文章编号: 1000-5862(2012)03-0253-04

一般非线性约束优化问题的信赖域法

夏红卫, 文传军

(常州工学院理学院, 江苏 常州 213022)

摘要: 通过引进松弛变量和极小化增广 Lagrange 函数的方法, 将等式约束的非线性优化问题推广到不等式约束和一般约束的情形, 同时将滤子技巧和信赖域法相结合, 提出一种求解非线性约束优化问题的信赖域新算法, 扩大了算法的适用范围, 提高了算法的计算效率, 并通过数值试验说明算法的有效性.

关键词: 一般非线性约束; 信赖域法; 滤子技巧; Matlab 程序

中图分类号: O 224.2　　文献标志码: A

0　引言

考虑如下等式和不等式约束的非线性优化问题:

$$\begin{aligned}&\min f(\boldsymbol{x})\\ &\text{s.t. } c_i(\boldsymbol{x})=0, i\in\mathcal{E}=\{1,2,\cdots,m_e\},\\ &\qquad c_i(\boldsymbol{x})\geqslant 0, i\in\mathcal{I}=\{m_e+1,\cdots,m\},\end{aligned}\tag{1}$$

其中 $f:\mathbf{R}^n\to\mathbf{R}, c:\mathbf{R}^n\to\mathbf{R}^m$ 都是2次连续可微的函数.

问题(1)在经济、工程技术、国防、管理及自动化等领域有着广泛应用, 因此求解问题(1)具有十分重要的现实意义. 但目前为止还没有特别有效的方法直接得到最优解, 人们普遍采用迭代的方法求解: 首先选择 1 个初始点, 利用当前点的相关信息, 产生下一个迭代点, 一步一步地逼近最优解. 这就是求解问题(1)的迭代算法. 利用迭代法求解最优化问题的方法一般有: 一种是利用目标函数和约束函数构造增广目标函数, 借此将约束最优化问题转化为无约束最优化问题, 然后利用求解无约束优化问题的方法求得新目标函数的局部最优解或者稳定点, 如罚函数法和乘子法等[1]. 另一种是在可行域内使目标函数下降的迭代点法, 如可行点法[2]. 此外, 近些年来形成的序列 2 次规划算法和信赖域法也引起了人们极大的关注. 尤其是信赖域法, 它和线搜索法并列为目前求解非线性优化问题的 2 类主要数值方法. 信赖域法思想新颖、算法可靠、具有很强的收敛性. 它不仅能较快地解决良态问题, 而且也能有效地求解病态的优化问题. 因此信赖域法成为近 20 年来非线性优化领域的一个十分重要的研究方向[3-5]. 本文在等式约束的非线性优化问题的信赖域法[6]基础上, 将等式约束推广到不等式约束和一般约束的情况, 同时将滤子技巧和信赖域法相结合, 扩大了算法的适用范围, 提高了算法的计算效率, 借助 Matlab 工具, 进行了数值试验, 从而说明算法的有效性.

1　算法

前面已经讨论了等式约束的非线性优化问题的信赖域法. 下面先讨论不等式约束的非线性优化问题的方法, 把求解等式约束的非线性优化问题的方法推广到不等式约束的情况, 其基本思想是先引进松弛变量, 把不等式约束转化为等式约束, 然后利用最优性条件消去松弛变量.

先考虑如下的不等式约束优化问题:

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

引进松弛变量 $z_i(i=1,2,\cdots,m)$ 后, 化成等价的等式约束优化问题:

$$\begin{aligned}&\min f(\boldsymbol{x})\\ &\text{s.t. } c_i(\boldsymbol{x})-z_i^2=0, i=1,2,\cdots,m.\end{aligned}$$

构造增广 Lagrange 函数如下:

收稿日期: 2012-01-20
基金项目: 国家自然科学基金(60863266)和常州工学院校级基金(YN1010)资助项目.
作者简介: 夏红卫(1968-), 男, 江苏金坛人, 讲师, 硕士, 主要从事非线性优化的研究.

$$M(\boldsymbol{x},z,\boldsymbol{\lambda},\sigma)=f(\boldsymbol{x})-\sum_{i=1}^{m}\lambda_i\left[c_i(\boldsymbol{x})-z_i^2\right]+\frac{\sigma}{2}\sum_{i=1}^{m}\left[c_i(\boldsymbol{x})-z_i^2\right]^2,$$

为消去松弛变量 z_i，将增广 Lagrange 函数极小化.即令 $\nabla_z M(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda},\sigma)=0$ 得

$$z_i\left[\lambda_i-\sigma(c_i(\boldsymbol{x})-z_i^2)\right]=0, i=1,2,\cdots,m,$$

也就是

$$z_i\left[\sigma z_i^2-(\sigma c_i(\boldsymbol{x})-\lambda_i)\right]=0, i=1,2,\cdots,m.$$

当 $\sigma c_i(\boldsymbol{x})-\lambda_i\geqslant 0$ 时,

$$z_i^2=-\frac{\lambda_i}{\sigma}+c_i(\boldsymbol{x})=\frac{1}{\sigma}(\sigma c_i(\boldsymbol{x})-\lambda_i),$$

否则,

$$\lambda_i-\sigma(c_i(\boldsymbol{x})-z_i^2)=\sigma z_i^2+(\lambda_i-\sigma c_i(\boldsymbol{x}))>0,$$

从而 $z_i=0$.

合并 2 种情况得

$$z_i^2=\frac{1}{\sigma}\max\{0,\sigma c_i(\boldsymbol{x})-\lambda_i\}, i=1,2,\cdots,m.$$

所以增广 Lagrange 函数修正为

$$M(\boldsymbol{x},\boldsymbol{z},\boldsymbol{\lambda},\sigma)=f(\boldsymbol{x})-\frac{1}{2\sigma}\sum_{i=1}^{m}\left\{\left[\max\{0,\sigma c_i(\boldsymbol{x})-\lambda_i\}\right]^2-\lambda_i^2\right\},$$

乘子迭代公式修正为

$$(\lambda_{k+1})_i=\max\{0,(\lambda_k)_i-\sigma c_i(x_k)\}, i=1,2,\cdots,m,$$

终止准则变更为

$$\left(\sum_{i=1}^{m}\left[\min\left\{c_i(x_k),\frac{(\lambda_k)_i}{\sigma}\right\}\right]^2\right)^{1/2}<\varepsilon.$$

现在求解一般非线性优化问题(1), 引进松弛变量 $s_i, i\in\mathcal{I}$. 将不等式约束转化为等式约束和界约束:

$$\begin{aligned}&\min f(\boldsymbol{x})\\&\text{s.t. } c_i(\boldsymbol{x})=0, i\in\mathcal{E},\\&\quad c_i-s_i=0, i\in\mathcal{I},\\&\quad s_i\geqslant 0, i\in\mathcal{I}.\end{aligned}\tag{2}$$

一般求解非负约束的松弛变量的方法是内点法,即用对数罚函数来代替不等式约束[7-9]. 本文使用另一种方法, 保持不等式约束不变, 而用极小化相应的增广 Lagrange 函数来代替等式约束, 即

$$\min_{\boldsymbol{x},\boldsymbol{s}} f(\boldsymbol{x})-\sum_{i\in\mathcal{E}}\lambda_i c_i(\boldsymbol{x})-\sum_{i\in\mathcal{I}}\lambda_i(c_i(\boldsymbol{x})-s_i)+\sigma\sum_{i\in\mathcal{E}}(c_i(\boldsymbol{x}))^2+\sigma\sum_{i\in\mathcal{I}}(c_i(\boldsymbol{x})-s_i)^2$$
$$\text{s.t. } s_i\geqslant 0, i\in\mathcal{I}.$$

对一个给定的 $\boldsymbol{x}$, (2)式中的目标函数对每一个松弛变量 s_i 是凸 2 次函数, 很容易找到每一个 s_i 的最优值:

$$s_i=\max\{c_i(\boldsymbol{x})-\lambda_i/(2\sigma),0\}, i\in\mathcal{I}.$$

将上面的关系式代入(2)式中, 得到一个关于 $\boldsymbol{x}$ 的无约束最小化问题:

$$\min_{\boldsymbol{x}}\Phi_{\mathcal{A}(\boldsymbol{x},\boldsymbol{\lambda},\sigma)}(\boldsymbol{x},\boldsymbol{\lambda},\sigma),\tag{3}$$

其中

$$\min_{\boldsymbol{x}}\Phi_{\mathcal{A}(\boldsymbol{x},\boldsymbol{\lambda},\sigma)}(\boldsymbol{x},\boldsymbol{\lambda},\sigma)=f(\boldsymbol{x})-\sum_{i\in\mathcal{A}}\lambda_i c_i(\boldsymbol{x})+\sigma\sum_{i\in\mathcal{A}}(c_i(\boldsymbol{x}))^2,$$

$\mathcal{A}(\boldsymbol{x},\boldsymbol{\lambda},\sigma)=\mathcal{A}=\mathcal{E}\bigcup\left\{i\in\mathcal{I}\,\middle|\,c_i(\boldsymbol{x})<\lambda_i/(2\sigma)\right\}$ 是有效集.

在当前迭代点 $\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)},\sigma^{(k)}$ 处, 定义如下工作集:

$$\mathcal{A}^{(k)}=\mathcal{A}(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)},\sigma^{(k)})=\mathcal{E}\bigcup\left\{i\in\mathcal{I}\,\middle|\,c_i^{(k)}<\lambda_i^{(k)}/(2\sigma^{(k)})\right\},$$

这样通过极小化(3)式的 2 次近似来计算尝试步.

第 k 次迭代的子问题为

$$\begin{aligned}&\min_{\boldsymbol{d}} Q_{\mathcal{A}}^{(k)}(\boldsymbol{d}),\\&\text{s.t. } \|\boldsymbol{d}\|\leqslant\Delta^{(k)},\end{aligned}\tag{4}$$

其中

$$Q_{\mathcal{A}}^{(k)}(\boldsymbol{d})=(\boldsymbol{g}^{(k)}-\boldsymbol{A}_{\mathcal{A}}^{(k)}\boldsymbol{\lambda}_{\mathcal{A}}^{(k)})^{\mathrm{T}}\boldsymbol{d}+\frac{1}{2}\boldsymbol{d}^{\mathrm{T}}\boldsymbol{B}_{\mathcal{A}}^{(k)}\boldsymbol{d}+\sigma^{(k)}\left\|\boldsymbol{c}_{\mathcal{A}}^{(k)}+(\boldsymbol{A}_{\mathcal{A}}^{(k)})^{\mathrm{T}}\boldsymbol{d}\right\|^2,$$

$$\boldsymbol{c}_{\mathcal{A}}^{(k)}=(c_i(\boldsymbol{x})^k\,|\,i\in\mathcal{A}^{(k)})^{\mathrm{T}},$$

$$\boldsymbol{\lambda}_{\mathcal{A}}^{(k)}=(\lambda_i^k\,|\,i\in\mathcal{A}^{(k)})^{\mathrm{T}},$$

$$\boldsymbol{A}_{\mathcal{A}}^{(k)}=(\nabla c_i(\boldsymbol{x})^k\,|\,i\in\mathcal{A}^{(k)}).$$

$\boldsymbol{B}_{\mathcal{A}}^{(k)}$ 是 Lagrange 函数基于 $\mathcal{A}^{(k)}$ 的 Hessian 阵的近似.

设子问题(4)的解为 $\boldsymbol{d}^{(k)}$, 用 $\boldsymbol{d}^{(k)}$ 作为尝试步.

关于 Lagrange 乘子的修正, 与等式约束的情况相同, 但要求所有的乘子对不等式约束是非负的, 因此当约束不在有效集中时, 就取那些乘子为 0. 这样得到下面的 Lagrange 乘子的修正公式:

$$\lambda_i^{trial}=\begin{cases}\lambda_i^k-2\sigma^{(k)}(\nabla c_i^{(k)\mathrm{T}}\boldsymbol{d}^{(k)}+c_i^{(k)}), & i\in\mathcal{E},\\ \max\left\{\lambda_i^k-2\sigma^{(k)}(\nabla c_i^{(k)\mathrm{T}}\boldsymbol{d}^{(k)}+c_i^{(k)}),0\right\}, & i\in\mathcal{A}^{(k)}\setminus\mathcal{E},\\ 0, & \text{否则}.\end{cases}\tag{5}$$

与等式约束问题相同, 也使用滤子技巧, 只是

将其中 $h(\boldsymbol{x})$ 的定义修正为

$$h(\boldsymbol{x})=\sqrt{\sum_{i\in\mathcal{A}}(c_i(\boldsymbol{x}))^2+\sum_{i\in\mathcal{I}}[\min(c_i(\boldsymbol{x}),0)]^2}.$$

根据上面的讨论, 提出求解一般非线性约束优化问题的信赖域滤子算法.

算法 **1**

步 0: 初始化.

给定始点 $\boldsymbol{x}^{(0)}\in\mathbf{R}^n,\Delta^{(0)}>0,\boldsymbol{\lambda}^{(0)}\in\mathbf{R}^m,\sigma^{(0)}>0$, 取定常数: $0<\eta_1<\eta_2<1,\gamma_{\mathcal{g}}\in(0,1),\varepsilon_s,\varepsilon_c$, 计算 $f(\boldsymbol{x}^{(0)})$, $c(\boldsymbol{x}^{(0)})$.

初始工作集: $\mathcal{A}^{(0)},\boldsymbol{B}^{(0)}$. 初始滤子集 $\mathcal{F}=\{(c_0,f_0),(10c_0,-\infty)\}$, $k=0$;

步 1: 步长计算.

求解(4)式得到尝试步 $\boldsymbol{d}^{(k)}$;

步 2: 终止性检验.

如果 $\left\|\boldsymbol{d}^{(k)}\right\|>\varepsilon_s$, 转到步 3; 如果 $\boldsymbol{h}^{(k)}<\varepsilon_c$, 停止且返回 $\boldsymbol{x}^{(k)}$ 作为解;

取 $\sigma^{(k)}=10\sigma^{(k)}$, 转到步 3;

步 3: 尝试点的接受.

(a) 尝试点相关量的计算.

取 $\boldsymbol{x}^{trial}=\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)}$. 计算

$$f(\boldsymbol{x}^{trial}),c(\boldsymbol{x}^{trial}),\mathcal{A}(\boldsymbol{x}^{trial})$$

及由公式(5)确定的相应的 Lagrange 乘子 $\boldsymbol{\lambda}^{trial}$;

(b) 罚函数的接受.

计算实际下降量与预测下降量之间的比值:

$$\rho^{(k)}=\frac{\Phi_{\mathcal{A}^{(k)}}(\boldsymbol{x}^{(k)},\boldsymbol{\lambda}^{(k)},\sigma^{(k)})-\Phi_{\mathcal{A}^{(trial)}}(\boldsymbol{x}^{(trial)},\boldsymbol{\lambda}^{(trial)},\sigma^{(trial)})}{Q^{(k)}_{\mathcal{A}^{(k)}}(\boldsymbol{0})-Q^{(k)}_{\mathcal{A}^{(k)}}(\boldsymbol{d}^{(k)})},$$

如果 $\rho^{(k)}>0$, 取 $\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{trial}$, 转到步 4;

(c) 滤子的接受.

如果 $\boldsymbol{x}^{trial}$ 被滤子集接受, 令 $\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)}$, 升级滤子集, 并转到步 4, 否则令 $\boldsymbol{x}^{(k+1)}=\boldsymbol{x}^{(k)}$, 转到步 4;

步 4: 参数升级.

(a)Lagrange 乘子升级: $\boldsymbol{\lambda}^{(k+1)}=\boldsymbol{\lambda}^{trial}$;

(b)工作集升级: $\mathcal{A}^{(k+1)}=\mathcal{A}^{trial}$;

(c)罚因子升级:

$$\boldsymbol{\sigma}^{(k+1)}=\begin{cases}\max\left(2\boldsymbol{\sigma}^{(k)},2\left\|\boldsymbol{\lambda}^{(k+1)}\right\|\right), h(\boldsymbol{x}^{(k)}+\boldsymbol{d}^{(k)})\geqslant 0.5h(\boldsymbol{x}^k),\\ \max\left(\boldsymbol{\sigma}^{(k)},2\left\|\boldsymbol{\lambda}^{(k+1)}\right\|\right),\ \text{否则};\end{cases}$$

(d)近似 Hessian 阵的升级:

$$\boldsymbol{B}^{(k+1)}=\nabla^2 f(\boldsymbol{x}^{(k)})-\sum_{i\in\mathcal{A}^{(k)}}\lambda_i^{(k)}\nabla^2 c_i(\boldsymbol{x}^{(k)}),$$

$\boldsymbol{B}^{(k+1)}$ 的升级还有其他方法[10];

(e)信赖域半径的升级:

$$\Delta^{(k+1)}=\begin{cases}\max\left(2\Delta^{(k)},2\left\|\boldsymbol{d}^{(k)}\right\|\right), & \rho^{(k)}\geqslant\eta_2,\\ \Delta^{(k)}, & \rho^{(k)}\in[\eta_1,\eta_2],\\ \min\left(0.5\Delta^{(k)},0.5\left\|\boldsymbol{d}^{(k)}\right\|\right), & \rho^{(k)}<\eta_1,\end{cases}$$

$k=k+1$; 转到步 1.

2　数值试验

用 Matlab 语言编写了新算法的程序, 并进行了数值试验, 算法中的参数如下:

$$\Delta_0=1,\sigma_0=1,\eta_1=0.1,\eta_2=0.9,$$
$$\gamma_{\mathcal{g}}=0.000\,1,\varepsilon_s=0.000\,01,\varepsilon_c=0.000\,01.$$

测试问题全部取自于 CUTEr 集, HS10 表示该测试问题集的第 10 个问题, 下同. 为了检验新算法的有效性, 还运行了基于最小化增广 Lagrange 函数的著名的软件包 LANCELOT[11]. 计算了相同的问题, 并比较它们的结果, “n”, “m”分别表示测试问题和约束条件的维数, "f,c"和"g,A"分别表示函数值和梯度值的迭代次数. 计算结果列于表 1.

表 **1**　新算法和 **LANCELOT** 软件的计算结果

Problem	Pro.Dim		LANCELOT		新算法	
	n	m	f,c	g,A	f,c	g,A
HS10	2	1	21	21	25	18
HS11	2	1	19	19	10	10
HS14	2	2	15	15	10	10
HS22	2	2	11	11	8	8
HS29	3	1	19	18	6	6
HS43	4	3	23	22	11	11
HS88	2	1	49	48	43	35
HS89	3	1	63	58	39	35
HS113	10	8	37	35	14	14
HS268	5	5	21	21	5	5

从表 1 可以看出, 新算法的计算效率较高.

3　参考文献

[1] 倪勤. 最优化方法与程序设计 [M]. 北京: 科学出版社, 2009.

[2] Lawrence C T, Tits A L. A computationally efficient feasible se-

quential quadratic programming algorithm [J]. SIAM Journal on Optimization, 2001, 11(4): 1092-1118.

[3] Yuan Yaxiang. A review of trust region algorithms for optimization [C]// Ball J M, Hunt J C R. ICM99: Proceedings of the 4th International Congress on Industial and applied mathematics. Edinburgh: Oxford University Press, 2000: 271-282.

[4] Powell M J D. Convergence properties of a class of minimization algorithms [C]//Mangassarian O L, Meyer R R, Robinson S M. Nonlinear Programming. New York: Academic Press, 1975: 1-27.

[5] 夏红卫, 陈荣军. 简单界约束非线性方程组的滤子信赖域法 [J]. 江西师范大学学报: 自然科学版, 2009, 33(6): 661-664.

[6] 夏红卫, 陈荣军. 一类等式约束非线性优化问题的信赖域新算法 [J]. 数学的实践和认识, 2010, 40(20): 131-137.

[7] Chen Lifeng, Goldfarb D. Interior-point l_2-penalty methods for nonlinear programming with strong global convergence properties [J]. Math Programming, 2006, 108(1): 1-36.

[8] Conn A R, Gould N I M, Orban D, et al. A primal-dual trust region algorithm for non-convex nonlinear programming [J]. Math Programming, 2000, 87(2): 215-249.

[9] Fletcher R, Leyffer S. Nonlinear programming without a penalty function [J]. Math Programming, 2002, 91(2): 239-269.

[10] Conn A R, Gould N I M, Toint P L. Trust-region methods [M]. Philadephia: SIAM, 2000.

[11] Conn A R, Gould N I M, Toint P L. Lancelot: a fortran package for large-scale nonlinear optimization [M]. New York: Springer, 1992.

The Trust-Region Method for General Nonlinear Constrained Optimization Problem

XIA Hong-wei, WEN Chuan-jun

(School of Science, Changzhou Institute of Technology, Changzhou Jiangsu 213022, China)

Abstract: By introducing the slack variables and minimizing the augmented Lagrange functions, the equality constrained nonlinear optimization are extended to the inequality constraints and the general constraint. At the same time the filter technique and trust region method are combined, a new algorithm for nonlinear constrained optimization problems with trust region algorithm is proposed, which the scope of application of the algorithm is expanded and computational efficiency of this algorithm is improved. The numerical experiment shows that the method is quite efficient.

Key words: general nonlinear constrained optimization; trust-region method; filter technique; Matlab procedures

(责任编辑: 曾剑锋)