

A RATIONAL LOGARITHMIC MODEL FOR UNCONSTRAINED NON-LINEAR OPTIMIZATION

By

Huda K. Mohammed

Nidhal H. Al Assady

Department of Mathematics

College of Science

Mosul University

نموذج جديد لحل مسائل الأمثلية اللاخطية وغير المقيدة

هدى محمد و نضال الأسدي

قسم الرياضيات، كلية العلوم، جامعة الموصل - العراق

يتركز هذا البحث على اقتراح نموذج جديد أكثر عمومية من النماذج التربيعية لحل مسائل الأمثلية اللاخطية وغير المقيدة والذي يحور ويطور خوارزميات التدرج المترافق الكلاسيكية وله نفس خواص خوارزميات التدرج المترافق المطبقية على دالة تربيعية النموذج الجديد المستخدم في خوارزميات التدرج المترافق الكلاسيكية هو النموذج اللوغاريتمي الجديد والذي تم اشتقاقه وحله عددياً لاستخدام بعض الدوال المختارة لاختيار كفاءة الخوارزمية الجديدة، وبصورة عامة النتائج تشير الى ان الخوارزمية الجديدة هي تحسين للخوارزميات السابقة.

Key Words : Logarithmic Model, Non-Linear Optimization.

ABSTRACT

A new non-quadratic model is proposed for solving unconstrained optimisation problems which modifies and develops the classical conjugate gradient methods. The technique has the same properties as the classical conjugate gradient method that can be applied to a quadratic function. An algorithm is derived and evaluated numerically for some standard test functions. The results indicate that in general the new algorithm is an improvement on the previous methods .

INTRODUCTION

A more general model than the quadratic one is proposed in this paper as a basis for a conjugate algorithm. If $q(x)$ is a quadratic function, then a function f is defined as a nonlinear scaling of $q(x)$ if the following condition holds:

$$f = F(q(x)), dF/dq = F' > 0 \text{ and } q(x) > 0 \tag{1}$$

where x^* is the minimizer of $q(x)$ with respect to x , Spedicato [1].

The following properties are immediately derived from the above condition:

- i) every contour line of $q(x)$ is a contour line of f ,
- ii) if x^* is a minimizer of $q(x)$, then it is a minimizer of f ,
- iii) that x^* is a global minimum of $q(x)$ does not necessarily mean that it is a global minimum of f .

Various authors have published related work in this area (see Al Assady, et. al., [2], Al Assady and Al Bayati, [3], Al Bayati, [4] and Hu et. at., [5]).

A conjugate gradient method which minimizers the function $f(x) = (q(x))^p$, $p > 0$ and $x \in R^n$ in at most n step had been described by Fried [6] and the special case,

$$F(q(x)) = \epsilon_1 q(x) + 1/2 \epsilon_2 q^2(x).$$

where ϵ_1 and ϵ_2 are scalars, has been investigated by Boland et. al [7], [8].

Tassopoulos and Storey [9] and [10] have proposed a specific model, vis, a rational model that is denoted by T/S. It is as follows:

$$F(q(x)) = (\epsilon_1 q(x) + 1) / \epsilon_2 q(x), \epsilon_2 < 0.$$

where ϵ_1 and ϵ_2 are scalars and $q(x) = 1/2 x^T G x + b^T x + c$ is a quadratic function.

In this paper a new logarithmic model is investigated and tested on a set of standard test functions, on the assumption that condition (1) holds. An extended conjugate gradient algorithm is developed which is based on this new model which scales $q(x)$ by the natural log function for the rational $q(x)$ functions.

$$F(q(x)) = \log [\epsilon_1 q(x) / (\epsilon_2 q(x) + 1)], \epsilon_2 < 0. \tag{2}$$

We first observe that $q(x)$ and $F(q(x))$ given by (2) have identical contours, though with different function values, and they have the same unique minimum point denoted by x^* .

For any f satisfying the condition (1) it is shown in [7],

[8] that the updating process given below generates identical conjugate directions and the same sequence of approximations x_i to the minimiser x^* , as does the original method of Fletcher-Reeves [11] (F/R) when applied to

$$f(x) = q(x).$$

In order to modify the property (iii) in the following way:

“that x^* is a global minimizer of $q(x)$ implies that it is a global minimiser of f ”, we have suggested a new logarithmic model defined in equation (2) based on the Renpu Ge’s Theorem, [12] which is illustrated below:

Suppose $F(x)$ has the form

$$F(x) = f_1(x_1) / g_2(x_2) \tag{3}$$

where $x^T = (x_1^T, x_2^T)$ and

$$f_1(x_1) > 0 \text{ and } g_2(x_2) > 0 \tag{4}$$

it follows from equations (3) and (4) that:

$$\log F(x) = \log f_1(x_1) / -\log g_2(x_2) \tag{5}$$

is a separable function. Thus according to the following theorem which states:

Theorem (1): $x^{*T} = (x_1^{*T}, x_2^{*T}, \dots, x_n^{*T})$ is a global minimizer of a separable function $F(x)$ if and only if every x_i^* , ($i = 1, 2, \dots, n$) is a global minimizer of $f_i(x_i)$.

Proof: See, [12].

We can conclude that x^* is a global minimizer of $\log F(x)$ if and only if x_1^* and x_2^* are respectively global minimizers of $\log f_1(x_1)$ and $-\log g_2(x_2)$. Furthermore, the monotonicity of $\log t$ implies that x^* is a global minimizer of $F(x)$ if and only if x_1^* and x_2^* are respectively global minimizers of $f_1(x_1)$ and $g_2(x_2)$.

The Algorithm:

Given $x_0 \in R^n$ an initial estimate of the minimizer x^*

Step (1): Set $d_0 = -g_0$.

Step (2): For $i = 1, 2 \dots$

$$\text{compute } x_i = x_{i-1} + \lambda_{i-1} d_{i-1}$$

where λ_{i-1} is the optimal step size obtained by the line search procedure.

Step (3): Calculate $\exp(f) = 1 + f + \frac{f^2}{2!} + \frac{f^3}{3!} + \dots$

and define $n = n + \lambda_{i-1} g_{i-1}^T d_{i-1} / 2$

$$w = \exp(f_i) - \exp(f_{i-1})$$

$$c = w - n \exp(f_{i-1})$$

Step (4): If $|w| \leq 0.1 \text{ E-}5$ or $|c| \leq 0.1 \text{ E-}5$, then set $\rho_i = 1.0$ and go to step (6).

Else go to step (5).

Step (5): Compute

$$\rho_i = [\exp(f_i) / \exp(f_{i-1})] [n \exp(f_{i-1}) / w]^2$$

where the derivation of scaling ρ_i will be presented below.

Step (6): Calculate the new direction

$$d_i = -g_i + \beta_i d_{i-1}$$

where β_i is defined by different formulae according to variation and it is expressed as follows:

$$\beta_i = \rho_i (\|g_i\|^2 / \|g_{i-1}\|^2) \text{ (modified Fletcher and Reeves [11])}$$

$$\beta_i = g_i^T (\rho_i g_i - g_{i-1}) / d_{i-1}^T (g_i - g_{i-1}) \text{ (modified Hestenes and Stiefle [13])}$$

$$\beta_i = g_i^T (\rho_i g_i - g_{i-1}) / g_{i-1}^T (g_i - g_{i-1}) \text{ (modified Polak and Ribiera [14])}$$

Conjugate gradient methods are usually implemented by restarts in order to avoid an accumulation of errors affecting the search directions. It is therefore generally agreed that restarting is very helpful in practice, so we have used the following restarting criterion in our practical investigations.

If the new direction satisfies:

$$d_i^T g_i \geq -0.8 \|g_i\|^2 \tag{6}$$

then a restart is also initiated. This new direction is sufficiently downhill.

The Derivation of π for the New Model:

The implementation of the extended conjugate gradient method has been performed for general functions $F(q(x))$ of the form of equations (2). The unknown quantities ρ_i were expressed in terms of available quantities of the algorithm (i.e function and gradient values of the objective function). It is first assumed that neither ϵ_2 , nor ϵ_1 is zero in equation (2). Solving equation (2) for $q(x)$, then:

$$q = -\exp(f) / \epsilon_2, [\exp(f) - \epsilon_1 / \epsilon_2] \tag{7}$$

and using the expression for ρ_i (henceforth $\rho_i \text{ NEW}$)

$$\rho_i \text{ NEW} = f'_{i+1} / f'_i$$

$$= [\exp(f_i) / \exp(f_{i-1})] [(\exp(f_{i-1}) - \epsilon_1 / \epsilon_2) / \exp(f_i) - \epsilon_1 / \epsilon_2)^2] \tag{8}$$

the quantity which has to be determined explicitly is ϵ_1 / ϵ_2

During every iteration ϵ_1 / ϵ_2 must be evaluated as a function of known available quantities

From the relation:

$$g_i = f_i G(x_i - x^*) \tag{9}$$

$$g_{i-1} = f'_{i-1} G(x_{i-1} - x^*) \tag{10}$$

where G is the Hessian matrix and x^* is the minimum point, we have:

$$\rho_i \text{ NEW} = f'_{i-1} / f'_i$$

$$= (g_{i-1}^T / g_i^T) (x_{i-1} - x^*) (x_i - x^*) \tag{11}$$

Furthermore,

$$g_{i-1}^T (x_{i-1} - x^*) = g_{i-1}^T + (x_{i-1} + \lambda_{i-1} d_{i-1} - x^*)$$

$$= g_{i-1}^T (x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1}$$

and

$$g_i^T (x_{i-1} - x^*) = g_i^T (x_i - \lambda_{i-1} d_{i-1} - x^*)$$

$$= g_i^T (x_i - x^*)$$

Since $g_i^T d_{i-1} = 0$

Therefore we can express $\rho_i \text{ NEW}$ as follows:

$$\rho_i \text{ NEW} = (g_{i-1}^T (x_{i-1} - x^*) + \lambda_{i-1} g_{i-1}^T d_{i-1}) / (g_i^T (x_i - x^*)) \tag{12}$$

From (9) and (10), we get:

$$\rho_i \text{ NEW} = (f_{i-1} (x_{i-1} - x^*) T G (x_{i-1} - x^*) + \lambda_{i-1} f_{i-1} g_{i-1}^T d_{i-1}) / (f_i (x_i - x^*) T G (x_i - x^*))$$

Therefore,

$$\rho_i \text{ NEW} = (2f'_{i-1} g_{i-1} + \lambda_{i-1} g_{i-1}^T d_{i-1}) / 2f'_i q_i$$

$$= \rho_i \text{ NEW} (q_{i-1} / q_i) + \lambda_{i-1} g_{i-1}^T d_{i-1} / 2f'_i q_i \tag{13}$$

The quantities (q_{i-1} / q_i) and $f'_i q_i$ can be rewritten as:

$$(q_{i-1} / q_i) = (1 / \sqrt{\rho_i \text{ NEW}}) [\sqrt{\exp(f_{i-1}) / \exp(f_i)}] \tag{14}$$

$$f'_i q_i = 1 / q_i (\epsilon_2 q_i + 1) (q_i)$$

$$= 1 / (\epsilon_2 q_i + 1)$$

$$= -\epsilon_2 (\exp(f_i) - \epsilon_1 / \epsilon_2) / \epsilon_1 \tag{15}$$

Substituting (14) and (15) in (13), gives:

$$\rho_i \text{ NEW} = \sqrt{\rho_i \text{ NEW}} [\sqrt{\exp(f_{i-1}) / \exp(f_i)}]$$

$$(\lambda_{i-1} g_{i-1}^T d_{i-1} / 2) (\epsilon_1 / \epsilon_2) / [\exp(f_i) - \epsilon_1 / \epsilon_2] \tag{16}$$

Using the transformation : $-\lambda_{i-1} gT_{i-1} d_{i-1} = 2n$

From (8) and (16), it follows that:

$$[\exp(f_i) / \exp(f_{i-1})] \{[\exp(f_{i-1}) - \epsilon_1 / \epsilon_2] / [\exp(f_i) - \epsilon_1 / \epsilon_2]\}^2 =$$

$$[\exp(f_{i-1}) / \exp(f_i)] \{[\exp(f_{i-1}) - \epsilon_1 / \epsilon_2] / [\exp(f_i) - \epsilon_1 / \epsilon_2]\} - n(\epsilon_1 / \epsilon_2) / [\exp(f_i) - \epsilon_1 / \epsilon_2] \quad (17)$$

$$[\exp(f_{i-1}) - \epsilon_1 / \epsilon_2]^2 =$$

$$[\exp(f_{i-1}) / \exp(f_i)] \{[\exp(f_{i-1}) - \epsilon_1 / \epsilon_2][\exp(f_i) - \epsilon_1 / \epsilon_2] - n(\epsilon_1 / \epsilon_2) [\exp(f_{i-1}) / [\exp(f_i)] [\exp(f_i) - \epsilon_1 / \epsilon_2]\} \quad (18)$$

By solving the equation (18), we have:

$$\epsilon_1 / \epsilon_2 = \exp(f_{i-1}) -$$

$$n[\exp(f_i) - \exp(f_{i-1})] \exp(f_{i-1}) / [\exp(f_i) - \exp(f_{i-1}) - n \exp(f_{i-1})]$$

Using the following transformations; $\exp(f_i) - \exp(f_{i-1}) = w$

The above equation can be rewritten as:

$$\epsilon_1 / \epsilon_2 \exp(f_{i-1}) - [n w \exp(f_{i-1})] / [w - n \exp(f_{i-1})] \quad (19)$$

and substituting equation (19) in equation (8), we get :

$$\rho_i \text{ NEW} = [\exp(f_i) / \exp(f_{i-1})] [n \exp(f_{i-1}) / w]^2$$

Defining ρ_i and assuming that f_i 'is positive for the new model, ρ_i must be positive

$$\rho_i = \left| [\exp(f_i) / \exp(f_{i-1})] [n \exp(f_{i-1}) / w]^2 \right|$$

The Numerical Experiments:

In order to test the effectiveness of the new algorithm (henceforth NEW) that have been used to extend the conjugate gradient methods, a number of functions has been chosen and solved numerically by utilizing the new and established methods.

The same line search was employed for all the methods. This was the Cubic interpolation procedure described in [14].

After the results of the NEW method being obtained, they were compared with the previous published results achieved by some scholars specialized in this field such as T/S and TTR (see [15]), methods.

It is found that the NEW method which modifies CG-algorithms is better than the previous published method such as the T/S and TTR methods as shown in Tables (1) and (2) below.

As shown from the following series that has been chosen for the test:

$$\exp f = 1 + f + f^2/2i + f^3/3i + \dots \quad (20)$$

the more terms are used the better the results will be. Consequently, we selected the program which uses three

and six terms out of the series (20) for the computations.

Table (1), which uses the F/R formula, gives a comparison between the results of the NEW method, that utilizes three terms out of the series (20), and the results of the previous published T/S method. This table shows that the NEW method has less Fes (i.e number of function evaluations (NOF) plus n times the number of gradient evaluations) than the T/S method. Thus, we can see from this computation in this table that the NEW method is better than the T/S method for F/R formula.

Table-1 The comparison between the different ECG-methods by using the F/R formula.

Test-Functions	N dimension	NEW (3-terms) Fes	T/S Fes
Cubic	2	144	162
Rosen	2	219	201
Wood	4	305	530
Wood	20	2247	2331
Wood	200	28140	70149
Wood	400	56140	184460
Powell	4	600	595
Powell	80	19359	23895
Powell	200	93063	113597
Miele	4	890	2340
Miele	20	2457	5649
Miele	80	19521	33120
Total	Fes	223085	437028

It is observed that from this table above that the NEW method has less Fes than the T/S method. So the NEW is better than the T/S method for F/R formula. Moreover, in this table (1) above, we have considered Fes as a criterion for the comparison between the algorithms as the published results in Tassopoulous and Storey [9].

Table (2), which uses the P/R formula, presents a comparison between the results of the previous published TTR method and the NEW method, that utilizes three terms out of the series (20). This table shows that the NEW method has less (NOI) and (NOF) than the TTR method. So the NEW method is also better than the T/S method for P/R formula.

Table -2 The comparison between the different ECG-methods by using the P/R formula.

Test-Functions	N dimension	NEW (3-terms) NOI (NOF)	T/S Fes NOI (NOF)
Cubic	2	17 (48)	18 (58)
Nrosen	20	24(59)	36(107)
OSP	50	24(79)	30(113)
Dixon	10	22(46)	58(125)
Rosen	2	31(73)	15(43)
Rosen	100	26(62)	34(106)
Miele	4	74(225)	26(76)
Miele	20	25(63)	39(108)
Powell	4	77(183)	66(198)
Wood	4	23(49)	38(93)
Wood	200	32(66)	35(82)
Total	NOI (NOF)	375(953)	395(1109)

Table-3 The comparison between the different ECG-methods by using the H/S formula.

Test-Functions	N dimension	NEW (6-terms) NOI (NOF)	Classical CG NOI (NOF)
Cubic	2	17 (48)	17 (48)
Nrosen	20	24(61)	24(61)
OSP	50	29(71)	24(79)
Dixon	10	21(44)	22(46)
Wood	4	36(75)	28(61)
Wood	20	34(70)	52(107)
Wood	100	76(154)	69(140)
Wood	200	48(98)	69(140)
Wood	400	34(270)	69(140)
Powell	4	51(109)	69(140)
Powell	40	54(120)	50(114)
Powell	80	95(194)	72(158)
Powell	200	215(444)	112(239)
Rosen	2	31 (73)	229(463)
Rosen	80	19 (46)	31(73)
Rosen	100	20(49)	23(56)
Miele	4	57(178)	57(178)
Miele	20	42(105)	46(114)
Miele	80	101(233)	102(241)
Fun1	10	12(31)	11(31)
Fun2	10	3(9)	3(9)
Fun3	10	2(14)	2(14)
Fun3	30	2(15)	2(15)
Total	NOI(NO F)	1123(2511)	1137(2583)

Furthermore, to show that when we increase the number of terms of the series (20) in the NEW method the best the results will be, we have used the NEW method, that utilises six terms out of the series (20) instead of the three terms. Also we have used the H/S formula since this formula is the best formula among the others.

This procedure is calculated as follows:

Suppose that,

the total results of the previous algorithm = A

and the total results of the new algorithm = B

Then the procedure is $((A-B) / A \times 100)$

Table (3), which uses the H/S formula, presents a comparison between the results of the NEW method, that utilizes six terms out of the series (20) and the classical CG-method. So we can show that the NEW method (6-terms) has less (NOI)

APPENDIX

1- Cubic Function:

$$F(x) = 100(x_2 - x_1^3)^2 + (1 - x_1)^2, x_0 = (-1.2, 1.0)^T.$$

2- Dixon Function:

$$F(x) = (1 - x_1)^2 + (1 - x_{10})^2 + \sum_{i=1}^9 (x_i - x_{i+1}),$$

$$x_0 = (-1.0; \dots)^T$$

3- Non-Diagonal Variant of Rosenbrock Function:

$$F(x) = \sum_{i=2}^n [100 (x_i - x_i^2)^2 + (1 - x_i)^2]; n > 1,$$

$$x_0 = (-1.0; \dots)^T$$

4- OSP Oren and Spedicato Powell Function:

$$F(x) = \left[\sum_{i=1}^n i x_i^2 \right]^2, x_0 = (1.0; \dots)^T$$

5- Wood Function:

$$F(x) = \sum_{i=1}^{n/4} 100 \left[\begin{array}{l} (x_{4i-2} + x_{4i-3}^2)^2 + (1 - x_{4i-3})^2 + 90(x_{4i} - x_{4i-1}^2)^2 \\ + (1 - x_{4i-1})^2 + (10.1(x_{4i-2} - 1))^2 + (x_{4i} - 1)^2 \\ + 19.8 (x_{4i-2} - 1) (x_{4i} - 1) \end{array} \right],$$

$$x_0 = (-3.0; -1.0; -3.0; 1.0; \dots)^T$$

6- Generalised Powell Quartics Functions:

$$F(x) = \sum_{i=1}^{n/4} \left[\begin{array}{l} (x_{4i-3} + 10x_{4i-2})^2 + 5(x_{4i-1} - x_{4i})^2 \\ + (x_{4i-2} - 2x_{4i-1})^4 + 10(x_{4i-3} - x_{4i})^4 \end{array} \right],$$

$$x_0 = (3.0; -1.0; 0.0; 1.0)^T$$

7- Rosenbrock Function:

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

$$x_0 = (-1.2; 1.0; \dots)^T$$

8- Miele Function:

$$F(x) = \sum_{i=1}^{n/4} [\exp(x_{4i-3}) - x_{4i-2}]^2 + 100 (x_{4i-2} - x_{4i-1})^6$$

$$+ [\tan(x_{4i-1} - x_{4i})]^4 + x_{4i-3}^8 + (x_{4i-1})^2,$$

$$x_0 = (1.0; 2.0; 2.0; 2.0; \dots)^T$$

9. Fun1. Function:

$$F(x) = \sum_{i=1}^{30} f^2(x)_i,$$

$$f_i(x) = 420 x_i + (i-15)^3 + \sum_{j=1, j \neq i}^{30} (x_j^2 + i/j)^{(1/2)}$$

$$[\sin^5 \log(x_j^2 + i/j)^{(1/2)} + \cos^5 \log(x_j^2 + i/j)^{(1/2)}]$$

$$x = -2.8742711 \sum_{i=1}^{30} e_i f_i(0)$$

10. Fun2. Function:

$$F(x) = \sum_{i=1}^{30} \left[y_i - \sum_{j=1}^{30} (a_{ij} \sin x_i + b_{ij} \cos x_i) - \right]^2$$

$$y_i - \sum_{j=1}^{30} (a_{ij} \sin \zeta_j + b_{ij} \cos \zeta_j)$$

a_{ij}, b_{ij} – random coefficients uniformly distributed within the interval (-100, +100)

ζ, δ – random coefficients uniformly distributed within the interval $(-\pi, +\pi)$

$$x = \zeta + 0.1 \delta$$

11. Fun3 Function:

$$F(x) + 1 - \exp(-1/60 \sum_{i=1}^{30} x_i^2)$$

$$x = \sum_{i=1}^{30} (-1)^i (1+i/30) e_i$$

REFERENCES

- [1] Spedicato, E., 1976. A Variable Metric Method for Function Minimization Derived From Invariancy to Non-linear Scaling, *JOTA*, 20.
- [2] Al-Assady, N. H.; Dawood, A. I. and Aziz, M.M., 1993 Generalized CG. Based on Quasi-Quadratic Model, *J. Ed. and Sc.*, NO. 13, pp 210-218.
- [3] Al-Assady, N. H.. and Al-Bayati, A. Y., 1994. Minimization of Extended Quadratic Functions with Inexact Line searches, *JOTA*, V, 82, No.1, pp 139-147.
- [4] Al Bayati, A. Y., 1992. A New Non-Quadratic Model for Unconstrained Non-Linear Optimization, *Mu'tah Journal for Research and Studies, Natural and Applied Sciences Series*, Jordan.
- [5] Hu, Y.F., Khoda; K.M.; Liu, Y; Storey, C; and Touati-Ahmed, C., 1994. Some Recent Research on Conjugate Gradient Methods for Unconstrained Optimization, *TR, N*, A217.
- [6] Fried, I., 1971. N-step Conjugate Gradient Minimization Scheme for Non-Quadratic Functions, *AIAA*, J.,9.
- [7] Boland, W. R. and Kowalik, J. S., 1979. Extended Conjugate Gradient Methods with Restarts, *Journal of Optimization Theory and Applications* 28, pp. 1-9.
- [8] Boland, W. R; Kamgnia, E. R and Kowalik, J. S., 1979. A conjugate Gradient Method Invariant to Non-Linear Scaling, *Journal of Optimization Theory and Applications* 27, pp. 221-230.
- [9] Tassopoulos, A. and Storey, C., 1984. A Conjugate Direction Method Based on a Non-Quadratic Model, *JOTA*, 43.
- [10] Tassopoulos, A. and Storey, C., 1984. Use of a Non-Quadratic Model in a Conjugate Gradient Searches, *JOTA*, 43.
- [11] Fletcher, R. and Reeves, C. M., 1964. Function Minimization by Conjugate Gradients", *Computer Journal*, 7.
- [12] Renpu Ge., 1989. A Parallel Global Optimization Algorithm for Rational Separable Factorable Function, *Journal of Applied Mathematics and Computation*, Vol. 32, No.1, pp 61-72.
- [13] Hestenes, M.R. and Stiefle, E., 1952. Methods of Conjugate Gradients for Solving Linear Systems, *J. Res. N.B.S.*, 49.
- [14] Bunday, B., 1984. *Basic Optimization Methods*, Edward Arnold, Bedford Square, London.
- [14] Polak, E and Ribiere, G. 1969. Note Sur La Convergence des Methodes de Directions Conjugues, *Rev. Fr. Inf. Rech, Oper.* 16-R1.
- [15] Nazareth.L, 1977. A Conjugate Direction Algorithm Without Line Searches, *JOTA*, 27.