New Hybrid CG-Projected Bfgs for Minimization

Abbas Y. AL Bayati **Osama H. Ta'ani**
College of Computers and
Mathematical Science
Mosul University

(Received 13/5/2001 , Accepted 26/6/2001)

ABSTRACT

In this paper an hybrid CG direction and a modified direction of the Storey's projected version of the BFGS Quasi-Newton (QN) update of optimization is investigated theoretically and experimentally.

The new proposed algorithm is compared with Storey's projected BFGS method on a large number of standard test functions with dimensionality varies between $4 \le n \le 1000$. The new algorithm is found to be superior to the Storey's projected BFGS method overall but it showed a considerable superiority on some of test functions.

خوارزمية جديدة هجينية لـ BFGS الاسقاطية مع التدرج المترافق

الملخص

في هذا البحث تم استحداث خوارزمية جديدة في مجال تهجين خط بحث التدرج المترافق المولد مع خط البحث للمصفوفة المستحدثة لمصفوفة BFGS الاسقاطية وتم التحقق من الخوارزمية نظرياً و عملياً.

تم مقارنة الخوارزمية المقترحة مع خوارزمية BFGS الاسقاطية باستعمال عدد كبير من الدوال اللاخطية ذات الابعاد 4≤n≤1000 وجدنا إن الخوارزمية الجديدة ذات كفاءة عالية مقارنةً بخوارزمية Story وهي متفوقة جداً في بعض الدوال المستخدمة في هذا البحث.

INTRODUCTION

The most important algorithms to find the minimum of unconstrained function are the Conjugate Gradient (CG) and Quasi-Newton (QN) methods. QN methods have the direction $d_k = -H_k g_k$ where H_k is k-th approximation matrix to the Hessian inverse of f, and g_k is the gradient vector ∇f.

There is different ways to obtain H, and the best is the BFGS method. Hu and Storey in 1990 developed the BFGS in projected version, where the theoretical explanation and numerical experience show that the projected QN method better than the QN method, the CG- algorithms have the direction

$d_k=-g_k+\beta_k d_k$.. (1)
where β_k, is a scalar obtained by different ways as we will see later.

On the other hand Buckley in 1978 introduced the idea of interleaving CG-direction by a QN-direction to accelerate the efficiency of the CG-algorithm. Nazareth in 1979 developed a precondition CG to accelerate the standard CG-algorithm which has the following direction

$d_k=-H_k g_k+\beta_k d_k$……………………………………………………… (2)

$\beta_k = g_{k+1}^T H_k g_{k+1} / g_k^T g_k$

Storey and Touati-Ahmed in (1990) developed the hybrid CG-algorithm by choosing a suitable scalar, β_k,

CONJUGATE GRADIENT METHODS

We consider the unconstrained minimization problem

$$\min_{x\in R^n} f(x) \quad \text{.. (1)}$$

where f is twice continuously differentiable function from R^n to R.

CG methods are iterative methods generates, respectively, three sequences $\{x_k\}$, $\{g_k\}$, $\{d_k\}$,for k =1,2, …. until termination, where x_k estimates a local solution x^* of equation (IA), g,; is the gradient of f ; and d_k is the search direction, which has the descent property

$d_k^T g_k \prec 0$.. (2)

where $g_k \neq 0$.

The original CG-methods proposed by (Fletcher and Reeves, 1964)is given by

$d_{k+1}^{FR} = -g_{k+1} + \beta_k g_k$ .... (3)

$x_{k+1} = x_k + \lambda_k d_k$ .... (4)

Such that d_{k+1} is "conjugate to" d_k in the sense that

$d_k^T G d_j = 0 \qquad for \;\; j \neq k$ …………………………………… (5)

where G is the Hessian matrix and λ_k is the step-length parameter, satisfying

$f(x_k + \lambda_k d) \leq f(x) + \sigma_1 \lambda_k g_k^T d_k$ .... (6)

and $\left| g_{k+1}^T d_k \right| \leq \sigma_2 g_k^T d_k$ ……………………………………………… (7)

here $0 < \sigma_1 < \sigma_2 < 1$ and where

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad \text{……………………………………………… (8)}$$

Other choices of the parameter β_k in (3) give rise to distinct algorithms for nonlinear problems the most famous of them are

$$\beta_k^{Hs} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_k^T \|g_k\|^2} \quad \text{………………………………………… (9)}$$

which is due to Hestenes and Stiefel (1952).

$$\beta_k^{FR} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{\|g_k\|^2} \quad \text{... (10)}$$

which is due to Polak and Ribeier (1969).

$$\beta_k^D = \frac{-\|g_{k+1}\|^2}{d_k^T g_k} \qquad (11)$$

which is due to Dixon (1972).

$$\beta_k^{BA} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{d_k^T g_k} \qquad (12)$$

which is due to Al-Bayati and A1 Assady (1986) .

Indeed, extensive numerical experience has shown that the Polak-Ribier method is more efficient than the original Fletcher-Reeves method (Polak, 1969).

There is a theoretical explanation which shows that the PR-formula is better than FRformula. On non-quadratic functions it can happen (Fletcher, 1987) that the search direction d_k becomes almost orthogonal to $-g_k$ and hence little progress can be made. In this event $x_{k-1}=x_k$ and g_{k+1}-g_k so FR-method then gives

$$d_{k+1} \cong -g_{k+1} + d_k \qquad (13)$$

while the PR-method becomes

$$d_{k+1} \cong -g_{k+1} \qquad (14)$$

So that in these circumstances the PR-algorithm tends to reset automatically to the steepest descent direction, thus, it seems that this formula should be used when solving large problems. Indeed CG-algorithms are usually implements errors.

Fletcher in his standard method suggested to restart his algorithm with the steepest descent direction every n or n+l iterations. In (Crowder and Wolfe, 1972) it has been shown that if restarting is not employed the convergence of the algorithm will be linear only.

PROJECTED QN- METHOD

The basic method to solve the problem (1) is the Newton's method where the search direction is given by

$$d_k = G_k^{-1} g_k \qquad (1)$$

where G_k is the Hessian matrix of f at x_k.

If we can afford the computer storage space, QN-methods are generally an improvement on the performance of the CG-methods and Newton's method (see Beale, 1988).

This type of the method is like Newton's method with line search, except that the Hessian matrix G_k is approximated by a symmetric positive definite matrix H_k which is corrected from iteration to iteration. Thus the k-th iteration has the following basic structure :-

I) Set $d_k = -H_k g_k$

II) Do a line search along d_k to get

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

III) Update H_k by a correction matrix to get.

$$H_{k+1} = H_k + Q_k$$

Where Q_k is a correction positive definite matrix and satisfies the QN-condition that is $H_{k+1} y_k = v_k$ .... (2)

where $y_k = g_{k+1} - g_k$; $v_k = x_{k+1} - x_k$.

There are different forms for H_k suggested by different researches and we consider some of them latter in this work.

In the first case, H_k is taken to be always the identity matrix but for Newton's method H_k is selected to be the inverse Hessian matrix of the function f(x) at x_k .

From the convergence viewpoint, Newton's method is the ideal in this class since it converges quadratically, but it is still requires the inversion of an n x n matrix.

The most general approach is the one published by (Haung, 1970). He considered an algorithm with exact line search that uses a general updating formula of the form

$$H_{k+1}^{H} = H_k + \omega_k v_k v_k^T + \pi_k H_k y_k y_k^T H_k + \rho(v_k y_k^T H_k + H_k y_k^T v_k) \text{ .... } (3)$$

where ω_k, π_k, ρ_k are scalars and different choices of the parameters ω_k, π_k, and ρ_k will product different QN-algorithms.

Huang proved that in order to obtain conjugate search directions in n-step convergence, for the quadratic case, $H_{k+1}\, y_k$ has to be multiple of v_k. Using this condition in formula (3) gives the following updating formula:

$$H_{k+1}^{Hung} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \theta_k \frac{v_k v_k^T}{v_k^T y_k} + \phi_k w_k w_k^T \text{ .... } (4a)$$

where

$$w_k = \left(y_k^T H_k y_k\right)^{\frac{1}{2}} \left(\frac{v_k v_k^T}{v_k^T y_k} + \frac{H_k y_k}{y_k^T y_k} \right) \text{ .... } (4b)$$

and θ_k, ϕ_k, are scalars.

If θ_k is fixed number, say θ, then for the quadratic case with Hessian G, we have $H_n = \theta G^{-1}$ For this reason, most of the algorithms derived for use $\theta_k = 1$ for all k to satisfy the exact QN-condition. Formula (4) actually includes all the well-known updating formula ; however, the updating formula which is obtained from (4) by setting θ_k = 1and ϕ_k= 0 is the DFP-method. Also the very interesting updating formula which has discovered independently by Broyden, Fletcher, Goldfart and Shanno in 1970 is the BFGS formula which is obtained from (4) by setting $\theta_k = 1$ and , $\phi_k = 1$, we get

$$H_{k+1}^{BFGS} = H_k - \frac{H_k y_k y_k^T H_k}{y_k^T H_k y_k} + \frac{v_k v_k^T}{v_k^T y_k} + w_k w_k^T \text{ .... } (5)$$

where wk is a vector defined in (4b).

Since QN-methods at the (k+l)-th iteration form an approximation matrix $H_{k+1} \in R^n$ to the inverse of the Hessian of f at x_{k+1}, where H_{k+1}, satisfies the so called the exact QN-condition defined in (2) and since (2) alone does not completely specify the matrix H_{k+1} there is a lot of freedom in choosing it. In order to, restrict this freedom, one can require H_{k+1}, to be symmetric positive definite so that d_{k+1} will be a descent direction. In many practical problems, the evaluation of functions and gradients is very expensive so that this part of the calculator overwhelms the routine part. To avoid the previous problems Storey in 1990 developed BFGS-update to projected version as follows :

He used extra restrictions to BFGS-update to reduce the function and gradient evaluations therefore he proved that to obtains suitable symmetric QN matrix H_{k+1} it is must be satisfied QN-condition and some previous equation , so that :

$$\begin{cases} H_{k+1}y_k = v_k \\ H_{k+1}y_i = H_k y_i & i = 1,2,.....,k-1 \\ H_{k+1}^T = H_{k+1} \end{cases} \quad \text{...(6)}$$

here we assume $1 \le k \le n$, if we set $\eta_k = v_k - H_k y_k$ then we can write BFGS-update as follows :

$$\beta_{k+1}^{BFGS} = H_k + \frac{v_k \eta_k^T + \eta_k v_k^T}{v_k^T y_k} - \frac{\eta_k^T y_k}{\left(v_k^T y_k\right)^2} v_k v_k^T \quad \text{.. (7)}$$

If we denote $E = H_{k+1} - H_k$

$$\begin{cases} Ey_k = \eta_k \\ Ey_i = 0 & i = 1,2,......,k-1 \\ E^T = E \end{cases} \quad \text{.. (8)}$$

Now the solution of the problem

$$\min \|E\|_F \quad \text{.. (9)}$$

Subject to (8) is

$$E = \frac{\hat{v}\hat{v}^T}{\hat{v}^T \hat{y}}\left(1 + \frac{\hat{y}^T H_k \hat{y}}{\hat{v}^T \hat{y}}\right) - \frac{\hat{v}\hat{v}^T H_k + H_k \hat{y}\hat{v}^T}{\hat{v}^T \hat{y}} \quad \text{......................................(10)}$$

Where $\| \ \|_F$ is the Frobenius norm and

$$\hat{y}_k = y_k - \sum_{i=1}^{k-1} \frac{\hat{v}_i^T y_k}{\hat{v}_i^T \hat{y}_i} \hat{y}_i \quad \text{.. (11)}$$

$$\hat{v}_k = v_k - \sum_{i=1}^{k-1} \frac{\hat{v}_i^T v_k}{\hat{v}_i^T \hat{y}_i} \hat{v}_i \quad \text{.. (12)}$$

Clearly this is a projected analogue of the BFGS update.

Algorithm 1:-Projected BFGS algorithm:

Step 1: Let k = 1 k_0= 1,H_1 = I

Step 2: d_k =-H_k g_k , line search along d_k to get $x_{k+1} = x_k + \lambda_k d_k$

Step 3: If at x_{k+1} the stopping criterion $\|g_{k+1}\| \le 10^{-5} \max\{1, \|x_{k+1}\|\}$ is satisfied, then terminate.

Step 4: From $\hat{v}_k, \hat{y}_k$ using no more than m +1 past vector pairs v_i , y_i k-m<i<k, if m>n set m=n

Step 5: If $\hat{v}_k^T \hat{y}_k < 0$ then k_o= k+l, $\hat{v}_k = v_k$, $\hat{y}_k = y_k$

Step 6: Form H_{k+1} by (10).

Step 7: Update H_{k+1} with BFGS formula

Step 8: Set k=k+1 , go to step 2.

PRECONDITION CG-METHODS

In CG methods if we have a quadratic function

$$f(x) = 1/2(x - x_{\min})^T G(x - x_{\min}) \qquad (1)$$

with gradient g(x) and given point x_l we set k=1 and d_1=-g_1= -g(x_1)and then iterate on the following steps

$$x_{k+1} = x_k + \lambda_k d_k \qquad (2a)$$

$$d_{k+1} = -g_k + \beta_k d_k \qquad (2b)$$

Here λ_k is determined by an exact line search. Now it is well known that this algorithm has quadratic termination property, but if d, is not equal to -g, the termination property is usually lost .However, a modification of the algorithm (2) if equation (4.2) can regain termination for other choices of d_1, (see Buckley, 1978) one choice a positive definite matrix H such that

$$H = LL^T \qquad (3)$$

where L is a real lower triangular and non-singular matrix and define the transformation of variables x in X-space to a new vector space Z-space as follows

$$x = L_Z \qquad (4)$$

and then the CG-algorithm is applied in Z- space. Of course, all of the standard properties of the CG-algorithm then hold in the Z-space, so in particular finite termination still obtains. In this work we have use preconditioned PR-direction which is

$$d_{k+1} = -H_k g_{k+1} + \frac{g_{k+1}^T H_k (g_{k+1} - g_k)}{g_k^T g_k} d_k \qquad (5)$$

Here in the preconditioned PR-algorithm does start with a direction different from the steepest descent direction -g_1 which is d_1= -Hg, , now the preconditioned PR-algorithm can be outlined as:

Algorithm 2:- (PCG-algorithm):

Step 1: Defined initial matrix H_1=1,and initial point x_l ,for k=1,2,3,...... iterate .

Step 2: Set $d_k = -H_K g_k$ line search along d_k to get $x_k = x_k + \lambda_k d_k$

Step 3: If at x_{k+1}, the stopping criterion $\|g_{k+1}\| \le 10^{-5}$ is satisfied, then terminate.

Step 4: Compute $H_{k+1} = H_k + Q_k$ where Q_k is a correction matrix.

Step5:Check if criterion a general restarting is satisfied then set: $d_{k+1} = -H_k g_k + \beta_k d_k$ where β_k is a scalar defined by $\beta_k^{PR} = \frac{g_{k+1}^T H_k (g_{k+1} + g_k)}{g_k^T g_k}$

else set $d_{k+1} = -H_k g_{k+1}$ and continue.

HYBRID CONJUGATE GRADIENT METHOD

Despite the numerical superiority of PR-method over FR-method the later has better theoretical properties than the former Under certain conditions FR-method can be shown to have global convergence with exact line search (Powell, 1983) and also with inexact line search satisfying the strong Wolfe-Powell condition (Al-Baali, 1985). This anomaly leads to speculation on the best way to choose β_k .

Touati-Ahmed and Storey in 1990 proposed the following hybrid method :

Step 1: If $\lambda \|g_{k+1}\|^2 \le (2\mu)^{k+1}$; with $^1/2 > \mu > \sigma$ and $\lambda > 0$ go to step 2.

Otherwise set $\beta_k=0$.

Step 2 : If $\beta_k^{PR} < 0$ set $\beta_k = \beta_k^{FR}$ Otherwise go to step 3.

Step 3 : If $\beta_k^{PR} \le (1/2\mu)\|g_{k+1}\|^2 / \|g_k\|^2$. With $\mu > \sigma$, set $\beta_k = \beta_k^{PR}$

Otherwise set $\beta_k = \beta_k^{PR}$.

Here μ,σ and λ, user supplied parameters. This hybrid was shown to be globally convergent under both exact and inexact line searches and to be quite competitive with PR-method and FR-method. .

The rate this proposed method was investigated by Hu and Storey in 1991 proved the following result concerning global convergence of conjugate gradient methods.

Suppose that f is twice continuously differentiable with bounded level sets and that the string Wolfe-Powell conditions are satisfied. Supposed that β_k, for all k , satisfies the following conditions :

There exist constants c > 0 and $\sigma \in (0,1/2)$ such that

$$\sigma \left|\beta_k / \beta_k^{FR}\right| \le ck$$

$$\zeta = \sum_{j=1}^{k}\left[\prod_{i=j}^{k-1} \gamma_i\right]$$

$$\gamma_i = \left|\beta_k / \beta_k^{FR}\right|^2$$

It then follows that $\lim_{x \to \infty} \inf \|g(x_k)\| = 0$.

Notice that this result allows β_k to be negative,(see Touati-Ahmed and Storey, 1990) .

Algorithm 3:-New hybrid algorithm:

Step l: Let k=l, $k_0=1, H_1 = I$

Step 2: $d_k = -H_k g_k$, line search along d_k; to get $x_{k+1} = x_k \lambda_k d_k$

Step 3: If at x_{k+1} the stopping criterion $\|g_{k+1}\| \le 10^{-5}$ is satisfied, then terminate.

Step 4: From $\hat{v}_k$, $\hat{y}_k$, using no more than m + 1 past vector pairs v_i , y_i ,

. k-m<_i<-k, if m>n set m=n

Step 5: If $\hat{v}_k^T \hat{y} < 0$ then k_0= k+l , $\hat{v}_k = v_k$, $\hat{y}_k = y_k$

Step 6: Form H_{k+1} by (18).

Step 7: Update H_{k-1} with BFGS formula

Step 8: Check if the following switching criterion

$\left|d_k^T y_k\right| < 0.0015 \|y_k\| . \|d_k\|$ is satisfied , then go to step 9.

else set $d_{k+1} = -H_k g_{k+1}$ and go to step (2).

Step 9: If $\lambda \|g_{k+1}\|^2 \le (2\mu)^{k+1}$, with$^1/2 > \mu > \sigma$ and $\lambda > 0$ go to step 10.

otherwise set $\beta_k=0$.

Step 10: If $\beta_k^{PR} < 0$ set $\beta_k = \beta_k^{FR}$. otherwise go to step 11.

Step 11: If $\beta_k^{PR} \le (1/2\mu)\|g_{k+1}\|^2 / \|g_k\|^2$. With $\mu > \sigma$, set $\beta_k = \beta_k^{PR}$

otherwise set $\beta_k = \beta_k^{PR}$.

Step 12:Compute $d_{k+1} = - H_k g_{k+1} + \beta_k d_k$

Step 13: Set k=k+l go to step 2.

NUMERICAL COMPARISON

Forty test functions were tested with a different dimensions ($4 \le n \le 1000$).Our programs were written in Fortran 90 language and for all cases the stopping criterion was taken to be

$$\|g_{k+1}\| \le 1\times10^{-5}$$

The line search routine used was a cubic interpolation which uses function and gradient values and it is an adaptation of the routine published by Bundey (1984).

The following two algorithms were tested in tables (l) and (2).The first (Storey) corresponds to the projected BFGS-algorithm, the second is the new proposed hybrid algorithm.

The comparative performances for this two algorithms were evaluated by considering both the total number of iterations (NOI) and the total number of function evaluations (NOF), calls quoted required to reduce the value of f(x) below $1x10^{-5}$.

We used in the new hybrid algorithm Dixon. switching criterion defined by

$$|d_k^T y_k| < 0.0015\|y_k\|.\|d_k\|$$

to keep the holding of the quadratic termination property .

In both algorithms and tables we take m=3 where m is the number of past vectors pairs of v; and y;.

In table (1) we have implemented test functions in low and medium dimensions ($4 \le n \le 80$). The new algorithm beats Storey's algorithm in 20 over 20 cases in both NOI and NOF. If we take the Storey's algorithm as 100% NOI and NOF , there is an improvement
of about 62% NOI and 64% NOF in the new hybrid algorithm.

Table (2) contains the result for the high size test functions ($l00 \le n \le 1000$), we have observed that the new hybrid algorithm beats Storey's algorithm in 20 over 20 cases for both NOI and NOF. taking Storey's algorithm as 100% NOI and NOF , we have an improvement of about 24% NOI-and also 24% NOF. Numerical experiences show that the best case obtained numerically is when $m \le 3$.

Table 1: Comparison between Storey's method and new proposed hybrid method for $4 \le n \le 80$

Test Function	**dimension**	**Storey's method**		**Hybrid method**	
		NOI	**NOF**	**NOI**	**NOF**
Powell	4	20	76	12	65
SUM	4	7	22	1	5
Wolfe	4	7	19	1	7
Wood	4	67	236	10	45
Powell	20	33	105	17 ~	66
Dixon	20	23	65	16	33
SUM	20	102	397	96	346
Wood	20	276	849	31	93
Tri	40	37	90	31	63
Miele	40	30	102	25	86
Powell	40	41	118	31	107
Wood	I 40	I 213	I 660	I 49	I 134
Wood	60	247	741	63	158
Miele	60	40	137	25	85
Central	60	22	108	19	84
Tri	60	50	121	44	89
Dixon	80	29	82	23	47
Miele	80	41	141	25	87
Wood	80	320	958	66	160
Tri	80	61	152	55	I11
Total		1666	5179	640	1871
NOI		%100		%38	
NOF		%100		%38	

Table 2: Comparison between Storey's method and new proposed hybrid method for $100\leq n\leq 1000$

Test Function	**dimension**	**Storey's method**		***Hybrid method***	
		NOI	**NOF**	**NOI**	**NOF**
Miele	100	43	147	36	120
Central	100	22	108	19	84
Dixon	100	29	82	23	47
Tri	100	71	161	65	131
Wolfe	200	75	153	69	143
Tri	200	109	219	103	207
Powel	200	41	114	32 ~	110
Dixon	200	29	81	24	49
Miele	400	40	138	40	126
Dixon	400	29	84	24	49
Powell	400	42	123 I	33	115
Wolfe	400	82	167	32	95
Dixon	800	29	83	24	49
Wolfe	800	88	177	30	97
Central	800	21	105	20	89
Powell	800	43	117	21	53
Dixon	1000	29	82	24	49
Miele	1000	44	148	40	132
Wolfe	1000	95	191	75	156
Powell	1000	50	144	37	111
Total		1011	2624	771	2012
NOI		%100		%76	
NOF		%100		%76	

APPENDIX

Test Functions

1. *Generalized powell function* :

$$f(x)=\sum_{i=1}^{n/4}(x_{4i-3}-10x_{4i-2})^2+5(x_{4i-1}-x_{4i})^2+(x_{4i-1}-2x_{4i})^2+10(x_{4i-9}-x_{4i})^4+(x_{4i-2}-2x_{4i-1}-x_{4i})^2)$$

Starting point : $(3,1,0,1,......................)^T$

2. *Generalized Wood function* :

$$f(x)=\sum_{i=1}^{n/4}100(x_{4i-2}+x_{4i-3}^2)^2+(1-x_{4i-3})^2+90(x_{4i}-x_{4i-1}^2)^2+(1-x_{4i-1}^2)^2+1.0$$

Starting point : $(-3,-1,-3,-1,...............)^T$

3. Generalized Miele Function:

$$f(x)=\sum_{i=1}^{n/4}\left[\exp(x_{4i-3}-x_{4i-1})^2+100(x_{4i-2}-x_{4i-1})^6+(\tan(x_{4i-1}-x_{4i}))^4+x_{4i-3}^8+(x_{4i}-1)^2\right]\quad X_0=$$

$(1,2,2,2,\ldots)^T$

4. Generalized Cantrel Function:

$$F(X)=\sum_{i=1}^{\frac{n}{4}}\ [(\exp(X_{4i-3})-X_{4i-2})^4+100(X_{4i-2}-X_{4i-1})^6+(\arctan(X_{4i-1}-X_{4i}))^4+X_{4i-}$$

5. *Dixon function* :

$$f(x)=(1-x_1)^2+(1-x_0)^2+\sum_{i=2}^{9}(x_i^2-x_{i-1})^2$$

Starting point : $(-1,......................)^T$

6. *Welfe function* :

$$f(x)=(-x_1(3-x_1/2)+2x_2-1)^2+\sum_{i=1}^{n-1}(x_{i-1}-x_i(3-x_i(3-x_i/2)+2x_{i+1}-1)^2+(x_{n+1}-x_n(3x_n/2-1)^2\quad 7.$$

TRI Function :

$$f(x)=\sum_{i=1}^{n}\left(ix_i^2\right)^2 \qquad x_0=(-1,...)^T$$

REFERENCES

Al- Baali, M., 1985. Descent Property and Global Convergence of the Fletcher Reeves Method with Inexact Line search. IMA Journal of Numerical Analysis, Vo1.5,pp.121-124.

Al- Bayati and Al- Assadi, 1988. Technical Report" Research Report Series No.89-25, School of Computer studies.(Leeds University).U.K.

Beale, E.M.L., 1988. Introduction to Optimization. Wiley Inter Science Series in Discrete Mathematics and Optimization.

Buckley, A., 1987. Extending The Relationship Between The ConjugateGradient and BFGS Algorithms. Math. Prog.15.1978, pp.343-348.

Bundey, B.D., 1984. Basic Optimization Method. Edward Arnold, Bedford Square London.

Crowder, H.P. and Woif, P., 1972. Linear Convergence of the Conjugated. Gradient Method. IBMJ. Res. Develop. 16, pp.431-433.

Dixon, L.C.W., 1972. Non-Linear Optimization. English Universities Press, London.

Fletcher, R. and Reeves, C.M., 1964. Function Minimization by ConjugateGradients. Computer Journal, 7, pp.149-54.

Fletcher, R., 1987. Practical Methods of Optimization. John Wiley and Sons, Chichester, New York, Brisbane, Toronto, and Singapore. Hu, Y. F. and Storey, C. JOTA,71, pp.399-405.

Huang, H.Y., 1970. Unified Approach to Quadratically Convergent Algorithms for Function Minimization. JOTA, 5, pp.405-423.

Nazareth, L., 1979. A Relationship Between BFGS and Conjugate Gradient Algorithms and Its Implementations for New Algorithms. SIAM Journal Numerical Analysis, 16, pp.794-800.

Polak, E., 1969. Computational Methods in Optimization a Unified Approach. Academic Press, New York.

Powell, M.J.B., Dampt Cambridge, Report NA 14 (1983).

Storey, C. and HU, Y.F., 1990. Projected Quasi-Newton Updates. Mathematics Report Number A 135, September, 1990. TOUATI-AHMED and STOREY, C. 1990. Some Recent Research On Conjugate Gradient Methods for Unconstrained Optimization. J. Fac. Sc. U. A. E. Univ. Vol. 10, pp.123-131.