Iraqi Journal of Statistical Science (17) 2010
The Second Scientific Conference of Mathematics-Statistics& Information
pp[33 -46]

A new multiplier for Lagrange interpolation in constrained non linear optimization

ABBAS Y. AL-BAYIATI* **EMAN T.HAMED****
HAMSA Th. CHILMERANE***

Abstract

In this paper, we have investigated a new multiplier for the Lagrange interpolation by modifying the initial value of the multiplier in order to reduce the errors and avoid the use of arbitrary values for the initial λ. The new procedure improves the rate of convergence of the Lagrange interpolation. Numerical results indicate that the new approach yields very effective numerical results depending on the number of iterations ; number of gradient and the number of function calls.

مضروب جديد لمتعدد حدود لاكرانج في الامثلية المقيدة غير الخطية

الملخص

في هذا البحث تم استحداث خوارزمية جديدة بالاعتماد على نقطة بداية دالة مضروب لاكرانج تحدد فيه القيمة الابتدائية لمضروب لاكرانج مع تحديث جديد لهذه ألداله لتقليل الخطأ الناتج من فرض القيمة الابتدائية بصوره عشوائية . ادت الخوارزمية الجديدة إلى زيادة سرعة تقارب ألطريقه اذ أثبتت النتائج كفاءتها بالاعتماد على عدد التكرارات وعدد حسابات الدالة .

*Prof./ Department of Mathematics//College of Computers Sciences & Mathematics
**Lecture/Department of Mathematics//College of Computers Sciences & Mathematics
***Lecture/Department of Mathematics//College of Computers Sciences & Mathematics
Received:1/7 /2009 ____________________Accepted: 6 /12 / 2009

1- Introduction

Constrained optimization problems are interesting because they arise naturally in engineering, science, operations research, etc. In general, a constrained numerical optimization problem is defined as:

$$\text{minimize } f(x)$$

$$\text{subject to } h_k(x) = 0 \qquad k = 1,2,\ldots,l \qquad \ldots\ldots\ldots\ldots\ldots\ldots(1)$$

Without loss of generality we may transform any optimization problem to one of the minimization algorithms and we, therefore, develop our discussion in such terms. Constraints define the feasible region, meaning that if the vector x complies with all constraints $h_k(x) = o$ then it belongs to the feasible region. Traditional methods relying on calculus demand that the functions and constraints have very particular characteristics (continuity, differentiability, second order derivability, etc.)

Most optimization problems have constraints. The solution or set of solutions which are obtained as a final result of an evolutionary search must necessarily be feasible, that is, satisfy all constraints. A taxonomy of constraints can be considered and composed of (a) number, (b)metric, (c) criticality and (d) difficulty. A first aspect is number of constraints, ranging upwards from one. Sometimes problems with multiple objectives are reformulated with some of the objectives acting as constraints. Difficulty in satisfying constraints will increase (generally more than linearly) with the number of constraints. A second aspect of constraints is their metric, either continuous or discrete, so that a violation of the constraint can be assessed in distance from satisfaction using that metric. A third consideration is the criticality of the constraint, in terms of absolute satisfaction. A constraint is generally formulated as hard (absolute) when in fact, it is often somewhat soft. That is, small violations would be considered for the final solution if the solution is otherwise superior to other solutions. Evolutionary algorithms are especially capable of handling soft constraints since a population of solutions is returned at each point in the search. This allows the user to select a solution which violates a soft constraint (technically infeasible) over a solution which would be the best, technically feasible solutions found. A final aspect of constraints to be considered is the difficulty of satisfying the constraints. This difficulty can be characterized by the size of the feasible region (*F*) compared to the size of the sample space (*S*). The difficulty may not be known *a priori*, but can be gauged in two ways. The first way is how simple it is to change a solution which violates the constraint to a solution which does not violate the constraint. The second way is the probability of violating the constraint during search **.INT [1]**

2- The Lagrange method

The area of Lagrange multiplier methods for constrained minimization has undergone a radical transformation starting with the introduction of augmented Lagrangian functions and methods of multipliers in 1968 by Hestenes and Powell. The initial success of these methods in computational practice motivated further efforts aimed at understanding and improving their properties. At the same time their discovery provided impetus and a new perspective for reexaminations of Lagrange multiplier methods proposed and nearly abandoned several years earlier. These efforts, aided by fresh ideas based on exact penalty functions, have resulted in a variety of interesting methods utilizing Lagrange multiplier iterations and competing with each other for solution of different classes of problems .INT [2].

We consider the equality constrained minimization problem:

$$minimize\ f(x)$$

$$\text{subject to } h_k(x) = 0 \qquad k = 1,2,\ldots,l \qquad \ldots\ldots(2)$$

The classical method of solving this problem is due to Lagrange. The method removes the equality constraints by considering the function

$$L(x,\lambda) = f(x) + \sum_{k=1}^{l} \lambda_k (h(x_k)) \qquad \ldots\ldots(3)$$

where . $\lambda_k = [\lambda_1,\ldots\lambda_l]^T$ denotes the set of Lagrange multipliers for this problem **[4]**,where λ is an $(1 * m)$ vector of Lagrange multipliers, one for each constraint. In general, we can set the partial derivatives to zero to find the minimum:

$$\frac{\partial L}{\partial x_k} = 0 \qquad k = 1,2,\ldots\ldots.l \qquad \ldots\ldots(4)$$

and

$$\frac{\partial L}{\partial \lambda_k} = 0 \qquad k = 1,2,\ldots\ldots.l \qquad \ldots\ldots(5)$$

Theorem(1):

If f and c are convex functions , X is a convex set , and x^* is an optimal solution to any problem ,then there exists a Lagrange multiplier $\lambda \in R$ such that $L(x^*, \lambda) \le L(x, \lambda)$ for $x \in X$.

Proof: **INT[3]**

Theorem(2):

Suppose that there exists a point $x^* \in X$ and $\lambda^*, \in R^n$ such that x^* maximize $L(x, \lambda^*)$ over all $x \in X$ and $h(x^*) = b$ solve the problem.

Proof :**INT[3]**

2. 1- The augmented Lagrange for inequality constrained

The augmented Lagrange method combines the classical Lagrange method with the penalty function method. We use the augmented Lagrange method to tackle equality constraints for the problem:

Minimize *f(x)*

$$\text{Subject to} \quad \begin{matrix} h_k(x) = 0 & \text{k} = 1,2,\ldots\ldots\ldots\text{i} \\ h_k(x) \neq 0 & \text{k} = i+1,2,\ldots\ldots\text{l} \end{matrix} \qquad \ldots\ldots(6)$$

The augmented Lagrange is now defined by

$$ALM(x,\lambda,r) = f(x_k) + \sum_{k=1}^{i} \lambda_k \left(h(x_k)\right) + \sum_{i+1}^{l} r\left(h(x_k)\right)^2 \quad \ldots\ldots(7)$$

where $\lambda_k = 2rh(x_k)$ ……(8)

2.2- outlines of the standard augmented Lagrange algorithm [5]

1. choose a tolerance $\varepsilon = 10^{-5}$,starting point $x_0 = 0$, initial penalty parameter $r_0 = 1$, and initial Lagrange multipliers $\lambda_0 = 0$
2. Perform unconstrained optimization on the augmented Lagrangian function of eq.7
3. set $\lambda_k = 2rh(x_k)$
5. Check the convergence criteria. If $\left\| x_{k+1}^* - x_k^* \right\| \prec \varepsilon$, then stop. Otherwise, set $x_0 = x_k^*$ and return to Step 2.

Theorem(3):

Suppose the objective function *f(x)* of the constraint optimization problem has a local minimum at the feasible point *x** and N is an open set. if *f(x)* and the various constraint function are continuously differentiable near x*.then there exists unit vectors of Lagrange multiplier λ_1, λ_2,…………….., λ_k and r_1, r_2,……………… such that

$$\lambda_0 \nabla f(x^*) + \sum_{k=1}^{i} \lambda_k \nabla h_k(x^*) + \sum_{k=i+1}^{l} r_k h_k(x^*) = 0 \quad \text{…………………………………………(8)}$$

each of the multipliers of the $\lambda_0, \mu_1, \mu_2, \ldots\ldots\ldots, \mu_k$ is nonnegative and $\mu_k = 0 \quad if \quad g_j \prec 0$.

Proof :INT [4]

3- New proposed constrained optimization algorithm

In this section we are going to produce a new procedure to calculate the parameter λ_k .This technique yields an updating to λ_k at each iteration instate of taking the value as a constant parameter to obtain a new value of λ_{k+1}

let $L(x_k, \lambda_k) = f(x_k) + \lambda_k \, h(x_k) + (1 - \lambda_k) \, h^2(x_k)$ ………………………………(9)

optimal values of λ_k occur

$$f(x_k) + h^2(x_k) + \lambda_k \left(h(x_k) - h^2(x_k)\right) = 0$$

$$-\left(f(x_k) + h^2(\mathrm{x_k})\right) = \lambda_k \left(h(x_k) - h^2(x_k)\right)$$

$$\lambda_k = \frac{-f(x_k) - h^2(x_k)}{h(x_k) - h^2(x_k)} \quad \text{………………………………………………………(10)}$$

3.1- The outline of new proposed algorithm

1. choose a tolerance $\varepsilon = 10^{-5}$,starting point $x_0 = 0$
2. Perform unconstrained optimization ($f_{\min}$ search) on the augmented Lagrangian function of eq.9
3. set $\lambda_k = \dfrac{-f(x_k) - h^2(x_k)}{h(x_k) - h^2(x_k)}$
4. Check the convergence criteria. If $\|x_{k+1}^* - x_k^*\| \prec \varepsilon$, then stop. Otherwise, set $x_0 = x_k^*$ and return to Step 2.

3.2- Theoretical Review

3.2.1- Augmented Lagrange with equality linear constrained

We consider the problem of finding a local minimize of the function

$$f(x)$$

Where x is required to satisfy general equality constraints

$$h_i(x) = 0\,,\quad 1 \le i \le m$$

and the linear inequality constraints

$$Ax - b \ge 0$$

Here f and h_i map $R^n \to R$, A is a n-by-n matrix and $b \in R^p$, and make the following assumptions:

As_1: The region $D = \{x / Ax - b \ge 0\}$ is nonempty.

As_2: the functions *f(x)* and *c(x)* are twice differentiable for all $x \in D$.

As_3: The iterates $\{x_k\}$ considers lie within a close ,bounded domain .

As_4**:** The matrix $j(x^*)$ has a column rank no smaller than *m* at any limit point x^*, of the sequence $\{x_k\}$.

Lemma (1): [Conn et al.(1993), Lemma 4.1]

Suppose that the parameter μ_k converges to zero as k increases. Then the product $\mu_k \|\lambda_k\|$ converges to zero.

Lemma(2):

Let $\{x_k\}, k \in K$, be a sequence which converges to the point x_* and suppose that $\sigma_{k \leq} w_k$, where the w_k are positive scalar parameters. Then , there is a positive constant k_1 and an integer k_0 such that $\left\| Z_*^T \nabla_x \phi_k \right\| \leq k_1 w_k$

for all $k \geq k_0$, $(k \in K)$

Theorem(4):

Assume that $As_1 - As_2$ hold, Let x^* be any limit point of the sequence $\{x_k\}$ generated by algorithm. For which As_3 and As_4 hold and let K be the set of indices of an infinite subsequence of the x_k whose limit x^*. Finally $\lambda_* = \lambda(x_*)$.

See[2]

3.2.2- Augmented Lagrangian with equality constraints in Hilbert spaces

Let $f : x \to R$ and $h : X \to Y$ be twice continuously X and Y be Hilbert space , We consider problem

$$\min f(x)$$

$$\text{Subject to } \mathrm{h(x)} = 0 \quad x \in \mathrm{X}$$

assume that $x^* \in X$ solve and that $h'(x_*)(\cdot)$ is susceptive and make the following assumptions

Ac_1**:** The mappings $f : x \to R$ and $h : X \to Y$ are twice continuously differentiable.

Ac_2:The iterates $\{x_k\}$ lie within a compact set.

Ac_3:At any limit point x_* of $\{x_k\}_{k \in N}$ the operator $h'(x_*)(\cdot)$ is susceptive

Ac_4: The convergence tolerances are $\omega_* = \eta_* = 0$

3.3 Global convergence analysis

Lemma(3):

Assume $Ac_1 - Ac_2$ to be valid. Let $(x_k)_{k\in K}$ be a subsequence with convergence to x^* .Further let $\lambda(x^*)$ be defined , assume that $(\lambda_k)_{k\in K\subset Y}$ is any sequence of vectors and that $(\mu_k)_{k\in K}$ from anon-creasing sequence of scalars. Suppose further ,that the iterates satisfy $\|\nabla_X \phi(x_k,\lambda_k,\mu_k)\| \prec \varepsilon$ where positive scalar parameters which converge to zero as $k\in K$ increase then there are positive constants k_1 and k_2 such that

$$\|\lambda^-(x_k,\lambda_k,\mu_k) - \lambda(x^*)\| \le k_1\omega_k + k_2\|x_k - x_*\|$$

$$\|\lambda(x_k) - \lambda(x^*)\| \le k_2\|x_k - x_*\|$$

$$\|c(x_k)\| \le k_1\omega_k\mu_k + \mu_k\|\lambda_k - \lambda(x_*)\| + k_2\mu_k\|x_k - x_*\|$$

for all $k\in K$ sufficiently large .Hence ,the sequence $(\lambda^-(x_k,\lambda_k,\mu_k))_{k\in K}$ and $(\lambda(x_k))_k$

$$\lim_{k\in K}\nabla_X \phi(x_k,\lambda_k,\mu_k) = \nabla_x L(x_*,\lambda(x_*)) = 0$$

Lemma(4):

Let $(x_k)_{k\in IN}$, $(\lambda_k)_{k\in IN}$, $(\eta_k)_{k\in IN}$ and $(w_k)_{k\in IN}$ be sequences generated by Algorithm. Further assume that step 3b is executed infinitely often ($\lim_{k\to\infty}\mu_k = 0$) then the product $\mu_k\|\lambda_k\|$ converges to zero.

Theorem(5):

let $Ac_1 - Ac_4$ be valid, let x_k be a subsequence whose limit is x_*. then x_* is a Kuhn Tucher (first order stationary point), and $\lambda(x_*)$ is the corresponding Lagrange multiplier

$\nabla L(x_*, \lambda(x^*)) = \nabla f(x_*) + h'(x_*) * (\lambda(x_*)) = 0, \qquad h(x_*) = 0$

Where the sequence $(\bar{\lambda}(x_k, \lambda_{k,} \mu_k))k \in K$ and $\lambda(x_k)_{k \in K}$, convergence to $\lambda(x_*)$ and the gradient convergence to $\nabla_x L(x_*, \lambda(x_{.*})) = 0$ for $k \in K$

see[3]

3.4- Augmented Lagrangian with equality constraints in feasible spaces

The optimization problem can be stated as follows:

$$\underset{x \in R^n}{\text{Minimize}} \quad f(x) \quad \text{subject to} \{h_k(x) = 0, \quad k \in \tau\} \qquad (11)$$

where f and the functions c_k, are all smooth, real-valued functions on a subset of R^n, and τ are finite sets of indices. As before, we call *f* the objective function, the $h_k, k \in \tau$ are the equality constraints .We define the feasible set ψ to be the set of points x that satisfy the constraints; i.e.,

$\psi = \{x : h_k(x) = 0,\ k \in \tau;\}$ ***and make the following assumptions***:

Ap_1: The region $\psi = \{x : h_k(x) = 0,\ k \in \tau;\}$ is nonempty.

Ap_2: the functions *f(x)* and *c(x)* are twice differentiable for all $x \in \psi$.

Ap_3 :The iterates $\{x_k\}$ considered lie within a closed ,bounded domain .

Ap_4**:** The convergence tolerances are $\varepsilon_* = 0.00000001$

New theorem (6):

let $Ap_1 - Ap_4$ be valid , let x_k a subsequence whose limit is x_* .Then x_* is a Kuhn Tucher (first order stationary point), and $\lambda(x_*)$ is the corresponding Lagrange multiplier

$\nabla L(x_*, \lambda(x^*)) = \nabla f(x_*) + h'(x_*) * (\lambda(x_*)) = 0, \qquad h(x_*) = 0$

Where the sequence $(\bar{\lambda}(x_k, \lambda_{k,}))k \in K$ and $\lambda(x_k)_{k \in K}$, convergence to $\lambda(x_*)$ and the gradient $\nabla_x L(x_*, \lambda_k, 1 - \lambda_k) = 0$ convergence to $\nabla_x L(x_*, \lambda(x_{.*})) = 0$ for $k \in K$

Proof: the proof is a combination of the theorem (4) and theorem (5).

4.Numerical Results:

In order to assess the performance of the proposed algorithm, the standard and new algorithms are tested over (8) non-linear constrained test functions (see appendix)with $1 \le n \le 5$ and $1 \le h(x) \le 3$.

All the results are obtained using (Pentium 4 computer). All programs are written in FORTRAN 90 language and for all cases the stopping criterion taken to be:

$$\|h_i\| \le \varepsilon$$

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

The comparative performance for all of these algorithms is evaluated by considering NOF , NOI , number of gradient NOG .

Table (1)

Comparison between the new method and standard

Lagrange multipliers.

Test Function	New algorithm NOF (NOG)NOI	standard Lagrange multipliers algorithm NOF (NOG)NOI
1	31(3)2	41(10)3
2	43(4)2	61(7)3
3	31(3)2	39(8)3
4	34(6)1	38(7)3
5	23(3)2	62(11)3
6	38(3)2	157(29)3
7	55(3)2	121(13)3
8	65(3)2	38(22)5
Total	320(28)15	557 (107)26

From the above table it is clear that the new algorithm **has an improvement of** $43\,\%\,NOF$, $75\,\%\,NOG$ and $43\,\%\,NOI$ compared with the standard algorithm .

5- References:

[1] Bunday, B.D.(1984) ”Basic Optimization Methods”, Edward Arnold, London.

[2] Conn, A.R" ,Global Convergence of two a augmented Lagrangian algorithms for optimization with a combination of general equality and linear constraints" ,1993.

[3] Jan, H. Maruhn ,An augmented Lagrangian algorithm for optimization with equality constraints in Hilbert spaces, April 30, 2001.

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

[5] Shuonan, Dong" Methods for Constrained Optimization"18.086 Project 2 ,Spring 2006.

INT [1]www.cs.cinvestav.mx/~constraint/papers/chapter.pdf

INT [2] www.athenasc.com/lmultpreface.html.

INT [3]www.cs.cinvestay.mx/ constraint /papers/kuri.pdf.

INT [4]www.neuro informat ik .rub.de.

Appendix

1. **min** $f(x) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$
 s.t
 $x_1^2 - x_2 = 0$

2. **min** $f(x) = x_1 x_2^2 + 2$
 s.t
 $x_1^2 - x_2^2 = -2$

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

5-**min** $f(x) = 0.5x_1^2 + 2.5x_2^2$
 s.t
 $x_1 - x_2 - 1 = 0$

6-**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
 $x_1 + x_2^2 + x_3^3 + 2 - 3\sqrt{2} = 0$
 $x_2 + x_3^2 + x_4 + 2 - 2\sqrt{2} = 0$
 $x_1 x_5 + 2 = 0$

7-**min** $f(x) = \exp(x_1 x_2 x_3 x_4 x_5)$
 s.t
 $x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2 = 1$
 $x_1^3 + x_2^3 = 1$
 $x_2 x_3 - 5x_4 x_5 = 0$

8-**min** $f(x) = (x_1 - 1)^2 + (x_1 - x_2)^2 + (x_3 - 1)^2 + (x_4 - 1)^4 + (x_5 - 1)^6$
 s.t
 $x_1^2 x_4 + \sin(x_4 - x_5) = 2\sqrt{2}$
 $x_2 + x_3^4 x_4^2 = 8 + \sqrt{2}$