Journal of Kirkuk University –Scientific Studies , vol.4, No.2,2009

A Modified estimation for the steplenght of a descent nonlinear algorithm

Abbas Y. AL-Bayati* and Ivan Subhi Latif**
****College of Computers Sciences and Mathematics-University of Mosul***
*****College of Scientific Education-University of Salahaddin, Erbil***

Abstract

In this paper, we have presented a numerical algorithm for the step-size estimation for minimization problems. Global convergence results are derived for descent algorithms in which the line search step is replaced by a step whose length is determined by step-size estimation formula. Numerical results show that the new estimation step-size required less storage and greatly speeded up the convergence of the gradient algorithm for large-scale unconstrained optimization problems. Also the new proposed algorithm seems to converge better and superior to other similar algorithms in many situations.

Introduction

Let $\Re^n$ be an n-dimensional Euclidean space and let $f:\Re^n \to \Re$ be continuously differentiable function. Line-search methods for solving the unconstrained minimization problem

$$\text{Min } f(x) \qquad x \in \Re^n, \tag{1}$$

have the form defined by the equation

$$x_{k+1} = x_k + \alpha_k d_k \qquad k = 1,2,3,... \tag{2}$$

where $x_1 \in \Re^n$ is an initial point , d_k is a descent direction of $f(x)$ at x_k ,and α_k is the step-size.

Let x_k be the current iterative point $k = 1,2,3,...$ and x^* be a stationary point which satisfies $\nabla f(x^*) = 0$ we denoted the gradient $\nabla f(x_k)$ by g_k, the function value $f(x_k)$ by f^* choosing the search direction d_k and determining the step-size α_k along the search direction at each iteration are the main tasks in line search methods. The search direction d_k is generally required to satisfy:

$$g_k^T d_k < 0 \tag{3}$$

which guarantees that d_k is a descent direction of $f(x)$ at x_k .In order to guarantee the global convergence, we required some times that d_k satisfies the sufficient descent condition

$$g_k^T d_k \le -c\|g_k\|^2 \tag{4}$$

where $c > 0$ is a constant .Moreover, we need to choose d_k to satisfy the angle property

$$\cos\langle -g_k , d_k \rangle = \frac{-g_k^{\mathrm{T}} d_k}{(\|g_k\| . \|d_k\|)} \geq \eta_0 \tag{5}$$

where $\eta_0 \in (0,1]$ is a constant and $\langle -g_k , d_k \rangle$ denotes the angle between the vectors $-g_k$ and d_k.

The commonly used line search rules are as follows see(Armijo,1966;Fiacco&Cormick,1990;Yuan,2006 and Yaun&Sun,1997):

(a) Minimization Rule: At each iteration, α_k is selected so that

$$f(x_k + \alpha_k d_k) = \min_{\alpha > 0} f(x_k + \alpha d_k) \tag{6}$$

(b) Approximate Minimization Rule:At each iteration, α_k is selected so that

$$\alpha_k = \min\left\{\alpha \middle| g(x_k + \alpha d_k)^{\mathrm{T}} d_k = 0 , \alpha > 0\right\} \tag{7}$$

(c) Armijo Rule: Set scalar $S_k , \beta, L > 0, \sigma$ with

$$S_k = \frac{-g_k^{\mathrm{T}} d_k}{\left(L\|d_k\|^2\right)} \quad , \beta \in (0,1), \sigma \in (0,1/2) .$$

Let α_k be the larges α in $\{s_k , \beta s_k , \beta^2 s_k ,...\}$ such that

$$f_k - f(x_k + \alpha d_k) \geq -\sigma\alpha g_k^{\mathrm{T}} d_k \tag{8}$$

(d) Limited minimization Rule: set $S_k = \frac{-g_k^{\mathrm{T}} d_k}{\left(L\|d_k\|^2\right)}$ where α_k is defined by

$$f(x_k + \alpha_k d_k) = \min_{\alpha \in [0, s_k]} f(x_k + \alpha d_k) \tag{9}$$

where $L > 0$ is a constant

(e) Goldsten Rule:A fixed scalar $\sigma \in (0,1/2)$ is selected and α_k is chosen to satisfy

$$\sigma \leq \frac{\left[f(x_k + \alpha_k d_k) - f_k \right]}{\alpha_k g_k^{\mathrm{T}} d_k} \leq 1 - \sigma \tag{10}$$

It is possible to show that, if f is bounded below, there exists an interval of step-size α_k for which the relation above is satisfied, there are fairly simple algorithms for finding such a step-size thorough a finite number of arithmetic operations.

(f) Strong Wolfe Rule: α_k is chosen to satisfy simultaneously

$$f_k - f(x_k + \alpha d_k) \geq -\sigma\alpha_k g_k^{\mathrm{T}} d_k \tag{11}$$

$$\left| g(x_k + \alpha d_k)^{\mathrm{T}} d_k \right| \leq -\beta g_k^{\mathrm{T}} d_k \tag{12}$$

where σ and β are some scalars with $\sigma \in (0,1/2)$ and $\beta \in (\sigma,1)$

(g) Wolfe Rule: α_k is chosen to satisfy (11) and

$$g(x_k + \alpha d_k)^T d_k \geq \beta g_k^T d_k \qquad (13)$$

some important global convergent result for various method using the above mentioned specific line-search procedures have been given (Armijo,1966;More,Garbow&Hilstrom,1981) .In fact, the above mentioned line-search methods are monotone descent for unconstrained optimization see (Fiacco&McCormick,1990;Goldstein&Price,1967). Non monotone line-search methods have been investigated also by many authors see (More,Garbow&Hilstrom,1981;Nocedal&Wright,1999) is a non monotone descent method which is an efficient algorithm for solving some special problems. Shi & Shen (Shi & Shen ,2006) described some descent algorithm without line-search.

In this paper we describe a new algorithm without line-search. The basic idea is to estimate the line-search step which based on a known parameter. The algorithm is compared with similar published algorithms that may be more effective than Shi & Shen algorithm (Shi & Shen ,2006).

1. Shi & Shen Algorithm without line-search (SSW)

Shi & Shen assumed that some property holds to find descent algorithm without line-search

(H1) $f(x)$ is bound below denoted that

$$L(x_0) = \{x \in \Re^n \mid f(x) \leq f(x_0)\} \qquad (14)$$

(H2) the gradient $g(x)$ is uniformly continuous on an open convex set B that contains L_0.Some times further assumed that the following condition holds.

(H3)the gradient $g(x)$ is lipschitc continuous on an open convex set B that contains the level set $L(x_0)$,i.e ,there exists L such that

$$\|g(x) - g(y)\| \leq L\|x - y\| \ , \ \forall x, y \in B \qquad (15)$$

Obviously, (H3) implies (H2)

1.1. Outlines of the SSW Algorithm without line search :

Corresponding SSW gradient descent algorithm without line-search may be listed as follows:

Step1: choose an initial point $x_0 \in \Re^n$, $v \in (0,2)$ and $M > L_0 > 0$ set $K = 1$.

Step2: if $\|g_k\| = 0$ then stop! Go to step3.

Step3: Estimate $L_k \in [L_0, M]$

Step4: $x_{k+1} = x_k - \frac{v}{L_k} g_k$;

Step5: set $k = k + 1$ and go to step2

The following formula for $\alpha_k = \dfrac{v}{L_k} \quad (k > 1)$

1. $L_k = \min(M, \max\left\{ L_{k-1}, \dfrac{y_{k-1}^{\mathrm{T}} V_{k-1}}{\|V_{k-1}\|^2} \right\}$ (16)

2. $L_k = \min(M, \max\left\{ L_{k-1}, \dfrac{\|y_{k-1}\|^2}{y_{k-1}^{\mathrm{T}} V_{k-1}} \right\}$ (17)

3. $L_k = \min(M, \max\left\{ L_{k-1}, \dfrac{\|y_{k-1}\|}{\|V_{k-1}\|} \right\})$ (18)

4. $L_k = \min(M, \dfrac{2(f_k - f_{k-1} + \alpha_{k-1}\|g_{k-1}\|^2)}{\alpha_{k-1}^2 \|g_{k-1}\|^2}$ (19)

where $V_{k-1} = x_k - x_{k-1}$, $y_{k-1} = g_k - g_{k-1}$ and $\|.\|$ denotes Euclidean norm.

1.2. Some properties of the SSW Algorithm:

The following lemma can be found in (Nocedal,1992).

Lemma1.2.1: (mean value theorem).Suppose that the objective function $f(x)$ is continuously differentiable on an open continuously differentiable on an open convex set B, then

$$f(x_k + \alpha_k d_k) - f_k = \alpha \int_0^1 d_k^{\mathrm{T}} g(x_k + t\alpha d_k dt) \tag{20}$$

where $x_k, x_k + \alpha d_k \in B$ and $d_k \in \Re^n$ further, if $f(x)$ is twice continuously differentiable on B then

$$g(x_k + \alpha d_k) - g_k = \alpha \int_0^1 \nabla^2 f(x_k + t\alpha d_k) \alpha_k^r dt \tag{21}$$

and

$$f(x_k + \alpha_k d_k) - f_k = \alpha g_k^{\mathrm{T}} d_k + \alpha^2 \int_0^1 (1-t) d_k^{\mathrm{T}} \nabla^2 f(x_k + t\alpha d_k) d_k dt \tag{22}$$

1.3. Convergence of the SSW Algorithm:

Theorem2.3.1

If (H1) and (H2) hold SSW algorithm generates an infinite sequence $\{x_k\}$ and $\rho \in (\frac{v}{2}, 1)$, $L_k \ge \rho L$

$$\sum_{k=0}^{+\infty} \frac{1}{L_k^2} = +\infty \tag{23}$$

Then

$$\lim_{k\to\infty} \inf \|g_k\| = 0 \tag{24}$$

For the proof see (Shi&Shen,2006)

□

In order to analyze the convergence rate of the algorithm, used of assumption (H4) below

(H4) $\{x_k\} \to x^*$,$(k \to \infty)$, $f(x)$ is twice continuously differentiable on $N(x^*,\varepsilon)$ and $\nabla^2 f(x^*)$ is positive definite.

Lemma1.3.1: Assume that (H4) holds. Then (H1),(H3) and (H2) hold automatically for k sufficiently large, and there exists $0 < m' < M'$ and $\varepsilon_0 < \varepsilon$ such that

$$m'\|y\|^2 \le y^{\mathrm{T}} \nabla^2 f(x) y \le M'\|y\|^2 \tag{25}$$

$$\frac{1}{2} m'\|x - x^*\|^2 \le f(x) - f(x^*) \le \frac{1}{2} M'\|x - x^*\|^2, \forall x, y \in N(x^*,\varepsilon) \tag{26}$$

$$M'\|x - y\|^2 \ge (g(x) - g(y))^{\mathrm{T}}(x - y) \ge m'\|x - y\|^2, \forall x, y \in N(x^*,\varepsilon) \tag{27}$$

and thus

$$M'\|x - x^*\|^2 \ge g(x)^{\mathrm{T}}(x - x^*) \ge m'\|x - x^*\|^2, \forall x \in N(x^*,\varepsilon) \tag{28}$$

and

$$\|g(x) - g(y)\| \le \|x - y\|, \forall x, y \in N(x^*,\varepsilon) \tag{29}$$

see [14],[15]

lemma1.3.2: If (H1) and (H3) hold and SSW Algorithm with $L_k < M$ and $L < M$ generates an infinite sequence $\{x_k\}$, then there exists $\eta > 0$ such that

$$f_k - f_{k+1} \ge \eta \|g_k\|^2, \forall k \tag{30}$$

see [9].

2. Anew Proposed Algorithm without line-search (NSSW)

In this section we propose a new algorithm without line-search based on estimate step-size. We assume that (H1), (H2), (H3) holds we shall implicitly assume that the constant L in (H3) is easy to estimate. We must estimate L_k at each iteration. Certainly if the Lipschitz constant L of the gradient of objective function is known apriore, then we can take $L_k \equiv L$ in the algorithm. We estimate Lipschitz constant L and find an approximation L_k to L we defined

$$L_k = \left\{ L_{k-1}, \frac{\|V_k\|^2}{y_{k-1}^{\mathrm{T}} V_{k-1}}, \frac{y_{k-1}^{\mathrm{T}} V_{k-1}}{\|y_{k-1}\|}, \frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{\|V_{k-1}\|}, \frac{2[f(x_{k-1}) - f(x_k)] + (g(x_k) - g(x_{k-1}))^{\mathrm{T}} V_{k-1}}{\|V_{k-1}\|^2} \right\} \quad (31)$$

where $V_{k-1} = x_k - x_{k-1}, y_{k-1} = g_k - g_{k-1}$

2.1. Outlines of the NSSW Algorithm without line search :

The outliers of the new algorithm without line-search may be given as follows:

Step1: choose an initial point $x_0 \in \Re^n$, $v = 0.8$ and $M = 10^8$ set $K = 1$.

Step2: if $\|g_k\| = 0$ then stop! Go to step3.

Step3: Estimate $L_k \in [L_0, M], v = 0.8, M = 10^8$

Step4: $x_{k+1} = x_k - \frac{v}{L_k} g_k$;

Step5: if $f_k > f_{k-1}$ then go to step 6 else go to step7

Step 6: $V_k = (0.05)^{V_k}$ go to step 7

Step7: set $k = k + 1$ and if $\frac{v}{L_k} < 0$ then let $\frac{v}{L_k} = 1$ and go to step2

The formulae for $\alpha_k = \frac{v}{L_k}, (k > 1)$

$$L_k = \min\left(M, \max\left\{ L_{k-1}, \frac{\|V_k\|^2}{y_{k-1}^{\mathrm{T}} V_{k-1}}, \frac{y_{k-1}^{\mathrm{T}} V_{k-1}}{\|y_{k-1}\|}, \frac{g_k^{\mathrm{T}}(g_k - g_{k-1})}{\|V_{k-1}\|}, \frac{2[f(x_{k-1}) - f(x_k)] + (g(x_k) - g(x_{k-1}))^{\mathrm{T}} V_{k-1}}{\|V_{k-1}\|^2} \right\} \right) \quad (32)$$

2.2. Some Theoretical properties of the NSSW Algorithm:

The new formulae are useful because they arise from the classical quasi-Newton condition(Nocedal,1999) and from Barzilai and Borweins idea(Barziliai&Borwein,1988) .

Some recent observations on Barzilai and Borweins method are very exciting (Fletcher,2001;Raydoan,1997;Shi&Shen,2006;Diao&Liao,2002).

Remark: the above theorem shows that we can set a large L_k and small v to guarantee the global convergence.

However, if L_k is very large then α_k will be very small and will slow the convergence rate of descent methods.

Theorem2.2.1 Assume that the hypotheses of theorem 1.2.1 hold. Denote the exact step by α_k^* (including exact line search rule (a) and (b)) then

$$\alpha_k^* \geq \frac{\rho}{v} \alpha_k \quad (33)$$

Proof: Where L_k defined (1)

$$L_k = \frac{\|g_{k+1} - g_k\|}{\|x_{k+1} - x_k\|}$$

$$= \frac{(g_{k+1} - g_k)^{\mathrm{T}}(g_{k+1} - g_k)}{V_k^{\mathrm{T}} V_k}$$

$$= \frac{g_{k+1}^{\mathrm{T}} g_{k+1} - 2g_{k+1}^{\mathrm{T}} g_k + g_k^{\mathrm{T}} g_k}{g_k^{\mathrm{T}} g_k \alpha_k^2}$$

$$= \frac{\|g_{k+1}\| + \|g_k\|}{\|g_k\| \alpha_k^2}$$

$$= \frac{1}{\alpha_k^2}[\frac{\|g_{k+1}\|}{\|g_k\|} + 1]$$

$$= \frac{1}{\alpha_k^2}[\beta_{FR} + 1]$$

for the line search rules (a)and (b),(H3)and Cauchy-Schwartz inequality, we have

$$\alpha_k^* L\|g_k\|^2 \geq \|g_{k+1} - g_k\| . \|g_k\|$$

$$\geq -(g_{k+1} - g_k)^{\mathrm{T}} g_k$$

$$= \|g_k\|^2$$

Therefore

$$\alpha_k^* \geq \frac{1}{L}$$

noting that $L_k \geq \rho L$ we have

$$\alpha_k^* \geq \frac{1}{L} \geq \frac{\rho}{L_k} = \frac{v}{L_k} \frac{\rho}{v} = \alpha_k \frac{\rho}{v}$$

In order to analyze the convergence rate of the new algorithm we use assumption (H4) and lemma (2.2.1) hold.

Lemma2.2.1: If (H1) and (H3) hold and NSSW Algorithm with $L_k < M$ and $L < M$ generates an infinite sequence $\{x_k\}$, then there exists $\eta > 0$ such that

$$f_k - f_{k+1} \geq \eta \|g_k\|^2, \forall k \qquad (34)$$

Proof: By lemma2.2.1 and (H3) we have

$$f(x_k + \alpha_k d_k) - f_k = \alpha \int_0^1 d_k^{\mathrm{T}} g(x_k + t\alpha d_k dt)$$

$$= \alpha g_k^{\mathrm{T}} d_k + \alpha \int_0^1 d_k^{\mathrm{T}} (g(x_k + t\alpha d_k) - g_k) dt$$

$$\leq \alpha g_k^{\mathrm{T}} d_k + \alpha \int_0^1 \|d_k\| \|g(x_k + t\alpha d_k - g_k)\| dt$$

$$\leq \alpha g_k^T d_k + \alpha^2 L \int_0^1 t \|d_k\|^2 dt$$

$$= \alpha g_k^T d_k + \frac{1}{2}\alpha^2 L \|d_k\|^2$$

Taking $d_k = -g_k$ and $\alpha = \frac{v}{L_k}$ in the above formula,we have

$$f_k - f_{k+1} \leq -\alpha_k \|g_k\|^2 + \frac{1}{2}\alpha_k^2 \|g_k\|^2$$

$$= \frac{v}{L_k}\|g_k\|^2 - \frac{v^2 L}{2L_k^2}\|g_k\|^2$$

$$\leq -\left(\frac{v}{L_k} - \frac{v^2 L}{2L_k^2}\right)\|g_k\|^2$$

thus

$$f_k - f_{k+1} \geq \left(\frac{v}{L_k} - \frac{v^2 L}{2L_k^2}\right)\|g_k\|^2$$

$$\geq \frac{(2\rho - v)vL}{2M^2}\|g_k\|^2$$

$$\eta = \frac{(2\rho - v)vL}{2M^2}$$

we obtain the desired result.

3. Numerical results

All the two algorithms described in this paper were coded in double precision newly-program FORTRAN programmed.

In comparisons of algorithms the function evaluation is normally assumed to be the most costly factor in iteration hence, the cost of solving a problem is normally presented in terms of the number of functions evaluation (NOF), is valuable in comparing similar algorithms, and is also presented here. The actual convergence criterion employed was $\|g_{k+1}\| < 1\times10^{-5}$ for all two algorithms, fourteen well-known test functions (Appendices 1 and 2) see (More,Garbow&Hillstrom,1981) and with dimensionality ranging (12-10000) are employed in the comparison.

The complete set of results is given in Tables (1), (2) while Table (3) gives the percentage of NOF for each function was solved using the following algorithms;

1-Shi & Shen Algorithm without line-search (SSW)

2- The New Algorithm without line-search (NSSW)

The important thing is that the new algorithm without line-search needs fewer evaluations of $f(x)$ and $g(x)$ than SSW algorithm. We can see that other algorithm may fail in some case while the new algorithm without line-search always converges. Moreover numerical experiments also show that the new algorithm always convergence stably.

It is clear from the Tables (1), (2) of the numerical results that the new proposed algorithm without line-search is very efficient and superior on the standard SSW algorithm namely there are approximately about (94 -95) improves of NOF for all dimensions .

Table (1):Comparison between the NSSW algorithm and SSW algorithms using different value of 12< N <4320 for the 1st class of test functions.

N.of Test	TEST FUNCTION	SSW (NOF)							NSSW (NOF)						
		12	36	360	1080	4320	8640	10000	12	36	360	1080	4320	8640	10000
1	GEN-Dixon	F	F	F	F	F	F	F	88	88	95	92	95	96	100
2	TPQ	F	F	F	F	F	F	F	87	151	511	1117	4191	4567	5663
3	Per Quadratic	F	F	F	F	F	F	F	66	137	538	952	2855	5568	5363
4	Quadratic	F	F	F	F	F	F	F	40	93	350	611	611	1450	2945
5	dqdric	F	F	F	F	F	F	F	327	277	311	318	320	312	312
6	GEN-start	F	F	F	F	F	F	F	367	408	405	423	398	435	431
7	BD1	F	F	F	F	F	F	F	90	92	99	101	104	104	107

Table (2):Comparison between the NSSW algorithm and SSW algorithms using different value of of 12< N <4320 for the 2nd class of test function

N.of Test	TEST FUNCTION	SSW (NOF)							NSSW (NOF)						
		12	36	360	1080	4320	8640	10000	12	36	360	1080	4320	8640	10000
1	Recip	14834	15520	16959	17645	18511	18944	19021	521	552	616	647	687	706	710
2	GEN-edger	697	724	781	888	842	860	863	42	43	45	46	48	48	49
3	GEN-Q1	599	599	599	599	599	599	599	57	59	60	61	64	66	66
4	GEN-Q2	194	194	194	194	194	194	194	63	63	64	63	63	66	65
5	Digonal4	1232	1287	1401	1456	1525	1559	1567	46	47	50	52	54	56	56
6	Digonal6	1220	1275	1389	1444	1513	1548	1582	27	28	32	33	36	37	37
7	GEN-Shallow	3301	3460	3795	3955	4157	4258	4279	438	445	458	471	462	457	433
General TOTAL of 7 functions		22077	23059	25118	26181	27341	27962	28105	1194	1237	1325	1373	1414	1436	1416

Table(3): Percentage performance of the NSSW algorithm against SSW algorithm for 100% in NOF.

N	Costs	NEW
12	NOF	94.59
36	NOF	94.64
360	NOF	94.73
1080	NOF	94.76
4320	NOF	94.83
8640	NOF	94.86
10000	NOF	94.96

4. Conclusions

In this Paper, a new step-size estimation numerical algorithm without line search is proposed .Global convergence is derived for gradient method in which the line-search is replaced by a step-size estimation whose length is determined by a formula. Where the gradient method with L_k defined by (32) to construct gradient methods without line-search be the best algorithm for test problems.

As a result, we should choose carefully L_k in practical computations in order to satisfy both global convergence and the fast convergence rate. Our numerical results supports our claim and also indicate that the new algorithm without line-search may be converge faster and is more efficient than SSW algorithm in many situations.

References

- Armijo, L. , (1966):Minimization of function having Lipschitz continuous first partial derivatives . pacific.J.Math., Vol.16, pp: 1-13.
- Barzilai,J. and Borwein, J.M., (1988): Two-point size gradient methods. IMA J.Numer.Anal.,Vol.8,PP.141-148,.
- Dai, Y.H. and Liao, L.Z., (2002): R-linear convergence of the Barzilai and Borwein gradient method. IMA J.Numer.Anal.,Vol.22,PP.1-10.
- Fletcher, R., (2001): The Barzilai Borwein method-steepest descent method resurgent. Report in the international Workshop on Optimization and control with Application,Erice,Italy,July,PP.9-17.

- Fiacco, A.V.and McCormick, G.P.,(1990): Nonlinear programming. Sequential Unconstrained Minimization Techniques,SIAM,Philadelphia.
- Goldstein, A.A., Price, J.F., (1967): An effective algorithm for minimization. Numer.Math., Vol.10, pp: 184-189.
- More, J.J., Garbow, B.S. and Hillstrom, K.E. , (1981):Testing unconstrained optimization software. ACM Trans,Math.Software,Vol.7, pp:17-41.
- Nocedal, J., (2005): Unconstrained Optimization Test function Research Institute for Informatics. Center for advanced Modeling and Optimization, Bucharest1, Romania. April 28,.
- Nocedal, J., (1992): Theory of Algorithms for Unconstrained Optimization . Acta.Numer., Vol.1, pp:199-242.
- Nocedal, J. and Wright, J.S.,(1999): Numerical Optimization.Spring-Verlag NewYork,.
- Raydan, M.,(1997): The Barzilai Borwein method for large scale unconstrained minimization problem . SIAM J.Optim.,Vol.7, pp:26-33.
- Raydan, M., (1993): On the Barzilai Borwein gradient choice of steplength for the gradient method. IAM J.Numer.Anal.,Vol.13, pp:321-326.
- Shi, Z.J.& Shen, J.,(2006): Step-size Estimation for Unconstrained Optimization Methods. J.Math.Apl.Comput.,Vol.24, PP:408-418.
- Yuan, Y., (1993): Numerical Methods for Nonlinear Programming. Shanghai Scientific&Technical Publishers,.
- Yuan,Y. and Sun, W.Y., (1997): Optimization Theory and Methods..Science Press,Bijing.

Appendix1

All the test functions used in Table (1) for this paper are from general literature:

1. Generalized Recip Function:

$$f(x)=\sum_{i=1}^{n/3}\left[(x_{3i-1}-5)^2+x_{9i-1}^2+\frac{x_{3i}^2}{(x_{3i-1}-x_{3i}-2)^2}\right],\ x_0=[2.,5.,1.,...,2.,5.,1.].$$

2. Generalized Edger Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}-2)^4+(x_{2i-1}-2)^2x_{2i}^2+(x_{2i}+1)^2\ ,\ x_0=[1.,0.,...,1.,0.].$$

3. Generalized quartic Function GQ1

$$f(x)=\sum_{i=1}^{n-1}x_i^2+(x_{i+1}+x_i^2)^2\ ,\ x_0=[1.,1.,...,1.].$$

4. Generalized Quartic Function GQ2:

$$f(x)=(x_1^2-1)^2+\sum_{i=2}^{n}(x_i^2-x_{i-1}-2)^2\ ,\ x_0=[1.,1.,...,1.,1.].$$

5. Diagonal4Function:

$$f(x)=\sum_{i=1}^{n/2}\tfrac{1}{2}\left(x_{2i-1}^2+cx_{2i}^2\right),\ x_0=[1,1,...,1]\ ,\ \mathrm{c}=100.$$

6. Diagonal 6 Function:

$$f(x)=\sum_{i=1}^{n}(\exp(x_i)-(1+x_i))\ ,\ x_0=[1.,1.,...,1.,1.].$$

7Generalized Shallow Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2-x_{2i})^2+(1-x_{2i-1})^2\ ,\ x_0=[-2.,-2.,...,-2.,-2.].$$

Appendix2

All the test functions used in Table (2) for this paper are from general literature:

1. Generalized Dixon function:

$$f(x)=(1-x_1)^2+(1-x_n)^2+\sum_{i=2}^{n/2}(x_i^2-x_{i+1})^2\ ,\ x_0=[-1.,..,-1.]$$

2. Tridiagonal Perturbed Quadratic Function:

$$f(x)=x_i^2+\sum_{i=2}^{n-1}ix_i^2+(x_{i-1}+x_i+x_{i+1})^2\ ,\ x_0=[0.5,0.5.,...,0.5,0.5].$$

3. Perturbed Penalty Function:

$$f(x)=\sum_{i=1}^{n}ix_i^2+\tfrac{1}{100}\left(\sum_{i=1}^{n}x_i\right)^2\ ,\ x_0=[0.5,0.5,...,0.5].$$

4. Quadratic Function QF2:

$$f(x)=\tfrac{1}{2}\sum_{i=1}^{n}i(x_i^2-1)^2-x_n\ ,\ x_0=[0.5,0.5,...,0.5,0.5].$$

5. Dqudrtic Function (CUTE):

$$f(x)=\sum_{i=1}^{n-2}\left(x_i^2+cx_{i+1}^2+dx_{i+2}^2\right),\ x_0=[3.,3.,...,3.,3.]\ ,\ \mathrm{c}=100,\mathrm{d}=100.$$

6. Generalized Strail Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2-x_{2i})^2+100(1-x_{2i-1})^2\ ,\ x_0=[-2.,...,-2.].$$

7. Extended Block-Diagonal BD1 Function:

$$f(x)=\sum_{i=1}^{n/2}(x_{2i-1}^2+x_{2i}^2-2.)^2+(\exp(x_{2i-1}-1.)-x_{2i}.)^2\ ,\ x_0=[0.1,0.1,...,0.1].$$

تخمين جديد لخط البحث في خوارزمية الانحدار اللاخطية

عباس يونس البياتي*   و   الانسة ايفان صبحي لطيف**
***كلية علوم الحاسبات والرياضيات ــ جامعة الموصل**
****كلية التربية العلمية ــ جامعة صلاح الدين/اربيل**

الخلاصة

في هذا البحـــث تم استحداث خوارزمية عددية لتخمين خط البحث في مسائل الامثلية. تم استنتاج التقارب الامثل من خلال اشتقاق خط بحث لخوارزمية الانحدار ،حيث تم استبدال خط البحث بأيجاد خطوة مخمنه لها .

النتائج العملية للخوارزمية المقدمة كفؤه و يحتاج الى خزن اقل ولسرعة تقارب اكبر في خوارزميات التدرج ذات القياس العالي في الامثلية اللاخطية . وكذلك تم اثبات ان الخوارزمية الجديده ذات تقارب أفضل مقارنة بالخوارزميات المماثلة في مثل هذه الحالات .