DOI: 10.13656/j.cnki.gxkx.2009.02.009
广西科学 Guangxi Sciences 2009, 16(2): 105～108

A Perturbed Feasible SQCQP Algorithm*
一个摄动的可行 SQCQP算法

TANG Chun-ming, JIAN Jin-bao, ZHENG Hai-yan
唐春明,简金宝,郑海艳

(College of Mathematics and Information Science, Guangxi University, Nanning, Guangxi, 530004, China)
(广西大学数学与信息科学学院,广西南宁 530004)

Abstract By introducing a new perturbation strategy, a perturbed feasible sequential quadratically constrained quadratic programming (SQCQP) algorithm is proposed. The algorithm is globally and superlinearly convergent, and the uniformly positive definiteness assumption in the global convergence analysis of traditional SQCQP algorithms is removed.

Key words SQCQP, perturbation strategy, uniformly positive definiteness, global convergence, superlinear convergence

摘要:通过引入新的摄动策略,提出一个摄动的可行序列二次约束二次规划(SQCQP)算法.该算法全局和超线性收敛并且去掉了传统 SQCQP算法全局收敛性分析中的一致正定性假设.

关键词: SQCQP 摄动策略 一致正定性 全局收敛 超线性收敛
中图法分类号: O221.2 文献标识码: A 文章编号: 1005-9164(2009)02-0105-04

Consider the inequality constrained nonlinear programming problem:

$$\min_{x\in R^n} f_0(x)$$

$$\text{s.t. } f_i(x)\leqslant 0, i\in I \overset{\text{def}}{=} \{1,2,\cdots,m\}, \qquad (1)$$

where $f_i, i\in \{0\}\cup I: R^n\to R$ are smooth functions. Denote the feasible set for problem (1) by $F=\{x\in R^n: f_i(x)\leqslant 0, i\in I\}$.

During the past several decades, the researches of sequential quadratic programming (SQP) methods for solving problem (1) have been greatly improved, but when the problems to be solved are highly nonlinear, SQP methods usually show slow convergence, or even fail. For this, in recent years a so-called sequential quadratically constrained quadratic programming (SQCQP) method has been proposed[1～4]. In particular, Jian[1] proposed a feasible norm-relaxed SQCQP algorithm, and under weaker conditions, the global, superlinear, and quasi-quadratic convergence are obtained. Compared with SQP methods, at each iteration SQCQP methods only need to solve a quadratically constrained quadratic programming (QCQP) subproblem, and without any correctional directions the Maratos effect will not occur and therefore the superlinear convergence can be obtained. QCQP subproblem is a quadratic approximation of the original problem, so it is a better approximation than the quadratic programming (QP) subproblem solved in SQP methods, therefore the expected numerical performance of SQCQP may be better than SQP.

Generally, in the global convergence analysis, the positive definiteness or even uniformly positive definiteness of the (approximate) Hessian matrix G_0 of the objective function f_0 is needed to be assumed, which is considered as a strong assumption. In order to overcome this shortcoming, Liu[5] presented a perturbed QCQP subproblem by adding a positive multiple of identity matrix to G_0. By suitably adjusting the perturbation parameters, the algorithm in Reference [5] (belongs to penalty function type SQCQP algorithms) is proved to be globally convergent only under the assumption that G_0 is positive semidefinite. Motivated by the perturbation idea in Reference [5], we propose a new perturbation strategy (which is simpler than that given in Reference [5]) to modify the feasible SQCQP algorithm in

收稿日期: 2008-11-17
作者简介:唐春明(1979-),男,助教,主要从事最优化算法研究.
* Supported by China NSF (10771040), Guangxi NSF (064001, 0728006), Guangxi University RF(X071088).

Reference [1]. By using the newly proposed strategy to update the perturbation parameters, we remove the assumption of uniformly positive definiteness used in the global convergence analysis of the algorithm in Reference [1]. The new algorithm is also proved to be globally and superlinearly convergent.

1 Algorithm

We begin this section by making the following basic hypothesis.

Assumption 1 The functions $f_i, i \in \{0\} \cup I$, are first-order continuously differentiable.

For simplicity, we denote the gradient vectors $\nabla f_i(x)$ by $g_i(x), i \in \{0\} \cup I$. For the current feasible iterate $x^k \in F$, we consider the QCQP subproblem as follows, which is perturbed from that in Reference [1].

$$\min_{(z,d) \in R^{n+1}} V_0 z + \frac{1}{2} d^T (G_0^k + M_k E) d$$

$$\text{s.t. } g_0(x^k)^T d \leqslant V_0 z,$$

$$f_i(x^k) + g_i(x^k)^T d + \frac{1}{2} d^T G_i^k d \leqslant V_i e_k z, i \in I,$$

$$\frac{1}{2} e_k \|d\|^2 \leqslant c e_k^{(2f+1)}, \quad (2)$$

where M_k and e_k are positive parameters, E is an $n \times n$ identity matrix, $V_i (i \in \{0\} \cup I), c, f$ are positive constants, and $G_i^k (i \in \{0\} \cup I)$ are symmetric, positive semidefinite matrices. The QCQP subproblem (2) is convex and therefore can be cast as a second-order cone program and then be solved efficiently by using interior point algorithms[6].

We call (z_k, d^k) a KKT (Karush-Kuhn-Tucker) point for subproblem (2) if there exist multipliers $\__k$, $v_k \in R, u^k = (u_i^k, i \in I) \in R^m$ such that

$$(G_0^k + M_k E) d^k + \__k g_0(x^k) + \sum_{i \in I} u_i^k (g_i(x^k) + G_i^k d^k) + e_k v_k d^k = 0, \quad (3)$$

$$V_0 = V_0 \__k + \sum_{i \in I} u_i^k V_i e_k, \quad (4)$$

$$0 \leqslant \__k \perp (V_0 z_k - g_0(x^k)^T d^k) \geqslant 0, \quad (5)$$

$$0 \leqslant u_i^k \perp (V_i e_k z_k - f_i(x^k) - g_i(x^k)^T d^k - \frac{1}{2} (d^k)^T G_i^k d^k) \geqslant 0, i \in I, \quad (6)$$

$$0 \leqslant v_k \perp (c e_k^{(2f+1)} - \frac{1}{2} e_k \|d^k\|^2) \geqslant 0, \quad (7)$$

where notation $x \perp y$ means $x^T y = 0$.

The following lemmas show that the QCQP subproblem (2) is well defined, and their proofs are similar to those in Reference [1] due to the fact that $G_0^k + M_k E$ is a positive definite matrix.

Lemma 1 Suppose that Assumption 1 holds. Then (i) QCQP subproblem (2) has an optimal solution, (ii) (z_k, d^k) is an optimal solution of subproblem (2) if and only if it is a KKT point for subproblem (2).

Lemma 2 Suppose that Assumption 1 holds, and that (z_k, d^k) is an optimal solution of subproblem (2). Then (i) $z_k \leqslant -\frac{1}{2V_0} (d^k)^T (G_0^k + M_k E) d^k \leqslant -\frac{1}{2V_0} M_k \|d^k\|^2$, (ii) $z_k = 0 \Leftrightarrow d^k = 0 \Leftrightarrow x^k$ is a Fritz John point for problem (1), (iii) if $z_k < 0$, then d^k is a feasible descent direction of problem (1) at x^k.

Now we present the perturbed feasible SQCQP algorithm as follows.

Algorithm A

Step 0 Initialization. Choose $V_i > 0, i \in \{0\} \cup I$, $M_0 > 0$, $e_0 > 0$, $f > 1, c > 0$, $T \in (0, 0.5)$, $U, Z, \theta \in (0, 1)$, $x^0 \in F$, $G_i^0 (i \in \{0\} \cup I)$ symmetric and positive semidefinite. Set $k := 0$.

Step 1 Solve QCQP. Solve subproblem (2) to obtain a KKT point (z_k, d^k). If $k = 0$, set $d = \|d^k\|$. If $z_k = 0$, stop.

Step 2 Line search. Compute the step size λ_k, the first value of λ in the sequence $\{1, U, U^2, \cdots\}$ that satisfies

$$f_0(x^k + \lambda d^k) \leqslant f_0(x^k) + T \lambda g_0(x^k)^T d^k, \quad (8)$$

$$f_i(x^k + \lambda d^k) \leqslant 0, i \in I. \quad (9)$$

Step 3 Generate perturbation parameter. If $\frac{\|d^k\|}{d} \leqslant Z$, set $M_{k+1} = \theta M_k$ and $d = \|d^k\|$; otherwise set $M_{k+1} = M_k$.

Step 4 Update. Set $x^{k+1} := x^k + \lambda_k d^k$. Compute new symmetric positive semidefinite matrices $G_i^{k+1} (i \in \{0\} \cup I)$ as well as positive parameters e_{k+1}. Set $k := k + 1$ and go to step 1.

Remark From Lemma 2(iii), we know that the line search in step 2 can be terminated in a finite number of computations, and therefore Algorithm A is well-defined. The updating strategy of perturbation parameters in step 3 is new and different from (more precisely, simpler than) that in Reference [5].

2 Global and superlinear convergence

We first establish the global convergence of the proposed algorithm without assuming that G_0^k is (uniformly) positive definite. Once Algorithm A stops finitely at x^k, it follows from Lemma 2 that x^k is a Fritz John point for problem (1). Now assume that an infinite sequence $\{x^k\}$ of iterates is generated by Algorithm A, we will show that there exists an accumulation point x^* of $\{x^k\}$ such that it is a KKT point for problem (1).

Assumption 2 (i) The sequences $\{x^k\}$ and $\{G_i^k\}, i \in \{0\} \cup I$ are bounded. (ii) The parameter sequence $\{e_k\}$ is strictly positive and bounded.

Define the index set of shrinking the perturbation parameters by $S = \{k: M_{k+1} = \theta M_k\}$.

Lemma 3 Suppose that Assumptions 1 and 2 hold. Then the sequences $\{\_{k}\}$, $\{d^k\}$ and $\{z_k\}$ are bounded.

Proof The boundedness of $\{\_{k}\}$ is obvious from formula (4). From the last constraint of subproblem (2), it follows that $\{d^k\}$ is bounded. This together with the boundedness of $\{x^k\}$ and the inequality: $0 \geqslant z_k \geqslant \frac{1}{V_0} g_0(x^k)^T d^k \geqslant -\frac{1}{V_0} \|g_0(x^k)\| \cdot \|d^k\|$ shows that $\{z_k\}$ is bounded.

Lemma 4 Suppose that Assumptions 1 and 2 hold. Then there exists an infinite index set K such that $\lim_{k \in K} d^k = 0$.

Proof We divide the proof into two cases.

Case 1 S is an infinite set. From step 3 of Algorithm A and $Z \in (0,1)$, it is obvious that there exists an index set K such that $\lim_{k \in K} d^k = 0$.

Case 2 S is a finite set. From step 3 of Algorithm A, we know that there exists a constant $\bar{M} > 0$ such that $M_k \geqslant \bar{M}$ for all k. Without loss of generality, suppose by contradiction that there exists a constant $\bar{k} > 0$ such that

$$\|d^k\| \geqslant \bar{k}, \text{ for all } k. \qquad (10)$$

This together with the last constraint of subproblem (2) shows that

$$0 < e_{\inf} \overset{\text{def}}{=} \inf\{e_k\} \leqslant e_k, \text{ for all } k. \qquad (11)$$

Next we will prove that there is a constant $\bar{\lambda} > 0$ such that the step size $\lambda_k \geqslant \bar{\lambda}$ for all k.

Analyze inequality (8). From Taylor expansion, formula (5), Lemma 2(i) and formula (10), we have

$f_0(x^k + \lambda d^k) - f_0(x^k) - T\lambda g_0(x^k)^T d^k = (1 - T)\lambda g_0(x^k)^T d^k + o(\lambda) \leqslant (1 - T)\lambda V_0 z_k + o(\lambda) \leqslant -\frac{1}{2}(1 - T)\lambda M_k \|d^k\|^2 + o(\lambda) \leqslant -\frac{1}{2}\bar{M}(1 - T)\bar{k}^2\lambda + o(\lambda)$.

This shows that inequality (8) holds for all k and $\lambda > 0$ small enough.

Analyze inequalities (9). From Taylor expansion, formula (6), Lemma 2(i), formula (10) and formula (11), we have for $i \in I$,

$f_i(x^k + \lambda d^k) = f_i(x^k) + \lambda g_i(x^k)^T d^k + o(\lambda) \leqslant (1 - \lambda) f_i(x^k) + \lambda V_i e_k z_k - \frac{1}{2}\lambda (d^k)^T G_i^k d^k + o(\lambda) \leqslant -\frac{1}{2V_0}\lambda V_i e_k M_k \|d^k\|^2 + o(\lambda) \leqslant -\frac{1}{2V_0} V_i e_{\inf} \bar{M}\bar{k}^2 \lambda + o(\lambda)$.

This together with $\frac{1}{2V_0} V_i e_{\inf} \bar{M}\bar{k}^2 > 0$ show that the inequalities (9) hold for all k and $\lambda > 0$ small enough.

Summarizing the analysis above, we can conclude that there exists a $\bar{\lambda} > 0$ such that $\lambda_k \geqslant \bar{\lambda}$ hold for all k. So from inequality (8), formula (5), Lemma 2(i) and formula (10), we have

$f_0(x^{k+1}) \leqslant f_0(x^k) + T\lambda_k g_0(x^k)^T d^k \leqslant f_0(x^k) + T\lambda_k V_0 z_k \leqslant f_0(x^k) - \frac{1}{2} T\lambda_k M_k \|d^k\|^2 \leqslant f_0(x^k) - \frac{1}{2} T\bar{\lambda}\bar{M}\bar{k}^2$,

for all k. Thus the sequence $\{f_0(x^k)\}$ is decreasing. Since $\{x^k\}$ is bounded, there exist an infinite set K_1 and a point x^* such that $\lim_{k \in K_1} x^k = x^*$, and therefore $\lim_{k \in K_1} f_0(x^k) = f_0(x^*)$. So we have $\lim_{k \to \infty} f_0(x^k) = f_0(x^*)$. This contradicts formula (12), and the proof is completed.

Corollary 1 Suppose that Assumptions 1 and 2 hold. Then $\lim_{k \to \infty} M_k = 0$.

Lemma 5 Suppose that Assumptions 1 and 2 hold. If $\lim_{k \in K} d^k = 0$, then $\lim_{k \in K} z_k = 0$.

Proof From formula (5) and Lemma 2(i), we have $g_0(x^k)^T d^k / V_0 \leqslant z_k \leqslant 0$, for all k. This together with $\lim_{k \in K} d^k = 0$ and the boundedness of $\{x^k\}$ shows that $\lim_{k \in K} z_k = 0$.

In order to obtain the global convergence, we further make the following assumptions.

Assumption 3 Suppose that the Mangasarian-Fromovitz constraint qualification holds at each limit point x^* of $\{x^k\}$, i. e., there exists a vector $d \in R^n$ such that $g_i(x^*)^T d < 0, \forall i \in I(x^*) \overset{\text{def}}{=} \{i \in I: f_i(x^*) = 0\}$.

Assumption 4 The multiplier sequence $\{v_k\}$ is bounded.

Theorem 1 Suppose that Assumptions 1~4 hold. Then Algorithm A either stops finitely at a Fritz John point x^k for problem (1) or generates an infinite sequence $\{x^k\}$ of iterates such that there exists an accumulation point x^* which is a KKT point for problem (1).

Proof From Lemma 4, Lemma 5 and the boundedness of $\{x^k\}$, we have that there exist an index set K and a point x^* such that $(x^k, d^k, z_k) \to (x^*, 0, 0), k \in K$. So following the proof of Theorem 3.2 in Reference [1], we can conclude that x^* is a KKT point for problem (1).

Assumption 5 (i) $f_i, i \in \{0\} \cup I$, are the third-order continuously differentiable. (ii) The matrices $G_i^k, i \in \{0\} \cup I$, are chosen as $G_i^k = \nabla^2 f_i(x^k), i \in \{0\} \cup I$, if k is sufficiently large, and the parameter sequence $\{\epsilon_k\}$ satisfies $\lim_{k\to\infty} \epsilon_k = 0$. (iii) The sequence $\{G_0^k\}$ of matrices is uniformly positive definite, i.e., there exist two positive constants a and b such that $a\|d\|^2 \leqslant d^T G_0^k d \leqslant b\|d\|^2, \forall d \in R^n, \forall k$.

Using Corollary 1, similar to Theorem 4.2 in Reference [1], we can prove the following result.

Theorem 2 Suppose that Assumptions 2~5 hold. Then Algorithm A is superlinearly convergent, i.e., $\|x^{k+1} - x^*\| = o(\|x^k - x^*\|)$.

References

[1] Jian J B. New sequential quadratically constrained quadratic programming norm-relaxed method of feasible directions [J]. Journal of Optimization Theory and Applications, 2006, 129: 109-130.

[2] Solodov M V. On the sequential quadratically constrained quadratic programming methods [J]. Mathematics of Operations Research, 2004, 29: 64-79.

[3] Fukushima M, Luo Z Q, Tseng P. A sequential quadratically constrained quadratic programming method for differentiable convex minimization [J]. SIAM Journal on Optimization, 2003, 13: 1098-1119.

[4] Anitescu M. A superlinearly convergent sequential quadratically constrained quadratic programming algorithm for degenerate nonlinear programming [J]. SIAM Journal on Optimization, 2002, 12: 949-978.

[5] Liu T W. BFGS method and its applications in solving constrained optimization problems [D]. Changsha: Ph D Thesis, Hunan University, 2007.

[6] Lobo M S, Vandenberghe L, Boyd S, et al. Applications of second-order cone programming [J]. Linear Algebra and Its Appl, 1998, 284: 193-228.

(责任编辑: 尹　闯)

(上接第 104页 Continue from page 104)

Table 1 Iterations\ function evaluations

Problem	dim	N-WYL1	N-WYL2	WYL
1	2	35\ 70	35\ 70	40\ 84
2	4	37\ 74	37\ 74	40\ 77
3	2	16\ 38	13\ 30	19\ 33
4	2	7\ 20	5\ 14	6\ 21
5	2	16\ 38	11\ 28	F
6	2	25\ 50	23\ 47	F
7	4	31\ 62	30\ 61	31\ 65
8	2	18\ 32	13\ 28	10\ 21
9	4	27\ 54	25\ 50	F
10	6	53\ 108	50\ 103	55\ 100
11	100	285\ 574	217\ 438	292\ 580
12	2	26\ 52	25\ 50	27\ 58
13	4	27\ 54	26\ 52	F
14	100	79\ 161	78\ 158	82\ 166
15	2	70\ 141	69\ 140	65\ 131
16	2	272\ 544	267\ 534	260\ 551
17	20	99\ 201	98\ 196	103\ 220
18	200	629\ 1261	647\ 1295	650\ 1302
19	5	73\ 146	72\ 144	F
20	11	75\ 150	74\ 146	74\ 148

4 Conclusion

In this paper, a nonmonotone line search has been proposed for guaranteeing the global convergence of WYL conjugate gradient method. It needs to estimate the Lipschitz constant but the estimation is easy and available in practical. In particular, if $m = 0$, then the new nonmonotone line search will reduce to a monotone line search and the WYL conjugate gradient method with the monotone line search has also global convergence. The Numerical experiments show that WYL method with the nonmonotone line search is available and efficient in practical computation.

References

[1] Fletcher R, Reeves C. Function minimization by conjugate gradients [J]. Comput J, 1964(7): 149-154.

[2] Polyak B T. The conjugate gradient method in extreme problems [J]. USSR Comput Math Phys, 1969(9): 94-112.

[3] Hestenes M R, Stiefel E. Method of conjugate gradient for solving linear systems [J]. J Res, Nat Bur Stand, 1952, 49: 409-436.

[4] Armijo L. Minimization of functions having Lipschits continuous partial derivatives [J]. Pacific J Math, 1966(16): 1-3.

[5] Goldstein A A. On steepest descent [J]. SIAM J Control, 1965(3): 147-151.

[6] Wolfe P. Convergence conditions for ascent methods [J]. SIAM Rev, 1969(11): 226-235.

[7] Grippo L, Lucidi S. A globally convergent version of the Polak-Ribiere conjugate gradient method [J]. Math Prog, 1997, 78: 375-391.

[8] SHI Z J. Restricted PR conjugate gradient method and its global convergence [J]. Adv Math, 2002, 31(1): 47-55.

[9] WEI Z X, YAO S W. The convergence properties of some new conjugate gradient methods [J]. Appl Math Comput, 2006, 183: 1341-1350.

[10] SHI Z J, Shen J. Convergence of PRP method with new nonmonotone line search [J]. Appl Math Comput, 2006, 181: 423-431.

(责任编辑: 尹　闯)