Investigated Non-Conic Model for Constrained Optimization

Eman Tarik Al-Haj Saeed   **Huda Issam Ahmed**
College of Computer Sciences and Mathematics
University of Mosul, Iraq

ABSETRACT

In this search, we develop the nonlinear constrained optimization by investigation a new region of solution depending on extended conic model to non-conic model by using conjugate gradient method. The new method is too effective when compared with other established algorithms to solve standard constrained optimization problems it performance from evaluations were the number of function (NOF),number of iteration (NOI) and the number of constrained (NOC).

Keywords: nonlinear constrained optimization, non-conic model, conjugate gradient method.

استحداث نموذج غير مخروطي للأمثلية المقيدة

هدى عصام احمد   **ايمان طارق الحاج سعيد**
كلية علوم الحاسوب والرياضيات، جامعة الموصل

تاريخ الاستلام: 2007/11/25   **تاريخ القبول: 2008/04/20**

الملخص

تم في هذا البحث تطوير تقنيات الامثلية المقيدة غير الخطية باستحداث منطقة جديدة للحل تعتمد على توسيع النموذج المخروطي إلى النموذج غير المخروطي باستخدام طريقة التدرج المترافق وان الطريقة ذات فعالية عالية عند مقارنتها مع مثيلاتها من الطرائق السابقة و تظهر من خلال حساب عدد مرات احتساب قيمة دالة الهدف وحساب عدد مرات استدعاء القيود.

الكلمات المفتاحية: الامثلية المقيدة غير الخطية، نموذج غير مخروطي، طريقة التدرج المترافق.

1. Constrained Optimization Problem

Constrained optimization problem is defined by:

$$\min \ f(x) \qquad \qquad \dots(1)$$

subject to
$$\begin{cases} z_i(x) \le 0, & \text{for } i = 1,\dots p \\ h_i(x) = 0, & \text{for } i = 1,\dots m \end{cases} \qquad \dots(2)$$

the variable x is called the design vector, $f(x)$ is the objective function, $z_i(x)$ are inequality constraints and $h_i(x)$ are the equality constraints whose feasible region we denote by:

$$f = \left\{ x \in R^n \ / \ z_i \le 0, i = 1, \dots p, h_i = 0, i = 1, \dots m \right\}, \text{ see [4]}$$

To solve this problem we use a method based on the combination of interior and exterior point methods. We describe these methods in the

following sections. This section explains both interior and exterior point methods in the context of their development. It illustrates how the two methods are related and why their integration is a reasonable approach for solving nonlinear programming problems. In the past two decades interior point methods have proven to be efficient and are widely used for solving linear and nonlinear programming problems. The interior point methods are closely related to the sequential unconstrained minimization technique (sumt) developed by Fiacco and McCormick for solving constrained optimization problem with inequality constraints, see[8].

Barrier and penalty methods are designed to solve P by instead of solving a sequence of specially constructed unconstrained optimization problems.

In a penalty method, the feasible region of P is expanded from F to all of n, but a large cost or "penalty" is added to the objective function for points that lie outside of the original feasible region F, see[3].

In a barrier method, we presume that we are given a point x_0 that lies in the interior of the feasible region F, and we impose a very large cost on feasible points that lie ever closer to the boundary of F, there by creating a "barrier" to exiting the feasible region.

2. Penalty methods

Penalty methods transform constrained problems into unconstrained. Appropriate methods can then be used. We distinguish so called *interior* and *exterior penalty methods* see [6]. With exterior penalty methods some function is added to the objective to penalize infeasible solutions. On the other hand such a function prevents interior penalty methods from leaving the feasible domain. That is the reason why the interior penalty method is also called *barrier function method*. In both cases the penalty factor adjusts the effect. As the factor approaches infinity the solution is increasingly better approximated while the problem condition worsens. As a consequence numerical problems occur if the factor is too large. The advantage of the interior method is that intermediate solutions are always feasible, the disadvantage is that equality constraints cannot be handled. The latter is no problem for exterior methods, these, however, converge from the exterior, i.e. the infeasible, domain towards the solution. Intermediate solutions can only be used if they are scaled back to the boundary of the feasible domain. The main advantage of both methods is that they can be handled very easily and make only little use of complicated theory, see [11]. They are very successful in practical application for all kind of problems. A well-known procedure is *SUMT, sequential unconstrained minimization technique*, which generates a series of solutions with increasing penalty factor which allows for controlled approach to the optimum until the

solution is accepted or the problem condition collapses. In this context, the penalty factor can be understood as a continuation parameter, a feature which this rather old method has in common with the inner point methods. These approach from the interior domain to the solution as interior penalty function methods do but are on a strong mathematical basis in contrast to the more intuitive motivation of SUMT , see [9].

There are two approaches for such a problem:

- Indirect Methods:

They transform the constrained problem to unconstrained problem and uses the unconstrained numerical techniques to solve the problem. Very popular in the industry. These methods are called Sequential Unconstrained Minimization Techniques (SUMT). The following represents some of the families of these methods:

1- Exterior penalty function methods.
2- Interior penalty function methods.
3- Mixed penalty function methods.

See[4] and [10].

-Direct Methods:

These methods handle the complete problem as formulated. They are based on linearizing the functions involved. The objective function however in some cases is expanded as a quadratic. They are currently popular see [5]. There are several methods and many of them have minor differences. Among these are:

1- Sequential Linear Programming.
2- Sequential Quadratic Programming.
3- Generalized Reduced Gradient Method. (INT[1]).

3. Conjugate Gradient Method for Conic Model

The conic model method for constrained optimization has been studied by Davidon and Sun & Yuan see[2] , [12].

Conic model methods are generalization of quadratic methods, they have more freedom in the model they can choose the model to approximate the original problem better such as using more interpolation conditions.

Atypical conic model for constrained optimization is:

$$f(x) = f(x_0 + s) = f_0 + \frac{g_0^T s}{1 - a_0^T s} + \frac{1}{2} \frac{s^T G s}{(1 - a_0^T s)^2} = c(x) \qquad \ldots(3)$$

where $x_0 \in R^n$ as $x = x_0 + s$, $g_0, a_0 \in R^n$ and G is an $n \times n$ positive definite and symmetric matrix. the vector a_k is the associated vector for the collinear scaling in the k-th iteration and it is normally called the horizontal vector .

If $a_o = 0$, the conic model reduces to a quadratic model, since the term $\frac{s}{1 - a_0^T s}$ appears once in the second term on the right- hand side of (3), and twice in the third term, it is clear that by letting:

$$w = \frac{s}{1 - a_0^T s} \quad, \qquad \qquad \ldots(4)$$

the conic function becomes a quadratic in the variable w,

$$f(x) = f(x_0 + s) = f_0 + g_0^T w + \frac{1}{2} w^T G w = c(w) \qquad \ldots(5)$$

We can express s in term of w:

$$s = \frac{w}{1 + a_0^T w} \qquad \qquad \ldots(6)$$

To simplify the formulas, we define

$$\gamma(x) = 1 - a_0^T s = \frac{1}{1 + a_0^T w} \quad, \qquad \qquad \ldots(7)$$

where $x = x_0 + s$, so that $w = s/\gamma$. We call $E = \{x : 1 - a_0^T s = 0\}$, the singular hyper plane, and note that if $\gamma(x)\gamma(y) < 0$, then x and y lie on opposite sides of E, see [7].

We will need to relate the derivative of f to that of c, since

$$f(x) = f(x_0 + s) = f(x_0 + \frac{w}{1 + a_0^T w}) = c(w) \qquad \ldots(8)$$

It follows from the chain rule that

$$c'(w) = \gamma(x)(I - a_0 s^T) g(x), \qquad \qquad \ldots(9)$$

where g denotes the gradient of f. The conic function (3) is useful as a model function for minimization only if it has a unique minimize. This is ensured by the conditions

$$G > 0 \text{ and } a_0^T G^{-1} g_0 \neq 1. \qquad \qquad \ldots(10)$$

If eq.(10) holds, f will be called a normal conic function see [2].

The horizon vector a can be computed using function and gradient values at any three collinear points. In particular, consider the iterates x_k and x_{k+1} and let x_t be given by $x_t = x_k + \lambda_t d_k$, where $\lambda_t \neq \lambda_k$. Then

$$a_0 = \frac{(\lambda_t - \lambda_i)\gamma_i^2 \, g_i + \lambda_i \gamma_t^2 g_t - \lambda_t \gamma_{i+1}^2 \, g_{i+1}}{(\lambda_t - \lambda_i)\gamma_i^2 \, s_i^T g_k + \lambda_i \gamma_i^2 \, s_t^T g_t - \lambda_t \gamma_{i+1}^2 s_{i+1}^T g_{i+1}}, \qquad \ldots(11)$$

where as before $s_i = x_i - x_0, i = k, t, k+1$ and where γ_i is defined by:

$$\gamma_{i+1} = \frac{-g_i^T (x_{i+1} - x_i)}{f_i - f_{i+1} + \varsigma_{i,i+1}} \gamma_i, \qquad \qquad \ldots(12)$$

with $\gamma_0 = 1$, and

$$\varsigma_{i,i+1} = [(f_i - f_{i+1})^2 - g_{i+1}^T(x_{i+1} - x_i)g_i^T(x_{i+1} - x_i)]^{1/2} \qquad \ldots(13)$$

For more detail see [7].

3-1 Outline of Conic Model

Step1: Set x_0, $\in$, (initial point, scalar termination) and let $\beta_0 = 0$, $d_{-1} = 0$, for k=0,1,…,n.

Step2: Set $d_0 = -c_0'$.

Step3: Set $x_{i+1} = x_i + \lambda_i d_i$, $i \geq 0$, where λ_i is obtained from the line search procedure.

Step4: Compute $a_0 = \dfrac{(\lambda_t - \lambda_i)\gamma_i^2 g_i + \lambda_i \gamma_t^2 g_t - \lambda_t \gamma_{i+1}^2 g_{i+1}}{(\lambda_t - \lambda_i)\gamma_i^2 s_i^T g_i + \lambda_i \gamma_i^2 s_t^T g_t - \lambda_t \gamma_{i+1}^2 s_{i+1}^T g_{i+1}}$.

Step5: The new search direction compute $d_i = -c_i' + \beta_i d_{i-1}$, with $\beta_i = \dfrac{\|c_{i+1}'\|^2}{\|c_i'\|^2}$, Whereis the conjugacy coefficient of conic model. β_i

Step6: Check for convergence if $\|c_{i+1}'\| < \in$, where $\in = 10^{-5}$ then stop. Otherwise go to step 7.

Step7: Check for restarting criterion i.e. if $d_{i+1}^T c_{i+1}' \geq -0.8|c_{i+1}'^T c_{i+1}'|$, is satisfied go to step 2, otherwise set i=i+1 and go to step 3.

4. New Conjugate Gradient Method for Extended Conic Models (Non-Conic Method)

The original conjugate direction methods were developed in such a way that they find a minimum of a conic function after a finite number of steps making use of perfect line searches. This paper describes the conjugate direction methods which minimize extended conic functions after a finite number of steps.

A new broad class of globalization approach for solving constrained optimization is the class. The main idea of non-conic model is, at a current iterate x_k to calculate a new point compute $x_{i+1} = x_i + \lambda_i d_i$ where λ_i is minimizer of $f(x_i + \lambda d_i)$. Conic models may not always be adequate to incorporate all the information which might be needed to represent the objective function $f(x)$ successfully. In order to obtain better global rate of convergence for minimization algorithms when applied to more general functions than the Conic when applied in this paper several new algorithms, which are invariant to nonlinear scaling of Conic functions. There is some

precedent for this approach, if $c(x)$ is conic function then a function f is defined as nonlinear scaling of $c(x)$ if the following condition holds

$$f = F(c(x)) \quad \frac{df}{dc} = F' \succ 0 \quad \text{And} \quad c(x) \succ 0 \qquad \ldots(14)$$

Where x^* is minimizer of $c(x)$ with respect to x.

The following properties are immediately derived from above condition.

        i- Every contour line of $c(x)$ is a contour line of the f ;

        ii- if x^* is minimizer of $c(x)$ then it is a minimizer of f ;

        iii- that x^* is global minimum of $c(x)$ does not necessarily mean that it is a global minimum of f .

5. Derivation of the New Non-Conic Model

      The implementation of the extended method has been performed for general function $F(c(x))$ of the form (14). The unknown quantities ρ were expressed in term of available quantities of the algorithm (i.e. function and gradient value of the objective function) using the expression for ρ :

$$\rho = \frac{F'_{i-1}}{F'_i} \qquad \ldots(15)$$

From the relations

$$\bar{g}_{i-1}(w) = F'_{i-1} \, c'_{i-1}(w)$$
$$\bar{g}_i(w) = F'_i \, c'_i(w) \qquad \ldots(16)$$

since $c'(w)$ is defined by eq(9) so we have:

$$\bar{g}_{i-1}(w) = F'_{i-1} \gamma_{i-1}(x)(I - a_0 s_{i-1}^T) g_{i-1}(x) \qquad \ldots(17a)$$
$$\bar{g}_i(w) = F'_i \, \gamma_i(x)(I - a_0 s_i^T) g_i(x) \qquad \ldots(17b)$$

then we have:

$$\rho = \frac{F'_{i-1}}{F'_i} = \frac{\bar{g}_{i-1}^T \gamma_i(x)(I - a_0 s_i^T) g_i(x)}{\bar{g}_i^T \gamma_{i-1}(x)(I - a_0 s_{i-1}^T) g_{i-1}(x)} \qquad \ldots(18)$$

since $s_i^T = s_{i-1}^T + \lambda d_{i-1}^T \qquad \ldots(19a)$

then $s_{i-1}^T = s_i^T - \lambda d_{i-1}^T \qquad \ldots(19b)$

substitute (19) in (18) then we get

$$\rho = \frac{\bar{g}_{i-1}^T(w) \, \gamma_i(x)(I - a_0(s_{i-1}^T - \lambda d_{i-1}^T)) \, g_i(x)}{\bar{g}_i^T(w) \, \gamma_{i-1}(x)(I - a_0(s_i^T - \lambda d_{i-1}^T)) g_{i-1}(x)} \qquad \ldots(20)$$

since for any exact line search $d_{i-1}^T g_i = 0$ therefore

$$\rho = \frac{\bar{g}_{i-1}^T(w) \, \gamma_i(x)(I - a_0 s_{i-1}^T) \, g_i(x)}{\bar{g}_i^T(w) \, \gamma_{i-1}(x)(I - a_0 s_i^T + \lambda a_0 d_{i-1}^T)) g_{i-1}(x)} \qquad \ldots(21)$$

From (17) we get

$$\rho = \frac{F'_{i-1}\,\gamma_{i-1}(x)(I - a_0 s_{i-1}^T)\,g_{i-1}(x)\,\gamma_i(x)\,(I - a_0 s_i^T)\,g_i(x)}{F'_i\,\gamma_i(x)\,(I - a_0 s_i^T)\,g_i(x)\,\gamma_{i-1}(x)\,[(I - a_0 s_i^T) + \lambda\,a_0 d_{i-1}^T]\,g_{i-1}(x)} \quad \ldots(22)$$

Therefore

$$\rho = \frac{F'_{i-1} g_{i-1}^T(x)\,(I - a_0 s_{i-1}^T)^2\,g_i(x)}{F'_i\,g_i^T(x)\,(I - a_0 s_i^T)\,[(I - a_0 s_i^T) + \lambda\,a_0 d_{i-1}^T]\,g_{i-1}(x)} \quad \ldots(23)$$

5.1 A Sine Function as a conic Functional Model

In this paper a new triangular function model is investigated and tested on a set of standard non-linear constrained test function, it is assumed that condition (3) holds. The extended conjugate gradient algorithm is developed which is based on this new model is as follows:

$$F(c(x)) = \sin(c(w)) \quad \ldots(24)$$

If we express F'_{i-1} and F'_i as follows, using the derivation of the sine function,

$$F'_i = \cos c_i(w) \quad \ldots (25a)$$

$$F'_{i-1} = \cos c_{i-1}(w) \quad \ldots (25b)$$

Solving (24) for $c(w)$ we have,

$$c_i = \sin^{-1}(F_i)$$
$$c_{i-1} = \sin^{-1}(F_{i-1}) \quad \ldots (26)$$

By substituting (26) in (23), we have

$$\rho = \frac{\cos(\sin^{-1}(F_{i-1}))g_{i-1}^T(x)\,(I - a_0 s_{i-1}^T)^2\,g_i(x)}{\cos(\sin^{-1}(F_i))\,g_i^T(x)\,(I - a_0 s_i^T)\,[(I - a_0 s_i^T) + \lambda\,a_0 d_{i-1}^T]\,g_{i-1}(x)} \quad \ldots (27)$$

5.2 Outline of Non-Conic Model for constrained problems

Step1: Set $x_0, \in, \mu_0$ (initial point, scalar termination and initial parameter).

Step2: $d_0 = -c'_0$.

Step3: Set $x_{i+1} = x_i + \lambda_i d_i$, $i \geq 0$, where λ_i is obtained from the line search procedure.

Step4: Compute $a_0 = \dfrac{(\lambda_t - \lambda_i)\gamma_i^2 g_i + \lambda_i \gamma_t^2 g_t - \lambda_t \gamma_{i+1}^2 g_{i+1}}{(\lambda_t - \lambda_i)\gamma_i^2 s_i^T g_i + \lambda_i \gamma_t^2\, s_t^T\, g_t - \lambda_t \gamma_{i+1}^2\, s_{i+1}^T\, g_{i+1}}$.

Step5:Compute $\rho = \dfrac{\cos(\sin^{-1}(F_{i-1}))g_{i-1}^T(x)\,(I - a_0 s_{i-1}^T)^2\,g_i(x)}{\cos(\sin^{-1}(F_i))\,g_i^T(x)\,(I - a_0 s_i^T)\,[(I - a_0 s_i^T) + \lambda\,a_0 d_{i-1}^T]\,g_{i-1}(x)}$.

Step6: Compute the new search direction defined by $d_{i+1} = -c'_{i+1} + \rho \beta_i d_i$ where β_i is the conjugacy coefficient.

Step7: Check for convergence if $\dfrac{1}{\mu_i}\sum\limits_{i=1}^{l} h(x_i) < \in$ then stop. Otherwise set

$\mu_{i+1} = \mu_i \times 10$ and take $x = x^*$ as a new starting point.

Step8: Check for restarting criterion if $d_{i+1}^T c'_{i+1} \ge -0.8 \left| c_{i+1}^{\prime T} c'_{i+1} \right|$, is Satisfied go to step 2, otherwise set $i = i + 1$ and go to step 3.

6. Numerical Results

In order to explain the new algorithm which is highly effective when compared with other standard optimization. Two minimization algorithms are tested over (9) non-linear equality constrained problems see (Appendix) with $1 \le n \le 5$ and $1 \le h_i(x) \le 7$.

FORTRAN 2001 programs were written to implement the suggested and the previous algorithms. All-test results quoted here were obtained using (Pentium 4 computer). For all cases the stopping criterion taken to $\dfrac{1}{\mu_i}\sum\limits_{i=1}^{l} h(x_i) < 1 \times 10^{-5}$.

All the algorithms in this paper use the same ELS strategy which is the cubic fitting technique fully described from Bunday see [1]

The comparative performance for all of these algorithms are evaluated by considering NOF, NOI and NOC, where NOF is the number of function evaluations, NOI is the number of iteration and NOC is the number of constrained evaluations.

1- Conic model

2- Non Conic model

Our numerical results are presented in table (1) confirm that the Non Conic model is superior to Quadratic model , Non-quadratic model and Conic model with respect to the total of NOF , NOI and NOC.

Table (1)

Test	Conic algorithm			New Non-Conic algorithm		
Fn.	NOI	NOF	NOC	NOI	NOF	NOC
1-	131	39	43	107	35	38
2-	265	65	68	204	54	59
3-	131	40	42	89	38	40
4-	29	14	15	29	14	15
5-	849	177	181	717	93	98
6-	50	18	20	42	18	19

7-	33	15	18	36	17	19
8-	58	22	25	43	18	21
9-	74	21	22	60	20	21
Total	1620	411	434	1327	307	330

7. Appendix

1. **min** $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$,

   s.t

   $x_1^2 - x_2 = 0$.

2-**min** $f(x) = (x_1 - 1)^2 + (x_2 - 2)^2 + (x_3 + 2)^2$,

   s.t

   $x_1^2 + 4x_2^2 + x_3^2 + 36 = 0$.

3-**min** $f(x) = (x_1 - 2)^2 + (x_1 - 2x_2)^2$,

   s.t

   $x_1^2 - x_2^2 = -4$.

4-**min** $f(x) = x_1^2 + x_2^2$,

   s.t

   $1 - x_1^2 - x_2^2 = 0$.

5-**min** $f(x) = (x_1 - 1)^2 + (x_1 - x_2)^2 + (x_2 - x_3)^2 + (x_3 - x_4)^4 + (x_4 - x_5)^4$,

   s.t

   $x_1 + x_2^2 + x_3^3 + 2 - 3\sqrt{2} = 0$,

   $x_2 + x_3^2 + x_4 + 2 - 2\sqrt{2} = 0$

   $x_1 x_5 + 2 = 0$

6-**min** $f(x) = x_1^2 + x_2^2$,

   s.t

   $(x_1 - 1)^2 + x_2^2 = 0$.

7-**min** $f(x) = 0.5x_1^2 + 2.5x_2^2$,

   s.t

   $x_1 - x_2 - 1 = 0$.

8-**min** $f(x) = x_1^2 + x_2^2$,

   s.t

   $x_1^2 - x_2^2 = -1$.

9-**min** $f(x) = x_1^2 + x_2^2$,

   s.t

   $(x_1 - 1)^3 - x_2^2 + 4 = 0$

REFERENCES

[1] Bundy, B. (1984), "Basic Optimization Method" Edward Arnold, Bedford Square, London, U.K.

[2] Davidon, W.C. (1980),"conic approximations and collinear scaling for optimizers", SIAM J. Numer. Anal. Vol. 17,No. 2,pp.268-281.

[3] Dong, S., (2006), " Methods for Constrained Optimization", Springer Science+Business Media, Inc. New York: 2005.

[4] Igor Griva ,(2004)" Numerical experiments with an interior-exterior point method for nonlinear programming" Princeton University, Department of ORFE, Princeton NJ 08544, Email:igriva@princeton.edu.

[5] Kwon. T.H. (2000)" Computer Aided Design" department of mechanical engineering, university of SCIENCE & TECHNOLOGY.

[6] Nesterov, Y. and Nemirovski, A. (1994)," Interior-Point Polynomial Algorithms in Convex Programming", Philadelphia, PA: SIAM Studies in Applied Mathematics.

[7] Nocedal J. and Liu C.D.(1989),"Algorithms with conic termination for nonlinear optimization"SIAM J.SC. Computer, Vol. 10 ,No. 1, PP. 1-17.

[8] Omojokum, E. O. (1989), " Trust Region Algorithms for Optimization with Non-Linear Equality and Inequality Constraints", ph. D. Thesis University of Colorado at Boulder.

[9] Patel R.V. (2000), " Class Notes for ECE468b / ES626b: System Optimization "Department of Electrical & Computer Engineering, University of Western Ontario, London, Ontario, Canada.

[10] Renegar, J. (2001)," A Mathematical View of Interior-Point Methods in Convex Optimization", Philadelphia, PA: SIAM (MPS-SIAM Series on Optimization).

[11] Robert M. Freund February, (2004)" Penalty and Barrier Methods for Constrained Optimization" 2004 Massachusetts Institute of Technology.

[12] Sun, W. and Yuan, J. (1997),"A Conic Model Trust Region Method for Nonlinearly Constrained Optimization, Submitted to Mathematics f Computation.

[13] **Internet Rep.**

   INT (1) **"Constrained Optimization (2004)"**

   http://www.rit.edu/~penveme/emem620/consrainedoptimization.html