Quadratic Regularization for Global Optimization

Anatolii Kosolap

anivkos@ua.fm

Abstract. In this paper we present a novel global optimization method for solving continuous nonlinear optimization problems. This method is based on an exact quadratic regularization (EQR). It allows the given problems to convert to equivalent problem maximization the norm of a vector on a convex set. Such problems are easier to solve than general nonlinear optimization problems. For the solution of these problems we use only primal-dual interior point method and a dichotomy search. The comparative numerical experiments have proved that EQR method to be very efficient.

INTRODUCTION

Many problems referring to economy, finance, project optimization, planning, computer graphics, management, scheduling, sensor networks of difficult systems can be transformed to optimization problems in finite-dimensional space. Such problems contain many local minima and belong to NP-difficult class. It is necessary to develop new methods of global optimization for the solution of these problems.

The existing methods in global optimization can be classified as deterministic and probabilistic. Deterministic ones include: Lipschitzian, Branch and Bound, Cutting Plane, Difference of Convex Function, Outer Approximation, Reformulation-Linearization, Semidefinite programming, Interval methods [1]. They demand the exponential number of iterations for finding global extremum. The probabilistic methods include random search, genetic and evolutionary methods [2]. However, these methods allow us to find global extremum only with some probability.

In this paper we propose a method of exact quadratic regularization for deterministic global optimization. This method can be used for the solution of a wide class of multiextreme problems. The EQR method includes methods of local optimization and of a dichotomy search. Particularly we use effective primal-dual interior point methods [3] for of local optimization.

EXACT QUADRATIC REGULARIZATION

The exact quadratic regularization allows us to construct the most effective method in global optimization. We consider the problem

$$\min\{f_0(x) \mid f_i(x) \le 0, i = 1,\ldots,m, x \in E^n\}, \tag{1}$$

where all functions $f_i(x)$ are twice differentiable, x is a vector in n-dimensional Euclidean space E^n. Let the solution of a problem (1) exists, its feasible domain is bounded and x^* is the point of global minimum (1). We converted the problem (1) to the following one

$$\min\{d \mid g_i(z) \le d, i = 0,\ldots,m, r\|z\|^2 \le d\} \tag{2}$$

or

$$\max\{\|z\|^2 \mid g_i(z) \le d, i = 0,\ldots,m\}. \tag{3}$$

where $z = (x, x_{n+1})$ and all $g_i(z), i = 0,\ldots,m$ are strong convex functions

$$g_0(z) = f_0(z) + s + (r-1)\|z\|^2, g_i(z) = f_i(z) + r\|z\|^2, i = 1,...,m.$$

The value s is chosen so that

$$f_0(x^*) + s \geq \|x^*\|^2. \tag{4}$$

Thus the problem (1) is converted to the maximization of a norm of a square vector. The value $r > 0$ exists so that all functions $f_i(z) + r\|z\|^2$ are convex on the bounded feasible domain of the problem (1). It follows from the fact that Hessians of these functions are positively defined matrixes (matrixes with a dominant main diagonal).

If the condition (4) holds, then the first constraint of the problem (3) is active. We will show that the minimum value $f_0(x)$ is arrived for the minimum value of d in of the problem (3).

Let z^1 and z^2 are two local maxima of a problem (3) and $\|z^2\| > \|z^1\|$ then we must prove that $f_0(z^2) < f_0(z^1)$. Consider the equalities

$$f_0(z^1) + s + (r-1)\|z^1\|^2 = d,$$

$$f_0(z^2) + s + (r-1)\|z^2\|^2 = d.$$

It follows that

$$f_0(z^2) - f_0(z^1) = (r-1)(\|z^1\|^2 - \|z^2\|^2) < 0$$

and $f_0(z^2) < f_0(z^1)$.

Let z^1 and z^2 are two local maxima and the conditions

$$r\|z^1\|^2 = d_1,$$

$$r\|z^2\|^2 = d_2,$$

hold and $d_2 < d_1$ then

$$f_0(z^1) + s + (r-1)\|z^1\|^2 = d_1,$$

$$f_0(z^2) + s + (r-1)\|z^2\|^2 = d_2.$$

These conditions are easily transformed to the equality

$$f_0(z^2) - f_0(z^1) + (r-1)(\|z^2\|^2 - \|z^1\|^2) = d_2 - d_1$$

or

$$f_0(z^2) - f_0(z^1) = \|z^2\|^2 - \|z^1\|^2 < 0.$$

Thus a point of global maximum of a problem (3) corresponds to the minimal value of d. We find this minimum value d using a dichotomy algorithm for variable d.

Let d_0 is the solution of the convex problem (2). We use an iterative procedure by a dichotomy algorithm

$$d_k = d_{k-1} + \alpha(d_{k-1} - r\|z\|^2), \tag{5}$$

where $\alpha \in (0, 1]$, for the search minimal value of d. In particular we can get the solution of problem (1) by solving only the convex problem (2).

Let (z^0, d_0) is the solution of the convex problem (2). If $r\|z^0\|^2 = d_0$ holds, then z^0 is the solution of the problem (1). Otherwise, we will solve a problem (3) for the different fixed values of a variable d to the following algorithm.

Algorithm

1. Choose r, s and ε.
2. Solve a convex problem (2). Let (z^0, d_0) – is its solution. If $|r\|z^0\|^2 - d_0| \leq \varepsilon$ then stop, the solution of problem (1) is x^0.
3. Choose α and compute d_k by using the formula (5). Solve problem (3) for $d = d_k$. Let z^k – is its solution.
4. If $|r\|z^k\|^2 - d_k| \leq \varepsilon$ and d_k is the minimal value then stop, the solution of problem (1) is x^k. Otherwise go back to 3.

We can specify the values of parameters r, s at the iterations of algorithm. We use next law for the choice of parameter α. If an objective function $f_0(x)$ begins to increase, then we decrease a value of parameter α. The number of iterations is determined the algorithm of dichotomy. Following two simple examples demonstrate complexity of algorithm.

Example 1. Consider the optimization problem

$$\min\{\|x\|^2 | -4x_1^2 - x_2^2 - x_1x_2 - 2x_1 - 22x_2 + 54 \leq 0,\ 3x_1^2 - 2x_2^2 + 8x_1x_2 + 64x_1 - 30x_2 + 102 \leq 0\}.$$

This problem has 3 local minima. Here s = 0, r = 5. The solution of a convex problem

$$\min\{d \mid x_1^2+4x_2^2+5x_3^2-x_1x_2-2x_1-22x_2+54\le d,\ 8x_1^2+3x_2^2+5x_3^2+8x_1x_2+64x_1-30x_2+102\le d\}$$

it is at a point

x^0 = (– 0.293853, 2.26301, 0.00052), d_0 = 26.037704 and determines the global minimum x^* = (– 0.293853, 2.26301) of the initial problem.

Example 2. Consider the optimization problem

$$\min\{\|x\|^2 \mid -4x_1^2+2x_2^2+5x_1-6x_2+10\le 0,\ -2x_1^2-4x_2^2-8x_1+6x_2+5\le 0\}.$$

This problem also has 3 local minima. But in this case the converted problem

$$\max\{\|x\|^2 \mid x_1^2+7x_2^2+5x_3^2+5x_1-6x_2+10\le 0,\ 3x_1^2+x_2^2+5x_3^2-8x_1+6x_2+5\le 0\}$$

also has 3 local maxima. For this problem we must use a dichotomy method. This method allows us to compute only a point of local minimum. However, displacement the feasible domain in this problem allows us to get a point of global minimum.

Further we consider such particularly a case for problem (3), then a convex feasible domain of the problem (3) is a regular polyhedron or rectangular parallelepiped. In these cases we can easy to find the solution of the problem (3).

Let S is a feasible domain of the problem (3). The convex set S can be circumscribed of a sphere.

Theorem 1. *Let S is convex set,* $S\in\{x \mid \|x-c\|^2\le r^2\}$ *and* x^* *– is the solution of convex problem*

$$\max\{c^T x \mid x\in S\}$$

then x^* *– is the point of global maximum of problem*

$$\max\{\|x\|^2 \mid x\in S\}$$

if $\|x^*-c\|^2=r^2$.

The proof follows from this, that the side $(0,x^*)$ of the triangle $(0,x^*,c)$ is more than the side $(0,x^0)$ of the triangle $(0,x^0,c)$ for any $x^0\in S$.

This theorem shows that the solution of problem (3) we must to search in the direction of center of feasible domain.

The convex set can be circumscribed of a parallelepiped P. The faces of the parallelepiped can be found by solving the sequence of problems

$$\min\{(z^*)^T z \mid z\in S_1\},\quad \max\{(z^*)^T z \mid z\in S_1\} \tag{6}$$

$$\min\{(a^i)^T z \mid z\in S_1\},\ i=1,\dots,n,\ \max\{(a^i)^T z \mid z\in S_1\},\ i=1,\dots,n, \tag{7}$$

where

$$a^i=(-z_1^*,\dots,z_{i-1}^*,\frac{\|z^*\|^2}{z_i^*}-z_i^*,-z_{i+1}^*,\dots,z_n^*),\ i=1,\dots,n.$$

The center of a parallelepiped is found at the point

$$c=\left(\frac{(b_0+b_1)z_1^*}{\|z^*\|^2},\dots,\frac{(b_0+b_{n-1})z_{n-1}}{\|z^*\|^2},(b_n-\sum_{i=1}^{n-1}\frac{(b_0+b_i)(z_i^*)^2}{\|z^*\|^2})/z_n^*\right)$$

where b_i is mean value of solutions the problems (6–7) $(b_i=(\min+\max)/2)$. Then a solution of the problem

$$\max\{c^T x \mid x\in P\}$$

gives an upper bound of a solution of the problem (3).

The solution of problem (3) can be simplified and by means of displacement of a feasible domain. We will show it on a simple example. The problem in example 2 is equivalent to problem

$$\min\{\|x-3\|^2 \mid -4(x_1-3)^2+2(x_2-3)^2+5(x_1-3)-6(x_2-3)+10\le 0,\ -2(x_1-3)^2-4(x_2-3)^2-8(x-3)_1+6(x_2-3)+5\le 0\}.$$

We use EQR to convert this problem to

$$\max\{\|x\|^2\|\|x-3\|^2+1+4(x_1^2+x_2^2+x_3^2)\le d,$$
$$-4(x_1-3)^2+2(x_2-3)^2+5(x_1-3)-6(x_2-3)+10+5(x_1^2+x_2^2+x_3^2)\le d,$$
$$-2(x_1-3)^2-4(x_2-3)^2-8(x-3)_1+6(x_2-3)+5+5(x_1^2+x_2^2+x_3^2)\le d\},$$

where s = 1 and r = 5. This problem is unimodal and we can easily solve it by using our algorithm.

When we use the displacement of the feasible domain then the curvature of the convex surface of the feasible domain does not change and the curvature of the sphere $r\|x\|^2 = d$ tends to zero. It allows us to decrease the number of local maximums in the problem (3). Thus, the displacement of the feasible domain often entails unimodality of the transformed problem. We used the displacement of space at the solution of most test problems.

NUMERICAL EXPERIMENTS

In this section, we report some numerical results to verify the performance of the proposed algorithm EQR. We have found the solutions in more than 300 difficult test problems using EQR method (see example: http://www.gamsworld.org/global/globallib.htm (GL)). The best solutions are obtained for more than 50 difficult test problems. Some results are shown in table 1.

Table I: Numerical experiments

Problem	***n***	***m***	**EQR method glob. min.**	**The best known glob. min.**	**Ref.**
Egg holder	100	0	**-89948.532**	– 89938	[4]
Rana	100	0	**-50855.784**	– 41047.18	[5]
Ex2_1_8	24	10	**15639**	15990	GL
Ex7_3_5	13	15	**7.684E-06**	1.2069	GL
Ex8_4_3	52	25	**-12,018963**	-3,25611	GL
Ex8_5_6	6	7	**-2.264901975**	-0.998832628	GL
Ex8_4_7	63	41	**26.99430909**	29.0473	GL
Harker	20	7	**-1020.242976**	-986.513	GL
Haverly	12	9	**-3.27297594**	900	GL
g16	5	38	**-1.9146086**	-1.9046617	[6]
Nie	50	0	**-93.999987**	-86.118	[7]
Nie	49	1	**-0.98284629**	-0.5322069	[7]
Charles	16	21	**156.2196293**	174.788	[8]

These known test problems were solved by different methods during many years. None of the existing methods shows best results.

We give optimal solutions for of test problems, where the value $n \le 20$. The optimal value of x^* for the problem Ex7_3_5 is equal

x^* = (0.175001, 0.199999, 10, 7.683588E-06, 0.1387446, 0.2254086, 0.1440133, 0.062608, 0.0151109, 0.0024137, 0.000184, 0.0000181, 4,71142769E-07).

We found optimal solution for of the problem Ex8_5_6, where the optimal value of x^* is

x^* = (0.98, 0.01, 0.01, 0.9611, 0.148065, 0.082172).

For the Harker's problem the optimal value of x^* is

x^* = (1.882576, 0.968872, –1, –1, –1, –1, 1.882576, –1, –1, 0.968872, 3.851448, –1, –1, –1, 10.95713, 35.65575, 31.66667, 15.80857, 30.8043, 31.66667).

The Haverly's problem we have

x^* = (96.72703, 100, 0.612072, 4.084774, 2.769821, 2, 5.466667, 1, 1.769821, 1, 3.696846, 1.260631).

We found optimal solution for the of problem g16, where the optimal value of x^* is

x^* = (707.335699, 68.6, 102.9, 282.0252, 84.1988).

For the Charles's problem optimal solution is equal

x^* = (8.03773, 9, 9, 9, 10, 0.09163, 1, 1.156863, 1.156863, 1.156863, 18.13995, 0.0000001, 1, 50, 50, 0.0000001).

CONCLUSION

We have solved many difficult optimizing problems in optimal designing, clustering, sensor networks and chemistry. The comparative numerical experiments have shown that method EQR are very efficient and promising.

REFERENCES

1. Encyclopedia of Optimization, editors by C.A. Floudas, P.M. Pardalos (Springer, 2009).
2. V.P Kenneth, R.M. Storn and J.A. Lampinen. *Differential Evolution. A Practical Approach to Global Optimization* (Berlin, Heidelberg: Springer-Verlag, 2005).
3. J. Nocedal and S.J. Wright. *Numerical optimization* (Springer, 2006).
4. A.P. Piotrowski and J. J. Napiórkowski. The grouping differential evolution algorithm for multi-dimensional optimization problems. Control and Cybernetics, vol. **39,** No. 2, 527–550 (2010).
5. S. Wu and T.W.S. Choy. Self-Organizing and Self-Evolving Neurons: A New Neural Network for Optimization. IEEE transactions on neural networks, vol. **18**, no. 2, 385–396 (2007).
6. M. K. Dhadwal, S. N. Jung, C. J. Kim. Advanced particle swarm assisted genetic algorithm for constrained optimization problems. Comput Optim Appl., **58**, 781–806 (2014).
7. J. Nie. Regularization Methods for Sum of Squares Relaxations in Large Scale Polynomial Optimization. (University of California, 2009).
8. C. Audet, P. Hansen, B. Jaumard, G. Savard. A branch and cut algorithm for nonconvex quadratically constrained quadratic programming. Math. Program., Ser. A **87**, 131–152 (2000).