\varrho(x^k, \alpha^k, \beta^k)$ 和通过 (2.4) 计算积极识别集 $I_k = I(x^k, \alpha^k, \beta^k)$ 及 $J_k = J(x^k, \alpha^k, \beta^k)$. 如果 $\varrho(x^k, \alpha^k, \beta^k) = 0$, 则终止， x^k 是问题 (1.8) 的稳定点.

步骤 2. 求解子问题 (2.5) 得到最优解 (z_k, d^k) 和对应的 KKT 乘子向量 $u_{I_k}^k$ 和 $v_{J_k}^k$. 令 $u^k = (u_{I_k}^k, 0_{I \backslash I_k^k})$, $v^k = (v_{J_k}^k, 0_{J \backslash J_k^k})$. 如果 $d^k = 0$, 则 x^k 是 (1.8) 的稳定点，终止.

步骤 3. 通过解线性方程组 (2.14) 计算高阶修正方向 $(\tilde{d}^k, s^k)$. 如果 (2.14) 无解或者 $\|\tilde{d}^k\| > \|d^k\|$, 则令 $\tilde{d}^k = 0$.

步骤 4. 计算数列 $\{1, \beta, \beta^2, \ldots, \}$ 中满足

$$F(x^k + td^k + t^2\tilde{d}^k) \leq F(x^k) - \alpha t (d^k)^T B_k d^k, \tag{2.17}$$

$$g_j(x^k + td^k + t^2\tilde{d}^k) \leq 0, \ \forall j \in J \tag{2.18}$$

的最大值 (步长) t_k.

步骤 5. 令 $x^{k+1}=x^k+t_kd^k+t_k^2\tilde{d}^k$, 按 $\alpha^{k+1}=u^k$, $\beta^{k+1}=v^k$ 更新乘子向量. 计算新的对称正定阵 B_{k+1}, 并令 $k:=k+1$, 返回步骤 1.

注 2.7 由于 d^k 是 (1.8) 的可行方向, 且

$$F'(x^k;d^k)\le\gamma_0 z_k\le-\frac{1}{2}(d^k)^TB_kd^k<-\alpha(d^k)^TB_kd^k,$$

因此步骤 4 中 Armijo 线搜索可通过有限次试算后完成. 此外, 不等式 (2.17) 等价于

$$f_i(x^k+td^k+t^2\tilde{d}^k)\le F(x^k)-\alpha t(d^k)^TB_kd^k,\ \forall\, i\in I.$$

2.4 本章小结

在本章中给出了主搜索方向和避免 Maratos 效应的高阶修正方向的构造, 并且给出了本学位论文中的算法的具体描述.

第 3 章 全局收敛性

3.1 全局收敛性证明准备

如果算法 2.6 有限步终止于点 x^k, 由引理 2.5 和步骤 2 知, 当前迭代点 x^k 是 (1.8) 的稳定点. 本节的主要目的是分析论证算法 2.6 具有全局收敛性, 即算法 2.6 产生的迭代点列 $\{x^k\}$ 的任意一个极限点都是问题 (1.8) 的稳定点. 为此需要如下基本假设.

假设 3.1 (i) 算法 2.6 产生的迭代点列 $\{x^k\}$ 有界且矩阵序列 $\{B_k\}$ 一致正定, 即存在两个正的常数 a 和 b 满足

$$a\|d\|^2 \le d^T B_k d \le b\|d\|^2, \ \forall\ d \in R^n, \ \forall\ k. \tag{3.1}$$

(ii) 参数序列 $\{\theta_k\}$, $\{\varepsilon_k\}$ 有界, 且当 k 充分大时, $\theta_k + \varepsilon_k \ge \underline{c}$, 其中 $\underline{c}$ 为正的常数.

引理 3.2 设假设 2.1, 2.2 及 3.1 均成立, 则

(i) 序列 $\{d^k\}$, $\{z_k\}$ 和 $\{\tilde{d}^k\}$ 都有界.

(ii) 子问题 (2.5) 的 KKT 乘子序列 $\{u^k\}$ 及 $\{v^k\}$ 均有界.

证明: (i) 由 (2.5) 的约束条件, 假设 2.1 、 3.1, 引理 2.5(ii) 可知, 存在两个常数 $q, q' > 0$, 使得

$$\begin{aligned} 0 &\ge \gamma_0 z_k + \tfrac{1}{2}(d^k)^T B_k d^k \\ &\ge f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) + \tfrac{1}{2}(d^k)^T B_k d^k \\ &\ge -q - q'\|d^k\| + \tfrac{1}{2}a\|d^k\|^2, \ \forall\ i \in I_k \ne \emptyset \end{aligned}$$

成立, 这说明 $\{d^k\}$ 有界. 又由 $0 \ge z_k \ge \frac{1}{\gamma_0}\nabla f_{i_k}(x^k)^T d^k$ 可知 $\{z_k\}$ 有界. 进而由算法 2.6 的步骤 3 ($\|\tilde{d}^k\| \le \|d^k\|$) 可知 $\{\tilde{d}^k\}$ 有界.

(ii) 由 (2.8) 有

$$\gamma_0 = \gamma_0 \sum_{i\in I_k} u_i^k + \sum_{j\in J_k} v_i^k(\gamma_j\theta_k - 2\varepsilon_k z_k) \ge \gamma_0 \sum_{i\in I_k} u_i^k, \ \ \sum_{i\in I_k} u_i^k \le 1, \ \ u_i^k \ge 0.$$

此说明 $\{u^k\}$ 有界.

接下来，用反证法证明 $\{v^k\}$ 的有界性. 设 $\{v^k\}$ 无界，则存在一个无穷指标集 K, 使得 $\|v^k_{L'_k}\| \to \infty$. 由 (2.8) 有

$$(d^k)^T B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k)^T d^k + \sum_{j\in L'_k} v_j^k \nabla g_j(x^k)^T d^k = 0,$$

此结合 (2.7) 有

$$(d^k)^T B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k)^T d^k + \sum_{j\in L'_k} v_j^k (\gamma_j\theta_k z_k - \varepsilon_k z_k^2 - g_j(x^k)) = 0,$$

即

$$-\left((d^k)^T B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k)^T d^k + \sum_{j\in L'_k} v_j^k (\gamma_j\theta_k z_k - \varepsilon_k z_k^2)\right) = \sum_{j\in L'_k} v_j^k (-g_j(x^k)).$$

结合乘子的非负性和 $g_j(x^k) \le 0$，有

$$-\left((d^k)^T B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k)^T d^k + \sum_{j\in L'_k} v_j^k (\gamma_j\theta_k z_k - \varepsilon_k z_k)\right) = \sum_{j\in L'_k} |v_j^k| \cdot |g_j(x^k)|. \tag{3.2}$$

又由 (2.8) 可知

$$z_k\gamma_0 - z_k \sum_{i\in L_k} u_i^k \gamma_0 + \sum_{j\in L'_k} v_j^k (2\varepsilon_k z_k^2 - \gamma_j\theta_k z_k) = 0,$$

即

$$\sum_{j\in L'_k} v_j^k (\gamma_j\theta_k z_k - \varepsilon_k z_k^2) = z_k\gamma_0 - z_k \sum_{i\in L_k} u_i^k \gamma_0 + \sum_{j\in L'_k} v_j^k \varepsilon_k z_k^2. \tag{3.3}$$

于是由 (3.2) 和 (3.3), 有

$$-\left((d^k)^T B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k)^T d^k + z_k\gamma_0 - z_k \sum_{i\in L_k} u_i^k \gamma_0\right) = \sum_{j\in L'_k} |v_j^k|\cdot|g_j(x^k)| + \sum_{j\in L'_k} v_j^k \varepsilon_k z_k^2.$$

另外，$\{(x^k, d^k, z_k, u^k)\}$ 的有界性及 (3.1) 显示，上式的左边有界，进而存在一个正的常数 M, 使得

$$\sum_{j\in L'_k} |v_j^k| \cdot |g_j(x^k)| + \sum_{j\in L'_k} v_j^k \varepsilon_k z_k^2 \le M.$$

由于 $\{x^k\}$ 有界，不妨假设 $x^k \to x^*$，$L'_k \equiv L'$，$k \in K$. 因此

$$\sum_{j\in L'\setminus J(x^*)} |v_j^k|\cdot|g_j(x^k)| + \sum_{j\in L'\setminus J(x^*)} v_j^k \varepsilon_k z_k^2 \le \sum_{j\in L'} |v_j^k|\cdot|g_j(x^k)| + \sum_{j\in L'_k} v_j^k \varepsilon_k z_k^2 \le M.$$

此说明序列 $\{v^k_{L'\setminus J(x^*)}\}_K$ 有界. 进而由 $\|v^k_{L'}\| \to \infty$, 有 $\|v^k_{L'\cap J(x^*)}\| \to \infty$.

另一方面，由 (2.8) 有

$$B_k d^k + \sum_{i\in L_k} u_i^k \nabla f_i(x^k) + \sum_{j\in L'\setminus J(x^*)} v_j^k \nabla g_j(x^k) + \sum_{j\in L'\cap J(x^*)} v_j^k \nabla g_j(x^k) = 0. \tag{3.4}$$

将 (3.4) 两边同时除以 $\|v^k_{L'\cap J(x^*)}\|$, 得

$$\begin{aligned} &\frac{1}{\|v^k_{L'\cap J(x^*)}\|} B_k d^k + \sum_{i\in L_k} \frac{u_i^k}{\|v^k_{L'\cap J(x^*)}\|}\nabla f_i(x^k) + \sum_{j\in L'\setminus J(x^*)} \frac{v_j^k}{\|v^k_{L'\cap J(x^*)}\|}\nabla g_j(x^k) \\ &+ \sum_{j\in L'\cap J(x^*)} \frac{v_j^k}{\|v^k_{L'_k\cap J(x^*)}\|}\nabla g_j(x^k) = 0. \end{aligned} \tag{3.5}$$

由于序列 $\{\frac{v^k_{L'\cap J(x^*)}}{\|v^k_{L'\cap J(x^*)}\|}\}$ 有界，且其范数为 1, 不妨设

$$\frac{v_j^k}{\|v^k_{L'\cap J(x^*)}\|} \to v'_j,\ \ j \in L'\cap J(x^*),\ 0 \le (v'_j,\ \ j\in L'\cap J(x^*)) \ne 0.$$

于是在 (3.5) 中对 $k \in K$ 和 $k \to \infty$ 取极限，有 $\sum\limits_{j\in L'\cap J(x^*)} v'_j \nabla g_j(x^*) = 0$. 这与弱 MFCQ 矛盾，故 $\{v^k\}$ 有界性获证.

引理 3.3 如果假设 2.1, 2.2 及 3.1 均成立，令 K 是使得 $\lim\limits_{k\in K} d^k = 0$, $\lim\limits_{k\in K} z_k = 0$ 及 $\lim\limits_{k\in K} x^k = x^*$ 成立的指标集，则 x^* 是问题 (1.8) 的稳定点. 此外，存在一个无限指标集 $K' \subseteq K$, 使得序列 $\{\bar{u}^k, \bar{v}^k\}_{K'}$ 收敛到相应于点 x^* 的 KKT (稳定点) 乘子.

证明：首先，由积极集 L_k 和 L'_k 的有限选取性以及 $\{u^k\}$ 和 $\{v^k\}$ 有界性可知，存在一个无限的指标集 $K' \subseteq K$, 使得

$$\begin{aligned} &L_k \equiv L,\ \ L'_k \equiv L',\ \ u^k = (u_i^k,\ \ i\in I) \to u^* = (u_i^*,\ \ i \in I),\ \ k \in K' \\ &v^k = (v_j^k,\ \ j\in J) \to v^* = (v_j^*,\ \ j\in J),\ \ (d^k, z_k, \theta_k) \to (0, 0, \theta_*),\ \ k\in K'. \end{aligned}$$

于是，在 (2.8) 中令 $k \in K'$，$k \to \infty$, 有

$$\begin{cases} \sum\limits_{i\in L} u_i^* \nabla f_i(x^*) + \sum\limits_{j\in L'} v_j^* \nabla g_j(x^*) = 0, \\ \sum\limits_{i\in L} u_i^* \gamma_0 + \sum\limits_{j\in L'} v_j^* \gamma_j \theta_* = \gamma_0, \\ 0 \le u_i^* \perp (f_i(x^*) - F(x^*)) \le 0, \ \forall\, i \in L, \\ 0 \le v_j^* \perp g_j(x^*) \le 0, \ \forall\, j \in L'. \end{cases} \tag{3.6}$$

注意到 $L' \subseteq J(x^*)$，$v^* \ge 0$，由 (3.6) 的前两个关系式及假设 2.2 可推知 $\sum\limits_{i\in L} u_i^* > 0$. 因此 (2.10) 可表为

$$\begin{cases} \sum\limits_{i\in L} \bar{u}_i^* \nabla f_i(x^*) + \sum\limits_{j\in L'} \bar{v}_j^* \nabla g_j(x^*) = 0;\ \sum\limits_{i\in L} \bar{u}_i^* = 1; \\ 0 \le \bar{u}_i^* \perp (f_i(x^*) - F(x^*)) \le 0, \ \forall\, i \in L; \\ 0 \le \bar{v}_j^* \perp g_j(x^*) \le 0, \ \forall\, j \in L', \end{cases} \tag{3.7}$$

其中，$\bar{u}_i^* = u_i^* / \sum\limits_{i\in L} u_i^*$，$i \in L$，$\bar{v}_j^* = v_j^* / \sum\limits_{i\in L} u_i^*$，$j \in L'$，$\bar{u}_i^* = 0$，$i \in I \setminus L$，$\bar{v}_j^* = 0$，$j \in J \setminus L'$. (3.7) 式说明 $(x^*, \bar{u}^*, \bar{v}^*)$ 是 (1.8) 的稳定点对.

引理 3.4 设假设 2.1, 2.2 及 3.1 均成立，如果 $\lim\limits_{k\in K} d^k = 0$, 则 $\lim\limits_{k\in K} z_k = 0$.

证明：由引理 2.5(ii) 和 (2.5) 的约束条件得

$$0 \ge \gamma_0 z_k \ge f_i(x^k) + \nabla f_i(x^k)^T d^k - F(x^k) = \nabla f_i(x^k)^T d^k,\ i \in I(x^k) \subseteq I_k.$$

由此式及 $\lim\limits_{k\in K} d^k = 0$ 有 $\lim\limits_{k\in K} z_k = 0$.

3.2 全局收敛性定理

基于本论文中 3.1 节中的假设条件和引理分析，可以得到本章的重要定理，即全局收敛性定理.

定理 3.5 设假设 2.1, 2.2 及 3.1 均成立, 则算法 2.6 或者有限步终止于问题 (1.8) 的稳定点 x^k, 或者产生一个无穷点列 $\{x^k\}$, 使得它的每个极限点 x^* 都是问题 (1.8) 的稳定点.

证明：如果算法在第 k 步迭代时终止，则由算法 2.6 步骤 2 和引理 2.5(iii) 可知，当前迭代点 x^k 是 (1.8) 的稳定点. 现假设算法 2.6 产生无穷点列 $\{x^k\}$, x^* 为 $\{x^k\}$ 的一个聚点. 由 $\{(u^k, v^k)\}$ 的有界性知，存在一个指标集 K, 使得 $x^k \to x^*$, $u^{k-1} \to u^*$, $v^{k-1} \to v^*$, $k \in K$.

用反证法，假设 x^* 不是稳定点. 则由引理 3.3、3.4 知 $\lim\limits_{k\in K} d^k \neq 0$, 进而存在一个指标集 $\bar{K} \subseteq K$ 以及常数 $\sigma > 0$, 使得当 $k \in \bar{K}$ 时, $\|d^k\| \geq \sigma$. 又注意到 (x^*, u^*, v^*) 不是 (1.1) 的稳定点对, 因此 $\varrho(x^k, \alpha^k, \beta^k) = \varrho(x^k, u^{k-1}, v^{k-1}) \to \varrho(x^*, u^*, v^*) > 0$, $k \in \bar{K}$. 结合 $\varrho(x^k, \alpha^k, \beta^k) > 0$ 可知，存在一个正的常数 $\zeta > 0$, 使

$$\varrho(x^k, \alpha^k, \beta^k) \geq \zeta, \ \forall\, k \in \bar{K}. \tag{3.8}$$

下面先证明存在一个常数 $\bar{t} > 0$, 使得 $t_k \geq \bar{t}$, $\forall\, k \in \bar{K}$, 为分析 (2.17), 记

$$\omega_k(t) = F(x^k + td^k + t^2\tilde{d}^k) - F(x^k) + \alpha t(d^k)^T B_k d^k.$$

由 $\{d^k\}$, $\{\tilde{d}^k\}$ 的有界性和 Taylor 展式,

$$\begin{aligned}\omega_k(t) &= \max\{f_i(x^k + td^k + t^2\tilde{d}^k) - F(x^k) + \alpha t(d^k)^T B_k d^k, \ i \in I\} \\ &= \max\{f_i(x^k) + t\nabla f_i(x^k)^T d^k + o(t) - F(x^k) + \alpha t(d^k)^T B_k d^k, \ i \in I\}.\end{aligned}$$

再记

$$b_{ki}(t) = f_i(x^k) + t\nabla f_i(x^k)^T d^k + o(t) - F(x^k) + \alpha t(d^k)^T B_k d^k, \ i \in I.$$

当 $i \in I_k$, $t > 0$, $k \in \bar{K}$ 充分小时，由 (2.5), 引理 2.5(ii), 假设 3.1 以及 $\alpha \in (0, \frac{1}{2})$, 得

$$\begin{aligned}b_{ki}(t) &\leq f_i(x^k) + t(\gamma_0 z_k + F(x^k) - f_i(x^k)) - F(x^k) + \alpha t(d^k)^T B_k d^k + o(t) \\ &= (1-t)(f_i(x^k) - F(x^k)) + t\gamma_0 z_k + \alpha t(d^k)^T B_k d^k + o(t) \\ &\leq (1-t)(f_i(x^k) - F(x^k)) + (\alpha - \tfrac{1}{2})t(d^k)^T B_k d^k + o(t) \\ &\leq (\alpha - \tfrac{1}{2})at\sigma^2 + o(t) \\ &\leq 0.\end{aligned}$$

因此，当 $i \in I_k$ 时 $b_{ki}(t) \leq 0$ 对 $k \in \bar{K}$ 及 $t > 0$ 充分小成立.

当 $i \in I \backslash I_k$, $k \in \bar{K}$ 时，由 (2.4) 知 $f_i(x^k) - F(x^k) < -\varrho(x^k, \alpha^k, \beta^k) \le -\zeta$, 则

$$\begin{aligned} b_{ki}(t) &= f_i(x^k) + t\nabla f_i(x^k)^T d^k - F(x^k) + \alpha t(d^k)^T B_k d^k + o(t) \\ &= f_i(x^k) - F(x^k) + O(t) \\ &\le -\zeta + O(t) \\ &\le 0. \end{aligned}$$

故当 $i \in I \backslash I_k$ 时 $b_{ki}(t) \le 0$. 因此 $\omega_k(t) \le 0$, 亦即 (2.17) 对 $k \in \bar{K}$ 及 $t > 0$ 充分小成立.

接下来分析 (2.18), 略去 "$k \in \bar{K}$" 的反复陈述. 当 $j \in J_k$, $t > 0$ 充分小时，利用 Taylor 展式，结合 $\{(d^k, \tilde{d}^k)\}$ 有界，由 (2.5) 得

$$\begin{aligned} g_j(x^k + td^k + t^2\tilde{d}^k) &= g_j(x^k) + t\nabla g_j(x^k)^T d^k + o(t) \\ &\le g_j(x^k) + t(\gamma_j \theta_k z_k - \varepsilon_k z_k^2 - g_j(x^k)) + o(t) \\ &= (1-t)g_j(x^k) + t(\gamma_j \theta_k z_k - \varepsilon_k z_k^2) + o(t) \\ &\le t(\gamma_j \theta_k z_k - \varepsilon_k z_k^2) + o(t). \end{aligned} \tag{3.9}$$

另外，由引理 2.5(ii) 及假设 3.1, 有 $z_k \le -\frac{a\sigma^2}{2\gamma_0}$, $-z_k^2 \le -\frac{a^2\sigma^4}{4\gamma_0^2}$, 此结合 (3.9), 可推得

$$g_j(x^k + td^k + t^2\tilde{d}^k) \le -\frac{a\sigma^2}{2\gamma_0}\eta t(\theta_k + \varepsilon_k) + o(t) \le -\frac{a\sigma^2}{2\gamma_0}\eta t\underline{c} + o(t) \le 0,$$

其中 $\eta = \min\{\gamma_j, \frac{a\sigma^2}{2\gamma_0}, j \in J\} > 0$. 因此，当 $j \in J_k$ 时，$g_j(x^k + td^k + t^2\tilde{d}^k) \le 0$ 成立.

当 $j \in J \backslash J_k$, $k \in \bar{K}$ 时，由 (2.4) 知 $g_j(x^k) < -\varrho(x^k, \alpha^k, \beta^k) \le -\zeta$, 故

$$g_j(x^k + td^k + t^2\tilde{d}^k) = g_j(x^k) + t\nabla g_j(x^k)^T d^k + o(t) = g_j(x^k) + O(t) \le -\zeta + O(t) \le 0.$$

综上所述，对于所有 $k \in \bar{K}$ 及充分小的正数 t, (2.17)-(2.18) 恒成立，故 $t_k \ge \bar{t} > 0$ $(\forall\, k \in \bar{K})$ 成立.

最后利用 $t_k \ge \bar{t} > 0$ $(k \in \bar{K})$ 导矛盾. 由算法 2.6 和假设 3.1 有

$$\begin{aligned} &F(x^{k+1}) - F(x^k) \le -\alpha t_k (d^k)^T B_k d^k < 0, \ \forall\, k, \\ &F(x^{k+1}) - F(x^k) \le -\alpha t_k (d^k)^T B_k d^k \le -\alpha \bar{t} a\sigma^2, \ \forall\, k \in \bar{K}. \end{aligned} \tag{3.10}$$

则 $\{F(x^k)\}$ 是下降的，结合 $\lim\limits_{k\in K} F(x^k) = F(x^*)$, 有 $\lim\limits_{k\to\infty} F(x^k) = F(x^*)$. 因此，在 (3.10) 中令 $k \in \bar{K}$, $k \to \infty$, 取极限得 $-\alpha\bar{t}a\sigma^2 \geq 0$, 这矛盾. 因此点 x^* 是 (1.8) 的稳定点.

3.3 本章小结

本章通过详细分析，证明出了算法 2.6 在假设 2.1, 2.2, 3.1 均成立的条件下，具有全局收敛性，即算法 2.6 产生的迭代点列 $\{x^k\}$ 的任意一个极限点都是问题 (1.8) 的稳定点. 在下一章中，本论文将研究分析算法的强收敛和超线性收敛性.

第 4 章 强收敛和超线性收敛性

4.1 强收敛性

本小节首先分析算法 2.6 的强收敛性，为此，需以下假设.

假设 4.1 (i) 函数 $f_i(x)$ $(i \in I)$, $g_j(x)$ $(j \in J)$ 在可行集 X 上都是二阶连续可微的，且参数序列 $\{\theta_k\}$ 满足 $\lim\limits_{k\to\infty} \theta_k = 0$.

(ii) 算法 2.6 产生的点列 $\{x^k\}$ 存在一个极限点 x^*, 使得在 x^* 处 LICQ 成立，即存在 $i_* \in I(x^*)$, 使得向量组 $\{\nabla f_i(x^*) - \nabla f_{i_*}(x^*),\ i \in I(x^*)\backslash\{i_*\};\ \nabla g_j(x^*),\ j \in J(x^*)\}$ 线性无关.

(iii) 稳定点对 $(x^*, \bar{u}^*, \bar{v}^*)$ 满足上层严格互补条件和强二阶充分条件，即

$$\bar{u}_i^* > 0,\ \forall\, i \in I(x^*);\ d^T\nabla_{xx}^2 L(x^*, \bar{u}^*, \bar{v}^*)d > 0,\ \forall\, d \in \Omega\backslash\{0\},$$

其中

$$\nabla_{xx}^2 L(x^*, \bar{u}^*, \bar{v}^*) = \sum_{i\in I} \bar{u}_i^* \nabla^2 f_i(x^*) + \sum_{j\in J} \bar{v}_j^* \nabla^2 g_j(x^*),$$

$$\Omega \triangleq \{d \in R^n : \nabla f_i(x^*)^T d = 0,\ \forall\, i \in I(x^*);\ \nabla g_j(x^*)^T d = 0,\ \forall\, j \in J_*^+\},\ J_*^+ = \{j \in J : \bar{v}_j^* > 0\}.$$

注 4.2 假设 4.1(ii) 与下面两个条件中的任何一个等价：

假设 4.1(ii-1). R^{n+1} 中的向量组 $\left\{ \begin{pmatrix} -1 \\ \nabla f_i(x^*) \end{pmatrix},\ i \in I(x^*);\ \begin{pmatrix} 0 \\ \nabla g_j(x^*) \end{pmatrix},\ j \in J(x^*) \right\}$ 线性无关.

假设 4.1(ii-2). 对任意的 $p \in I(x^*)$, $\{\nabla f_i(x^*) - \nabla f_p(x^*),\ i \in I(x^*)\backslash\{p\};\ \nabla g_j(x^*),\ j \in J(x^*)\}$ 线性无关.

引理 4.3 如果假设 2.2, 3.1 及 4.1 成立，若指标集 K 满足 $\lim\limits_{k\in K} x^k = x^*$, 其中 x^* 是指假设 4.1 中所云，则有 $\lim\limits_{k\in K} d^k = 0$, $\lim\limits_{k\in K} \tilde{d}^k = 0$ 及 $\lim\limits_{k\in K} z_k = 0$, 其中 $\tilde{d}^k$ 是算法 2.6 步骤 4 中的搜索方向.

证明：先证 $\lim_{k\in K} d^k = 0$. 由定理 3.5 知 x^* 是 (1.8) 的稳定点. 用反证法. 假设 $\lim_{k\in K} d^k \neq 0$, 从而存在指标集 $K' \subseteq K$ 和一个常数 $c > 0$, 使得当 $k \in K'$ 时，有 $\|d^k\| \geq c$. 由引理 3.2 、假设 3.1(ii) 及 $\theta_k \to 0$ 知，存在一个指标集 $K'' \subseteq K'$, 使得

$$\lim_{k\in K''} x^k = x^*, \quad \lim_{k\in K''} d^k = d^* \neq 0, \quad \lim_{k\in K''} \varepsilon_k = \varepsilon_* > 0,$$
$$\lim_{k\in K''} \alpha^k = \bar{\alpha}^*, \quad \lim_{k\in K''} \beta^k = \bar{\beta}^*, \quad I_k \equiv \bar{I}, \quad J_k \equiv \bar{J}, \ \forall\, k \in K''.$$

接下来证明 $I(x^*) \subseteq \bar{I}$, $J(x^*) \subseteq \bar{J}$.

如果 $(x^*, \bar{\alpha}^*, \bar{\beta}^*)$ 构成 (1.8) 的稳定点对，注意到 $(x^k, \alpha^k, \beta^k) \to (x^*, \bar{\alpha}^*, \bar{\beta}^*)$, $k \in K''$, 在假设 4.1 下，由文献 [21] 中定理 3.4 可知，$\bar{I} = I_k \equiv I(x^*)$, $\bar{J} = J_k \equiv J(x^*)$.

如果 $(x^*, \bar{\alpha}^*, \bar{\beta}^*)$ 不是 (1.8) 的稳定点对，则 $\varrho(x^*, \bar{\alpha}^*, \bar{\beta}^*) > 0$, 进而易知 $I(x^*) \subseteq \bar{I}$, $J(x^*) \subseteq \bar{J}$ 成立. 因此，$I(x^*) \subseteq \bar{I}$ 及 $J(x^*) \subseteq \bar{J}$ 恒成立.

另一方面，由引理 2.5(ii) 和假设 3.1 有

$$z_* \triangleq \lim_{k\in K''} z_k \leq \lim_{k\in K''} \left(-\tfrac{1}{2\gamma_0}(d^k)^T B_k d^k\right) \leq -\tfrac{1}{2\gamma_0} a\|d^*\|^2 < 0.$$

因此，对于 $i \in I(x^*) \subseteq I_k$, $j \in J(x^*) \subseteq J_k$, 由 (2.5) 的约束条件和 $\theta_k \to 0$, 有

$$\nabla f_i(x^*)^T d^* \leq \gamma_0 z_* < 0, \ \nabla g_j(x^*)^T d^* \leq -\varepsilon_* z_*^2 < 0, \tag{4.1}$$

又 $(x^*, \bar{u}^*, \bar{v}^*)$ 为稳定点对，得

$$\sum_{i\in I(x^*)} \bar{u}_i^* \nabla f_i(x^*) + \sum_{j\in J(x^*)} \bar{v}_j^* \nabla g_j(x^*) = 0, \ 0 \neq (\bar{u}^*, \bar{v}^*) \geq 0. \tag{4.2}$$

而 (4.1) 与 (4.2) 矛盾. 因此，假设不成立，即 $\lim_{k\in K} d^k = 0$. 由算法 2.6 的步骤 3 得 $\lim_{k\in K} \tilde{d}^k = 0$. 最后，由引理 3.4 可知 $\lim_{k\in K} z_k = 0$.

定理 4.4 如果假设 2.2, 3.1 及 4.1 均成立，则算法 2.6 产生的序列 $\{x^k\}$ 整列收敛到点 x^*, 即算法 2.6 是强收敛的.

证明：在假设 4.1 下，由文献 [10] 中定理 1.4.2 知，极限点 x^* 是 (1.8) 的孤立稳定点. 进而由定理 3.5 知 x^* 是 $\{x^k\}$ 的一个孤立极限点. 又由步骤 3 和引理 4.3 知，

对于任何满足 $x^k \to x^*$, $k \in K$ 的指标集 K, 有

$$\lim_{k\in K} \|x^{k+1} - x^k\| \le \lim_{k\in K}(t_k\|d^k\| + t_k^2\|\tilde{d}^k\|) \le \lim_{k\in K}(t_k\|d^k\| + t_k^2\|d^k\|) = 0.$$

因此由文献 [10] 中定理 1.1.7 可知, $\lim\limits_{k\to\infty} x^k = x^*$.

4.2 超线性收敛性

接下来分析算法的全局收敛性，首先得到如下引理.

引理 4.5 在假设 2.2, 3.1 及 4.1 下，有

$\lim\limits_{k\to\infty} d^k = 0$, $\lim\limits_{k\to\infty} z_k = 0$, $\lim\limits_{k\to\infty} \sum\limits_{i\in I} u_i^k = 1$, $\lim\limits_{k\to\infty}(\alpha^k, \beta^k) = \lim\limits_{k\to\infty}(u^{k-1}, v^{k-1}) = (\bar{u}^*, \bar{v}^*)$.

证明：首先由定理 4.4 和引理 4.3 立知 $d^k \to 0$, $z_k \to 0$ 成立. 其次，由 (2.8) 及算法步骤 2 知

$$\sum_{i\in I} u_i^k = \sum_{i\in I_k} u_i^k = 1 + \frac{1}{\gamma_0}\sum_{j\in J_k} v_j^k(2\varepsilon_k z_k - \gamma_j\theta_k).$$

故由 $\{v^k\}$ 及 $\{\varepsilon_k\}$ 的有界性和 $(\theta_k, z_k) \to (0, 0)$ 可知 $\sum\limits_{i\in I} u_i^k \to 1$. 又注意到乘子 $(\bar{u}^*, \bar{v}^*)$ 的唯一性 (因在点 x^* 处 LICQ 成立), 由引理 3.3 和 $\lim\limits_{k\to\infty}(x^k, d^k, z_k) = (x^*, 0, 0)$, 有 $\lim\limits_{k\to\infty}(\alpha^k, \beta^k) = \lim\limits_{k\to\infty}(u^{k-1}, v^{k-1}) = (\bar{u}^*, \bar{v}^*)$.

引理 4.6 若假设 2.2, 3.1 及 4.1 均成立，则当 k 充分大时，有

(i) $L_k = I_k \equiv I(x^*)$, $L_k' \subseteq J_k \equiv J(x^*)$;

(ii) (2.14) 的系数矩阵 V_k 非奇异，进而方程组 (2.14) 有唯一解，且存在常数 $\bar{p}$ 使得 $\|V_k^{-1}\| \le \bar{p}$.

证明：(i) 由 $\lim\limits_{k\to\infty}(x^k, d^k, z_k) = (x^*, 0, 0)$, 以及 $\{(\varepsilon_k, \theta_k)\}$ 的有界性易知, $L_k \subseteq I(x^*)$ 及 $L_k' \subseteq J(x^*)$ 成立. 由 $u_{I(x^*)}^k \to \bar{u}_{I(x^*)}^* > 0$ 得 $u_{I(x^*)}^k > 0$, 于是 $I(x^*) \subseteq L_k$, 因此 $L_k \equiv I(x^*)$. 另一方面, 由于 $(x^k, \alpha^k, \beta^k) \to (\bar{x}^*, \bar{u}^*, \bar{v}^*)$, 故由文献 [21] 定理 3.4 知道, $I_k \equiv I(x^*)$ 及 $J_k \equiv J(x^*)$ 在假设 4.1 下成立，即 $L_k = I_k \equiv I(x^*)$, $L_k' \subseteq J_k \equiv J(x^*)$

成立.

(ii) 注意到 $I_k \equiv I(x^*)$ 及 $J_k \equiv J(x^*)$, 在 LICQ 条件下，由文 [10] 中定理 1.1.10, 可知结论 (ii) 成立.

引理 4.7 若假设 2.2, 3.1 及 4.1 均成立，则

$$| z_k |= O(\|d^k\|),\ \|\tilde{d}^k\| = O(\max\{\|d^k\|^2,\ \varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k\}) = o(\|d^k\|), \tag{4.3}$$

其中 $(\tilde{d}^k, s^k)$ 由线性方程组 (2.14) 产生. 因此，当 k 充分大时，步骤 3 中的高阶修正方向恒由 (2.14) 产生.

证明：由 $z_k \le 0$ 及 (2.5) 的第一个约束条件，得

$$| z_k |\le \frac{1}{\gamma_0}(F(x^k) - f_i(x^k) - \nabla f_i(x^k)^T d^k) \le \frac{1}{\gamma_0}\|\nabla f_i(x^k)\| \cdot \|d^k\|,\ i \in I(x^k) \neq \emptyset.$$

因此，$| z_k |= O(\|d^k\|)$ 成立. 另一方面，由 (2.16) 和 Taylor 展式，有

$$\tilde{f_i}^k = f_i(x^k + d^k) - f_i(x^k) - \nabla f_i(x^k)^T d^k - \Delta_k = O(\|d^k\|^2),\ i \in I_k \backslash \{i_k\},$$

$$\begin{aligned}\tilde{g_j}^k &= g_j(x^k + d^k) - g_j(x^k) - \nabla g_j(x^k)^T d^k + \gamma_j\theta_k z_k - \varepsilon_k z_k^2 \\ &= O(\|d^k\|^2) + O(\varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k),\ j \in J_k.\end{aligned}$$

注意到 $|z_k| = O(\|d^k\|)$, $\tau \in (2, 3)$, $\theta_k \to 0$ 以及引理 4.6(ii), 由 (2.14)-(2.16) 有

$$\begin{aligned}\begin{pmatrix} \tilde{d}^k \\ s^k \end{pmatrix} &= O(\max\{\|d^k\|^\tau,\ -\bar{\gamma}\theta_k^\nu z_k\|d^k\| + \varepsilon_k z_k^2\|d^k\|^\nu\}) + \|\tilde{h}^k\| \\ &= O(\max\{\|d^k\|^\tau,\ -\bar{\gamma}\theta_k^\nu z_k\|d^k\| + \varepsilon_k z_k^2\|d^k\|^\nu\}) + O(\|d^k\|^2) + O(\varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k) \\ &= O(\max\{\|d^k\|^\tau,\ o(\varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k)\}) + O(\|d^k\|^2) + O(\varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k) \\ &= O(\max\{\|d^k\|^2,\ \varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k\}).\end{aligned}$$

因此，

$$\tilde{d}^k = O(\max\{\|d^k\|^2,\ \varepsilon_k z_k^2 - \bar{\gamma}\theta_k z_k\}) = o(\|d^k\|) \le \|d^k\|.$$

进而，由算法 2.6 的结构，可知当 k 充分大时，步骤 3 中的高阶修正方向由 (2.14) 产生.

为保证步长 1 能被算法 2.6 所接受，还需如下二阶逼近条件 (SOAC).

假设 4.8 假设 $\|P_k(\nabla^2_{xx}L(x^k, u^k, v^k) - B_k)d^k\| = o(\|d^k\|)$，其中 $P_k = E_n - A_k(A_k^T A_k)^{-1}A_k^T$，$A_k = (\nabla f_i(x^k) - \nabla f_{i_k}(x^k),\ i \in L_k\backslash\{i_k\})$，当 $L_k\backslash\{i_k\} = \emptyset$ 时，约定 $P_k = E_n$.

注 4.9 (i) 由于 $(x^k, \sum\limits_{i\in I} u_i^k, u^k, v^k) \to (x^*, 1, \bar{u}^*, \bar{v}^*)$，易见

$$\|P_k(\nabla^2_{xx}L(x^k, u^k, v^k) - B_k)d^k\| = o(\|d^k\|) \Leftrightarrow \|P_k(\nabla^2_{xx}L(x^k, \bar{u}^k, \bar{v}^k) - B_k)d^k\| = o(\|d^k\|)$$
$$\Leftrightarrow \|P_k(\nabla^2_{xx}L(x^*, \bar{u}^*, \bar{v}^*) - B_k)d^k\| = o(\|d^k\|).$$

(ii) 假设 4.8 是一个新型的 SOAC, 它比文献 [24] 中的 $\|(\nabla^2_{xx}L(x^k, \lambda^k) - B_k)d^k\| = o(\|d^k\|)$ 要弱，之所以能通过乘项 "P_k" 进行弱化，得益于假设 4.1(iii) 中的上层严格互补条件 $\bar{u}^*_{I(x^*)} > 0$.

(iii) 由 (2.6), $i_k \in I(x^*) = I_k$ 和 LICQ 条件，方向 d^k 可分为

$$d^k = P_k d^k + \bar{d}^k,\quad \|\bar{d}^k\| = O(\|h^k\|_1),\quad h^k = (F(x^k) - f_i(x^k),\ i \in L_k\backslash\{i_k\}). \tag{4.4}$$

定理 4.10 若假设 2.2, 3.1, 4.1 及 4.8 都成立，则当 k 充分大时，算法 2.6 的步长恒等于 1, 即 $t_k \equiv 1$.

证明：只需证明 (2.17) 及 (2.18) 对 $t = 1$ 和充分大的 k 成立. 为简便，在下面的证明中省略 "k 充分大" 陈述.

首先证明 (2.18) 对 $t = 1$ 成立，当 $j \in J(x^*) = J_k$ 时，有 $\lim\limits_{k\to\infty} g_j(x^k) = g_j(x^*) = 0$. 由 (2.14)'和 (2.15) 知

$$\nabla g_j(x^k)^T \tilde{d}^k = -\max\{\|d^k\|^\tau,\ -\bar{\gamma}\theta_k^\nu z_k\|d^k\| + \varepsilon_k z_k^2\|d^k\|^\nu\} - \tilde{g}_j^k,\quad j \in J(x^*).$$

进而，由 (2.16) 有

$$\begin{aligned} & g_j(x^k + d^k) + \nabla g_j(x^k)^T \tilde{d}^k \\ = {} & g_j(x^k) + \nabla g_j(x^k)^T d^k - \gamma_j\theta_k z_k + \varepsilon_k z_k^2 - \max\{\|d^k\|^\tau,\ \varepsilon_k z_k^2\|d^k\|^\nu - \bar{\gamma}\theta_k^\nu z_k\|d^k\|\}. \end{aligned}$$

因此，由 Taylor 展式，(4.3) 和引理 4.7, 有

$$
\begin{aligned}
g_j(x^k+d^k+\tilde{d}^k) &= g_j(x^k+d^k)+\nabla g_j(x^k+d^k)^T\tilde{d}^k+O(\|\tilde{d}^k\|^2) \\
&= g_j(x^k+d^k)+\nabla g_j(x^k)^T\tilde{d}^k+O(\|\tilde{d}^k\|^2)+O(\|d^k\|\cdot\|\tilde{d}^k\|) \\
&= g_j(x^k+d^k)+\nabla g_j(x^k)^T\tilde{d}^k \\
&\quad +O(\max\{\|d^k\|^3,\ \varepsilon_k z_k^2\|d^k\|-\bar{\gamma}\theta_k z_k\|d^k\|\}) \\
&= g_j(x^k)+\nabla g_j(x^k)^T d^k-\gamma_j\theta_k z_k+\varepsilon_k z_k^2 \\
&\quad -\max\{\|d^k\|^\tau,\ -\bar{\gamma}\theta_k^\nu z_k\|d^k\|+\varepsilon_k z_k^2\|d^k\|^\nu\}) \\
&\quad +O(\max\{\|d^k\|^3,\ \varepsilon_k z_k^2\|d^k\|-\bar{\gamma}\theta_k z_k\|d^k\|\}) \\
&\le 0.
\end{aligned}
\tag{4.5}
$$

当 $j\notin J(x^*)$ 时，有 $g_j(x^*)<0$. 于是由 $(x^k,d^k,\tilde{d}^k)\to(x^*,0,0)$，得 $g_j(x^k+d^k+\tilde{d}^k)\le 0$. 因此，(2.18) 对于 $t=1$ 时成立.

其次证明当 $t=1$ 时 (2.17) 亦成立. 对于 $i\in L_k\backslash\{i_k\}$, 有 $f_i(x^k)+\nabla f_i(x^k)^T d^k-F(x^k)-\gamma_0 z_k=0$. 由 (2.14) 有

$$
(\nabla f_i(x^k)-\nabla f_{i_k}(x^k))^T\tilde{d}^k=-\max\{\|d^k\|^\tau,\ \ -\bar{\gamma}\theta_k^\nu z_k\|d^k\|+\varepsilon_k z_k^2\|d^k\|^\nu\}-\tilde{f}_i^k,
$$

即

$$
\nabla f_i(x^k)^T\tilde{d}^k=\nabla f_{i_k}(x^k)^T\tilde{d}^k-\max\{\|d^k\|^\tau,\ -\bar{\gamma}\theta_k^\nu z_k\|d^k\|+\varepsilon_k z_k^2\|d^k\|^\nu\}-\tilde{f}_i^k,\ i\in L_k\backslash\{i_k\}.
\tag{4.6}
$$

注意到

$$
\begin{aligned}
f_i(x^k+d^k+\tilde{d}^k) &= f_i(x^k+d^k)+\nabla f_i(x^k+d^k)^T\tilde{d}^k+O(\|\tilde{d}^k\|^2) \\
&= f_i(x^k+d^k)+\nabla f_i(x^k)^T\tilde{d}^k+O(\|\tilde{d}^k\|\cdot\|d^k\|)+o(\|d^k\|^2) \\
&= f_i(x^k+d^k)+\nabla f_i(x^k)^T\tilde{d}^k+o(\|d^k\|^2),\ \ i\in L_k\backslash\{i_k\},
\end{aligned}
\tag{4.7}
$$

由 (4.7), (4.6), $|z_k|=O(\|d^k\|)$ 及 (2.16), 得

$$
\begin{aligned}
f_i(x^k+d^k+\tilde{d}^k) &= f_i(x^k+d^k)+\nabla f_{i_k}(x^k)^T\tilde{d}^k-\tilde{f}_i^k+o(\|d^k\|^2) \\
&= f_i(x^k)+\nabla f_i(x^k)^T d^k+\nabla f_{i_k}(x^k)^T\tilde{d}^k+\Delta_k+o(\|d^k\|^2) \\
&= F(x^k)+\gamma_0 z_k+\nabla f_{i_k}(x^k)^T\tilde{d}^k+\Delta_k+o(\|d^k\|^2),\ \ i\in L_k\backslash\{i_k\}.
\end{aligned}
\tag{4.8}
$$

注意到 $i_k \in I(x^*) \equiv L_k$, 由 (2.16) 及 (2.6) 有

$$\begin{aligned} f_{i_k}(x^k+d^k+\tilde{d}^k) &= f_{i_k}(x^k+d^k)+\nabla f_{i_k}(x^k)^T\tilde{d}^k+o(\|d^k\|^2) \\ &= f_{i_k}(x^k)+\nabla f_{i_k}(x^k)^T d^k+\nabla f_{i_k}(x^k)^T\tilde{d}^k+\Delta_k+o(\|d^k\|^2) \\ &= F(x^k)+\gamma_0 z_k+\nabla f_{i_k}(x^k)^T\tilde{d}^k+\Delta_k+o(\|d^k\|^2). \end{aligned} \tag{4.9}$$

(4.8) 和 (4.9) 意味着

$$f_i(x^k+d^k+\tilde{d}^k)=F(x^k)+\gamma_0 z_k+\nabla f_{i_k}(x^k)^T\tilde{d}^k+\Delta_k+o(\|d^k\|^2),\ \ \forall\, i\in L_k. \tag{4.10}$$

因此,

$$f_i(x^k+d^k+\tilde{d}^k)=f_j(x^k+d^k+\tilde{d}^k)+o(\|d^k\|^2),\ \ \forall\, i,\ \ j\in L_k. \tag{4.11}$$

又由于 $I(x^k+d^k+\tilde{d}^k)\subseteq I(x^*)\equiv L_k$, 有

$$F(x^k+d^k+\tilde{d}^k)=\max\{f_i(x^k+d^k+\tilde{d}^k),\ i\in L_k\}.$$

因此, 存在 $i_k^0\in L_k$, 使得 $F(x^k+d^k+\tilde{d}^k)=f_{i_k^0}(x^k+d^k+\tilde{d}^k)$. 进而由 (4.11)

$$F(x^k+d^k+\tilde{d}^k)=f_i(x^k+d^k+\tilde{d}^k)+o(\|d^k\|^2),\ \ \forall\, i\in L_k. \tag{4.12}$$

另一方面, 由 (2.12) 有 $\sum\limits_{i\in L_k}\bar{u}_i^k=1$, 于是由 Taylor 展式和 $\|\tilde{d}^k\|=o(\|d^k\|)$, 可得

$$\begin{aligned} F(x^k+d^k+\tilde{d}^k) &= \sum_{i\in L_k}\bar{u}_i^k F(x^k+d^k+\tilde{d}^k) \\ &= \sum_{i\in L_k}\bar{u}_i^k f_i(x^k+d^k+\tilde{d}^k)+o(\|d^k\|^2) \\ &= \sum_{i\in L_k}\bar{u}_i^k\left(f_i(x^k)+\nabla f_i(x^k)^T(d^k+\tilde{d}^k)+\tfrac{1}{2}(d^k)^T\nabla f_i(x^k)d^k\right)+o(\|d^k\|^2). \end{aligned}$$

又由 (2.8) 以及 $\|\tilde{d}^k\|=o(\|d^k\|)$, 知

$$\sum_{i\in L_k}\bar{u}_i^k\nabla f_i(x^k)^T(d^k+\tilde{d}^k)=-\sum_{j\in L_k'}\bar{v}_j^k\nabla g_j(x^k)^T(d^k+\tilde{d}^k)-\frac{1}{\sum\limits_{i\in L_k}u_i^k}(d^k)^T B_k d^k+o(\|d^k\|)^2,$$

注意到 $1 \geq \sum\limits_{i \in L_k} u_i^k \to 1$，故

$$\begin{aligned} F(x^k + d^k + \tilde{d}^k) &= \sum_{i \in L_k} \bar{u}_i^k f_i(x^k) - \frac{1}{\sum\limits_{i \in L_k} u_i^k}(d^k)^T B_k d^k - \sum_{j \in L'_k} \bar{v}_j^k \nabla g_j(x^k)^T (d^k + \tilde{d}^k) \\ &\quad + \frac{1}{2}(d^k)^T \left(\sum_{i \in L_k} \bar{u}_i^k \nabla^2 f_i(x^k) \right) d^k + o(\|d^k\|^2) \\ &\leq \sum_{i \in L_k} \bar{u}_i^k f_i(x^k) - (d^k)^T B_k d^k - \sum_{j \in L'_k} \bar{v}_j^k \nabla g_j(x^k)^T (d^k + \tilde{d}^k) \\ &\quad + \frac{1}{2}(d^k)^T \left(\sum_{i \in L_k} \bar{u}_i^k \nabla^2 f_i(x^k) \right) d^k + o(\|d^k\|^2). \end{aligned} \tag{4.13}$$

当 $j \in L'_k \subseteq J(x^*)$, 即 $g_j(x^k) + \nabla g_j(x^k)^T d^k - \gamma_j \theta_k z_k + \varepsilon_k z_k^2 = 0$, (4.5) 意味着

$$g_j(x^k + d^k + \tilde{d}^k) = o(\|d^k\|^2), \ \forall j \in L'_k.$$

因此,

$$\begin{aligned} o(\|d^k\|^2) &= \sum_{j \in L'_k} \bar{v}_j^k g_j(x^k + d^k + \tilde{d}^k) \\ &= \sum_{j \in L'_k} \bar{v}_j^k \left(g_j(x^k) + \nabla g_j(x^k)^T (d^k + \tilde{d}^k) + \frac{1}{2}(d^k)^T \nabla^2 g_j(x^k) d^k \right) + o(\|d^k\|^2). \end{aligned}$$

进而,

$$-\sum_{j \in L'_k} \bar{v}_j^k g_j(x^k)^T (d^k + \tilde{d}^k) = \sum_{j \in L'_k} \bar{v}_j^k \left(g_j(x^k) + \frac{1}{2}(d^k)^T \nabla^2 g_j(x^k) d^k \right) + o(\|d^k\|^2),$$

将其代入 (4.13), 得

$$\begin{aligned} & F(x^k + d^k + \tilde{d}^k) \\ \leq & \sum_{i \in L_k} \bar{u}_i^k f_i(x^k) - (d^k)^T B_k d^k + \sum_{j \in L'_k} \bar{v}_j^k g_j(x^k) + \frac{1}{2}(d^k)^T \nabla_{xx}^2 L(x^k, \bar{u}^k, \bar{v}^k) d^k + o(\|d^k\|^2) \\ = & \sum_{i \in L_k} \bar{u}_i^k f_i(x^k) + \sum_{j \in L'_k} \bar{v}_j^k g_j(x^k) - F(x^k) + \frac{1}{2}(d^k)^T (\nabla_{xx}^2 L(x^k, \bar{u}^k, \bar{v}^k) - B_k) d^k \\ & + F(x^k) - \alpha (d^k)^T B_k d^k + (\alpha - \frac{1}{2})(d^k)^T B_k d^k + o(\|d^k\|^2). \end{aligned}$$

注意到 $i_k \in L_k$, $f_{i_k}(x^k) = F(x^k)$, $L_k \equiv I(x^*)$, $\bar{u}_{I(x^*)}^k \to \bar{u}_{I(x^*)}^* > 0$, 由注 4.9 (i,

iii) 可知

$$\begin{aligned}
&\sum_{i\in L_k}\bar{u}_i^k f_i(x^k)+\sum_{j\in L'_k}\bar{v}_j^k g_j(x^k)-F(x^k)+\tfrac{1}{2}(d^k)^T(\nabla_{xx}^2L(x^k,\bar{u}^k,\bar{v}^k)-B_k)d^k\\
=&\sum_{i\in L_k\backslash\{i_k\}}\bar{u}_i^k(f_i(x^k)-F(x^k))+\sum_{j\in L'_k}\bar{v}_j^k g_j(x^k)+\tfrac{1}{2}(P_kd^k+\tilde{d}^k)^T(\nabla_{xx}^2L(x^k,\bar{u}^k,\bar{v}^k)-B_k)d^k\\
\le&-\tfrac{1}{2}\min\{\bar{u}_i^*,\ i\in I(x^*)\}\|h^k\|_1+o(\|h^k\|_1)+o(\|d^k\|^2)\\
\le&\, o(\|d^k\|^2).
\end{aligned}$$

因此，结合 $\alpha\in(0,\frac{1}{2})$, 有

$$\begin{aligned}
F(x^k+d^k+\tilde{d}^k)&\le F(x^k)-\alpha(d^k)^TB_kd^k+(\alpha-\tfrac{1}{2})(d^k)^TB_kd^k+o(\|d^k\|^2)\\
&\le F(x^k)-\alpha(d^k)^TB_kd^k+(\alpha-\tfrac{1}{2})a\|d^k\|^2+o(\|d^k\|^2)\\
&\le F(x^k)-\alpha(d^k)^TB_kd^k.
\end{aligned}$$

因此 (2.17) 对 $t=1$ 及充分大的 k 恒成立. 综上所述，定理证毕.

假设 4.8 等条件虽然确保算法 2.6 能克服 Matatos 效应，保证步长 1 被接受，但根据文献 [17] 中定理 3.2 超线性收敛之充要条件可知，这些条件还不足以保证算法的超线性收敛性. 为此，尚需以下二阶逼近条件.

假设 4.11 $\|\tilde{P}_k(\nabla_{xx}^2L(x^k,u^k,v^k)-B_k)d^k\|=o(\|d^k\|)$, 其中 $\tilde{P}_k=E_n-\tilde{A}_k(\tilde{A}_k^T\tilde{A}_k)^{-1}\tilde{A}_k^T$, $\tilde{A}_k=(A_k\ \nabla g_{L'_k}(x^k))=(\nabla f_i(x^k)-\nabla f_{i_k}(x^k),\ i\in L_k\backslash\{i_k\};\ \nabla g_j(x^k),\ j\in L'_k)$.

注 4.12 假设 4.8 加上假设 4.11 仍比 $\|(\nabla_{xx}^2L(x^k,u^k,v^k)-B_k)d^k\|=o(\|d^k\|)$ 弱，后者常在上下层严格互补条件均不成立下使用. 但假设 4.8 与假设 4.11 之间的内在联系有待进一步研究.

定理 4.13 设假设 2.2, 3.1, 4.1, 4.8 和 4.11 均成立，则算法 2.6 是超线性收敛的，即算法产生的点列 $\{x^k\}$ 满足 $\|x^{k+1}-x^*\|=o(\|x^k-x^*\|)$.

注意到高阶修正方向 $\tilde{d}^k=o(\|d^k\|)$, $|\ \gamma_j\theta_kz_k-\varepsilon_kz_k^2\ |=o(\|d^k\|)$, 基于引理 4.5, 4.6, 4.7 和定理 4.10, 类似于文献 [15] 中定理 4.3 的证明可完成该定理的证明，从略.

4.3 本章小结

在本章中提出了一个新的互补条件，即上层严格互补条件，在上层严格互补条件和强二阶充分条件成立的情况下，得到算法是强收敛的．接下来证明出了算法 2.6 是超线性收敛的．

第 5 章 数值试验

本节对算法 2.6 的有效性进行测试. 先将子问题 (2.5) 转化成一个二阶锥规划问题, 然后用 MOSEK [26] 中的二阶锥优化求解器求解. 数值试验程序是在 MATLAB 2009A 上编写执行的. 电脑配置为: Windows 7, intel 酷睿 i3 350M, 2G 内存. 矩阵 B_k 采用文献 [25] 中的 BFGS 公式进行修正, 即 $B_0 = E_n \in R^{n\times n}$, 并按以下公式修正:

$$B_{k+1} = B_k - \frac{B_k s^k (s^k)^T B_k}{(s_k)^T B_k s_k} + \frac{\tilde{y}^k (\tilde{y}^k)^T}{(s^k)^T \tilde{y}^k},\ \ k \geq 0,$$

其中

$$s^k = x^{k+1} - x^k,\ \ \tilde{y}^k = y^k + a_k(\iota_k s^k + R_k R_k^T s^k),\ \ \iota_k = \min\{\|d^k\|^2,\ \xi\},\ \ \xi \in (0,\,1),$$

$$y^k = \nabla_x L(x^{k+1},\,\bar{u}^k,\,\bar{v}^k) - \nabla_x L(x^k,\,\bar{u}^k,\,\bar{v}^k),$$

$$R_k = (\nabla f_i(x^k) - \nabla f_{i_k}(x^k),\ \ i \in I_k \backslash \{i_k\};\ \ g_j(x^k),\ \ j \in J_k),$$

$$\nabla_x L(x,\,\bar{u},\,\bar{v}) = \sum_{i\in I} \bar{u}_i \nabla f_i(x) + \sum_{j\in J} \bar{v}_j \nabla g_j(x),$$

$$a_k = \begin{cases} 0, & \text{如果 } (s^k)^T y^k \geq \zeta \|s^k\|^2,\ \ \zeta \in (0,\,1); \\ 1, & \text{如果 } 0 \leq (s^k)^T y^k < -\zeta \|s^k\|^2; \\ 1 + \frac{\iota_k \|s^k\|^2 - (s^k)^T y^k}{\iota_k \|s^k\|^2 + (s^k)^T R_k (R_k)^T s^k}, & \text{其他}. \end{cases}$$

在数值试验中, 选取参数如下:

$$\gamma_0 = 2,\ \ \gamma_i = 0.01,\ \ \ \tau = 2.5,\ \ \alpha = 0.2,\ \ \beta = 0.5,\ \ \delta = 0.5,\ \ \nu = 0.8,$$

$$\zeta = 0.5,\ \ \xi = 0.3,\ \ \varepsilon_k = 10^{-9},\ \ \theta_k = \begin{cases} 6.5, & k = 0; \\ \min\{100,\ \|d^{k-1}\|^{0.1}\}, & k \geq 1. \end{cases}$$

算法的终止准则为 $\|d^k\| \leq 10^{-6}$.

5.1 数值算例

在表 5-1 的第 1 组报告中，对退化为光滑约束优化 (即 $m=1$) 的问题 (1.8) 进行测试，并与文献 [10] 中的算法 12.4.7 (亦即 [9] 中的算法) 进行比较. 表中算例来源于文献 [27]. 算例表述如下：

算例 HS12:

$$\begin{aligned} \min \quad & \frac{x_1^2}{2}+x_2^2-x_1x_2-7x_1-7x_2 \\ \text{s.t.} \quad & 4x_1^2+x_2^2-25\le 0. \end{aligned}$$

算例 HS29:

$$\begin{aligned} \min \quad & -x_1x_2x_3 \\ \text{s.t.} \quad & x_1^2+2x_2^2+4x_3^2-48\le 0. \end{aligned}$$

算例 HS30:

$$\begin{aligned} \min \quad & x_1^2+x_2^2+x_3^2 \\ \text{s.t.} \quad & -(x_1^2+x_2^2-1)\le 0, \\ & -(x_1-1)\le 0, \\ & x_1-10\le 0, \\ & -(x_2+10)\le 0, \\ & x_2-10\le 0, \\ & -(x_3+10)\le 0, \\ & x_3-10\le 0. \end{aligned}$$

算例 HS31:

$$
\begin{aligned}
\min \quad & 9x_1^2 + x_2^2 + 9x_3^2 \\
\text{s.t.} \quad & -(x_1^2 + x_2^2 - 1) \le 0, \\
& -(x_1 - 1) \le 0, \\
& x_1 - 10 \le 0, \\
& -(x_2 + 10) \le 0, \\
& x_2 - 10 \le 0, \\
& -(x_3 + 10) \le 0, \\
& x_3 - 10 \le 0.
\end{aligned}
$$

算例 HS34:

$$
\begin{aligned}
\min \quad & -x_1 \\
\text{s.t.} \quad & -(x_2 - exp(x_1)) \le 0, \\
& -(x_3 - exp(x_2)) \le 0, \\
& x_1 - 100 \le 0, \\
& x_2 - 100 \le 0, \\
& x_3 - 10 \le 0, \\
& -x_1 \le 0, \\
& -x_2 \le 0.
\end{aligned}
$$

算例 HS43:

$$
\begin{aligned}
\min \quad & x_1^2 + x_2^2 + 2x_3^2 + x_4^2 - 5x_1 - 5x_2 - 21x_3 + 7x_4 \\
\text{s.t.} \quad & x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_1 - x_2 + x_3 - x_4 - 8 \le 0, \\
& x_1^2 + 2x_2^2 + x_3^2 + 2x_4^2 - x_1 - x_4 - 10 \le 0, \\
& 2x_1^2 + x_2^2 + x_3^2 + 2x_1 - x_2 - x_4 - 5 \le 0.
\end{aligned}
$$

算例 HS66:

$$\begin{aligned}
\min \quad & -0.8x_1 + 0.2x_3 \\
\text{s.t.} \quad & exp(x_1 - x_2) \le 0, \\
& exp(x_2 - x_3) \le 0, \\
& -x_1 \le 0, \\
& x_1 - 100 \le 0, \\
& -x_2 \le 0, \\
& x_2 - 100 \le 0, \\
& -x_3 \le 0, \\
& x_3 - 10 \le 0.
\end{aligned}$$

算例 HS76:

$$\begin{aligned}
\min \quad & x_1^2 + 0.5x_2^2 + x_3^2 + 0.5x_4^2 - x_1x_3 + x_3x_4 - x_1 - 3x_2 + x_3 - x_4 \\
\text{s.t.} \quad & x_1 \le 0, \\
& -x_2 \le 0, \\
& -x_3 \le 0, \\
& -x_4 \le 0, \\
& 3x_1 + x_2 + 2x_3 - x_4 - 4 \le 0, \\
& 3x_1 + x_2 + 2x_3 - x_4 - 4 \le 0, \\
& -(x_2 + 4x_3) + 1.5 \le 0.
\end{aligned}$$

算例 HS100:

$$\begin{aligned}
\min \quad & (x_1 - 10)^2 + 5 * (x_2 - 12)^2 + x_3^4 + 3(x_4 - 11)^2 + 10x_5^6 + 7x_6^2 + x_7^4 - 4x_6x_7 - 10x_6 - 8x_7 \\
\text{s.t.} \quad & 2x_1^2 + 3x_2^4 + x_3 + 4x_4^2 + 5x_5 - 127 \le 0, \\
& 7x_1 + 3x_2 + 10x_3^2 + x_4 - x_5 - 282 \le 0, \\
& 23x_1 + x_2^2 + 6x_6^2 - 8x_7 - 196 \le 0, \\
& 4x_1^2 + x_2^2 - 3x_1x_2 + 2x_3^2 + 5x_6 - 11x_7 \le 0.
\end{aligned}$$

算例 HS113:

$$
\begin{aligned}
\min \quad & x_1^2+x_2^2+x_1x_2-14x_1-16x_2+(x_3-10)^2+4(x_4-5)^2+(x_5-3)^2+2(x_6-1)^2+5x_7^2 \\
& +7(x_8-11)^2+2(x_9-10)^2+(x_{10}-7)^2+45 \\
\text{s.t.} \quad & -(105-4x_1-5x_2+3x_7-9x_8)\leq 0, \\
& -(-10x_1+8x_2+17x_7-2x_8)\leq 0, \\
& -(8x_1-2x_2-5x_9+2x_10+12)\leq 0, \\
& -(-3(x_1-2)^2-4(x_2-3)^2-2x_3^2+7x_4+120)\leq 0, \\
& -(-5x_1^2-8x_2-(x_3-6)^2+2x_4+40\leq 0, \\
& -(-0.5(x_1-8)^2-2(x_2-4)^2-3x_5^2+x_6+30)\leq 0, \\
& -(-x_1^2-2(x_2-2)^2+2x_1x_2-14x_5+6x_6)\leq 0, \\
& -(3x_1-6x_2)-12(x_9-8)^2+7x_{10})\leq 0.
\end{aligned}
$$

算例 HS225:

$$
\begin{aligned}
\min \quad & x_1^2+x_2^2 \\
\text{s.t.} \quad & -(x_1+x_2-1)\leq 0, \\
& -(x_1^2+x_2^2-1)\leq 0, \\
& -(9x_1^2+x_2^2-9)\leq 0, \\
& -(x_1^2-x_2)\leq 0, \\
& -(x_2^2-x_1)\leq 0.
\end{aligned}
$$

算例 HS264:

$$
\begin{aligned}
\min \quad & (x_1^2+x_2^2+2x_3^2+x_4^2-5x_1-5x_2-21x_3+7x_4)\cdot 0.01 \\
\text{s.t.} \quad & x_1^2+x_2^2+x_3^2+x_4^2+x_1-x_2+x_3-x_4-8\leq 0, \\
& x_1^2+2x_2^2+x_3^2+2x_4^2-x_1-x_4-9\leq 0, \\
& 2x_1^2+x_2^2+x_3^2+2x_1-x_2-x_4-5\leq 0.
\end{aligned}
$$

在表 5-2 的第 2 组试验报告中，对 minimax 约束优化 (1.8) 进行测试，并与文献 [18] 中的强次可行方向法 B 进行比较. 测试问题来源于文献 [28]. 现把算例表示如下：其中 OF 表示算例的目标函数， CF 表示算例的约束函数.

算例 01, 02, 03：其中目标函数为

$$\mathrm{OF}:\ f_j(x)=x_j^2,\ j=1,\ldots,n,$$

算例 01 的约束函数： $\mathrm{CF1}:\ f_{n+i}(x)=(3-2x_{i+1})x_{i+1}-x_i-2x_{i+2}+1,\ i=1,\ldots,n-2,$

算例 02 的约束函数： $\mathrm{CF2}:\ f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-2x_i-2x_{i+1}+1,\ i=1,\ldots,n-1,$

算例 03 的约束函数： $\mathrm{CF3}:\ f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}+1,\ i=1,\ldots,n-1,$

算例 04, 05, 06：其中目标函数为

$$\mathrm{OF}:\ f_1(x)=\sum_{j=1}^{n-1}(-x_j-x_{j+1}),\ f_2(x)=\sum_{j=1}^{n-1}(-x_j-x_{j+1}+(x_j^2+x_{j+1}^2-1),$$

算例 04 的约束函数： $\mathrm{CF4}:\ f_{2+i}(x)=(3-2x_{i+1})x_{i+1}-x_i-2x_{i+2}+2.5,\ i=1,\ldots,n-2,$

算例 05 的约束函数： $\mathrm{CF5}:\ f_{n+i}(x)=(3-2x_{i+1})x_{i+1}-x_i-2x_{i+2}+1,\ i=1,\ldots,n-2,$

算例 06 的约束函数： $\mathrm{CF6}:\ f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-1,\ i=1,\ldots,n-1,$

算例 07, 08, 09：其中目标函数为

$$\mathrm{OF}:\ f_1(x)=\sum_{j=1}^{n-1}(x_j^4+x_{j+1}^2),\ f_2(x)=\sum_{j=1}^{n-1}((2-x_j)^2+(2-x_{j+1})^2),\ f_3(x)=\sum_{j=1}^{n-1}(2e^{-x_j+x_{j+1}})\},$$

算例 07 的约束函数： $\mathrm{CF7}:\ f_{2+i}(x)=(3-2x_{i+1})x_{i+1}-x_i-2x_{i+2}+2.5,\ i=1,\ldots,n-2,.$

算例 08 的约束函数： $\mathrm{CF8}:\ f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-1,\ i=1,\ldots,n-1,$

算例 09 的约束函数： $\mathrm{CF9}:\ f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-2x_i-2x_{i+1}+1,\ i=1,\ldots,n-1,$

算例 10, 11, 12, 13：其中目标函数为

$$\text{OF}: f_1(x)=\sum_{j=1}^{n-1}(x_j^2+(x_{j+1}-1)^2+x_{j+1}-1),\ f_2(x)=\sum_{j=1}^{n-1}(-x_j^2-(x_{j+1}-)^2+x_{j+1}+1),$$

算例 10 的约束函数：CF10 : $f_{n+i}(x)=(3-2x_{i+1})x_{i+1}-x_i-2x_{i+2}+1,\ i=1,\ldots,n-2,$

算例 11 的约束函数：CF11 : $f_{n+i}(x)=x_i^2+x_{i+1}^2+x_ix_{i+1}-2x_i-2x_{i+1}+1,\ i=1,\ldots,n-1,$

算例 12 的约束函数：CF12 : $f_{n+i}(x)=(3-0.5x_{i+1})x_{i+1}-x_i-2x_{i+2}+1,\ i=1,\ldots,n-2.$

5.2 数值结果

在报告表中，Alg. 2.6— 本文算法 2.6, Alg. 12.4.7— 文献 [10] 中算法 12.4.7, Alg. B— 文献 [18] 中的强次可行方向法 B, Ni— 迭代次数， Nf— 目标函数的计算次数，Ng— 约束函数的计算次数， $F(x^*)$— 迭代终止时目标函数值， Time(s)— CPU 时间 (秒).

通过以上初步的数值试验可见，与同类算法相比，本文所建立的算法是有效的.

5-1: 光滑约束问题 (m=1) 的数值试验

Prob	n, m, m'	Initial point	Method	Ni	Nf	Ng	$F(x^*)$	Time(s)
HS12	2, 1, 1	(0,0)	Alg. 2.6	6	9	19	-2.999999967e+001	6.196359720e+001
			Alg. 12.4.7	6	9	19	-2.999999998e+01	——
HS29	3, 1, 1	(1,1,1)	Alg. 2.6	13	22	42	-2.262741699e+001	6.325840550e+001
			Alg. 12.4.7	13	22	42	-2.262741700e+01	——
HS30	3, 1, 7	(1,1,1)	Alg. 2.6	10	21	163	1.000000437e+000	6.470921480e+001
			Alg. 12.4.7	10	21	163	1.000000411e+00	——
HS31	3, 1, 7	(1,1,1)	Alg. 2.6	8	15	162	6.063799169e+000	6.578562170e+001
			Alg. 12.4.7	9	16	171	6.063799265e+00	——
HS34	3, 1, 8	(0,1.05,2.9)	Alg. 2.6	6	7	79	-2.996855003e-001	6.681522830e+001
			Alg. 12.4.7	6	7	79	-2.997007475e-01	——
HS43	4, 1, 3	(0,0,0,0)	Alg. 2.6	11	12	192	-1.499999975e+001	6.928004410e+001
			Alg. 12.4.7	11	12	192	-1.499999975e+01	——
HS66	3, 1, 8	(0,1.05,2.9)	Alg. 2.6	3	4	41	5.265699062e-001	7.032525080e+001
			Alg. 12.4.7	3	4	41	5.265698800e-01	——
HS76	4, 1, 7	(0.5,0.5,0.5,0.5)	Alg. 2.6	9	11	132	-4.681817808e+000	7.172925980e+001
			Alg. 12.4.7	9	11	132	-4.681817808e+00	——
HS100	7, 1, 4	(1,2,0,4,0,1,1)	Alg. 2.6	23	51	277	6.806300576e+002	7.341407060e+001
			Alg. 12.4.7	24	53	287	6.806300575e+02	——
HS113	10, 1, 8	(2,3,5,5,1,2,7,3,6,10)	Alg. 2.6	24	25	470	2.430620919e+001	7.500528080e+001
			Alg. 12.4.7	24	25	470	2.430620919e+01	——
HS225	2, 1, 5	(3,3)	Alg. 2.6	6	7	52	2.000000014e+000	7.600368720e+001
			Alg. 12.4.7	7	8	63	2.000000017e+000	——
HS264	4, 1, 3	(0,0,0,0)	Alg. 2.6	21	25	123	-4.399979017e-001	7.734529580e+001
			Alg. 12.4.7	20	24	117	-4.399978742e-001	——

5-2: Minimax 问题的数值试验

Prob	n,m,m'	Initial point	Method	Ni	Nf	Ng	$F(x^*)$	$\|d^k\|$	Time(s)
01	5,5,3	(1,1,1,1,1)	Alg. 2.6	3	48	40	0.5001	1.9485E-007	0.0776
			Alg. B	5	60	36	0.5	2.4926E-008	0.0982
	10,10,8	(1,1,...,1,1)	Alg. 2.6	5	110	120	0.5002	5.5518E-010	0.1804
			Alg. B	7	160	128	0.5	5.5518E-010	0.2592
02	5,5,4	(1,1,1,1,1)	Alg. 2.6	6	78	76	0.1056	1.4733E-007	0.1214
			Alg. B	9	100	80	0.11111	4.3645E-007	0.1680
	10,10,9	(1,1,...,1,1)	Alg. 2.6	7	138	125	0.1111	5.9056E-008	0.2178
			Alg. B	6	140	126	0.11111	6.9438E-008	0.2236
03	5,5,4	(0.5,0.5,0.5,0.5,0.5)	Alg. 2.6	17	190	152	1.2862E-012	5.6720E-007	0.2630
			Alg. B	17	190	152	1.2862E-012	5.6720E-007	0.2630
	10,10,9	(0.5,0.5,...,0.5,0.5)	Alg. 2.6	12	260	234	2.2685E-012	7.9805E-007	0.3359
			Alg. B	12	260	234	2.2685E-012	7.9805E-007	0.3359
04	5,2,3	(2,1,2,1,2)	Alg. 2.6	7	36	67	-2.7084	4.6562E-008	0.1356
			Alg. B	8	48	72	-3.3424	7.1599E-008	0.1302
	10,2,8	(2,1,...,2,1)	Alg. 2.6	11	64	240	-6.9799	2.7389E-008	0.1975
			Alg. B	9	56	224	-6.9799	8.8916E-007	0.2446
05	5,2,3	(2,2,2,2,2)	Alg. 2.6	11	78	117	-4.4812	8.9340E-008	0.1786
			Alg. B	11	78	117	-4.4812	4.3107E-008	0.1757
	10,2,8	(2,2,...,2,2)	Alg. 2.6	12	84	336	-1.9350	3.2923E-008	0.3134
			Alg. B	11	74	296	-1.9350	7.2578E-007	0.3059
06	5,2,4	(0,0,0,0,0)	Alg. 2.6	8	46	92	-4.6188	5.5460E-008	0.1368
			Alg. B	8	46	92	-4.6188	4.3647E-008	0.1317
	10,2,9	(0,0,...,0,0)	Alg. 2.6	6	36	162	-10.3923	3.7628E-011	0.1831
			Alg. B	6	36	162	-10.3923	4.0322E-011	0.1744
07	5,3,3	(2,1,2,1,2)	Alg. 2.6	15	165	165	10.2390	8.5759E-009	0.2150
			Alg. B	16	177	177	10.2390	2.4310E-007	0.2316
	10,3,8	(2,1,...,2,1)	Alg. 2.6	13	162	432	24.1074	4.1559E-007	0.3559
			Alg. B	15	174	464	24.1074	1.8294E-007	0.3872
08	5,3,4	(0.5,0.5,0.5,0.5,0.5)	Alg. 2.6	5	33	44	16.1915	8.6057E-007	0.0871
			Alg. B	7	48	64	16.1915	1.9231E-009	0.1107
	10,3,9	(0.5,0.5,...,0.5,0.5)	Alg. 2.6	4	27	81	36.4308	5.8544E-007	0.1333
			Alg. B	4	27	81	36.4308	5.6584E-007	0.1321
09	5,3,4	(0.5,0.5,0.5,0.5,0.5)	Alg. 2.6	10	93	124	8.0	8.3391E-008	0.1850
			Alg. B	11	105	140	8.0	3.5831E-007	0.2018
	10,3,9	(0.5,0.5,...,0.5,0.5)	Alg. 2.6	7	75	225	18.0	2.9088E-008	0.2436
			Alg. B	11	87	261	18.0	1.9480E-008	0.4072
10	5,2,3	(-1,-1,-1,-1,-1)	Alg. 2.6	5	54	110	0.7324	3.2420E-008	0.1978
			Alg. B	11	76	114	0.7324	3.2420E-008	0.1976
	10,2,8	(-1,-1,...,-1,-1)	Alg. 2.6	67	501	1872	3.3567	7.5496E-007	1.6188
			Alg. B	51	478	1912	3.3436	4.8727E-007	1.9282
11	5,2,4	(0.5,0.5,0.5,0.5,0.5)	Alg. 2.6	12	111	254	1.3480	2.2172E-008	0.4390
			Alg. B	17	126	252	1.3480	2.5237E-008	0.3191
	10,2,9	(0.5,0.5,...,0.5,0.5)	Alg. 2.6	26	175	810	3.4679	5.6557E-007	0.8827
			Alg. B	21	138	621	3.4679	4.6742E-007	0.7271
12	5,2,3	(2,2,2,2,2)	Alg. 2.6	11	76	114	1.4534	1.9721E-007	0.1689
			Alg. B	11	76	114	0.9396	2.1359E-008	0.1815
	10,2,8	(2,2,...,2,2)	Alg. 2.6	7	52	234	3.6860	8.2824E-007	0.3421
			Alg. B	14	90	360	16.1324	1.8729E-007	0.4126

5.3 本章小结

本章主要对算法进行一些初步的数值试验，从数值试验上看，本文的算法是有效的. 而且对于表 5-2 中的数值试验，从 CPU 时间上可以看出，本文算法的计算量相对来说是比较小的.

结论与展望

在本篇论文中提出了一个求解带非线性约束的 minimax 问题的新算法. 该算法在每次迭代的过程中, 可行下降方向通过解一个带简单二次约束的子问题. 用于避免 Maratos 效应高阶修正方向通过计算一个仅包括积极识别集中的约束函数及其梯度的线性方程组产生, 从而进一步减少高阶修正方向计算规模和计算量.

通过研究, 本论文得到如下的结论:

该文的算法在弱 Mangasarian-Fromovitz 约束规格 (MFCQ) 条件下具有全局收敛性. 在上层严格互补等较温和的条件下具有超线性收敛性. 最后, 进行了一些初步的数值试验. 数值试验表明, 对于所测试的问题, 本文的算法是有效的.

下面给出有待研究的问题和展望:

在本篇论文中研究的是初始点可行的可行方向法, 可以扩展到初始点任意的算法, 例如与强次可行方向法, 拟强次可行方向法等算法相结合, 得到应用背景更广泛的算法; 简单二次约束二次规划子问题 (2.5) 的特殊有效求解方法; 本文算法对中大规模实例的数值仿真试验有待于进一步研究深入.

参考文献

[1] Zoutendijk G. Methods of feasible direction[m]. Elsevier, Amsterdam, 1960.

[2] Wolfe P. Methods of fesible direction[m]. Elsevier, Amsterdam, 1960.

[3] Topkis D M, Veinott A F. On the convergence of some feasible direction algorithms for nonlinear programming[J]. SIAM Journal on Control and Optimization, 1967, 5: 268-279.

[4] Pironneau O, Polak E. Rate of convergence of a class of methods of feasibe direction[J]. SIAM Journal on Numerical Analysis, 1973, 10: 161-173.

[5] Cawood M E, Kostreva M M. Norm-relaxed method of feasible direction for solving the nonlinear programming problems[J]. Journal of Optimization Theory and Applications, 1994, 83: 311-320.

[6] Chen X, Kostreva M M. A generalization of the norm-relaxed method of feasible directions[J]. Applied Mathematics and Computation, 1999, 102: 257-273.

[7] Jian J B, Ke X Y, Zheng H Y, et al. A method combining norm-relaxed QP subproblems with systems of linear equations for constrained optimization[J]. Journal of Computation and Applied Mathematics, 2008, 223: 1013-1027.

[8] Maratos N. Exact penalty function algorithm for finite dimensional and control optimization problems[J]. Ph.D.thesis, Imperial College Science, Technology, University of London, 1978.

[9] Jian J B, Hu Q J, Tang C M, et al. A sequential quadratically constrained quadratic programming method of feasible directions[J]. Applied Mathematics Optimization, 2007, 56: 343-363.

[10] 简金宝. 光滑约束优化快速算法 —— 理论分析与数值试验 [M]. 北京：科学出版社，2010.

[11] Zhu Z B, Cai X, Jian J B. An improved SQP algorithm for solving minimax problems[J]. Applied Mathematics Letters, 2009, 22: 464-469.

[12] Wang F S, Zhang K C. A hybrid algorithm for nonlinear minimax problems[M]. Annals of Operations Research, 2008, 164: 167-191.

[13] Polak E, Womersley R S, Yin H X. An algorithm based on active sets and smoothing for discretized semi-infinite minimax problems[J]. Journal of Mathematical Analysis and Applications, 2008, 138: 311-328.

[14] Hu Q J, Chen Y, Chen N P, et al. A modified SQP algorithm for minimax problems[J]. Journal of Mathematical Analysis and Applications, 2009, 360: 211-222.

[15] Jian J B, Quan R, Zhang X L. Feasible generalized monotone line search SQP algorithm for nonlinear minimax problems with inequality constraints[J]. Journal of Computational and Applied Mathematics, 2007, 205: 406-429.

[16] Jian J B, Zhang X L, Quan R, et al. Generalized monotone line seach SQP algorithm for constrained minimax problems[J]. Optimization, 2009, 58(1): 101-131.

[17] Chao M T, Wang Zh X, Liang Y M, et al. Quadratically constraint quadratical algorithm model for nonlinear minimax problems[J]. Applied Mathematics and Computation, 205 (2008) 247-262

[18] Li J. A superlinearly convergent norm-relaxed method of quasi-strongly sub-feasibe direction for inequality constrained minimax problems[D]. Guangxi: Guangxi University, 2009.

[19] Facchinei F, Fischer A, Kanzow C. On the accurate identification of active constraints[J]. SIAM Journal on Optimization, 1998, 9(1): 14-32.

[20] Jian J B, Liu Y. A superlinearly convergent method of quasi-strongly sub-feasible directions with active set identifying for constrained optimization[J]. Nonlinear Analysis, 2011.

[21] Han D L, Jian J B, Li J. On the accurate identification of active set for constrained minimax problems[J]. Nonlinear Analysis, 2011, 74: 3022-3032.

[22] Luksan L, Matonoha C, Vlcek J. Primal interior point method for minimization of generalized minimax function[J]. Institute of Computer Science Academy of Sciences of the Czech Republic, 2007.

[23] Wrobel M. Minimax Optimization without second order information[m]. Eksamensprojekt nr.6, lyngby 2003.

[24] Jian J B, Quan R, Hu Q J. A new superlinearly convergent SQP algorithm for nonlinear minimax problems[J]. Acta Mathematicas Applicate Sinica, English Series, 2007, 23(3). 395-410.

[25] Powell M J D. A fast algorithm for nonlinearly constrained optimization calculations[J]. in Numerical analysis, G.A.Watson, ed., Lecture Notes in Math, Springer, Berlin, 1978, 630: 144-157.

[26] MOSEK ApS. The MOSEK optimization toolbos for MATLAB manual, Version 5.0 (Revision 79). http://www.mosek.com/.

[27] Hock W, Schittkowski K. Test Examples for Nonlinear Programming Codes[M]. Lecture Notes in Economics and Mathematical Systems, 187. Springer, Berlin 1981.

[28] Karmitsa N. Test problems for large-scale nonsmooth minimization[J]. Reports of the Department of Mathematical Information Technology, Series B.Scienctific Computing, B. 2007, 4.

致 谢

时光荏苒，岁月如梭，在不知不觉中三年的研究生生活已接近尾声，在这短短的三年硕士学习生涯中，我深切的感受到我是在导师简金宝教授的亲切关怀下，在众多师友们的帮助下一步一个脚印走过来的，心中的感激之情溢于言表.

首先我要衷心地感谢我的导师简金宝教授. 简老师无论是在学习上还是在生活上都殷切的关怀着我，在我论文的一次次修改过程中，简老师都废寝忘食，每次看到论文修改稿上那一个个鲜红的标注，大到论文理论的修改，小到标点符号的注示，简老师都一丝不苟的帮助我们修改. 每一次听到简老师那一声声“孩子们”的呼唤，都让我倍感温暖. 简老师用他那渊博的专业知识，敬业的工作作风，非凡的人格魅力，始终鼓舞着我、激励着我，让我不断的奋发图强，不断的在专业领域中学习. 简老师不仅授我以文，而且教我为人，让我养成了谨慎的工作态度. 我对简老师的感激之情是无法用言语来表达的，谨在此郑重说一声“谢谢”!

衷心感谢我的师母江秀丽医生，感谢她默默为团队的付出和关怀，感谢她对我们象母亲一样的关怀.

衷心感谢我的师姐韩道兰，师兄唐春明对我论文程序的帮助. 感谢李杰师兄，刘逸师兄对我的帮助. 感谢数信学院黎健玲老师，感谢我的师兄全然、晁绵涛和师姐曾友芳、柯晓艳等，感谢我的两位同门杨素敏和莫兴德，还有研一、研二的师弟、师妹们. 感谢他们在这三年里，对我无论是学习，还是生活的无私关怀和帮助. 这些支持与帮助当然离不开导师强有力的科研团队定期的研讨班，让我在每个星期短短的几个小时内吸收到更多的知识营养.

衷心感谢数信学院的领导和老师们，是他们为我提供了美好的学习环境. 衷心感谢数信学院 08 级的全体研究生.

感谢我最最挚爱的父母，感谢生我养我的美丽的家乡，感谢我的亲友和朋友，是他们让我无后顾只忧的学习. .

最后，要感谢导师主持的国家自然科学基金项目 (No. 71061002) 和广西自然科学基金项目 (No. 2100GXNSFD018022) 对论文的资助. 感谢本文中所引用文献的编、著、译者.

“书到用时方恨少”，心中有太多太多的感激之情要抒发，但发现情到却言轻，最后只好对所有我要感激的人说一声谢谢. 祝愿你们身体健康，快乐幸福！

攻读学位期间发表论文情况

简金宝, 杨素敏, 邱丽娟等, 约束优化一个初始点任意的 SSQCQP 算法, 袁亚湘, 胡晓东, 吴凌云, 刘德刚编辑, 中国运筹学会第十届学术交流会论文集, Global-Link Informatics Limited, Hong Kong, 2010, 80-85.