IOSR Journal of Engineering (IOSRJEN) **www.iosrjen.org**
ISSN (e): 2250-3021, ISSN (p): 2278-8719
Vol. 04, Issue 05 (May. 2014), ||V7|| PP 38-43

New Homotopy Conjugate Gradient for Unconstrained Optimization using Hestenes- Stiefel and Conjugate Descent

Salah Gazi Sharee, Hussein Ageel Khatab
Dept. of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region-Iraq

Abstract: - In this paper, we suggest a hybrid conjugate gradient method for unconstrained optimization by using homotopy formula , We calculate the parameter β_k as aconvex combination of β^{HS} (Hestenes Stiefel)[5] and β^{CD} (Conjugate descent)[3].

Keywords: *- Unconstrained optimization, line search, conjugate gradient method. homotopy formula.*

I. INTRODACTION

Suppose that f : $R^n \to R$ is continuously differentiable function whose gradient is denoted by $g(x)$. Consider the nonlinear following unconstrained optimization problem

$$\text{Min } f(x) \quad x \in R^n \tag{1.1}$$

The iterates for solving (1.1)are given by

$$x_{k+1} = x_k + \propto_k d_k \quad ,k=0,1,\ldots..,n \tag{1.2}$$

Where $\propto_k$is a positive size obtained by line search and d_k is a search direction. The search direction at the very first iteration is the steepest descent $d_0 = -g_0$, the directions along the iterations are computed according to;

$$d_{k+1} = -g_{k+1} + \beta_k d_K \,, \quad k \geq 0 \tag{1.3}$$

where $\beta_k \in R$ is known as conjugate gradient coefficient and different β_k will yield different conjugate gradient methods. Some well known formulas are given as follows:

$$\beta_k^{PR} = \frac{g_k^T(g_k - g_{k-1})}{||g_{k-1}||^2} \tag{1.4}$$

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{||g_k||^2} \tag{1.5}$$

$$\beta_k^{HS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{(g_{k+1} - g_k)^T d_k} \tag{1.6}$$

$$\beta_k^{DY} = \frac{g_{k+1}^T g_{k+1}}{(g_{k+1} - g_k)^T d_k} \tag{1.7}$$

$$\beta_k^{CD} = \frac{g_{k+1}^T g_{k+1}}{-d_k^T g_k} \tag{1.8}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{-d_k^T g_k} \tag{1.9}$$

Where g_{k+1} and g_k are gradients $\nabla f(x_{k+1})$ and $\nabla f(x_k)$ of $\nabla f(x)$ at the point $x_{k+1}$and x_k, respectively , $\|.\|$ denotes the Euclidian norm of vectors. The line search in the conjugate gradient algorithms often is based on the standard Wolfe conditions:

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{1.10}$$

$$g(x_k + \alpha_k d_k)^T d_k \geq \sigma g_k^T d_k \tag{1.11}$$

where 0 <δ<σ< 1.

The strong Wolfe line search corresponds to: that

$$f(x_k + \alpha_k d_k) - f(x_k) \leq \delta \alpha_k g_k^T d_k \tag{1.12}$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leq |\sigma g_k^T d_k| \tag{1.13}$$

where 0 <δ<σ< 1.[10]

II. HYBRID CONJUGATE GRADIENT ALGORITHMS

The hybrid conjugate gradient algorithms are combinations of different conjugate gradient algorithms. They are mainly purposed in order to avoid the jamming phenomenon.[1],these methods are an important class of conjugate gradient algorithms.[6],[9]

The methods of (FR)[4],(DY) [2] and (CD)[3] have strong convergence properties, but they may have modest practical performance due to jamming.
On the other hand, the methods of (PR)[8],(HS)[5] and (LS)[7] may not always be convergent, but the often have better computational performances.[1]

III. NEW HYBRID SUGGESTION

our suggestion generates iterates x_0 , x_1, x_2,… computed by means of the recurrence ($x_{k+1} = x_k + \propto_k d_k$), where the stepsize $\propto_k > 0$ is determined according to the Wolf line search condition (1.10) and (1.11), and the directions d_k are generated by the rule :

$$d_{k+1} = -g_{k+1} + \beta_k^{NEW} d_k \tag{3.1}$$

Where

$$\beta_k^{NEW} = (1-\theta_k)\beta_k^{HS} + \theta_k\beta_k^{CD} \quad , 0 \le \theta \le 1$$

$$\beta_k^{NEW} = (1-\theta_k)\frac{(g_{k+1}^T y_k)}{d_k^T y_k} - \theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} \tag{3.2}$$

Observe that , if $\theta_k = 0$, then $\beta_k^{NEW} = \beta^{HS}$, if $\theta_k = 1$, then $\beta_k^{NEW} = \beta^{CD}$,
On the other hand , if $0 < \theta_k < 1$, then we can find β_k^{NEW} as follows:
We know that

$$d_{k+1} = -g_{k+1} + (1-\theta_k)\frac{(g_{k+1}^T y_k)}{d_k^T y_k} d_k - \theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} d_k \tag{3.3}$$

Our motivation is to choose the parameter θ_k in such a way so that the direction d_{k+1} given (3.3) to be the Newton direction. Therefore

$$-\nabla^2 f(x_{k+1})^{-1} g_{k+1} = -g_{k+1} + (1-\theta_k)\frac{(g_{k+1}^T y_k)}{d_k^T y_k} d_k - \theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} d_k$$

Multiply both sides of above equation by $d_k^T \nabla^2 f(x_{k+1})$, we get

$$-d_k^T g_{k+1} = -d_k^T \nabla^2 f(x_{k+1}) g_{k+1} + (1-\theta_k) d_k^T \nabla^2 f(x_{k+1}) \frac{(g_{k+1}^T y_k)}{d_k^T y_k} d_k$$
$$-\theta_k d_k^T \nabla^2 f(x_{k+1}) \frac{\|g_{k+1}\|^2}{d_k^T g_k} d_k$$

$$-d_k^T g_{k+1} = -d_k^T \nabla^2 f(x_{k+1}) g_{k+1} + (1-\theta_k) \frac{(g_{k+1}^T y_k)}{d_k^T y_k} (d_k^T \nabla^2 f(x_{k+1}) d_k)$$
$$-\theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} (d_k^T \nabla^2 f(x_{k+1}) d_k)$$

Since $d_k^T \nabla^2 f(x_{k+1}) = y_k$, then we have

$$-d_k^T g_{k+1} = -y_k^T g_{k+1} + (1-\theta_k)(g_{k+1}^T y_k) - \theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} y_k^T d_k$$

$$-d_k^T g_{k+1} = -\theta_k (g_{k+1}^T y_k) - \theta_k \frac{\|g_{k+1}\|^2}{d_k^T g_k} y_k^T d_k$$

Implies that

$$\theta_k = \frac{d_k^T g_{k+1}}{(g_{k+1}^T y_k) + \frac{\|g_{k+1}\|^2}{d_k^T g_k} y_k^T d_k}$$

Or

$$\theta_k = \frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)} \tag{3.4}$$

3.1. Convergence of the new hybrid conjugate gradient algorithm

Theorem 3.1.1 : Assume that $\mathrm{d_k}$ is a descent direction and $\propto_k$ in algorithm (1.2) and (3.2) where θ_k is given by (2.4) ,is determined by the wolfe line search (1.10)and (1.11). If $0 < \theta < 1$, then the direction $\mathrm{d_{k+1}}$ given by (3.3) is a descent direction.

Proof:-

From (3.3) and (3.4) we have

$$d_{k+1} = -g_{k+1} + \left(1 - \frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)}{d_k^T y_k} d_k - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2}{d_k^T g_k} \quad (3.1.1)$$

Multiply both sides of (3.1.1) by g_{k+1}^T , we get

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \left(1 - \frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2 (g_{k+1}^T d_k)}{d_k^T g_k} \quad (3.1.2)$$

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2 (g_{k+1}^T d_k)}{d_k^T g_k} \quad (3.1.3)$$

The prove is complete if the step length $\propto_k$ is chosen by an exact line search which requires $d_k^T g_{k+1} = 0$.
Now, if the step length $\propto_k$ is chosen by an inexact line search which requires $d_k^T g_{k+1} \neq 0$,
We know that the first two terms of equation (3.1.3) are less than or equal to zero because the algorithm of Hestenes – Stiefel (HS) is satisfies the descent condition (i.e)

$$-\|g_{k+1}\|^2 + \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} \leq 0, \quad (3.1.4)$$

It remains to consider the third and fourth terms

$$-\frac{(d_k^T g_{k+1})(d_k^T g_k)(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{\left((g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)\right) d_k^T y_k} - \frac{(d_k^T g_{k+1})(d_k^T g_k)\|g_{k+1}\|^2 (g_{k+1}^T d_k)}{\left((g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)\right) d_k^T g_k} =$$

$$-\frac{(d_k^T g_{k+1})^2 (d_k^T g_k)(g_{k+1}^T y_k)}{\left((g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)\right) d_k^T y_k} - \frac{(d_k^T g_{k+1})^2 (d_k^T g_k)\|g_{k+1}\|^2}{\left((g_{k+1}^T y_k)(d_k^T g_k)^2 + \|g_{k+1}\|^2 (y_k^T d_k)(d_k^T g_k)\right)} =$$

$$\frac{-(d_k^T g_{k+1})^2}{d_k^T y_k} \left(\frac{(d_k^T g_k)(g_{k+1}^T y_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)} + \frac{(d_k^T g_k)\|g_{k+1}\|^2}{\frac{(g_{k+1}^T y_k)(d_k^T g_k)^2}{d_k^T y_k} + \|g_{k+1}\|^2 (d_k^T g_k)} \right)$$

$$= \frac{-(d_k^T g_{k+1})^2}{d_k^T y_k} \quad (3.1.5)$$

We know that $\left(d_k^T g_{k+1}\right)^2$ is greater than or equal to zero and
$d_k^T y_k > 0$. Consequently, we have

$$\frac{-\left(d_k^T g_{k+1}\right)^2}{d_k^T y_k} \leq 0$$

Implies that

$$g_{k+1}^T d_{k+1} \leq 0.$$

Then the proof is completed.■

Theorem 3.1.2 :- Assume that the conditions in theorem (3.1.1) hold and $\frac{(y_k^T g_{k+1})(d_k^T g_{k+1})}{y_k^T d_k} \leq \|g_{k+1}\|^2$. If there exists a constant $c_1 > 0$, such that $g_{k+1}^T d_{k+1} \leq -c_1 \|g_{k+1}\|^2$, then the direction d_{k+1} satisfies the sufficient descent condition.

Proof:

From (3.3) and (3.4) we have

$$d_{k+1} = -g_{k+1} + \left(1 - \frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_k\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)}{d_k^T y_k} d_k - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_k\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2}{d_k^T g_k} d_k \quad (4.2.1)$$

Multiply both sides of (4.2.1) by g_{k+1}^T we get

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \left(1 - \frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2 (g_{k+1}^T d_k)}{d_k^T g_k} \quad (3.1.6)$$

$$g_{k+1}^T d_{k+1} \leq -\|g_{k+1}\|^2 + \|g_{k+1}\|^2 - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{(g_{k+1}^T y_k)(g_{k+1}^T d_k)}{d_k^T y_k} - \left(\frac{(d_k^T g_{k+1})(d_k^T g_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)}\right) \frac{\|g_{k+1}\|^2 (g_{k+1}^T d_k)}{d_k^T g_k} \quad (3.1.7)$$

$$g_{k+1}^T d_{k+1} \leq \frac{-(d_k^T g_{k+1})^2}{d_k^T y_k} \left(\frac{(d_k^T g_k)(g_{k+1}^T y_k)}{(g_{k+1}^T y_k)(d_k^T g_k) + \|g_{k+1}\|^2 (y_k^T d_k)} + \frac{(d_k^T g_k)\|g_{k+1}\|^2}{\frac{(g_{k+1}^T y_k)(d_k^T g_k)^2}{d_k^T y_k} + \|g_{k+1}\|^2 (d_k^T g_k)} \right)$$

After some operations, we get

$$g_{k+1}^T d_{k+1} \leq \frac{-(d_k^T g_{k+1})(d_k^T g_{k+1})}{(d_k^T y_k)} \quad (3.1.8)$$

Multiply and divided right hand side of above inequality by $y_k^T g_{k+1}$, we get

$$g_{k+1}^T d_{k+1} \leq \frac{-(d_k^T g_{k+1})(d_k^T g_{k+1})(y_k^T g_{k+1})}{(y_k^T g_{k+1})(d_k^T y_k)} \quad (3.1.9)$$

By hypothesis , (3.1.9) gives

$$g_{k+1}^T d_{k+1} \leq \frac{-(d_k^T g_{k+1})}{(y_k^T g_{k+1})} \|g_{k+1}\|^2$$

Multiply and divided right hand side by $y_k^T g_{k+1}$ *we get*

$$g_{k+1}^T d_{k+1} \leq \frac{-(d_k^T g_{k+1})(y_k^T g_{k+1})}{(y_k^T g_{k+1})^2} \|g_{k+1}\|^2$$

Since $d_k^T g_{k+1} < d_k^T y_k$, then

$$g_{k+1}^T d_{k+1} \leq -\frac{(d_k^T y_k)(y_k^T g_{k+1})}{(y_k^T g_{k+1})^2} \|g_{k+1}\|^2 \quad (3.1.10)$$

Now, if $y_k^T g_{k+1} > 0$,

Let $c_1 = \frac{(d_k^T y_k)(y_k^T g_{k+1})}{(y_k^T g_{k+1})^2}$

Then ,(3.1.10) gives

$$g_{k+1}^T d_{k+1} \leq -c_1 \|g_{k+1}\|^2$$

If $y_k^T g_{k+1} < 0$ and we know that $d_k^T y_k > 0$, then,

$$(d_k^T y_k)(y_k^T g_{k+1}) < d_k^T y_k$$

,

Then, (3.1.10) gives

$$g_{k+1}^T d_{k+1} \leq -\frac{d_k^T y_k}{(y_k^T g_{k+1})^2} \|g_{k+1}\|^2$$

Let $c_1 = \frac{d_k^T y_k}{(y_k^T g_{k+1})^2}$

Hence

$$g_{k+1}^T d_{k+1} \leq -c_1 \|g_{k+1}\|^2$$

Then the proof is completed.■

3.2 Theorem of global convergence

Since the new hybrid conjugate gradient algorithm is satisfies the sufficient descent condition by using wolfe conditions , then the new hybrid conjugate gradient algorithm is satisfies the global convergence property **.**

3.3 Algorithm of New Hybrid Conjugate Gradient algorithm

step (1) :- set k=0 , select the initial point x_k .

step(2) :- $g_k = \nabla f(x_k)$, If $g_k = 0$,then stop .

else

set $d_k = - g_k$.

step (3) :- compute $\alpha_k > 0$ satisfying the wolfe lline search condition to mini mize $f(x_{k+1})$.

step (4) :- $x_{k+1} = x_k + \alpha_k d_k$.

step (5) :- $g_{k+1} = \nabla f(x_{k+1})$, If $g_{k+1} = 0$,then stop .

step (6):- compute θ_k as in (3.4).

Step (7):-if $0 < \theta_k < 1$, then compute β_k^{NEW} as in (3.2). If $\theta_k \geq 1$, then set $\beta_k^{NEW} = \beta^{CD}$. If $\theta_k \leq 0$, then set $\beta_k^{NEW} = \beta^{HS}$.

step (8) :- $d_{k+1} = - g_{k+1} + \beta_k^{NEW} d_k$.

step (9) :- If k=n then go to step 2,

else

k=k+1 and go to step 3.

3.4 NUMRICAL RESULTS:-

This section is devoted to test the implementation of the new formula . We compare the hybrid algorithm with standard Hestenes – Stiefel (HS) and conjugate direction (CD) ,the comparative tests involve well-known nonlinear problems (standard test function) with different dimension 4 ≤n≤5000, all programs are written in FORTRAN95 language and for all cases the stopping condition is $\|g_{k+1}\|_\infty \leq 10^{-5}$ The results are given in below table is specifically quote the number of functions NOF and the number of iteration NOI .experimental results in below table confirm that the new CG method is superior to standard CG method with respect to the NOI and NOF.

Table Comparative Performance of the three algorithms (Standard HS,CD and New formula)

Test fun.	N	Standard formula (HS)		Standard formula (CD)		New formula	
		NOI	NOF	NOI	NOF	NOI	NOF
Powell	4	65	170	994	2077	30	74
	100	105	276	3102	6942	108	242
	500	502	1062	134002	270117	502	1011
	1000	637	1332	*	*	241	532
	3000	879	1816	*	*	249	568
	5000	1008	2074	*	*	409	913
Wood	4	26	59	353	709	26	61
	100	27	61	928	2115	26	61
	500	28	63	1008	2277	27	6
	1000	28	63	561	1165	27	63
	3000	28	63	500	1038	27	63
	5000	28	63	517	1111	27	63
Cubic	4	15	43	540	1104	11	34
	100	14	40	801	2060	11	32
	500	14	40	1501	5287	11	32
	1000	14	40	*	*	11	32
	3000	14	40	*	*	11	32
	5000	14	40	*	*	11	32
Rosen	4	23	66	611	1262	23	64
	100	17	52	461	1374	21	62
	500	*	*	2063	6612	21	62
	1000	*	*	4036	13523	21	62
	3000	*	*	*	*	21	62
	5000	*	*	*	*	21	62

Mile	4	28	101	*	*	18	50
	100	142	346	*	*	149	355
	500	501	1108	*	*	501	1092
	1000	1001	2312	*	*	998	2290
	3000	1442	3252	*	*	1270	2834
	5000	1660	3688	*	*	1418	3130
Non Digonal	4	23	61	249	558	23	59
	100	22	60	*	*	18	51
	500	22	59	*	*	18	53
	1000	22	59	*	*	18	53
	3000	22	59	*	*	19	55
	5000	22	59	*	*	19	55
Woolf	4	12	27	49	99	12	27
	100	49	99	901	1879	49	99
	500	56	113	3893	8393	55	111
	1000	70	141	8001	19675	69	139
	3000	166	343	*	*	166	343
	5000	176	365	*	*	175	363

IV. CONCLUSION

In this paper we have presented a new hybrid conjugate gradient method in which a famous parameter β_k is computed as a convex combination of β_k^{HS} and β_k^{CD} and comparative numerical performances of a number of well known conjugate gradient algorithms Hestenes Stiefel (HS) and Conjugate decsent (CD). We saw that the performance profile of our method was higher than those of the well established conjugate gradient algorithms HS and CD.

REFRENCES

[1] N. Andrei, Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization , *J.Optim. Theory Appl., 141*, 2009, 249-264.

[2] Y.H. Dai, and Y. Yuan, A nonlinear Conjugate Gradient Method With a Strong Global Convergence Property, *SIAM J. Optimization , 10*,1999, Pp. 177-182.

[3] R. Fletcher, Practical Methods of Optimization,(Vol. 1: Unconstrined Optimization , John Wiley and Sons, New York, 1987).

[4] R. Fletcher, and C. Reeves, Function Minimization by Conjugate Gradients, *Comput. J.,7*, 1964 , Pp. 149-154.

[5] M.R. Hestenes, and E.L. Stiefel, Methods of Conjugate Gradients for Solving Linear Systems, *J.Research Nat. Bur. Standards ,49* ,1952, Pp. 409-436.

[6] Y.F. Hu and C. Storey, Global Convergence Result for Conjugate Gradient Methods, *J. Optim. , Theory Appl., 71*,1991, Pp. 399-405.

[7] Y. Liu , and C. Storey, Efficient generalized Conjugate Gradient Algorithms , *Part 1: Theory. JOTA,69*,1991, Pp.129-137.

[8] E. Polak , andG. Ribiere, Note sur Ia convergence de directions Conjugate , *Rev. Francaise Information, Recherche Operationelle, 3e Annee, 16*,1969, Pp. 35-43.

[9] D. Touati-Ahmed and C. Storey, Efficient Hybrid Conjugate Gradient Techniques , *J. Optim. , Theory Appl., 64*,1990, Pp.379-397.

[10] P. Wolfe, Convergence conditions for ascent methods, *SIAM Rev., 11*,1969, 226-235.