ISSN 1009 - 3095 Journal of Zhejiang University(SCIENCE) V.2, No.1, P.66 - 70, Jan. - Mar., 2001
http://www.chinainfo.gov.cn/periodical; http://www.zju.edu.cn;
http://lib.zju.edu.cn/English; http://www.zjupress.com

SOLVING CONVEX QUADRATIC PROGRAMMING BY POTENTIAL-REDUCTION INTERIOR-POINT ALGORITHM

LIANG Xi-ming(梁昔明)[1], MA Long-hua(马龙华)[2], QIAN Ji-xin(钱积新)[2]

([1] *College of Information Science & Engineering, Central South University, Changsha 410083, China*)

([2] *Institute of Systems Engineering, National Lab of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China*)

Received June 18, 1999; revision accepted Apr. 18, 2000

Abstract: The solution of quadratic programming problems is an important issue in the field of mathematical programming and industrial applications. In this paper, we solve convex quadratic programming by a potential-reduction interior-point algorithm. It is proved that the potential-reduction interior-point algorithm is globally convergent. Some numerical experiments were made.

Key words: potential-reduction interior-point algorithm, convex quadratic programming, convergence, numerical experiments
Document code: A **CLC number**: O224

INTRODUCTION

In this paper, we consider the following convex quadratic programming

$$\begin{cases} \min f(x) = \frac{1}{2}\boldsymbol{x}^{\mathrm{T}} Q \boldsymbol{x} + \boldsymbol{c}^{\mathrm{T}} \boldsymbol{x} \\ \text{s.t. } \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b},\ \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases} \tag{1}$$

where $\boldsymbol{c}$, $\boldsymbol{x}$ are n-vectors, $\boldsymbol{b}$ is an m-vector, $\boldsymbol{A}$ is a matrix and $\boldsymbol{Q}$ is a symmetric positive semidefinite $m \times n$ matrix. The form of (1) does not lose generality, because any p equality constraints $\sum_{j=1}^{n} a_{ij}x_j = d_i$, $i = 1, \cdots, p$, can be transformed equivalently into the $p + 1$ inequality constraints $\sum_{j=1}^{n} a_{ij}x_j \leqslant d_i$, $i = 1, \cdots, p$, $-\sum_{j=1}^{n}(\sum_{i=1}^{p} a_{ij})x_j \leqslant -\sum_{i=1}^{p} d_i$.

The solution of quadratic programming (QP) is in itself an important issue in mathematical programming and industrial applications. QP solvers are also the key components of successive quadratic programming algorithms for the general nonlinear programming problem. The earliest methods based on the simplex method for linear programming (LP), such as the Lemke pivoting algorithm, which solves the Karush - Kuhn - Tucker (KKT) necessary conditions for QP (Lucia and Xu, 1990). These methods can be used to identify feasible solutions, but they increase the number of variables by introducing slack variables to handle inequality constraints, which can cause a significant deceleration of the solution procedure. Another approach uses the active set concept and solves the KKT necessary conditions directly (Fletcher, 1987). These methods are generally robust, but a large number of inequality constraints might significantly slow down their convergence. A final major area is the reduced space quadratic programming approach, which has become of great interest in recent years and has been studied extensively. The key concept of this approach is to project the Hessian to the space of the degrees of freedom thus getting a significantly reduced number of free variables (Lucia et al., 1993). This holds if the number of degrees of freedom is small, as is the case in most practical chemical engineering applications (Vasantharajan et al., 1990). Kozlov, Tarasov and Khachiyan (1979) presented polynomial-time algorithms for convex quadratic programming problems based on the ellipsoid method. Recently, with the advent of the new interior point algorithm by Karmarkar (1984) for solving LP problems, some interior point methods had been studied for addressing the full - space QP problems. Kapoor and Vaidya (1986) and Ye and Tse (1986) presented interior point algorithms for convex QP problems based on Karmarkar's projective transformation. Monteiro and Adler (1989) described a primal-dual interior point algorithm for convex QP problems based on the path-following idea.

We will develop a potential – reduction interior – point algorithm for solving convex QP of the form (1) in this paper. At each iteration, a system of linear equations is solved to get a search direction. Then, along the search direction, we take a line search using Armijo's rule to reduce the value of a potential function. It is proved that for any initial point, the sequence generated by the algorithm is bounded and that any accumulation point of the iterative sequence is a solution of (1). Some numerical experiments were made.

In the remainder of the paper, we use R^n_+, R^n_{++} and e_n to denote the nonnegative orthant, positive orthant and the vector of all one's in R^n, respectively. For any vector $\boldsymbol{x}$, we use $\boldsymbol{X}$ to denote the diagonal matrix with diagonal entries x_i.

THE POTENTIAL-REDUCTION INTERIOR-POINT ALGORITHM

It is well known that, at the optimal solution to (1), the following necessary and sufficient conditions of optimality are satisfied:

$$\begin{aligned} &\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x} = \boldsymbol{v},\ \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{u}, \\ &\boldsymbol{x}^{\mathrm{T}}\boldsymbol{u} = \boldsymbol{0},\ \boldsymbol{y}^{\mathrm{T}}v = \boldsymbol{0}, \\ &x, y, u, v \geqslant \boldsymbol{0} \end{aligned} \tag{2}$$

For $z = (\boldsymbol{x}^{\mathrm{T}}, \boldsymbol{y}^{\mathrm{T}})^{\mathrm{T}} \in R^{n+m}$, we define $\boldsymbol{u} = \boldsymbol{Q}\boldsymbol{x} + \boldsymbol{c} + \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}$, $\boldsymbol{v} = \boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}$, $g(z) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{u} + \boldsymbol{y}^{\mathrm{T}}\boldsymbol{v}$ and

$$\begin{aligned} \Omega &= \{z \in R^{n+m} : \boldsymbol{x} \geqslant \boldsymbol{0},\ \boldsymbol{y} \geqslant \boldsymbol{0},\ \boldsymbol{u} \geqslant \boldsymbol{0}, \\ &\quad \boldsymbol{v} \geqslant \boldsymbol{0}\}, \\ \Omega^0 &= \{z \in R^{n+m} : \boldsymbol{x} > \boldsymbol{0},\ \boldsymbol{y} > \boldsymbol{0},\ \boldsymbol{u} > \boldsymbol{0}, \\ &\quad \boldsymbol{v} > \boldsymbol{0}\} \end{aligned}$$

We assume that Ω^0 be nonempty. Noticing that $g(z) > 0$ for all $z \in \Omega^0$ and that z^* satisfies (2) if and only if $z^* \in \Omega$ and $g(z^*) = 0$, we define a potential function $p(z)$ as follows:

$$p(z) = \gamma \log(z) - \sum_{i=1}^{n} \log(x_i u_i) - \sum_{j=1}^{m} \log(y_j v_j), z \in \Omega^0$$

where $\gamma > n + m$ is any given scalar. From the arithmetic-geometric mean inequality, we have

$$p(z) \geqslant (\gamma - n - m)\log(z) + (n + m)\log(n + m),\ \forall z \in \Omega^0 \tag{3}$$

The proposed algorithm for solving (1) is iterative. At each iterate z^k, we first obtain a search direction $\boldsymbol{d}^k$ by solving the following system of linear equations:

$$\nabla \boldsymbol{L}(z^k)\boldsymbol{d} + \boldsymbol{L}(z^k) = \sigma_k \beta(z^k) e_{n+m} \tag{4}$$

where $\boldsymbol{L} : R^{n+m} \rightarrow R^{n+m}$ is defined by $\boldsymbol{L}(z) = ((\boldsymbol{x} \circ \boldsymbol{u})^{\mathrm{T}}, (\boldsymbol{y} \circ \boldsymbol{v})^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{x} \circ \boldsymbol{u} = (x_1 u_1, \cdots, x_n u_n)^{\mathrm{T}}$ and $\boldsymbol{y} \circ \boldsymbol{v} = (y_1 v_1, \cdots, y_m v_m)^{\mathrm{T}}$ are the Hadamard products of vectors $\boldsymbol{x}$, $\boldsymbol{u}$ and $\boldsymbol{y}$, $\boldsymbol{v}$, respectively, $\beta(z) = \frac{g(z)}{n + m}$ and $\sigma_k \in [0, 1]$ is a centering parameter. Then, we determine a stepsize along the direction $\boldsymbol{d}^k$ to reduce the value of the potential function $p(z)$ and obtain a new iterate. Summarily, we solve the convex quadratic programming (1) by the following potential-reduction interior-point algorithm.

ALGORITHM PR

1. Give $\alpha, \rho \in (0, 1)$, $\bar{\sigma} \in [0, 1)$, $\beta > 0$ and choose $z^0 \in \Omega^0$ and. Set $k = 0$.

2. Solve the system (4) to get the search direction $\boldsymbol{d}^k$.

3. Let m_k be the smallest nonnegative integer m such that

$$z^k + \beta\rho^m \boldsymbol{d}^k \in \Omega^o,\ p(z^k + \beta\rho^m \boldsymbol{d}^k) - p(z^k) \leqslant - \alpha\beta\rho^m (1 - \sigma_k)(\gamma - n - m)$$

Set $z^{k+1} = z^k + \beta\rho^{m_k}\boldsymbol{d}^k$.

4. Terminate if z^{k+1} satisfies a prescribed stopping rule;

Otherwise pick $\sigma_{k+1} \in [0, \bar{\sigma})$ set $k = k + 1$ and go to Step 2.

(1) The initial point z^0 can be obtained via the method used by McCormick (1983).

(2) Suitable choice of the centering parameter σ_k is necessary. In the algorithm PR, σ_k is adjusted according to the following rule: if the stepsize at the $(k - 1)$-th iteration is not less than one, then $\sigma_k = \frac{1}{10}\sigma_{k-1}$; otherwise $\sigma_k = \sigma_{k-1}$.

(3) A natural choice for the stopping rule in Step 4 is to check if $g(z^{k+1})$ is less than a given tolerance ε (see the next section). If $g(z^{k+1}) \leqslant \varepsilon$, z^{k+1} is accepted as an approximate point satisfying the condition (2) and an approximate solution x^{k+1} to (1) is then obtained.

To prove the search direction d^k in Step 2 is unique, we first prove the following lemma.

Lemma 1 Let C and D be positive definite diagonal matrices and B be a positive semidefinite matrix. Then the matrix $C + DB$ is nonsingular.

Proof By the assumption, the matrix $E = C^{-1}D$ is a positive definite diagonal. Let $E = E^{0.5}E^{0.5}$, then $C^{-1}DB = E^{0.5}E^{0.5}BE^{0.5}E^{-0.5}$ and the eigenvalues of matrices $C^{-1}DB$ and $E^{0.5}BE^{0.5}$ are the same. Hence all eigenvalues of the matrix $I + C^{-1}DB$ are not less than one and the matrix $C + DB = C(I + C^{-1}DB)$ is nonsingular.

Theorem 1 The matrix $\nabla L(z)$ is nonsingular for all $z \in \Omega^0$ and then d^k in Step 2 is unique.

Proof Let $d = (d_1^T, d_2^T)^T \in R^{n+m}$ be any solution of the homogeneous system $\nabla L(z)d = \mathbf{0}$. Then $(U + XQ)d_1 + XA^Td_2 = \mathbf{0}$, $d_2 = V^{-1}YAd_1$ and $(U + X(Q + A^TV^{-1}YA))d_1 = \mathbf{0}$. By lemma 1 and the assumption that Q is positive semidefinite, the matrix $U + X(Q + A^TV^{-1}YA)$ is nonsingular. So $d = \mathbf{0}$ is the unique solution of $\nabla L(z)d = 0$ and the proof is completed.

Theorem 1 shows that for any $z \in \Omega^0$, the system of linear equations (4) has a unique solution d. The next theorem will show that the solution d is also a descent direction of the potential function $p(z)$ at z.

Theorem 2 Let $d = (d_1^T, d_2^T)^T \in R^{n+m}$ be the solution of (4) at z, then

$$\nabla p(z)^T d \leqslant -(1-\sigma)(\gamma - n - m) ; \quad (5)$$

Hence the integer m_k in Step 3 can be determined in a finite number of time.

Proof The function $p(z)$ is continuously differentiable on the set Ω^0 and by the arithmetic-geometric mean inequality, we obtain

$$\nabla p(z)^T d = \frac{\gamma}{g(z)} \cdot ((u^T + x^TQ)d_1 - y^TAd_1 + x^TA^Td_2 + v^Td_2) - \left(\frac{1}{x_1u_1}, \cdots, \frac{1}{x_nu_n}\right)((U + XQ)d_1 + XA^Td_2 - \left(\frac{1}{y_1v_1}, \cdots, \frac{1}{y_mv_m}\right)(-YAd_1 + Vd_2) \leqslant -(1-\sigma)(\gamma - n - m).$$

Thus d is also a descent direction of the potential function $p(z)$ at z. Because Ω^0 is an open set, there exists $\bar{\lambda} > 0$ such that, for all $\lambda \in (0, \bar{\lambda})$, $z + \lambda d \in \Omega^0$ and $p(z + \lambda d) - p(z) \leqslant -\alpha\lambda(1-\sigma)(\gamma - n - m)$. Hence, starting with $m = 0$ and increasing m by one after a trial fails, the integer m_k in Step 3 can be determined in a finite number of trials.

CONVERGENCE

The convergence result of the algorithm PR is stated as follows.

Theorem 3 The sequence $\{z^k\}$ generated by the algorithm PR is bounded, every accumulation point of $\{z^k\}$ satisfies the condition (2) and $\lim_{k\to\infty} g(z^k) = 0$.

Proof From the assumption that Q is positive semidefinite and (3), we have

$$\min(u_1^0, \cdots, u_n^0, v_1^0, \cdots, v_m^0)(\|x^k\|_1 + \|y^k\|_1) \leqslant x^{k^T}u^0 + y^{k^T}v^0 + x^{0^T}u^k + y^{0^T}v^k \leqslant x^{0^T}u^0 + y^{0^T}v^0 + x^{k^T}u^k + y^{k^T}v^k \leqslant 2\exp\left(\frac{p(z^0) - (n+m)\log(n+m)}{\gamma - n - m}\right)$$

This shows that the sequence $\{z^k\}$ is bounded.

Let z^* be an accumulation point of the sequence $\{z^k\}$. Then $z^* \in \Omega$ and there is a subsequence $\{z^{k_l}\}$ of $\{z^k\}$ such that $z^{k_l} \to z^*$. To prove that z^* satisfies the condition (2), we only need to show $g(z^*) = 0$. Suppose that $g(z^*) > 0$. Then, there is $\delta > 0$ such that $g(z^{k_l} \geqslant \delta$ holds for l large enough. Because $\{p(z^k)\}$ is bounded above by $p(z^0)$, $\{L(z^{k_l}\}$ is contained in a compact subset of R_{++}^{n+m} (Monteiro, 1994). Thus $L(z^*) > 0$ and $z^* \in \Omega^0$. By theorem 1, $\nabla L(z^*)$ is nonsingular. Hence the sequence $\{\nabla L(z^{k_l}\}$ converges to $\nabla L(z^*)^{-1}$. Since $0 \leqslant \sigma_k \leqslant \bar{\sigma} \leqslant 1$, without loss of generality, we assume that $\{\sigma_{k_l}\}$ converges to $\sigma^* \in [0,1)$.

Since $d^{k_l} = \nabla L(z^{k_l})^{-1}(-L(z^{k_l}) + \sigma_{k_l}\beta(z^{k_l}) e_{n+m})$, $\{d^{k_l}\}$ will converge to a vector d^* that satisfies the equation $\nabla L(z^*)d^* + L(z^*) = \sigma^* \beta(z^*) e_{n+m}$. Hence, by (5) and (0,1), we have

$$\nabla p(z^*)^T d^* < -\alpha(1-\sigma^*)(\gamma - n - m); \quad (6)$$

On the other hand, since $p(z^{k_l}) \to p(z^*)$, the continuity of $p(z)$ and the strictly decreasing of $\{p(z^k)\}$ imply that the sequence $\{p(z^k)\}$ converges to $p(z^*)$ and $\lim_{k\to\infty}(p(z^{k+1}) - p(z^k)) = 0$. Since $\{\sigma_k\}$ is bounded above by 1, it follows $\lim_{k\to\infty} \rho^{m^k} = 0$, i.e., $m_k \to \infty$. Since $\{z^k\} \subseteq \Omega^0$ and $\{d^{k_l}\}$ is bounded, for any l large enough, there is an $\bar{m}_l$ such that $\bar{m}_l \to \infty$, $z^l + \beta\rho^{\bar{m}_l} d^l \in \Omega^0$, but $p(z^l + \beta\rho^{\bar{m}_l} d^l) - p(z^l) > -\alpha\beta\rho^{\bar{m}_l}(1 - \sigma_l)(\gamma - n - m)$. Therefore we have

$$\nabla p(z^*)^T d^* \geqslant -\alpha(1-\sigma^*)(\gamma - n - m),$$

which contradicts (6), and hence $g(z^*) = 0$.

From the above conclusions, we deduce that $\{z^k\}$ is bounded and $g(\bar{z}) = 0$ for every accumulation point $\bar{z}$ of the sequence $\{z^k\}$, so $\lim_{k\to\infty} g(z^k) = 0$ holds.

The above theorem implies that $g(z^k) \leqslant \varepsilon$ can be used as a stopping rule where $\varepsilon > 0$ is a given tolerance and that for any starting point $z^0 \in \Omega^0$, the iterative sequence $\{z^k\}$ has accumulation point and every accumulation point z^* satisfies the condition (2) and hence x^* is a solution to (1).

NUMERICAL EXPERIMENTS

In this section, we make some numerical experiments on the algorithm PR. All the quadratic programming problems in Hock and Schittkowski (1981) and Schittkowski (1987) are solved. The initial points z^0 for all test problems are obtained via the method used by McCormick (1983). The parameters used in the experiments are chosen as follows:

$$\gamma = n + m + 10,\ \bar{\sigma} = \sigma^0 = 0.5,\ \rho = \alpha = 0.5,\ \beta = 2,$$

and the tolerance is $\varepsilon = 10e-6$. The computer codes are written in Turbo C 2.0 with double precision. We run the experiments on an AMD K6-2 computer. The results obtained are tabulated in Table 1, where the number of iterations NI, the number of potential function evaluations NP (this number is related to the total number of the line search steps), the number of objective function evaluations NF, the final objective value $f(x^*)$, the final value $p(x^*)$ of the potential function and the value of $g(x^*)$ are given. The information we provided should be a good indication about the general performance of the potential-reduction interior-point algorithm.

Table 1 Numerical results on the algorithm PR

Problem	n	m	NI	NF	NP	$f(x^*)$	$p(x^*)$	$g(x^*)$
HS3	2	1	8	9	9	2.97113e-13	-1.85427e2	1.58973e-9
HS21	2	3	9	10	10	-9.9960000e1	-1.38191e2	4.17337e-7
HS28	3	1	8	13	9	1.93287e+20	-1.59833e2	8.23015e-8
HS35	3	1	8	13	9	0.111111184	-1.52561e2	1.23435e-7
HS44	4	6	8	10	9	-1.5000000e1	-1.29559e2	2.36215e-7
HS48	5	2	6	8	7	3.27156e+15	-2.19993e2	1.56438e-10
HS51	5	3	5	7	6	5.81852e+21	-1.64427e2	5.19880e-8
HS52	5	3	6	8	7	5.326647564	-1.85251e2	6.47912e-9
HS53	5	8	7	10	8	4.093023256	-1.76228e2	1.27127e-8
HS76	4	3	11	14	12	-4.68181818	-1.69120e2	1.15763e-8
HS118	15	44	12	14	13	6.648205e+2	5.98163e1	1.41172e-8
HS224	2	6	7	8	8	-3.0400000e+2	-1.08764e2	3.03915e-7
HS268	5	5	5	7	6	4.87658e-17	-2.12502e2	6.10987e-6
HS392	30	45	9	12	10	-1.698880e7	-1.48766e2	5.95867e-6

We summarize the results of numerical experimentation in the following comments:

1. The algorithm PR solves all problems to the given tolerance and the number of iterations is small.

2. The cost for generating the search direction is low: only a system of linear equations is solved.

3. The algorithm PR is stable and robust for solving convex quadratic programming.

References

Fletcher, R., 1987. Practical Methods of Optimization, 2nd edition. Wiley Pub, New York, p.115 – 174.

Hock, W. and Schittkowski, K., 1981. Test Examples for Nonlinear Programming Codes. Springer, Berlin. p.26 – 140

Kappor, S. and Vaidya, P. M., 1986. Fast algorithms for convex quadratic programming and multicommodity flows, Proceedings of the 18th Annual ACM Symposium on Theory of Computing. California, p.147 – 159.

Karmarkar, N., 1984. A new polynomial time algorithm for linear programming. *Combinatorica*, **4**: 373 – 393.

Kozlov, M. K., Tarasov, S. P. and Khachiyan, L. G., 1979. Polynomial solvability of convex quadratic programming. *Soviet Mathematics Doklady*, **20**: 1108 – 1111.

Lucia, A. and Xu, J., 1990. Chemical process optimization using Newton-like methods. *Comput. Chem. Engng.*, **14**: 119 – 138.

Lucia, A., Xu, J. and D'Couto, G. C., 1993. Sparse quadratic programming in chemical process optimization. *Ann. Oper. Res.*, **42**: 55 – 83.

McCormick, G. P., 1983. Nonlinear Programming: Theory, Algorithms and Applications. John Wileg & Sons, Inc. Chichester, p. 47 – 132.

Monteiro, R. D. C., 1994. A globally convergent primal-dual interior point algorithm for convex programming. *Mathematical Programming*, **64**: 123 – 147.

Monteiro, R. D. C. and Adler, I., 1989. Interior path following primal-dual algorithms. Part II: Convex quadratic programming. *Mathematical Programming*, **44**: 43 – 66.

Schittkowski, K., 1987. More Test Examples for Nonlinear Programming Codes. Springer, Berlin, p. 48 – 210.

Vasantharajan, S., Viswanathan, J. and Biegler, L. T., 1990. Reduced successive quadratic programming implementation for large-scale optimization problems with smaller degrees of freedom. *Comput. Chem. Engng.*, **14**: 907 – 915.

Ye, Y. and Tse, E., 1986. A polynomial algorithm for convex programming, Working Paper, Department of Engineering-Economic Systems. Stanford University, Stanford, CA. p.214 – 227