Modification of Non-Linear Constrained Optimization

Ban Ahmed Mitras
dr.banah.mitras@gmail.com
College of Computer Sciences and Mathematics
University of Mosul, Iraq

Received: 13/05/2002 **Accepted: 12/10/2002**

ABSTRACT

In this paper we have investigated a new initial parameter, the new parameter is to make balance between interior suitable for inequality constrained exterior method (suitable for equality and inequality constrained) for non-linear constrained optimization. The new algorithm is programmed to solve some standard problems in non-linear optimization method. The results are too effective when compared with Barriar –Penalty algorithm.

Keyword: constrained optimization, penalty method, Barrier method.

تحسين للأمثلية المقيدة غير الخطية

بان احمد متراس
كلية علوم الحاسوب والرياضيات
جامعة الموصل

تاريخ استلام البحث: 2002/05/13 **تاريخ قبول البحث: 2002/10/12**

الملخص

في هذا البحث تم استحداث معلم ابتدائي جديد يربط بين الطريقتين الخارجية (المناسبة لقيود غير المساواة) والطريقة الداخلية (المناسبة لقيود المساواة وغير المساواة) في مجال ألا مثلية غير الخطية المشروطة. و إن التقنية الجديدة تمت برمجتها لحل بعض مسائل ألا مثلية القياسية المعروفة وتوصلنا إلى أن نتائجها اكثر كفاءة من طريقة Barriar -Penalty.

الكلمات المفتاحية: الأمثلية المقيدة، طريقة الجزاء، طريقة الحاجز.

1.Introduction:

We shall first state the most general from of the problem we are addressing, namely

Minimize $f(x), \qquad x \in R^n$ (1)

Subject to the general (possibly nonlinear) inequality constraints

$$c_j(x) \le 0 \qquad 1 \le j \le L \tag{2}$$

and (possibly nonlinear) equality constraints

$$c_j(x) = 0 \qquad L+1 \le j \le m \tag{3}$$

with the simple bounds

$$L_i \le x_i \le u_i, \qquad 1 \le i \le n \tag{4}$$

Where f and c_j are all assumed to be twice continuously differentiable and defined on E_n, x is a subset of E_n, and x is a vector of n components, $x_1, x_2, \ldots, x_n$.

The above problem must be solved for the values of the variables $x_1, x_2, \ldots, x_n$ that satisfy the restrictions and meanwhile minimize the function f. The function f is called the objective function and any of the bounded in eq.(4) may be infinite. (See Conn et al, 1994). The exterior-point method is suitable for equality and inequality constraints. The new objective function $\phi(x, \mu_k)$ is defined by:

$$\phi(x, \mu_k) = f(x) + \frac{1}{\mu_k}\alpha(x) \tag{5}$$

Where μ_k is a positive scalar and the remainder of the second term is the penalty function.

Interior-point method is suitable for inequality constraints. The new objective function $\phi(x, \mu_k)$ is defined by

$$\phi(x, \mu_k) = f(x) + \mu_k B(x) \tag{6}$$

Where μ_k is a positive scalar and the reminder of the second term is the Barrier function. (see Gottefred, 1973).

Although both exterior and interior-point methods have many points of similarly, they represent two different points of view. In an exterior-point procedure, we start from an infeasible point and gradually approach feasibility, while doing so, we move away from the unconstrained optimum of the objective function. In an interior-point procedure we start at a feasible point and gradually improve our objective function, while maintaining feasibility. The requirement that we begin at a feasible point and remain within the interior of the feasible inequality constrained region is the chief difficulty with interior-point methods. In many problems we have no easy way to determine a feasible starting point, and a separate initial computation may be needed. Also, if equality constraints are present, we do not have a feasible inequality constrained region in which to maneuver freely. Thus interior-point methods cannot handle equalities.

We may readily handle equalities by using a "mixed" method in which we use interior-point penalty functions for inequality constraints only. Thus, if the first *m* constraints are inequalities and constraints *(m+1)* to *n* are equalities, our problem becomes:

$$\textit{Minimize} \quad \phi(x,\mu_k) = f(x) + g(\mu_k)B(x) + \frac{1}{g(\mu_k)}\alpha(x) \qquad (7)$$

The function $\phi(x,\mu_k)$ is then minimized for a sequence of monotonically decreasing $\mu_k > 0$.

2. Mixed Exterior-Interior Methods:

We can solve the constrained problem given in eq.(1) to eq.(3) construct a new objective function $\phi(x,\mu_k)$ which is defined in eq.(7). Now our aim is to minimize the function $\phi(x,\mu_k)$ by starting form a feasible point x_0 and with initial value $\mu_0 = 1$ and the method reducing μ_k is simple iterative method such that:

$$\mu_{k+1} = \frac{\mu_k}{10}, \tag{8}$$

where μ_k is a constant equal to *10* and the search direction d_k in this case can be defined

$$d_k = -H_k g_k, \tag{9}$$

where H is a positive definite symmetric approximation matrix to the inverse Hessian matrix G^{-1} and g is the gradient vector of the function $\phi(x, \mu_k)$.

The next iteration is set to a further point

$$x_{k+1} = x_k + \lambda_k d_k, \tag{10}$$

where λ is a scalar chosen in such that $f_{k+1} < f_k$, we thus test $c_i(x_{k+1})$ to see that it is positive for all i. We find a feasible x_{k+1} and we can then proceed with the interpolation. Then the matrix H_k is updated by a correction matrix to get

$$H_{k+1} = H_k + \phi_k \tag{11}$$

where ϕ_k is a correction matrix which satisfies quasi-Newton condition namely $(H_{k+1} y_k = \rho v_k)$ where v_k and y_k are difference vector between two successive points and gradients respectively and ρ is any positive scalar.

The initial matrix H_0 chosen to be identity matrix I. H_k is updated through the formula of BFGS update. (see Bazarra et al, 2000).

Given some approximation H_k to the inverse Hessian matrix, we compute the search direction $d_k = -H_k g_k$, and we define $v_k = x_{k+1} - x_k$ and

$y_k = g_{k+1} - g_k = G(x_{k+1} - x_k) = G v_k$.

We now want to construct a matrix

$$H_{k+1} = H_k^{(1)} + H_k^{(2)} \tag{12}$$

where $H_k^{(2)}$ is some symmetric correction matrix that ensures that $v_1, v_2, \ldots, v_k$ are eigenvectors of $H_{k+1}G$ with unit eigenvalues. Hence

$$H_{k+1}y_k = v_k$$

This condition translates to the requirement that

$$H_{k+1}y_k = v_k - H_k y_k$$

This therefore, leads to the rank-two DFP (Fletcher and Powell, 1963) update via the correction term

$$H_k = \frac{v_k v_k^T}{v_k^T y_k} - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} \equiv H_k^{DFP} \qquad (13)$$

The Broyden updates suggest the use of the correction matrix $H_k = H_k^B$ given by

$$H_k^B = H_k^{DFP} + \frac{\theta \tau_k p_k p_k^T}{v_k^T y_k} \qquad (14)$$

where $p_k = v_k - (\frac{1}{\tau_k})H_k y_k$ and where τ_k is chosen so that the quasi-Newton condition holds by virtue of $p_k^T y_k$ being zero. Then

$$H_k^{BFGS} = H_k^B(\theta = 1) = \frac{v_k v_k^T}{v_k^T y_k}\left(1 + \frac{y_k^T H_k y_k}{v_k^T y_k}\right) - \left(\frac{H_k y_k v_k^T + v_k y_k^T H_k}{v_k^T y_k}\right) \qquad (15)$$

and terminate the method if

$$\left|x_k - x_{k-1}\right| < \varepsilon \qquad (16)$$

where $\varepsilon = 0.000001$.

3. Combined Barrier-Penalty Algorithm:

Step (1): Find an initial approximation x_0 in the interior of the feasible region for the inequality constraints i.e. $g_i(x_0)<0$.

Step (2): Set $k=1$ and $\mu_0 = 1$ is the initial value of μ_k.

Step (3): Set $\phi(x, \mu_k) = f(x) + \mu_k B(x) + \frac{1}{\mu_k}\alpha(x)$.

Step (4): Set $d_k=-H_kg_k$

Step (5): Set $x_{k+1} = x_k + \lambda_k d_k$, where λ is a scalar.

Step (6): Check for convergence i.e. if eq.(16) is satisfied then stop.

Step (7): Otherwise, set $\mu_k = \frac{\mu_k}{10}$ and take $x=x^*$ and set $k=k+1$ and go to Step 3.

4. The Initial Value of the Parameter:

The initial value μ_k can be important in reducing the number of iterations and the number of function calls to minimize $\phi(x,\mu_k)$, since the unconstrained minimization of $\phi(x,\mu_k)$ is to be carried out for a decreasing sequence of μ_k, it might appear that by choosing a very small value of μ_k, we can avoid an excessive number of minimization of the function $\phi(x,\mu_k)$. Also as $\phi(x,\mu_k)$ is close to $f(x)$, the method should converge more quickly. However, such a choice can cause serious computational problems. Also if μ_k is small, the function $\phi(x,\mu_k)$ will be changed rapidly in the vicinity of its minimum. This rapid change in the function can cause difficulties for a gradient based on methods. (see Bazaraa, 2000).

Al-Bayati and Hamed in (1997) suggested a new parameter of the Barrier function.

Al-Assady and Hamed in (2002) proposed a new initial parameter of Barrier-Penalty method.

5. Development of the New Algorithm:

Consider the problem stated in eq.(1) to eq.(4). The new objective function $\phi(x,\mu_k)$ defined in eq.(7) with a starting feasible point x_0 and with an initial value μ_0 which is derived as

$$\phi(x,\mu_k) = f(x) + \mu_k B(x) + \frac{1}{\mu_k}\alpha(x) \tag{17}$$

$$= f(x) + \mu_k \frac{1}{c(x)} + \frac{1}{\mu_k}[h(x)]^2 . \tag{18}$$

Then the gradient of $\phi(x, \mu_k)$ is

$$\nabla\phi(x, \mu_k) = \nabla f(x) - \mu_k \frac{\nabla c(x)}{[c(x)]^2} + \frac{2}{\mu_k} h(x)\nabla h(x) \tag{19}$$

such that

$$\nabla\phi(x, \mu_k) = 0$$

we have

$$\nabla f(x) - \mu_k \frac{\nabla c(x)}{[c(x)]^2} + \frac{2}{\mu_k} h(x)\nabla h(x) = 0 \tag{20}$$

Now, since $\mu_k > 0$, then we have

$$\mu_k \nabla f(x) - \mu_k^2 \frac{\nabla c(x)}{[c(x)]^2} + 2h(x)\nabla h(x) = 0 \tag{21}$$

Arranging eq.(21) and multiplying it by (-1), we have

$$\mu_k^2 \frac{\nabla c(x)}{[c(x)]^2} - \mu_k \nabla f(x) - 2h(x)\nabla h(x) = 0 \tag{22}$$

The optimum value of μ_k is then given by one of the following roots to eq.(22):

$$\mu_{\min} = \frac{\nabla f(x) \mp \sqrt{(\nabla f(x))^2 + 8h(x)\nabla h(x)\dfrac{\nabla c(x)}{(c(x))^2}}}{2\dfrac{\nabla c(x)}{(c(x))^2}} \tag{23}$$

In the above suggestion corresponding to the assumption for deriving a new parameter to make balance between the exterior-interior point method, we have suggested the following new algorithm.

6. The Outline of the New Algorithm:

Step (1): Find an initial approximation x_0 in the interior of the feasible region for the inequality constraints i.e. $g_i(x_0)<0$.

Step (2): Set $k=1$ and $\mu_0 = 1$ is the initial value of μ_k.

Step (3): Find the initial value of μ_k by using eq.(23), and compute

$$\phi(x,\mu_k) = f(x) + \mu_k B(x) + \frac{1}{\mu_k}\alpha(x).$$

Step (4): Set $d_k=-H_k g_k$

Step (5): Set $x_{k+1} = x_k + \lambda_k d_k$, where λ is a scalar.

Step (6): Check for convergence i.e. if eq.(16) is satisfied then stop.

Otherwise go to step 7.

Step (7): Set $\mu_{k+1} = \frac{\mu_k}{10}$

Step (8): Set $x=x^*$ and set $k=k+1$ and go to step 4.

7. Results and Calculation:

In order to test the effectiveness of the new algorithm that has been used to Barrier-Penalty function method, the comparative tests involving several well-known test function (see Appendix) have been chosen and solved numerically by utilizing the new and established method. So the new algorithm has been compared with Barriar -Penalty algorithm.

In table (1) we have compared the new algorithm with standard Barrier-Penalty algorithm for $1 \le n \le 3$ and $1 \le c_i(x) \le 7$ using (5) nonlinear test functions.

From table (2) it is clear that, taking the standard Barrier-Penalty algorithm as 100%, and the new algorithm has 75%, 76.8%, and 81.9% improvements on the standard Barrier-Penalty algorithm in bout number of iterations NOI and number of function evaluations NOF.

Table (1)
Comparison between Barrier-Penalty and new algorithms

Test function	Barrier-Penalty algorithm NOI (NOF)	New algorithm NOI (NOF)
1.	7 (61)	2 (17)
2.	8 (2141)	9 (1991)
3.	7 (141)	5 (72)
4.	10 (956)	5 (216)
5.	10 (2205)	9 (1955)
6.	10 (803)	9 (596)
Total	52 (6307)	39 (4847)

Table (2)

	Barrier-Penalty algorithm	New algorithm
NOI	100%	75
NOF	100%	76.8

8. Appendix:

Test functions:

1. $\min f(x) = (x_1 - 2)^2 + (x_2 - 1)^2$ ***s.p(7,9)***

s.t.

$$x_1 - 2x_2 = -1$$

$$\frac{-x_1^2}{4} + x_2^2 + 1 \geq 0$$

2. $\min f(x) = x_1 x_2$ ***(18,16)***

s.t.

$$25 - x_1^2 - x_2^2 = 0$$

$$x_1 + x_2 \geq 0$$

3. $\min f(x) = x_1^2 + x_2^2$ ***s.p.(0.9,2)***

s.t.

$x_1 + 2x_2 = 4$

$x_1^2 + x_2^2 \leq 5$

$x_i \geq 0$

4. $\min f(x) = (x_1 - 2)^2 + (x_2 - 3)^2$ ***(2,7)***

s.t.

$x_1 - 2x_2 = -1$

$-x_1^2 + x_2 \geq 0$

5. $\min f(x) = x_1 x_4 (x_1 + x_2 + x_3) + x_3$ ***s.p(4,3,3,3)***

s.t.

$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 40$

$x_1 x_2 x_3 \geq 25$

$5 \geq x_i \geq 1$

6. $\min \ f(x) = (x_1 - 3)^2 + (x_2 - 2)^2$ ***s.p(0.1,2)***

s.t.

$x_1^2 + x_2^2 \leq 3$

$x_1 + 2x_2 = 2$

$x_i \geq 0$

REFERECES

[1] Al-Assady, N.H. and Hamed, E.T. (2002), "Investigation A New Initial Parameter of Mixed Interior-Exterior Method for Nonlinear Optimization" Al-Rafideen J. No.5, Vol.14.

[2] Al-Bayati, A.Y. and Hamed, E.T. (1997), "Some Modifications in Non-Linear Constrained Optimization Algorithms", Ms.c. Thesis, Dep. of Math. College of Science, Univ. of Mosul.

[3] Bazaraa, S. and Shetty, C.M. (2000), "Nonlinear Programming: Theory and Algorithms", New York, Chichester, Brisbane, Toronto.

[4] Biggs, M.C. (1989), "Halfield Polytechnic Numerical Optimization", The Optima Package, Special Communication.

[5] Conn, A.R., Gould, N. and Toint, Ph.L., (1994), "Large scale Nonlinear Constrained Optimization", A current Survey, Report 94/1 IBM T.J. Watson Research Center, USA.

[6] Fiacco, A.V. and McCormick, G.P. (1968), "Extension of Sum for Non-Linear Programming Equality Constraints and Extrapolation", Management Science, Vol. 12, No. 11, pp. 816-828.

[7] Fletcher, R. and Powell, M.J.D., (1963), "A Rapidly Convergent Descent Method for Minimization", Computer Journal, 6, pp. 163-168.

[8] Gottefred, B.S. and Weisman, J., (1973), "Introduction to Optimization Theory", Prentice-Hall, Englewood Cliffs, N.J.