运筹学与控制论

DOI:10.11721/cqnuj20160413

线性等式约束非线性最优化问题的一个新算法*

申合帅[1], 李泽民[2]

(1. 郑州财税金融职业学院 基础部, 郑州 450000; 2. 重庆大学 数学与统计学院, 重庆 401331)

摘要:有文献给出了一般等式约束非线性最优化问题的一种求解途径。在此基础上将线性等式约束非线性最优化问题转化为非线性最小二乘问题求解,提出了求解最优化问题的一种新思路。然后利用 Gauss-Newton 法求解非线性最小二乘问题,在求解过程中引入非精确的一维搜索,提高了计算的效率,加快了算法收敛的速度,从而找到了具有线性等式约束非线性最优化问题的一个新算法,算法具有很好的收敛性,收敛速度是二阶的。最后经过数值实验证明新算法与 Matlab 优化工具箱计算的结果一致,是可行的、有效的。

关键词:非线性规划;线性等式约束;非线性最小二乘问题

中图分类号:O224　　**文献标志码**:A　　**文章编号**:1672-6693(2016)04-0006-04

最优化是研究决策问题的最佳选择之特性,构造寻求最佳解的计算方法。最优化问题广泛应用于经济计划、工程设计、生产管理、交通运输、国防等一系列重要领域,越来越多地受到政府部门、科研机构和产业部门的高度重视。随着应用越来越广泛,优化问题的规模越来越大,如何找到简单有效的计算方法具有非常重要的实际意义。传统的 Lagrange 乘子法求解,所得的方程有 $n+m$ 个,而利用的文献[1]中的途径进行求解最优化问题,所得的方程组仅含有 n 个方程,通过转化可以大大减少运算的复杂性,此类算法具有规模上的优势。在利用该途径求解最优化问题方面,已经做了大量的研究工作,但主要集中在二次规划问题的求解上。其中,童东付等人[2]将二次规划问题转化为线性方程组进行求解;王开荣[3]讨论了具有混合约束的二次规划问题,利用二次逼近,将混合约束转化为线性约束,从而把问题转化为二次规划问题求解;袁松琴等人[4]讨论了线性等式约束多目标规划问题,首先把多目标问题转化为单目标问题,通过对目标函数二次逼近,从而构成线性等式约束二次规划序列,最后利用二次规划问题方法求解;胡资骏[5]讨论了极值存在的最优性条件;杨懿等人[6]将算法和积极集法结合,解决了不等式约束的二次规划问题。本文利用该途径将最优化问题转化为非线性最小二乘问题解决。非线性最小二乘问题用途非常广泛,其算法的研究也越来越受到重视,具有非常重要的实际意义。其中文献[7-8]讨论了非线性最小二乘问题的迭代以及收敛性的相关证明。

1 预备知识

考虑最优化问题(ECP)

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

其中 $\boldsymbol{x}\in\mathbf{R}^n$, $f:\mathbf{R}^n\to\mathbf{R}$, $h:\mathbf{R}^n\to\mathbf{R}^m$ 连续可微,$m<n$,$\boldsymbol{h}(\boldsymbol{x})=\{h_1(\boldsymbol{x}),\cdots,h_m(\boldsymbol{x})\}$。

令 $l=n-m$,本文使用下列记号:

$$\boldsymbol{P}(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_1},\frac{\partial f(\boldsymbol{x})}{\partial x_2},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_l}\right)^{\mathrm{T}},\boldsymbol{Q}(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_{l+1}},\frac{\partial f(\boldsymbol{x})}{\partial x_{l+2}},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_n}\right)^{\mathrm{T}},$$

$$\nabla\boldsymbol{h}(\boldsymbol{x})=\begin{bmatrix}\frac{\partial h_1(\boldsymbol{x})}{\partial x_1}&\frac{\partial h_1(\boldsymbol{x})}{\partial x_2}&\cdots&\frac{\partial h_1(\boldsymbol{x})}{\partial x_n}\\\cdots&\cdots&\cdots&\cdots\\\frac{\partial h_m(\boldsymbol{x})}{\partial x_1}&\frac{\partial h_m(\boldsymbol{x})}{\partial x_2}&\cdots&\frac{\partial h_m(\boldsymbol{x})}{\partial x_n}\end{bmatrix},\boldsymbol{N}(\boldsymbol{x})=\begin{bmatrix}\frac{\partial h_1(\boldsymbol{x})}{\partial x_1}&\frac{\partial h_1(\boldsymbol{x})}{\partial x_2}&\cdots&\frac{\partial h_1(\boldsymbol{x})}{\partial x_l}\\\cdots&\cdots&\cdots&\cdots\\\frac{\partial h_m(\boldsymbol{x})}{\partial x_1}&\frac{\partial h_m(\boldsymbol{x})}{\partial x_2}&\cdots&\frac{\partial h_m(\boldsymbol{x})}{\partial x_l}\end{bmatrix},$$

* **收稿日期**:2015-11-29　**修回日期**:2016-05-19　**网络出版时间**:2016-07-07 16:34

作者简介:申合帅,男,讲师,研究方向为最优化理论,E-mail:48911355@qq.com

网络出版地址:http://www.cnki.net/kcms/detail/50.1165.N.20160707.1634.054.html

$$\boldsymbol{M}(\boldsymbol{x})=\begin{bmatrix}\frac{\partial h_1(\boldsymbol{x})}{\partial x_{l+1}} & \frac{\partial h_1(\boldsymbol{x})}{\partial x_{l+2}} & \cdots & \frac{\partial h_1(\boldsymbol{x})}{\partial x_n}\\ \cdots & \cdots & \cdots & \cdots\\ \frac{\partial h_m(\boldsymbol{x})}{\partial x_{l+1}} & \frac{\partial h_m(\boldsymbol{x})}{\partial x_{l+2}} & \cdots & \frac{\partial h_m(\boldsymbol{x})}{\partial x_n}\end{bmatrix}。$$

引理 1[1]　若 $\exists x_0\in\mathbf{R}^n$，$\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\cdots,\lambda_m)^{\mathrm{T}}$，满足：1) $\boldsymbol{h}(x_0)=0$；2) $\nabla\boldsymbol{f}(x_0)+\lambda^{\mathrm{T}}\nabla\boldsymbol{h}(x_0)=\boldsymbol{0}$；3) 矩阵 $\boldsymbol{L}(x_0)=\nabla^2\boldsymbol{f}(x_0)+\sum_{i=1}^{m}\lambda_i\nabla^2h_i(x_0)$ 在子空间 $T=\{\boldsymbol{x}\in\mathbf{R}^n\mid\nabla\boldsymbol{h}(x_0)^{\mathrm{T}}\boldsymbol{x}=\boldsymbol{0}\}$ 上正定，则 x_0 是(ECP)的严格局部极小点。

定理 1　若 $x_0\in\mathbf{R}^n$ 是方程组

$$\begin{cases}\boldsymbol{P}(\boldsymbol{x})=[\boldsymbol{M}(\boldsymbol{x})^{-1}\boldsymbol{N}(\boldsymbol{x})]^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x})\\ \boldsymbol{h}(\boldsymbol{x})=\boldsymbol{0}\end{cases}\tag{2}$$

的解，使得：1) $\boldsymbol{M}(x_0)$非奇异；2) 矩阵 $\boldsymbol{L}(x_0)$在切平面 T 上正定。则 x_0 是(ECP)的严格局部极小点。

证明　由假设，有 $\boldsymbol{h}(x_0)=\boldsymbol{0}$，$[\boldsymbol{M}(x_0)^{-1}\boldsymbol{N}(x_0)]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\boldsymbol{P}(x_0)$，则有 $\boldsymbol{N}(x_0)^{\mathrm{T}}[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\boldsymbol{P}(x_0)$。自然有 $\boldsymbol{M}(x_0)^{\mathrm{T}}[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\boldsymbol{Q}(x_0)$，即 $\begin{bmatrix}\boldsymbol{N}(x_0)^{\mathrm{T}}\\ \boldsymbol{M}(x_0)^{\mathrm{T}}\end{bmatrix}[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\begin{bmatrix}\boldsymbol{P}(x_0)\\ \boldsymbol{Q}(x_0)\end{bmatrix}$。从而有

$$[\boldsymbol{N}(x_0)\boldsymbol{M}(x_0)]^{\mathrm{T}}[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\nabla\boldsymbol{f}(x_0)^{\mathrm{T}},$$

即 $\nabla\boldsymbol{h}(x_0)[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\nabla\boldsymbol{f}(x_0)^{\mathrm{T}}$，$\nabla\boldsymbol{f}(x_0)^{\mathrm{T}}-\nabla\boldsymbol{h}(x_0)[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)=\boldsymbol{0}$。

令 $\boldsymbol{\lambda}=-[\boldsymbol{M}(x_0)^{-1}]^{\mathrm{T}}\boldsymbol{Q}(x_0)$，矩阵转置得 $\nabla\boldsymbol{f}(x_0)+\boldsymbol{\lambda}^{\mathrm{T}}\nabla\boldsymbol{h}(x_0)=\boldsymbol{0}$。从而，由引理 1 可证。　　证毕

对于矩阵 $\boldsymbol{P},\boldsymbol{Q},\boldsymbol{M},\boldsymbol{N}$ 中的变量 $x_1,\cdots,x_n$，它们的下标不要求连续，但是约束函数 $\boldsymbol{h}(\boldsymbol{x})$在点 x_0 的导数 $\nabla\boldsymbol{h}(x_0)$要求是行满秩的。

从定理 1 的证明来看，方程组(1)的解必是(ECP)中一个 K-T 点。因此(ECP)最优解的问题可转化为方程组(2)的求解问题，这就大大减少了工作量。

引理 2　(Cramer 法则)向量组$\{\boldsymbol{\alpha}_1,\boldsymbol{\alpha}_2,\cdots,\boldsymbol{\alpha}_p\}$线性无关的充分必要条件是该向量组的 Cramer 行列式不等于零。

2 等式约束最优化问题

考虑下面的非线性最优化问题(ECP_1)：

$$\min f(\boldsymbol{x}),$$
$$\mathrm{s.t.}\ \boldsymbol{Ax}=\boldsymbol{b}。$$

其中 $\boldsymbol{x}\in\mathbf{R}^n$，$f:\mathbf{R}^n\to\mathbf{R}$，$\boldsymbol{A}=(a_{ij})_{m\times n}$，$\boldsymbol{b}=(b_1,b_2,\cdots,b_m)^{\mathrm{T}}$，这时 $\boldsymbol{h}(\boldsymbol{x})=\boldsymbol{Ax}-\boldsymbol{b}=\boldsymbol{0}$，$\boldsymbol{A}$ 是 $m\times n$ 矩阵，并假设 $\boldsymbol{A}$ 是行满秩的，即秩 $\boldsymbol{A}=m$，则有 $\nabla\boldsymbol{h}(\boldsymbol{x})=\nabla(\boldsymbol{Ax}-\boldsymbol{b})=\boldsymbol{A}$。

令

$$\boldsymbol{P}(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_{j_1}},\frac{\partial f(\boldsymbol{x})}{\partial x_{j_2}},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_{j_p}}\right)^{\mathrm{T}},\boldsymbol{Q}(\boldsymbol{x})=\left(\frac{\partial f(\boldsymbol{x})}{\partial x_{i_1}},\frac{\partial f(\boldsymbol{x})}{\partial x_{i_2}},\cdots,\frac{\partial f(\boldsymbol{x})}{\partial x_{i_m}}\right)^{\mathrm{T}},$$

$$\boldsymbol{M}=\begin{bmatrix}a_{1i_1} & a_{1i_2} & \cdots & a_{1i_m}\\ a_{2i_1} & a_{2i_2} & \cdots & a_{2i_m}\\ \cdots & \cdots & \cdots & \cdots\\ a_{mi_1} & a_{mi_2} & \cdots & a_{mi_m}\end{bmatrix},\boldsymbol{N}=\begin{bmatrix}a_{1j_1} & a_{1j_2} & \cdots & a_{1j_p}\\ a_{2j_1} & a_{2j_2} & \cdots & a_{2j_p}\\ \cdots & \cdots & \cdots & \cdots\\ a_{mj_1} & a_{mj_2} & \cdots & a_{mj_p}\end{bmatrix}。$$

由定理 1，非线性最优化问题(ECP_1)可转化为方程组求解问题 $\begin{cases}\boldsymbol{P}(\boldsymbol{x})=[\boldsymbol{M}^{-1}\boldsymbol{N}]^{\mathrm{T}}\boldsymbol{Q}(\boldsymbol{x})\\ \boldsymbol{Ax}-\boldsymbol{b}=\boldsymbol{0}\end{cases}$。其中，关键是如何在矩阵 $\boldsymbol{A}$ 中找出一个 m 列阶非奇异子块 $\boldsymbol{M}$。

根据引理 1，引入非精确的一维搜索，由下列算法确定矩阵 $\boldsymbol{M}$。

算法 0　第一步，构造一个新的矩阵 $\boldsymbol{M}_1:=p_1$，即该矩阵仅含 $\boldsymbol{A}$ 的中的第一列；

第二步，对于 $\boldsymbol{A}$ 中的列 $p_i:i=2$，根据 Cramer 法则，判断 p_i 是否与 $\boldsymbol{M}$ 中的列线性无关：若线性无关，则将 p_i 作为新的一列加入 $\boldsymbol{M}_1$ 中；若线性相关，再对 $\boldsymbol{A}$ 中的不含于 $\boldsymbol{M}$ 中的列进行此操作，$i=i+1$，如此循环下去，直至

得出 m 个列线性无关。

当 $\boldsymbol{h}(\boldsymbol{x})$ 不是行满秩矩阵时,对列的操作换成对行的操作即可。

令 $\boldsymbol{W}=[\boldsymbol{M}^{-1}\boldsymbol{N}]^{\mathrm{T}}=(g_{ij})_{p\times m}$,则有 $\begin{cases}\boldsymbol{p}(\boldsymbol{x})-\boldsymbol{WQ}(\boldsymbol{x})=\boldsymbol{0},\\ \boldsymbol{Ax}-\boldsymbol{b}=\boldsymbol{0}。\end{cases}$

根据定理1,将非线性最优化问题(ECP_1)转化为非线性方程组(2)的求解问题。在此基础上,最终可转化为非线性最小二乘问题

$$(\mathrm{P}_1)\quad \min f(\boldsymbol{x})=\frac{1}{2}r(\boldsymbol{x})^{\mathrm{T}}r(\boldsymbol{x})=\frac{1}{2}\sum_{i=1}^{n}[r_i(\boldsymbol{x})]^2,\tag{3}$$

其中 $r_i(\boldsymbol{x})=\dfrac{\partial f(\boldsymbol{x})}{\partial x_{j_i}}-\sum\limits_{k=1}^{m}g_{ik}\dfrac{\partial f(\boldsymbol{x})}{\partial x_{i_k}},i=1,\cdots,n-m;r_i(\boldsymbol{x})=\sum\limits_{j=1}^{n}a_{ij}x_j-b_i,i=n-m+1,\cdots,n$。

设矩阵 $\boldsymbol{J}(\boldsymbol{x})$ 是 $\boldsymbol{r}(\boldsymbol{x})$ 的 Jacobi 矩阵,$\boldsymbol{J}(\boldsymbol{x})=\begin{bmatrix}\dfrac{\partial r_1}{\partial x_1} & \dfrac{\partial r_1}{\partial x_2} & \cdots & \dfrac{\partial r_1}{\partial x_p}\\ \cdots & \cdots & \cdots & \cdots\\ \dfrac{\partial r_m}{\partial x_1} & \dfrac{\partial r_m}{\partial x_2} & \cdots & \dfrac{\partial r_m}{\partial x_p}\end{bmatrix}$,用 Gauss-Newton 法[9]求解问题(P_1)。

定理2[10] 设 $\boldsymbol{f}(\boldsymbol{x})\in C^2$,$\boldsymbol{x}^*$ 为非线性最小二乘问题(3)的局部极小点,$\boldsymbol{J}(\boldsymbol{x}^*)^{\mathrm{T}}\boldsymbol{J}(\boldsymbol{x}^*)$ 正定。如果 $\boldsymbol{r}(\boldsymbol{x}^*)=\boldsymbol{0}$,则存在 $\varepsilon>0$,使得对于任何 $x_0\in N(\boldsymbol{x}^*,\varepsilon)$,由 Gauss-Newton 法产生的序列 $\{x_k\}$ 收敛到 x^*,且收敛速度是二阶。

下面给出求解规划问题(ECP_1)的算法。

算法1 第一步,利用算法0确定矩阵 $\boldsymbol{M}$;

第二步,确定出矩阵 $\boldsymbol{N},\boldsymbol{P},\boldsymbol{Q}$;

第三步,利用(3)式构造出非线性最小二乘问题(P_1);

第四步,用 Gauss-Newton 法[9]求解问题(P_1)。

3 数值实验

例 $\min f=x_1^3+x_1^2+x_1+x_2^3+2x_2^2+x_2+x_3^3+3x_3^2+x_3+x_4^3+4x_4^2+x_4+x_5^3+5x_5^2+x_5+x_6^3+6x_6^2+x_6$,

$$\text{s.t.}\begin{cases}x_1+x_2+x_3+x_4+x_5+x_6=1,\\ x_1-2x_2+3x_3+5x_4+2x_5+x_6=1,\\ x_1+x_3+x_4+6x_5+4x_6=1。\end{cases}$$

用 Matlab 优化工具箱计算的结果是 $\boldsymbol{x}=(0.360\,4,0.279\,0,0.163\,5,0.121\,1,0.025\,5,0.050\,4)^{\mathrm{T}}$。

再利用算法1求解。首先,利用算法0确定矩阵 $\boldsymbol{M}=\begin{bmatrix}1 & 1 & 1\\ 1 & -2 & 3\\ 1 & 1 & 0\end{bmatrix}$,从而确定 $\boldsymbol{N}=\begin{bmatrix}1 & 1 & 1\\ 5 & 2 & 1\\ 1 & 6 & 4\end{bmatrix}$,$\boldsymbol{P}(\boldsymbol{x})=(3x_4^2+8x_4+1,3x_5^2+10x_5+1,3x_6^2+12x_6+1)^{\mathrm{T}}$,$\boldsymbol{Q}(\boldsymbol{x})=(3x_1^2+2x_1+1,3x_2^2+4x_2+1,3x_3^2+6x_3+1)^{\mathrm{T}}$。最后转换为非线性最小二乘问题,利用 Gauss-Newton 法求得结果

$$\boldsymbol{x}=(0.360\,4,0.279\,0,0.163\,4,0.121\,0,0.025\,5,0.050\,5)^{\mathrm{T}}。$$

与 Matlab 优化工具箱计算的结果是一致的。

通过数值实验可看出该算法是可行的、有效的,并且具有很好的收敛性。相对于经典的 Lagrange 乘子法,该算法不需要计算乘子,而是将等式约束问题转化为非线性最小二乘问题,从而提出了一种新的思路。当最优化问题规模比较大时,通过转化减少运算的复杂性;并且由于引入了一维搜索,进而提高了计算的效率,加快了算法收敛的速度。

参考文献:

[1] 李泽民. 最优化的一种新途径[J]. 重庆建筑工程学院学报,1990,12(1):49-55.

Li Z M. A new approach of optimization[J]. Journal of Chongqing Institute of Architecture and Engineering,1990, 12(1):49-55.

[2] 童东付,李泽民. 二次规划问题的降维算法[J]. 重庆建筑大学学报,1999,21(5): 64-68.
Tong D F, Li Z M. Descending dimension algorithm for quadratic programming problem[J]. Journal of Chongqing Jianzhu University,1999,21(5):64-68.

[3] 王开荣. 具有混合约束二次函数的逼近方法[J]. 重庆大学学报:自然科学版,2004,27(1):131-134.
Wang K R. The approximate method of quadratic function with mixed constraint[J]. Journal of Chongqing University:Natural Science,2004,27(1):131-134.

[4] 袁松琴,李泽民. 线性等式约束多目标规划的一个降维算法[J]. 运筹学学报,2005,9(1):70-74.
Yuan S Q, Li Z M. Descending dimension algorithm for multi-objective programming with linear equality constraints[J]. Operations Research Transactions,2005,9(1): 70-74.

[5] 胡资骏. 一类向量极值问题的最优性条件[J]. 经济数学,2006,23(1):95-98.
Hu Z J. Optimality conditions for a kind of vector optimization problems[J]. Mathematics in Economics,2006,23(1): 95-98.

[6] 杨懿,李泽民. 不等式约束二次规划的一新算法[J]. 应用数学与计算数学学报,2005,19(2):55-60.
Yang Y,Li Z M. A new algorithm for quadratic programming problem[J]. Communication on Applied Mathematics and Computation,2005,19(2):55-60.

[7] 陈忠,黄惠. 求解非线性最小二乘问题的迭代法[J]. 武汉大学学报,2003,49(1):14-16.
Chen Z, Huang H. Iterative method for solving nonlinear least-squares problems[J]. Journal of Wuhan University, 2003,49(1):14-16.

[8] 陈宏彩,张友兰. 非线性最小二乘问题收敛性的证明[J]. 数学的实践与认识,2010,40(23):149-154.
Chen H C,Zhang Y L. A proof on convergence properies of nonlinear least squares problem[J]. Mathematics in Practice and Theory,2010,40(23):149-154.

[9] 胡毓达. 非线性规划[M]. 北京:高等教育出版社,1990.
Hu Y D. Nonlinear programming[M]. Beijing:Higher Education Press,1990.

[10] 袁亚湘,孙文瑜. 最优化理论与方法[M]. 北京:科学出版社,1997.
Yuan Y X, Sun W Y. Optimization theory and method [M]. Beijing:Science Press,1997.

Operations Research and Cybernetics

Descending Dimension Algorithm for Nonlinear Optimization Problem with Linear Equality Constraints

SHEN Heshuai[1], LI Zemin[2]

(1. Department of Basic, Zhengzhou Vocational College of Finance and Taxation, Zhengzhou 450000;
2. College of Mathematics and Statistics, Chongqing University, Chongqing 401331, China)

Abstract: In this paper, a method for solving nonlinear optimization problems with general equality constraints is proposed by [1]. On this basis, linear equality constrained nonlinear optimization problem is transformed into solving nonlinear least squares problem by this way. We find a new way to solve the optimization problem, and then the Gauss-Newton method is used to solve the nonlinear least squares problem. We introduce a inaccurate one dimension research search in the process of solving the problem, improving the calculation efficiency, accelerating the convergence speed of algorithm. A new algorithm for nonlinear optimization problems with linear equality constraints is proposed. Algorithm has very good convergence, convergence speed is second order. Finally, after numerical experiments, it is proved that the new algorithm is feasible and effective.

Key words: nonlinear programming; linear equality constraint; nonlinear least squares problem

(责任编辑　黄　颖)