

约束优化问题稳定序列二次规划方法研究综述*

An Overview of the Researches on Stabilized Sequential Quadratic Programming Methods for Constrained Optimization Problems

刘美杏, 简金宝**

LIU Meixing, JIAN Jinbao

(玉林师范学院, 复杂系统优化与大数据处理广西高校重点实验室, 广西玉林 537000)

(Guangxi Colleges and Universities Key Lab of Complex System Optimization and Big Data Processing, Yulin Normal University, Yulin, Guangxi, 537000, China)

摘要: 稳定序列二次规划(sSQP)方法由于在求解病态或退化约束优化问题获得理论与数值的突破性进展而备受关注, 重要成果频繁问世. 本文对近期国际上若干重要 sSQP 方法及其思想进行概述, 包括罚函数型 sSQP 方法, 滤子型 sSQP 方法和非精确恢复(IR)型 sSQP 方法等, 并对约束优化问题 sSQP 方法的进一步研究进行探索性思考.

关键词: 约束优化问题 稳定序列二次规划 收敛速度

中图分类号: O221.2 **文献标识码:** A **文章编号:** 1005-9164(2016)05-0385-07

Abstract: The stabilized sequential quadratic programming (sSQP) methods attract great attention with respect to the theoretical and numerical breakthrough for solving ill-posed or degenerate constrained optimization problems, and many important references about sSQP methods were published. This paper gives an overview on some important sSQP methods, which mainly include penalty function type sSQP methods, filter type sSQP methods and inexact restoration (IR) type sSQP methods, and a few exploratory considerations for further study on sSQP methods are given.

Key words: constrained optimization problems, stabilized sequential quadratic programming, convergence rate

收稿日期: 2016-07-01

修回日期: 2016-08-27

作者简介: 刘美杏(1987-), 女, 硕士, 主要从事最优化理论与算法研究.

* 国家自然科学基金项目(11271086), 广西自然科学基金项目(2014GXNSFFA118001), 广西高校科研项目(ZD201407)和复杂系统优化与大数据重点实验室开放基金项目(2015CSOBDP0203)资助.

** 通信作者: 简金宝(1964-), 男, 教授, 博士, 博士生导师, 主要从事最优化理论与算法及应用研究, E-mail: jianjb@gxu.edu.cn.

0 引言

稳定序列二次规划(sSQP)方法作为序列二次规划(SQP)方法的重要扩展, 致力于考虑子问题模型、假设条件、技术构造、收敛性质(全局收敛或收敛速度)与数值效果等方面的研究. 尤其是对相对弱条件下的收敛速度与数值效果. 值得注意的是, 对于退化约束优化问题, 当对应原始解的拉格朗日乘子不唯一时, 快速收敛性的实现难度极大. 自 1998 年, Wright^[1]首次提出求解退化不等式约束优化的 sSQP 方法以来, 以 Wright, Hager, Izmailov,

Fernández, Solodov, Gill 及 Robinson 等为代表的 sSQP 方法及理论研究发展迅速,取得系列成果. 文献[1]的方法在每次迭代时需求解一个目标函数二次的极大极小问题,可等价转化为原始对偶空间上的稳定二次规划(QP)问题. 该作者在二阶充分最优性条件(SOSC)、Mangasarian-Fromovitz 约束规格(MFCQ)和严格互补等条件下,证明了 sSQP 方法的局部二次收敛性. 文献[2]进一步改进文献[1]的 sSQP 方法,在 SOSC 和 MFCQ 等较弱的条件下,证明算法具备局部二次收敛性. Hager^[3]对收敛性假设条件进行研究,仅在 SOSC 条件下获得文献[1]中算法的局部收敛性,并提出用一对不等式约束来表示等式约束,从而可将方法推广到一般约束优化问题. 相比较,传统 SQP 方法对乘子唯一性有着较苛刻的要求,只能依靠严格 MFCQ 条件来保证. 特别地,在等式约束情形下,迫使线性无关约束规格(LICQ)成立^[4]. 文献[5]将 sSQP 方法推广到求解含等式与不等式约束的变分问题,当初始点充分靠近原始对偶稳定点时,仅需 SOSC 假设条件,实现了算法的超线性收敛性. 对于等式约束优化问题,文献[6]在非临界乘子的条件下建立了超线性收敛的 sSQP 方法. 文献[7]基于线性方程组,在无需 LICQ 和严格互补等较强的假设条件下,提出一个二次收敛的牛顿型 sSQP 方法,并给出了该方法超线性收敛的重要定理. 文献[8]将拟牛顿型 sSQP 方法拓广到变分不等式问题,通过对 Bregman 距离极小化来更新矩阵信息,这样仅需要 SOSC 便可证明算法的超线性收敛性.

上述文献主要聚焦在最优解的局部范围内构造有效算法,并研究不同假设条件对其收敛速度的影响,建立了一批局部超线性收敛或二次收敛的 sSQP 方法. 这些文献都侧重于局部收敛速度的分析,对算法全局优化策略(全局收敛性质)未做深入研究. 而此问题正是实际应用和优化学者们希冀解决的问题,也是衡量最优化算法有效性的重要指标之一. 近年来,尽管国内外不少学者对全局化 sSQP 方法潜心研究,但成果有限. 本文主要介绍如下几类的全局 sSQP 算法:罚函数型 sSQP 方法^[9-12]、滤子型 sSQP 方法^[13]和 IR 型 sSQP 方法^[14],并对约束优化问题 sSQP 方法的进一步研究进行探索性思考.

1 函数型 sSQP 方法

当前,对 sSQP 方法全局化策略的研究仍是一项极具挑战性的工作,其中约束优化 sSQP 方法全局化时罚函数的选取极其困难,这对于最优性或下降性影响甚大. Gill 等^[9]引入原始对偶广义增广拉格朗

日(AL)函数,提出了一个求解等式约束加简单界约束优化问题的全局收敛 sSQP 方法. Izmailov 等^[11]充分利用 AL 方法的鲁棒性,在没有任何积极约束识别策略下构造了一个 sSQP 方法,并证明了算法的全局收敛性和局部收敛速度. 随后,他们又在文献[12]中以原始对偶精确罚函数^[15]为效益函数,提出了一个求解等式约束优化问题的 sSQP 方法,并分析了算法的全局收敛性和超线性收敛速度. 下面参考文献[12],介绍一个具体的罚函数型 sSQP 算法.

考虑等式约束优化问题:

$$\begin{aligned} \min f(x) \\ \text{s. t. } h(x) = 0, \end{aligned} \quad (1)$$

其中 $f(x):R^n \rightarrow R, h(x) = (h_1(x), \dots, h_l(x))^T: R^n \rightarrow R^l$ 至少是二阶可微函数. 令问题(1)的 Lagrange 函数为

$$L(x, \lambda) = f(x) + \langle \lambda, h(x) \rangle, \quad (2)$$

其中,符号 $\langle \cdot, \cdot \rangle$ 表示向量内积运算. 则问题(1)对应的一阶最优性条件为

$$\frac{\partial L}{\partial x}(x, \lambda) = 0, h(x) = 0. \quad (3)$$

进而,定义函数 $\Phi:R^n \times R^l \rightarrow R^n \times R^l$,

$$\Phi(x, \lambda) = \left(\frac{\partial L}{\partial x}(x, \lambda), h(x) \right). \quad (4)$$

对一个给定的原始对偶迭代点 $(x, \lambda) \in R^n \times R^l$ 及稳定参数 $\sigma > 0$,考虑如下原始对偶空间上的 sSQP 子问题,并产生 sSQP 方向 (ξ, η) ,

$$\begin{aligned} \min_{(\xi, \eta)} \quad & \langle f'(x), \xi \rangle + \frac{1}{2} \langle \frac{\partial^2 L}{\partial x^2}(x, \lambda) \xi, \xi \rangle + \frac{\sigma}{2} \| \lambda + \eta \|^2 \\ \text{s. t. } \quad & h(x) + h'(x)\xi - \sigma\eta = 0, \end{aligned} \quad (5)$$

其中 $h'(x) = (\nabla h_1(x), \dots, \nabla h_l(x))^T$.

基于上述 sSQP 方向(搜索方向),文献[12]采用如下原始对偶效益函数^[15] $\varphi_{c_1, c_2}:R^n \times R^l \rightarrow R$,

$$\begin{aligned} \varphi_{c_1, c_2}(x, \lambda) = L(x, \lambda) + \frac{c_1}{2} \| h(x) \|^2 + \\ \frac{c_2}{2} \left\| \frac{\partial L}{\partial x}(x, \lambda) \right\|^2. \end{aligned} \quad (6)$$

此处选取适当罚参数 $c_1 > 0, c_2 > 0$,可以保证由于子问题(5)产生的 sSQP 方向是 $\varphi_{c_1, c_2}(x, \lambda)$ 的下降方向.

由文献[16]和文献[17]知,原始对偶效益函数(6)是一个精确罚函数. 即如果 $c_2 > 0$ 充分小, $c_1 > 0$ 充分大,则 $\varphi_{c_1, c_2}(x, \lambda)$ 任意稳定点都是问题(1)的稳定点;反之,如果最优性系统(3)的原始对偶解 (x, λ) 满足 LICQ 和 SOSC 等条件,则它必是罚函数 $\varphi_{c_1, c_2}(x, \lambda)$ 的严格局部极小解.

为便于分析和陈述算法的收敛性条件,下面复

述二阶充分条件(SOSC)(文献[12])以及临界 Lagrangian 乘子和非临界 Lagrangian 乘子^[6,18-19]等定义. 记 $\mathcal{M}(\bar{x})$ 是由问题(1)在稳定点 \bar{x} 处对应的 Lagrangian 乘子构成的集合. 记号 $\ker A, \text{im } A$ 分别表示矩阵 A 的零空间和列空间, 即 $\ker A = \{d: Ad = 0\}, \text{im } A = \{d: d = Ay\}$.

定义 1 称 $\bar{\lambda} \in \mathcal{M}(\bar{x})$ 满足二阶充分条件(SOSC), 如果

$$\left\langle \frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) \xi, \xi \right\rangle > 0, \forall \xi \in \ker h'(\bar{x}) \setminus \{0\}. \quad (7)$$

定义 2 称 $\bar{\lambda} \in \mathcal{M}(\bar{x})$ 是临界乘子, 如存在 $\xi \in \ker h'(\bar{x}) \setminus \{0\}$, 使得 $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}) \xi \in \text{im}(h'(\bar{x}))^\top$, 否则, 称 $\bar{\lambda}$ 是非临界乘子.

显然, 如 $\bar{\lambda} \in \mathcal{M}(\bar{x})$ 满足 SOSC, 则其是非临界乘子, 故非临界乘子条件弱于 SOSC.

算法 1 (文献[12]中 sSQP 算法)

初始步 选取参数 $\sigma > 0, c_1 > 0, c_2 > 0, C_1 > 0, C_2 > 0, \delta > 0, \rho > 0, q > 1, \epsilon, \theta \in [0, 1]$. 固定连续函数(除了 0 点外函数值几乎处处为正) $\phi_1, \phi_2, \phi_3: R_+ \rightarrow R_+$. 选取原始对偶初始点 $(x^0, \lambda^0) \in R^n \times R^l$, 置 $k := 0$.

步骤 1 由(4)式计算 $\Phi_k \triangleq \Phi(x^k, \lambda^k)$, 若 $\Phi_k = 0$, 终止.

步骤 2 令 $\sigma_k = \min\{\sigma, \|\Phi_k\|\}, \sigma = \sigma_k, (x, \lambda) = (x^k, \lambda^k)$, 求解子问题(5)的 KKT 系统得到稳定点 $d^k = (\xi^k, \eta^k)$. 若该系统无解, 则进入步骤 4.

步骤 3 (i) 如果

$$\langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle \leq -\rho \|d^k\|^q, \quad (8)$$

则进入步骤 5.

(ii) 如果

$$\|h(x^k)\| \geq \phi_1(\sigma_k) \quad (9)$$

且

$$\langle h(x^k), h'(x^k)\xi^k \rangle \leq -\phi_2(\|h(x^k)\|), \quad (10)$$

则令 $c_1 = \bar{c}_{1,k} + \delta$, 其中 $\bar{c}_{1,k} = \bar{c}_1(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k)$. 如果 $c_1 \leq C_1$, 进入步骤 5; 否则, 令 $c_1 = C_1$, 进入步骤 4.

(iii) 如果

$$\left\| \frac{\partial L}{\partial x}(x^k, \lambda^k) \right\| \geq \phi_3(\sigma_k), \quad (11)$$

则令 $c_2 = \bar{c}_{2,k} + \delta$, 其中 $\bar{c}_{2,k} = \bar{c}_2(\rho \|d^k\|^q; x^k, \lambda^k; \xi^k, \eta^k, c_1)$. 如果 $c_2 \leq C_2$, 进入步骤 5; 否则, 令 $c_2 = C_2$.

步骤 4 选取 $(n+l)$ 阶对称正定矩阵 Q_k , 计算方向 $d^k = -Q_k \varphi'_{c_1, c_2}(x^k, \lambda^k)$.

步骤 5 计算 $\alpha_k = \theta^j$, 此处 j 为满足不等式

$$\varphi_{c_1, c_2}(x^k, \lambda^k) + \theta^j d^k \leq \varphi_{c_1, c_2}(x^k, \lambda^k) + \theta^j \langle \varphi'_{c_1, c_2}(x^k, \lambda^k), d^k \rangle \quad (12)$$

的最小非负整数.

步骤 6 令 $(x^{k+1}, \lambda^{k+1}) = (x^k, \lambda^k) + \alpha_k d^k, k := k+1$, 返回步骤 1.

在算法 1 的迭代过程中, 常量值 C_1, C_2 的设置对该算法的执行有着重要的防御作用. 一般地, 这两个数值应该取足够大, 方能保证罚参数 c_1, c_2 的更新法则不受影响. 为了增强算法设计的有效性, 当子问题(5)KKT 系统无解或步骤 3 的任一情形都不被执行时, 算法 1 执行步骤 4 的拟牛顿步防御措施, 以修正 sSQP 方向使之具有下降性, 从而算法具有良好的稳定性和全局收敛性质. 当初始点充分靠近稳定点时, 仅需要非临界 Lagrangian 乘子假设, 即可证明算法 1 的超线性收敛. 对退化测试问题, 算法 1 表现出了良好的数值效果.

2 滤子型 sSQP 方法

滤子法是近 20 年提出的求解非线性约束优化的一种有效方法. 该方法的提出是为了避免在实际问题中设置罚参数. 早期滤子法的研究归功于 Fletcher 和 Leyffer^[20], 随后该方法因良好的数值效果而备受关注, 近期代表成果见文献[21-23]等. 文献[13]对稳定 QP 子问题进行修正, 并结合双滤子技术^[21]提出了求解一般约束优化问题的滤子型 sSQP 方法. 下面对文献[13]中的算法进行分析.

考虑如下一般非线性规划问题:

$$\begin{aligned} \min & f(x) \\ \text{s. t. } & h_\epsilon(x) = 0, \\ & h_I(x) \leq 0, \end{aligned} \quad (13)$$

其中, $h_\epsilon(x) = (h_i(x), i \in \epsilon), h_I(x) = (h_i(x), i \in I), \epsilon = \{1, 2, \dots, l\}, I = \{l+1, l+2, \dots, m\}, f: R^n \rightarrow R, h_i: R^n \rightarrow R, i \in \epsilon \cup I$ 为连续可微函数. 定义问题(13)的 Lagrangian 函数 $L(x, \mu)$ 为

$$L(x, \mu) = f(x) + \sum_{i \in \epsilon \cup I} \mu_i h_i(x). \quad (14)$$

这里 $\mu = (\mu_1, \mu_2, \dots, \mu_m)^\top$ 是 Lagrangian 乘子向量.

对于当前给定的原始对偶估计迭代点 $(x^k, \mu^{L,k})$, 求解如下修正的稳定 QP 子问题:

$$\begin{aligned} \min_{(d, \Delta \lambda)} & f'(x^k)^\top d + \frac{1}{2} d^\top B^k d + \frac{\sigma^{L,k}}{2} \|\mu^{L,k} + \Delta \lambda\|^2 \\ \text{s. t. } & h_\epsilon(x^k) + h'_\epsilon(x^k)^\top d - \sigma^{L,k} \Delta \lambda_\epsilon = 0, \\ & h_I(x^k) + h'_I(x^k)^\top d - \sigma^{L,k} \Delta \lambda_I \leq 0. \end{aligned} \quad (15)$$

其中, $\sigma^{L,k}$ 是稳定参数, B^k 是 $\nabla_{xx}^2 L(x^k, \mu^{L,k})$ 的近似对称正定矩阵. 当 $B^k = \nabla_{xx}^2 L(x^k, \mu^k)$, $\mu^{L,k} = \mu^k$, $\sigma^{L,k} = \sigma(x^k, \mu^k)$ 时, 其中

$$\sigma(x^k, \mu^k) = \left\| \begin{pmatrix} \nabla_x L(x^k, \mu^k) \\ h_\varepsilon(x^k) \\ \min\{-h_I(x^k), \mu_I^k\} \end{pmatrix} \right\|, \quad (16)$$

问题(15) 是一个传统的稳定 QP 子问题模型. 源于算法良好的收敛性质, 设计了两种乘子更新方式, 其中 $\mu^{L,k}$ 保证全局收敛, μ^k 则满足局部收敛的需求. 类似地, 考虑 B^k 和 $\sigma^{L,k}$ 的更新准则.

基于问题(15) 的解 $(d^k, \Delta\lambda^k)$, 定义原始对偶搜索方向 $(d^k, \Delta\mu^k)$ (此处 $\Delta\mu^k = \Delta\lambda^k + \mu^{L,k} - \mu^k$), 此方向对算法收敛性理论分析有着重要的作用, 其等价于求解如下相容的稳定 QP 子问题:

$$\begin{aligned} \min_{(d, \Delta\mu)} \quad & f'(x^k)^\top d + \frac{1}{2} d^\top B^k d + \frac{\sigma^{L,k}}{2} \|\mu^{L,k} + \Delta\mu\|^2 \\ \text{s. t.} \quad & h_\varepsilon(x^k) + h'_\varepsilon(x^k)^\top d - \sigma^{L,k}(\mu_\varepsilon^k + \Delta\mu_\varepsilon - \mu_\varepsilon^{L,k}) = 0, \\ & h_I(x^k) + h'_I(x^k)^\top d - \sigma^{L,k}(\mu_I^k + \Delta\mu_I - \mu_I^{L,k}) \leq 0. \end{aligned} \quad (17)$$

在设计非线性规划问题(13) 的有效算法时, 除需要对目标函数进行极小化, 同时要不断降低约束违反度函数:

$$\tilde{h}(x) = \sum_{i=1}^l |h_i(x)| + \sum_{i=l+1}^m \max\{h_i(x), 0\}. \quad (18)$$

然而, 搜索方向 $(d^k, \Delta\mu^k)$ 难以直接达到这样的要求. 因此, 引入如下辅助目标函数 $\Phi(x, \mu)$ 和松弛约束违反度函数 $p(x, \mu)$:

$$\Phi(x, \mu) = f(x) + \frac{\sigma^{L,k}}{2} \|\mu\|^2, \quad (19)$$

$$p(x, \mu) = \sum_{i=1}^l |h_i(x) - \sigma^{L,k}(\mu_i - \mu_i^{L,k})| + \sum_{i=l+1}^m \max\{h_i(x) - \sigma^{L,k}(\mu_i - \mu_i^{L,k}), 0\}. \quad (20)$$

进而引进双滤子技术(基于文献[21], 但有所改进), 包括“全局滤子”和“局部滤子”. 前者致力于算法收敛于 KKT 点或稳定点, 后者以实现其局部收敛速度. 对于第 k 次迭代, “全局滤子”和“局部滤子”分别记为 $\mathcal{F}_g^k, \mathcal{F}_l^k$.

定义 3^[13] 称点对 $(\hat{x}, \hat{\mu})$ 被“全局滤子”元 (x^l, μ^l) 或 $(p(x^l, \mu^l), \Phi(x^l, \mu^l))$ 接受, 如果它满足如下两个不等式之一:

$$p(x^l, \mu^l) - p(\hat{x}, \hat{\mu}) \geq \gamma_1 [p(\hat{x}, \hat{\mu})]^2, \quad (21)$$

$$\Phi(x^l, \mu^l) - \Phi(\hat{x}, \hat{\mu}) \geq \gamma_2 [p(\hat{x}, \hat{\mu})]^2, \quad (22)$$

此处 $\gamma_1, \gamma_2 \in (0, 1)$.

定义 4^[13] 称点对 $(\hat{x}, \hat{\mu})$ 为“全局滤子”元, 如果它被“全局滤子”所有元素 $(p(x^l, \mu^l), \Phi(x^l, \mu^l))$ 接受.

定义 5^[13] 称点对 $(\hat{x}, \hat{\mu})$ 被“局部滤子”元 (x^l, μ^l) 或 $(h(x^l), f(x^l))$ 接受, 如果它满足下面两个不等式之一:

$$\tilde{h}(x^l) - \tilde{h}(\hat{x}) \geq \max\{\gamma_1 \tilde{h}(\hat{x})^2, \min\{\gamma_3, \sigma(x^l, \mu^l)\}\}, \quad (23)$$

$$f(x^l) - f(\hat{x}) \geq \max\{\gamma_2 \tilde{h}(\hat{x})^2, \min\{\gamma_3, \sigma(x^l, \mu^l)\}\}, \quad (24)$$

此处 $\gamma_3 > 0$ 是一个常数.

定义 6^[13] 称点对 $(\hat{x}, \hat{\mu})$ 为“局部滤子”元, 如果它被“局部滤子”所有元素 $(h(x^l), f(x^l))$ 接受.

对于问题(13) 及当前迭代点对 (x^k, μ^k) , 求解稳定 QP 子问题(17) 得到搜索方向 $(d^k, \Delta\mu^k)$, 并沿着此方向和步长 α , 分别定义函数 Φ 的实际下降量和线性预测下降量:

$$\begin{aligned} \Delta\Phi^k(\alpha) &= \Phi(x^k, \mu^k) - \Phi(x^k + \alpha d^k, \mu^k + \alpha \Delta\mu^k), \\ \Delta l_\Phi^k(\alpha) &= -\alpha [f'(x^k)^\top d^k + \sigma^{L,k}(\mu^k)^\top \Delta\mu^k]. \end{aligned}$$

令 $\hat{x} = x^k + \alpha d^k, \hat{\mu} = \mu^k + \alpha \Delta\mu^k$, 当试探步长 $\alpha = 1$ 时, 定义 KKT 误差下降条件:

$$\vartheta \sigma(\hat{x}, \hat{\mu}) \leq \tau \sigma^{L,k}, \tau \in (0, \frac{1}{2}), \vartheta > 0. \quad (25)$$

此外, 文献[13] 结合“全局滤子”接受条件, 进一步利用回溯技术选择了恰当的步长 α^k , 判断条件和步长下界.

开关条件:

$$\Delta l_\Phi^k(\alpha^k) \geq \delta [p(x^k, \mu^k)]^2, \delta \in (0, \infty). \quad (26)$$

充分下降条件:

$$\begin{aligned} \Delta\Phi^k(\alpha^k) &\geq \tau \Delta l_\Phi^k(\alpha^k), \\ \alpha_{\min}^k &= \end{aligned} \quad (27)$$

$$\begin{cases} \min\{\gamma_1, -\frac{\min\{\gamma_2 p(x^k, \mu^k), \delta [p(x^k, \mu^k)]^2\}}{f'(x^k)^\top d^k + \sigma^{L,k}(\mu^k)^\top \Delta\mu^k}\}, \\ \text{若 } f'(x^k)^\top d^k + \sigma^{L,k}(\mu^k)^\top \Delta\mu^k < 0; \\ \underline{\alpha}, \text{ 否则.} \end{cases} \quad (28)$$

其中, 常数 $\underline{\alpha} \in (0, 1)$.

对于原始对偶迭代点对 $(\hat{x}, \hat{\mu})$, 如果它被局部滤子接受, 或者满足 KKT 误差下降条件(25), 或者满足如下近似松弛一阶条件:

$$\psi(\hat{x}, \hat{\mu}) \leq \epsilon_\sigma, \quad (29)$$

其中,

$$\psi(\hat{x}, \hat{\mu}) := \left\| \begin{pmatrix} \nabla_x L(\hat{x}, \hat{\mu}) \\ h_\varepsilon(\hat{x}) - \sigma^{L,k}(\hat{\mu}_\varepsilon - \mu_\varepsilon^{L,k}) \\ \min(-h_I(\hat{x}) + \sigma^{L,k}(\hat{\mu}_I - \mu_I^{L,k}), \hat{\mu}_I) \end{pmatrix} \right\|, \quad (30)$$

则序列 $\{\epsilon_g^k\}$ 下降趋于 0. 构造 Lagrange 乘子 $\mu^{L,k}$ 和稳定参数 $\sigma^{L,k}$ 的更新准则:

$$\mu^{L,k+1} := \begin{cases} \hat{\mu}, & \text{若 } \|\hat{\mu}\|_\infty \leq \mu_{\max}, \\ \mu^{L,k}, & \text{否则.} \end{cases} \quad (31)$$

$$\sigma^{L,k+1} = \min\{\beta\sigma^{L,k}, \vartheta\sigma(\hat{x}, \hat{\mu})\}, \quad (32)$$

其中, $\mu_{\max} > 0, \beta \in (0, 1)$ 是两个常数.

下面给出求解问题(13)的滤子型 sSQP 算法.

算法 2 (文献[13]中 sSQP 算法之外循环)

初始步 选取初始点 $x^0 \in R^n$, 参数 $\epsilon_g^0 > 0, u_i > 0, u_g > 0, \delta > 0, \gamma_3 > 0, \mu_{\max} > 0, \alpha \in (0, 1), \gamma_1 \in (0, 1), \gamma_2 \in (0, 1), r \in (0, 1), \tau \in (0, \frac{1}{2}), \beta \in (0, 1), \mu^{L,0} = 0, \mu^0 = 0, \vartheta > 0, \sigma^{L,0} > 0, \mathcal{F}_i^0 = \{(u_i, -\infty)\}, \mathcal{F}_g^0 = \{(u_g, -\infty)\}, \text{Flag} = \text{GLOBAL}, \text{置 } k := 0.$

步骤 1 如果 (x^k, μ^k) 是问题(13)的 KKT 点或不可行稳定点, 则终止.

步骤 2 求解稳定 QP 子问题(17), 产生原始对偶解 $(d^k, \Delta \mu^k)$, 令 $\alpha = 1, \hat{x} = x^k + \alpha d^k, \hat{\mu} = \mu^k + \alpha \Delta \mu^k$.

步骤 3 如果 $(\hat{x}, \hat{\mu})$ 被局部滤子接受或满足(25)式, 则由准则(31)和(32)更新乘子 $\mu^{L,k+1}$ 和稳定参数 $\sigma^{L,k+1}$, 置 $\text{Flag} = \text{LOCAL}, \mathcal{F}_g^{k+1} = \{(u_g, -\infty)\}$. 否则, 进入步骤 4.

步骤 4 启动算法 3(算法 2 之内循环), 获得点对 $(\hat{x}, \hat{\mu})$ 和步长 α . 如果 $\alpha < \alpha_{\min}^k$, 则令 $\mu^{k+1} = \mu^{L,k}$, 进入可行性恢复阶段并找到新的迭代点 x^{k+1} , 使得 (x^{k+1}, μ^{k+1}) 被全局滤子 \mathcal{F}_g^k 接受. $\mu^{L,k+1}, \sigma^{L,k+1}, \epsilon_g^{k+1}, B^{k+1}$ 保持不变. 如果 $p(x^k, \mu^k) > 0$, 则令 $k := k + 1$, 将 $(p(x^k, \mu^k), \Phi(x^k, \mu^k))$ 加入 \mathcal{F}_g^k .

如果点对 $(\hat{x}, \hat{\mu})$ 满足(29)式, 则由准则(31)和(32)更新 $\mu^{L,k+1}$ 和 $\sigma^{L,k+1}$, 令 $\mathcal{F}_g^{k+1} = \{(u_g, -\infty)\}, \epsilon_g^{k+1} = \epsilon_g^k / 10$. 否则, 令 $\mu^{L,k+1} = \mu^{L,k}, \sigma^{L,k+1} = \sigma^{L,k}$.

步骤 5 如果 $(\hat{x}, \hat{\mu})$ 被局部滤子接受, 则将 $(\bar{h}(\hat{x}), f(\hat{x}))$ 加入 \mathcal{F}_i^k , 更新局部滤子. 否则, 进入步骤 6.

步骤 6 令 $x^{k+1} = \hat{x}, \mu^{k+1} = \hat{\mu}, \alpha^k = \alpha, k := k + 1$, 且 $\text{Flag} = \text{GLOBAL}$, 利用拟牛顿公式更新矩阵 B^{k+1} , 返回步骤 1.

算法 3(算法 2 之内循环)

步骤 1 令 $\hat{x} = x^k + \alpha d^k, \hat{\mu} = \mu^k + \alpha \Delta \mu^k$. 如果 $\alpha \geq \alpha_{\min}^k$, 则进入步骤 2; 否则, 终止.

步骤 2 如果 $(\hat{x}, \hat{\mu})$ 被全局滤子接受, 则进入步骤 3. 否则令 $\alpha = r\alpha$, 返回步骤 1.

步骤 3 如果开关条件(26)不成立或充分下降条件(27)成立, 则终止. 否则令 $\alpha = r\alpha$, 返回步骤 1.

利用双滤子技术, 由算法 2 产生的所有迭代点均

能被滤子接受(全局滤子或局部滤子), 这对整个算法的收敛性质分析起到至关重要的作用. 在理论上, 不需要任何约束规格, 即可证明: 滤子型 sSQP 算法 2 不仅具有全局收敛性(存在收敛子列或收敛于 KKT 点, 或收敛于不可行稳定点), 而且在 SOSOC 下, 算法达到超线性收敛速度.

3 IR 型 sSQP 方法

sSQP 方法的研究最早可以追溯到 20 世纪末, 虽然在适当的假设条件或技术下早期的研究成果实现了算法快速的收敛速度, 但在理论上能实现全局收敛的成果较少. 滤子型 sSQP 方法需储存滤子信息, 会造成存储量大, 而且当迭代点远离可行域, 或者在迭代过程中当试探步长比既定下界要小时, 为寻找一个能被全局滤子接受的新迭代点, 需要执行可行性恢复阶段, 这无疑增加算法计算成本, 从而影响数值效果. 然而, IR 技术^[24-27]可以减少可行性恢复阶段的复杂性, 对 sSQP 算法的理论分析有着重要的作用. 文献[14]提出了求解含等式且变量有界约束优化的 IR 型 sSQP 方法. 下面详细介绍此算法.

考虑问题(1), 且满足约束条件 $x \in \Omega = \{x \in R^n \mid a \leq x \leq b\}$, 此处 $a, b \in R^n$. 定义问题(1)的自然残差 $\sigma: R^n \times R^l \rightarrow R$,

$$\sigma(x, \lambda) = \left\| \begin{bmatrix} P_\Omega(x - \frac{\partial L}{\partial x}(x, \lambda)) - x \\ h(x) \end{bmatrix} \right\|, \quad (33)$$

其中, P_Ω 表示在 Ω 上的正交投影, $L: R^n \times R^l \rightarrow R$ 是 Lagrangian 函数(2).

对于原始对偶迭代点对 (x^k, λ^k) , 罚参数 $\rho_k > 0$ 及 Q_k , 其中 Q_k 是问题(1)Lagrangian Hesse 的近似, 考虑如下拟牛顿型 sSQP 子问题:

$$\begin{aligned} \min_{(\xi, \lambda)} & \langle f'(x^k), \xi - x^k \rangle + \frac{1}{2} \langle Q_k(\xi - x^k), \xi - x^k \rangle + \frac{1}{2\rho_k} \|\lambda\|^2 \\ \text{s. t.} & \quad h(x^k) + h'(x^k)^\top(\xi - x^k) - \frac{1}{\rho_k}(\lambda - \lambda^k) = 0, \\ & \quad \xi \in \Omega. \end{aligned} \quad (34)$$

为构造全局收敛算法, 引入下面的辅助函数 $F_k(x, \lambda), H_k(x, \lambda): R^n \times R^l \rightarrow R$:

$$F_k(x, \lambda) = f(x) + \frac{1}{2\rho_k} \|\lambda\|^2, \quad (35)$$

$$H_k(x, \lambda) = h(x) - \frac{1}{\rho_k}(\lambda - \lambda^k). \quad (36)$$

选取点 $Y^k(x) = (x, \lambda^k + \rho_k h(x))$, 易知 $H_k(Y^k(x)) = 0, \forall x \in R^n$, 且对于 $Y^k := Y^k(x^k), Y := (x, \lambda)$, sSQP

子问题(34) 等价于如下 QP 子问题:

$$\min_Y \langle F'_k(Y^k), Y - Y^k \rangle + \frac{1}{2} \langle \begin{bmatrix} Q_k & 0 \\ 0 & \frac{1}{\rho_k} I \end{bmatrix} (Y - Y^k), Y - Y^k \rangle$$

$$\text{s. t. } H'_k(Y^k)^T (Y - Y^k) = 0, Y \in \Omega \times R^l. \quad (37)$$

注意到 $H_k(Y^k) = 0$. 不难发现上述 QP 子问题(37) 对应于优化问题:

$$\begin{aligned} \min_Y & F_k(Y) \\ \text{s. t. } & H_k(Y) = 0, Y \in \Omega \times R^l \end{aligned} \quad (38)$$

的 QP 近似子问题. 通过近似求解问题(37) 可获得 sSQP 方法的局部收敛性质^[3,5].

对每次迭代, IR 方法包括恢复阶段和极小化阶段. 在恢复阶段, 给定迭代点 X^k , 计算恢复点 Y^k , 避免了目标函数值以及可行性恶化. 在极小化阶段, 基于可行性和最优性构造了一个含罚参数的效益函数, 并沿着一阶可行方向进行线搜索.

下面给出 IR 型 sSQP 算法的具体步骤.

算法 4(文献[14] 中 sSQP 算法)

初始步 选取参数 $\bar{\gamma} \in (0, 1), \bar{r} \in (0, 1), \bar{\epsilon} > 0$, 序列 $\{\epsilon_k\}$ 满足 $\epsilon_k \searrow 0, \alpha_L, \alpha_U > 0$. 选取任意近似初始点 $X^0 = (x^0, \lambda^0) \in \Omega \times R^l, \rho_0 > 0, \psi_{-1} = \sigma(X^0)$, 置 $k := 0$. 对当前参数 ρ_k , 定义

$$\Pi_k(x, \lambda) = (x, \max\{-\alpha_L \sqrt{\rho_k} e, \min\{\lambda, \alpha_U \sqrt{\rho_k} e\}\}). \quad (39)$$

步骤 1 计算 $\sigma(X^k)$, 若 $\sigma(X^k) \leq \bar{\epsilon}$, 则终止.

步骤 2 (i) 令 $X^{k,0} = (x^{k,0}, \lambda^{k,0}) = (x^k, \lambda^k), \theta_0 \in (0, 1), Q_{k,0}$ 为对称正定矩阵, $j := 0$.

(ii) 计算 $Y^{k,j} = (x^{k,j}, \lambda^{k,j} + \rho_k h(x^{k,j}))$, 求解如下 QP 子问题, 得到最优解 $D^{k,j} \in R^n \times R^l$.

$$\min_D \langle F'_k(Y^{k,j}), D \rangle + \frac{1}{2} \langle \begin{bmatrix} Q_{k,j} & 0 \\ 0 & \frac{1}{\rho_k} I \end{bmatrix} D, D \rangle$$

$$\text{s. t. } H'_k(Y^{k,j})^T D = 0, Y^{k,j} = D \in \Omega \times R^l. \quad (40)$$

如果 $\|D^{k,j}\| < \epsilon_k$, 则令 $X^{k+1} = \Pi_k(Y^{k,j} + D^{k,j})$, 进入步骤 3.

(iii) 令 $\theta^{j+1} \in \{2^{-i} : i \in N \cup \{0\}\}$ 为满足下面不等式的最大值:

$$\Phi_k(Y^{k,j}, \theta^{j+1}) - \Phi_k(X^{k,j}, \theta^{j+1}) \leq \frac{\bar{r}-1}{2} \|H_k(X^{k,j})\|, \quad (41)$$

此处 $\Phi_k(X, \theta) = \theta F_k(X) + (1-\theta) \|H_k(X)\|$ 是效益函数.

(iv) 令 $t_j \in \{2^{-i} : i \in N \cup \{0\}\}$ 为满足下面不等式的最大值:

$$\Phi_k(Y^{k,j} + t_j D^{k,j}, \theta^{j+1}) - \Phi_k(X^{k,j}, \theta^{j+1}) \leq \frac{\bar{r}-1}{2} \|H_k(X^{k,j})\|. \quad (42)$$

(v) 令 $X^{k,j+1} = Y^{k,j} + t_j D^{k,j}$, 选取对称正定矩阵 $Q_{k,j+1}$, 令 $j := j + 1$, 返回步骤 2(ii).

步骤 3 更新罚参数, 令 $\psi_k = \min\{\psi_{k-1}, \sigma(X^k)\}$, 如果 $\|h(x^{k+1})\| > \bar{\gamma} \|h(x^k)\|$ 且 $\sigma(X^{k+1}) > \bar{\gamma} \psi_k$, 则令 $\rho_{k+1} = 10\rho_k$. 否则令 $\rho_{k+1} = \rho_k$. 令 $k := k + 1$, 返回步骤 1.

事实上, 基于 IR 方法的思想, 算法 4 无需实质性地修正 sSQP 子问题(34) 的结构, 只需求解子问题(37) 得到非精确解, 然后巧妙利用类似增广拉格朗日(AL) 罚参数更新策略(步骤 3), 即可获得算法的全局收敛性质. 在严格 MFCO 和 SOSOC 等条件下, 可保证罚参数序列是有界的, 进而算法继承了良好的局部收敛速度(线性收敛).

4 展望

本文对求解约束优化问题的若干 sSQP 方法的思想及算法作了一个较详细的概述, 主要介绍了早期 sSQP 方法的局部收敛成果; 近年来实现了全局收敛的罚函数型 sSQP 方法、滤子型 sSQP 方法和非精确恢复(IR) 型 sSQP 方法. 对 sSQP 方法还有如下一些问题值得思考和研究.

①如何修正原始对偶罚函数型 sSQP 算法, 在适当的假设条件下, 将方法扩展到一般约束优化问题, 并制成软件包使之得到广泛应用;

②现有的滤子型 sSQP 算法考虑内点法求解一个与传统 QP 子问题相当的修正稳定 QP 子问题, 无疑会在严格内点的选择和计算上遇到困难; 加之滤子技术需要进入可行性恢复阶段, 也会增加计算成本. 因此, 如何改进算法以减少计算量, 有待进一步探讨;

③目前基于 IR 技术构造 sSQP 方法时, 需要严格 MFCQ 和 SOSOC 等条件以实现算法线性收敛, 没有继承传统 sSQP 方法优点(具有快速收敛速度的鲁棒性), 如何修正罚参数以更新技术等, 建立超线性收敛的 IR 型 sSQP 方法值得进一步深入研究;

④求解大规模稀疏优化与工程优化(如庞杂的电气工程中的潮流问题、机组合问题) 的高效的 sSQP 方法有必要进一步探索.

参考文献:

- [1] WRIGHT S J. Superlinear convergence of a stabilized SQP method to a degenerate solution[J]. Computational Optimization and Applications, 1998, 11(3): 253-275.
- [2] WRIGHT S J. Modifying SQP for degenerate problems

- [J]. *SIAM Journal on Optimization*, 2002, 13(2): 470-497.
- [3] HAGER W W. Stabilized sequential quadratic programming[J]. *Computational Optimization and Applications*, 1999, 12(1/2/3): 253-273.
- [4] IZMAILOV A F, SOLODOV M V. *Newton-Type Methods for Optimization and Variational Problems*; Springer Series in Operations Research and Financial Engineering [M]. Switzerland: Springer International Publishing, 2014.
- [5] FERNÁNDEZ D, SOLODOV M. Stabilized sequential quadratic programming for optimization and a stabilized Newton-type method for variational problems[J]. *Mathematical Programming*, 2010, 125(1): 47-73.
- [6] IZMAILOV A F, SOLODOV M V. Stabilized SQP revisited[J]. *Mathematical Programming*, 2012, 133(1/2): 93-120.
- [7] LI D H, QI L. *A Stabilized SQP Method via Linear Equations*[R]. New South Wales; Mathematics Department, University of New South Wales, 2000.
- [8] FERNÁNDEZ D. A quasi-Newton strategy for the sSQP method for variational inequality and optimization problems[J]. *Mathematical Programming*, 2013, 137(1/2): 199-223.
- [9] GILL P E, ROBINSON D P. A globally convergent stabilized SQP method[J]. *SIAM Journal on Optimization*, 2013, 23(4): 1983-2010.
- [10] GILL P E, KUNGURTSSEV V, ROBINSON D P. A Globally Convergent Stabilized SQP Method; Superlinear Convergence[R]. UCSD Center for Computational Mathematics Technical Report CCoM-13-4, 2014.
- [11] IZMAILOV A F, SOLODOV M V, USKOV E I. Combining stabilized SQP with the augmented Lagrangian algorithm[J]. *Computational Optimization and Applications*, 2015, 62(2): 405-429.
- [12] IZMAILOV A F, SOLODOV M V, USKOV E I. Globalizing stabilized sequential quadratic programming method by smooth primal-dual exact penalty function [J]. *Journal of Optimization Theory and Applications*, 2016, 169(1): 148-178.
- [13] SHEN C G, ZHANG L H, LIU W. A stabilized filter SQP algorithm for nonlinear programming[J]. *Journal of Global Optimization*, 2016, 65(4): 677-708.
- [14] FERNÁNDEZ D, PILOTTA E A, TORRES G A. An inexact restoration strategy for the globalization of the sSQP method [J]. *Computational Optimization and Applications*, 2013, 54(3): 595-617.
- [15] DI PILLO G, GRIPPO L. A new class of augmented Lagrangians in nonlinear programming[J]. *SIAM Journal on Control and Optimization*, 1979, 17(5): 618-628.
- [16] BERTSEKAS D P. *Constrained Optimization and Lagrange Multiplier Methods*[M]. New York: Academic Press, 1982.
- [17] BERTSEKAS D P. Enlarging the region of convergence of Newton's method for constrained optimization[J]. *Journal of Optimization Theory and Applications*, 1982, 36(2): 221-252.
- [18] IZMAILOV A F, SOLODOV M V. On attraction of Newton-type iterates to multipliers violating second-order sufficiency conditions[J]. *Mathematical Programming*, 2009, 117(1/2): 271-304.
- [19] IZMAILOV A F, SOLODOV M V. Critical Lagrange multipliers; What we currently know about them, how they spoil our lives, and what we can do about it[J]. *TOP*, 2015, 23(1): 1-26.
- [20] FLETCHER R, LEYFFER S. Nonlinear programming without a penalty function[J]. *Mathematical Programming*, 2002, 91(2): 239-269.
- [21] SHEN C G, LEYFFER S, FLETCHER R. A nonmonotone filter method for nonlinear optimization[J]. *Computational Optimization and Applications*, 2012, 52(3): 583-607.
- [22] GOULD N I M, LOH Y, ROBINSON D P. A filter method with unified step computation for nonlinear optimization[J]. *SIAM Journal on Optimization*, 2014, 24(1): 175-209.
- [23] GOULD N I M, LOH Y, ROBINSON D P. Anonmonotone filter SQP method; Local convergence and numerical results [J]. *SIAM Journal on Optimization*, 2015, 25(3): 1885-1911.
- [24] MARTÍNEZ J M, PILOTTA E A. Inexact-restoration algorithm for constrained optimization[J]. *Journal of Optimization Theory and Applications*, 2000, 104(1): 135-163.
- [25] MARTÍNEZ J M, PILOTTA E A. Inexact-restoration methods for nonlinear programming: Advances and perspectives [M]// QI L Q, TEO K, YANG X Q (eds.). *Optimization and Control with Applications*. US: Springer, 2005: 271-291.
- [26] BIRGIN E G, MARTÍNEZ J M. Local convergence of an inexact-restoration method and numerical experiments[J]. *Journal of Optimization Theory and Applications*, 2005, 127(2): 229-247.
- [27] FISCHER A, FRIEDLANDER A. A new line search inexact restoration approach for nonlinear programming [J]. *Computational Optimization and Applications*, 2010, 46(2): 333-346.

(责任编辑:尹 闯)