A MODIFIED SQP ALGORITHM FOR MATHEMATICAL PROGRAMMING

Vasile MORARU, Sergiu ZAPOROJAN, Daniela ISTRATI
Technical University of Moldova
vasile.moraru@ia.utm.md zaporojan_s@yahoo.com daniela.istrati@ia.utm.md

Abstract **— Recent efforts in mathematical programming have been focused a popular sequential quadratic programming (SQP) method. In this paper, a method for mathematical programs with equalities and inequalities constraints is presented, which solves two subproblems at each iterate, one a linear programming subproblem and the other is a quadratic programming (QP) subproblem. The considered method assures that the QP subproblem, is consistent.**

Keywords — ***Sequential Quadratic Programming, constrained optimization, merit function, superlinear convergence.***

I. INTRODUCTION

We consider the mathematical programming problem with general equality and inequality constraints

$$\left.\begin{array}{l} f(x) \to \min \\ \text{subject to} \\ h(x) = 0, \\ g(x) \le 0, \end{array}\right\} \qquad (1.1)$$

where the objective function $f : \Re^n \to \Re$, and the constraint functions $h : \Re^n \to \Re^m, g : \Re^n \to \Re^p$ are assumed to be twice continuously differentiable.

We briefly will describe the notation used in this paper. All vectors are column vectors. The suberscript notation x_i referees to an element of the vector *x*. A supscript *k* is used to denote iteration numbers. Superscript "*T*" denotes transposition. $\Re^n$ denotes the space of *n*-dimensional real column vectors.

We denote by x^* a local solution of the problem (1.1). The Lagrangian function associated with the problem (1.1) is defined by $L(x,\lambda,\mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$, where $\lambda \in \Re^m, \mu \in \Re^p$ are vectors of Lagrange multipliers. Assume that a *Linear Independence Constraint Qualification* (*LICQ*) condition holds at x^*; then multipliers λ^* and $\mu^* \ge 0$ exist such that [1]:

$$\left.\begin{array}{l} \nabla_x L(x^*,\lambda^*,\mu^*) = 0, \\ (\lambda^*)^T g(x^*) = 0, \\ h(x^*) = 0, \\ g(x^*) \le 0. \end{array}\right\} \qquad (1.2)$$

A primal-dual solution (x^*,λ^*,μ^*) is said to be a Karush-Kuhn-Tucher (*KKT*) triple.

The basic idea of the typical sequential quadratic programming (*SQP*) is as follows [2]. Let the current *KKT* point be $(x^{(k)},\lambda^{(k)},\mu^{(k)})$. A new approximation $(x^{(k+1)},\lambda^{(k+1)},\mu^{(k)+1})$ to the solution is the procedure:

$$x^{(k+1)} = x^{(k)} + \alpha_k d^{(k)},\ \lambda^{(k+1)} = \lambda^{QP},\ \mu^{(k+1)} = \mu^{QP},$$

where $d^{(k)}$ is a search direction which minimizes a quadratic model subject to the linearized constraints

$$\left.\begin{array}{l} \frac{1}{2} d^T B_k d + \left[\nabla f(x^{(k)})\right]^T d \to \min \\ h_i(x^{(k)}) + \left[\nabla h_i(x^{(k)})\right]^T d = 0, i = 1,2,\dots,m, \\ g_i(x^{(k)}) + \left[\nabla g_i(x^{(k)})\right]^T d \le 0, i = 1,2,\dots,p, \end{array}\right\} \qquad (1.3)$$

and (λ^{QP},μ^{QP}) are taken as the Lagrange multipliers for (1.3). α_k is the step size along the direction chosen to reduce the value of the merit function [3-5]:

$$F_{c_k}(x) = f(x) + c_k \max\{0, |h_1(x)|,\dots,|h_m(x)|, g_1(x),\dots g_p(x)\},$$

where $c_k > \sum_{j=1}^{m}|\lambda_j^*| + \sum_{j=1}^{p}\mu_j^*$ is a penalty parameter.

The matrix B_k is a symmetric approximation to the Hessian of the Lagrangian function [6, 7]:

$$B_k \approx \nabla_{xx}^2 L(x^{(k)},\lambda^{(k)},\mu^{(k)}).$$

In traditional *SQP* method, the quadratic program (1.3) may be inconsistent; the feasible set of (1.3) may be empty. This is serious limitation of the *SQP* method. Several techniques for evitation of the inconsistency phenomen of the linearized constraints of the quadratic programming problem (1.3) were proposed [8-13]. Recently, in [11, 12], modifications of the SQP method were proposed where at each step two subproblems are resolved: one linear programming problem or one linear square problem and one quadratic programming problem.

The presented method in this paper was announced in [14] and is similar to [9] .At each iteration, two subproblems are solved – one is a linear programming; the other is a quadratic subproblem. Our algorithm is distinct from the one proposed in [9] in two important ways. Firstly, in both linear and quadratic programming problems, beside the inequality constraints of the problem (PNL), we also consider the equality ones. Secondly, at each iterate the linear programming subproblem is deferent from the one is [9]; we consider the local behavior of all constraints.

2. The modified SQP algorithm

We consider the following linear programming subproblem:

$$\left.\begin{array}{c}\varphi(y,z)=\sum_{i=1}^{m} y_i+\sum_{i=1}^{p} z_i \to \min\\ -y_i \le h_i\left(x^{(k)}\right)+\left[\nabla h_i\left(x^{(k)}\right)\right]^T d \le y_i, i=1,2,\ldots,m,\\ g_i\left(x^{(k)}\right)+\left[\nabla g_i\left(x^{(k)}\right)\right]^T d \le z_i, i=1,2,\ldots,p,\\ y_i \ge 0, z_i \ge 0, \forall i.\end{array}\right\} \quad (2.1)$$

Let $\tilde{d}^{(k)}, \tilde{y}_i^{(k)}, \tilde{z}_i^{(k)}$ be the solution of (2.1). If $x^{(k)}$ is feasible, we have $\tilde{d}^{(k)}=0$. Now we consider the following modified quadratic programming (*MQP*) subproblem:

$$\left.\begin{array}{c}\frac{1}{2} d^T B_k d+\left[\nabla f\left(x^{(k)}\right)\right]^T d \to \min\\ -\tilde{y}_i^{(k)} \le h_i\left(x^{(k)}\right)+\left[\nabla h_i\left(x^{(k)}\right)\right]^T d \le \tilde{y}_i^{(k)}, i=1,2,\ldots,m,\\ g_i\left(x^{(k)}\right)+\left[\nabla g_i\left(x^{(k)}\right)\right]^T d \le \tilde{z}_i^{(k)}, i=1,2,\ldots,p.\end{array}\right\} \quad (2.2)$$

Notice that $\tilde{d}^{(k)}$ is feasible solution of (2.2) so, the feasible region of this subproblem is nonempty. Let $d^{(k)}$ be the solution of *MQP* (2.2). If matrix B_k is positive definite, $d^{(k)}$ is unique and is a descendent direction of $F_{c_k}(x)$.

We now describe the proposed algorithm.

Step 0. *Given the initial approximate* $x^{(0)} \in \Re^n$, $\lambda^{(0)} \in \Re^m, \mu^{(0)} \in \Re^p$ *a* $n \times n$ *symmetric positive definite matrix* B_0, *an initial penalty parameter* $c_0 > 0$ *and the scalars* $\beta \in \left(0, \frac{1}{2}\right), \gamma \in (0,1)$; $k := 0$;

Step 1. *Solve subproblem (2.1) to obtain* $\tilde{d}^{(k)}, \tilde{y}_i^{(k)}, \tilde{z}_i^{(k)}$ *.If* $\tilde{d}^{(k)}=0$ *and* $\exists i\ \tilde{y}_i^{(k)} > 0$ *or*, $\tilde{z}_i^{(k)} > 0$, *stop;*

Step 2. *Solve subproblem (2.2) to generate* $d^{(k)}$. *If* $d^{(k)}=0$, *stop*;

Step 3. *Choose the penalty parameter* c_k *such that* $c_k > \sum_{j=1}^{m}\left|\lambda_j^{MPQ}\right|+\sum_{j=1}^{p}\mu_j^{MPQ}$;

Step 4. *Select the smallest positive integer s such that*
$F_{c_k}\left(x^{(k)}+\gamma^s d^{(k)}\right) \le F_{c_k}\left(x^{(k)}\right)-\beta\gamma^s\left(d^{(k)}\right)^T B_k d^{(k)}$.
Let $\alpha_k=\gamma^s$ *and*
$x^{(k+1)}=x^{(k)}+\alpha_k d^{(k)}, \lambda^{(k+1)}=\lambda^{MQP}, \mu^{(k+1)}=\mu^{MQP}$;

Step 5. *Choose a symmetric positive definite matrix* B_{k+1}. *Set* $k := k+1$. *Go to Step 1.*

The matrix B_k can be calculated using the technique from [6]. This guaranty that they are positive definite and they approximate the Hessian matrix $\nabla_{xx}^2 L\left(x^{(*)}, \lambda^{(*)}, \mu^{(*)}\right)$ on the "tangent" subspace of active constraints, and that the $\left\{x^{(k)}\right\}$ superlinear converge to $\left\{x^{(*)}\right\}$.

The efficiency of proposed SQP algorithm depends on the efficiency of the algorithm of solving quadratic programming sub problems (2.2). There are a great number of algorithms of solving quadratic programming problems. A relative complete bibliography of these methods can be found in [14].

References

1. Bertsekas D.P. Constrained optimization and Lagrange multiplier methods. Academic Press, 1982.
2. Boggs P., Tolle J. Sequential quadratic programming. Acta Numerica, 4, 1995, pp. 1-51
3. Pshenichniy B. N., Danilin Iu. M. Numerical methods in extremal problems. (In Russian) –M.: Nauka, 1975, 319 p.
4. Han S. P. A globally convergent method for nonlinear programming. Journal of Optimization Theory and Applications 15, 1977, pp.319-342.
5. Gould N., Toint Ph. SQP methods for large-scale nonlinear programming. In System modeling and optimization, methods, theory and applications (Powel M. and Scholtes S. editors), Kluwer, 2000, pp.149-178.
6. Moraru V. Quasi-Newton methods for solving nonlinear programming problems. Computer Science Journal of Moldova, 3, No.3, 1995, pp.263-277.
7. Schittkowski K. The nonlinear programming method of Wilson, Han and Powell with an augmented Lagrange type line search function. Part 1 : convergence analysis. Numer. Math. 38, 1981, pp. 83-114.
8. Burke J. V., Han S. P. A robust SQP method. Mathematical Programming, 43, 1989, pp. 277-303.
9. Zhang J., Zhang X. A robust SQP method for optimization with inequality constraints. Journal of Computational Mathematics, 21, No. 2, 2003, pp. 247-256.
10. Liu X., Yuan Y. A robust algorithm for optimization with general equality and inequality constraints. SIAM Journal on Scientific and Computing, 22, No. 2, 2000, pp. 517-534.
11. Wan Z. Global convergence of a modified SQP method for mathematical programs with inequalities and equalities constraints. Journal of Inequalities in Pure and Applied Mathematics, 5, No.2, 2004, Art. 37.
12. Byrd R.H., Gould N., Nocedal J., Waltz R. On the convergence of successive linear-quadratic programming algorithms. Mathematical Programming, 100, No. 1, 2004, pp. 27-48.
13. Moraru V. (1997) An Algorithm for Solving Quadratic Programming Problems. Computer Science Journal of Moldova, vol.5, No. 2, pp.223-235.
14. Moraru V. Primal-Dual Method for Solving Convex Quadratic Programming Problems. Computer Science Journal of Moldova, Volume 8, No.2, 2000, pp.209-220.