Indonesian Journal of Electrical Engineering and Computer Science
Vol. 27, No. 3, September 2022, pp. 1635~1641
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v27.i3.pp1635-1641

Using a new coefficient conjugate gradient method for solving unconstrained optimization problems

Ranen M. Sulaiman, Basim A. Hassan
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul, Mosul, Iraq

Article Info

Article history:

Received Mar 31, 2022
Revised Jun 14, 2022
Accepted Jul 5, 2022

Keywords:

Conjugate gradient
Descent property
Global converg-ence
Numerical results
Optimization

ABSTRACT

The conjugate gradient technique is a numerical solution strategy for finding minimization in mathematics. We present a simple, straightforward, efficient, and resilient conjugate gradient technique in this study. To address the convergence difficulty and descent property, the new technique is built on the quadratic model. Under some assumptions, the new improved approach meets the convergence characteristics and the adequate descent criterion. The suggested unique strategy is substantially more efficient than the classic FR method, according to our numerical analysis. The number of function evaluations, iterations and restarts are all included in the numerical results. The computational efficiency of the proposed approach is proved by comparative results.

Corresponding Author:

Ranen M. Sulaiman
Department of Mathematics, College of Computers Sciences and Mathematics, University of Mosul
Mosul, 00964, Iraq
Email: ranen.sulyman@uomosul.edu.iq

1. INTRODUCTION

Use the following formula to get the minimum of a continuously differentiable function:

$$\text{Min } F(x),\ x \in R^n \tag{1}$$

the iterative methods we use are iterative approaches of the following form:

$$x_{k+1} = x_k + \lambda_k d_k \tag{2}$$

in (2) shows that various stepsizes λ_k and directions d_k result in distinct approaches, as shown in [1]. For example, in the quadratic case, λ_k is an accurate step size as (3):

$$\lambda_k = -\frac{g_k^T d_k}{d_k^T Q d_k} \tag{3}$$

for further information, see [2]. After that, the step length λ_k is selected to meet the Wolfe conditions, which are as (4) and (5):

$$F_{k+1} \le F_k + \delta \lambda_k g_k^T d_k \tag{4}$$

$$d_k^T g_{k+1} \geq \sigma d_k^T g_k \tag{5}$$

where $0 < \delta < \sigma < 1$, see [3]. The search directions in conjugate gradient algorithms can be specified recursively:

$$d_{k+1} = -g_{k+1} + \beta_k s_k \tag{6}$$

where β_k is selected in a way that d_k and d_{k+1} must satisfy the conjugacy property. To compute the scalar β_k, a number of formulae have been presented. Fletcher and Reeves (FR) [4] and Dai and Yuan (DY) [5] are two well-known formulae. They're provided by:

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}, \beta_k^{DR} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k} \tag{7}$$

The other nonlinear conjugate gradient techniques, for example, are the subject of a lot of study in this area (e.g. see [6]-[9]). Many writers have researched the nonlinear conjugate gradient technique in recent years, particularly from the perspective of global convergence. The nonlinear conjugate gradient technique was often studied independently because its characteristics might vary greatly depending on β_k (see Powell [10]).

We'd want to close the new search direction to the quasi-Newton direction later because of the theoretical usefulness of quasi-Newton approaches. In this situation, we're looking for a parameter that will allow us to:

$$-Q_{k+1}^{-1} g_{k+1} = -g_{k+1} + \beta_k s_k \tag{8}$$

where Q_{k+1} is the Hessian matrix. (See [11], [12]) for a useful resource for research describing the most recent CG coefficients with notable results and numerous β_k adjustments. The approaches are efficient in reality, according to numerical findings, and the methods' convergence guarantees are comparable to the classical variations. Our key contribution is a novel coefficient derivation based on the second-order Taylor's series, which we utilized to build an inverse Hessian matrix for computing the search direction and ensuring global convergence.

2. OUR NEW COEFFICIENT CONJUGATE

A second order Taylor series is used to derive the new coefficient conjugate. Let we clarify:

$$F(x) = F(x_{k+1}) + g_{k+1}^T (x - x_{k+1}) + \frac{1}{2} (x - x_{k+1})^T Q(x_{k+1})(x - x_{k+1}) \tag{9}$$

the $F(x)$ has the following gradient:

$$g_{k+1} = g_k + Q(x_{k+1}) s_k \tag{10}$$

putting (3) in (9) and using ELS search, we get:

$$s_k^T Q(x_{k+1}) s_k = (F_{k+1} - F_k) + 3/2\, \lambda_k g_k^T g_k \tag{11}$$

the yielded matrix $Q(x_{k+1})$ can be as:

$$Q(x_{k+1}) = \frac{(F_{k+1} - F_k) + 3/2 \lambda_k g_k^T g_k}{s_k^T s_k} I_{nxn} \tag{12}$$

putting $Q(x_{k+1})$ in (8) we get:

$$\beta_k = \left(1 - \frac{s_k^T s_k}{(F_{k+1} - F_k) + 3/2 \lambda_k g_k^T g_k}\right) \frac{g_{k+1}^T y_k}{s_k^T y_k} \tag{13}$$

We will do some algebra manipulations on (14), in order to achieve an ideal direction of descent:

$$\beta_k = \frac{1}{s_k^T y_k}\left(y_k - \omega \frac{\|y_k\|^2}{s_k^T y_k} s_k\right)^T g_{k+1} \tag{14}$$

where:

$$\omega = \frac{(s_k^T y_k)}{\|y_k\|^2}\left[\frac{s_k^T y_k}{s_k^T s_k} * \frac{s_k^T s_k}{(F_{k+1} - F_k) + 3/2\lambda_k g_k^T g_k}\right] \tag{15}$$

this is the formale that will be utilized to do the convergence analysis. The application of optimization theory and methods to new formulations is a vast field of applied mathematics. The method described by is denoted by New (13). As follows, we suggest a novel conjugate gradient method.

New Algorithm:
1): Give $x_1 \in R^n$. Set $k = 1$ and $d_1 = -g_1$. If $\|g_1\| \leq 10^{-6}$, then, stop.
2): Evaluate $\lambda_k > 0$ satisfying a (4-5).
3): Let $x_{k+1} = x_k + \lambda_k d_k$. If $\|g_{k+1}\| \leq 10^{-6}$, then come to a halt.
4): Evaluate β_k by the formulae (14) and d_{k+1} by (6).
5): Let $k = k + 1$ and continue with step 2.

Theorem (2.2): Consider the CG technique (2), (4), (5), and the descent direction d_{k+1} provided by (6) with (14) is adequate.

Proof: Since $d_0 = -g_0$ we get $d_0^T g_0 = -\|g_0\|^2 \leq 0$. Suppose that $g_k^T d_k < 0$ for all $k \in n$. To finish the proof, we must prove that the theorem holds for all $k + 1$. Because $Q(x_{k+1})$ is the quadratic model's search direction matrix, we may write it like this:

$$Q(x_{k+1}) = \frac{(F_{k+1} - F_k) + 3/2\lambda_k g_k^T g_k}{s_k^T s_k} I_{nxn} = \omega\, I_{nxn} \tag{16}$$

now, we have to prove that $\omega > 0$. Using Wolfe's condition for determining the value of ω, we have:

$$\omega = \frac{\lambda_k \delta d_k^T g_k + 3/2\lambda_k g_k^T g_k}{s_k^T s_k} \tag{17}$$

since $d_k^T g_k < -c g_k^T g_k$, then, using the equation above, we obtain:

$$\omega = \frac{3/2\lambda_k g_k^T g_k + \lambda_k \delta g_k^T g_k}{s_k^T s_k} \tag{18}$$

because the first half of (18) is greater than the second, we get:

$$\omega > 0 \tag{19}$$

we may describe the search directions of the new approach as follows using (6) and (14) and some algebraic manipulations:

$$d_{k+1} = -Q_{k+1}^{-1} g_{k+1} = -\frac{s_k^T s_k}{(F_{k+1} - F_k) + 3/2\lambda_k g_k^T g_k} g_{k+1} \tag{20}$$

multiplying (20) by g_{k+1}, we have:

$$d_{k+1}^T g_{k+1} = -\frac{s_k^T s_k}{(F_{k+1} - F_k) + 3/2\lambda_k g_k^T g_k}\|g_{k+1}\|^2 = -\omega\|g_{k+1}\|^2 \tag{21}$$

since $\omega > 0$, from (17) we obtained:

$$d_{k+1}^T g_{k+1} = -\omega^{-1}\|g_{k+1}\|^2 \leq -c\|g_{k+1}\|^2 \tag{22}$$

3. CONVERGENCE ANALYSIS

To do this, establish the “global convergence” of new Algorithm is one of the most property of numerical algorithms, the assumptions must be made: The $\Omega = \{x \in R^n / F(x) \leq F(x_1))\}$ is a confined level set. The gradient of function g is Lipschitz continuous in some neighborhood Λ of Ω, i.e., there exists, a constant, $L > 0$ such that:

$$\|g(o) - g(\tau)\| \leq L\|o - \tau\|, \forall o, \tau \in \Lambda \tag{23}$$

for more details see [8]. We show why the Dai *et al.* [13] theorem is crucial for determining global convergence.

Lemma (3.1): Let x_k be produced by (2), d_k satisfy descent property and α_k be satisfy (4-5). If:

$$\sum_{k\geq 0}^{\infty} \frac{1}{\|d_{k+1}\|^2} = \infty \tag{24}$$

then:

$$\lim_{k\to\infty} \inf \|g_{k+1}\| = 0 \tag{25}$$

we began by stating our paper's key theorem.

Theorem (3.2): Assume that $F(x)$ meets Assumptions 1 and 2. Let $\{x_k\}$ be the sequence that (6) generates (14). If Wolfe criteria (4) and (5) are satisfied by step size λ_k, then:

$$\lim_{k\to\infty} \inf \|g_k\| = 0 \tag{26}$$

Proof: Using (14) as an example of β_k in (6), we get:

$$\|d_{k+1}\| = \|-g_{k+1} + \beta_k^{New} d_k\| \leq \|g_{k+1}\| + \left\|\left(y_k - \omega \frac{\|y_k\|^2}{s_k^T y_k} s_k\right)\right\| \frac{\|y_k\|}{\|s_k\|\|y_k\|} \|d_k\| \tag{27}$$

and combining $s_k = \lambda_k d_k$, we get:

$$\|d_{k+1}\| \leq \|g_{k+1}\| + \frac{\|y_k\|\|g_{k+1}\| + \omega\|g_{k+1}\|\|y_k\|}{\lambda_k\|d_k\|\|y_k\|} \|d_k\| \leq \left[\frac{\lambda_k + 1 + \omega}{\lambda_k}\right] \|g_{k+1}\| \tag{28}$$

which results in:

$$\sum_{k\geq 1} \frac{1}{\|d_k\|^2} \geq \left[\frac{\lambda_k}{\lambda_k + 1 + \tau}\right] \frac{1}{T} \sum_{k\geq 1} 1 = \infty \tag{29}$$

we may deduce from Lemma 1 that $\lim_{k\to\infty} \inf \|g_k\| = 0$ is identical to $\lim_{k\to\infty} \|g_k\| = 0$ for a uniformly convex function.

4. NUMERICAL RESULTS

On a series of unconstrained optimization test problems, this section shows the computing efficiency of a Fortran implementation of the novel CG technique and the FR-Algorithm. Readers who are interested can access the papers and references listed below ([14], [15]). The unconstrained concerns in [16] are the test problems. We investigated numerical experiments with 100 and 1000 variables for each test function for 15 large scale unconstrained optimization problems in extended or generalized form. Many papers have proposed this method for optimization problems [17]-[20]. As for the papers, it is concerned with the convergence feature [21]-[24]. As a termination condition, we employ the inequality $\|g_{k+1}\| \leq 10^{-6}$. The $\delta = 0.001$ and $\sigma = 0.9$ were used to evaluate both methods.The numerical findings are reported in Table 1. “The following are the definitions for each column: NI: the total number of iterations, NR: the total number of restart, NF: the total number of evelation functions”. Table 1 shows how many issues these algorithms have solved in terms of iterations (NI), restart (NR) and function evaluations (NF).

Table 2 shows that, overall, the tools score 100 percent for the FR-method, whereas the new approach (New) scores 22 percent NI, 23 percent NR, and 52 percent NF. “The results are shown in Figures 1, 2 and 3. Note that a performance measure introduced by Dolan and More [25], [26] method was employed”.

Table 1. The FR and new methods' numerical results

P. No.	n	FR algorithm NI	NR	NF	New algorithm NI	NR	NF
Extended Rosenbrock	100	47	18	93	43	23	89
	1000	78	45	131	37	21	82
Extended Beale	100	32	15	52	18	11	34
	1000	22	10	42	19	11	35
Penalty	100	10	6	27	12	8	31
	1000	24	16	191	22	13	47
Extended PSC1	100	15	9	31	11	7	23
	1000	8	6	17	8	6	17
Extended Maratos	100	89	32	174	73	44	171
	1000	107	40	211	69	44	157
Extended Q. Penalty	100	32	12	65	24	15	56
	1000	53	22	116	37	20	90
Quadratic QF2	100	130	49	196	117	36	184
	1000	364	119	593	367	106	580
ARWHEAD (CUTE)	100	9	4	18	10	6	18
	1000	12	7	82	9	6	37
NONDIA (CUTE)	100	13	7	25	12	7	23
	1000	15	7	29	12	6	25
Partial Quadratic	100	74	21	123	85	25	134
	1000	370	88	616	329	85	543
Broyden Tridiagonal	100	30	10	49	28	6	49
	1000	34	10	63	38	11	67
EDENSCH (CUTE)	100	69	50	1202	34	17	56
	1000	98	82	1967	42	30	450
LIARWHD (CUTE)	100	23	11	45	24	11	29
	1000	27	11	55	22	16	43
DENSCHNA (CUTE)	100	20	11	33	18	11	30
	1000	19	11	35	19	11	32
DENSCHNC (CUTE)	100	49	22	80	16	10	30
	1000	129	67	166	14	10	27
Total		**2002**	**818**	**6527**	**1569**	**633**	**3189**

Table 2. The Performance percentage for the new algorithm compared with FR method

	NI	NR	NF
FR	.100 %	.100 %	.100 %
New	78.37 %	77.38 %	48.85 %

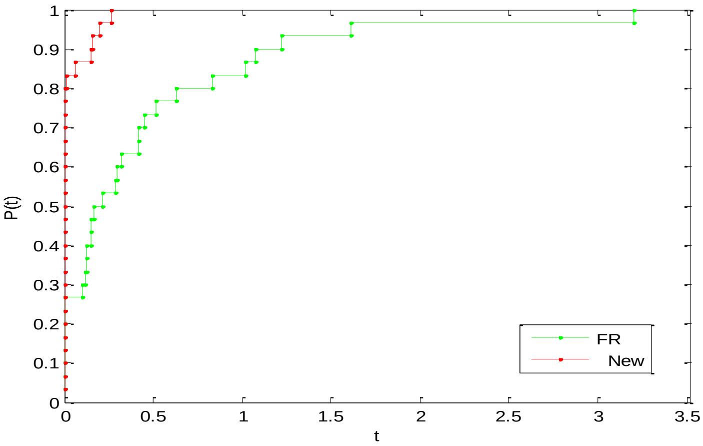

Figure 1. Performance measure based on the NI

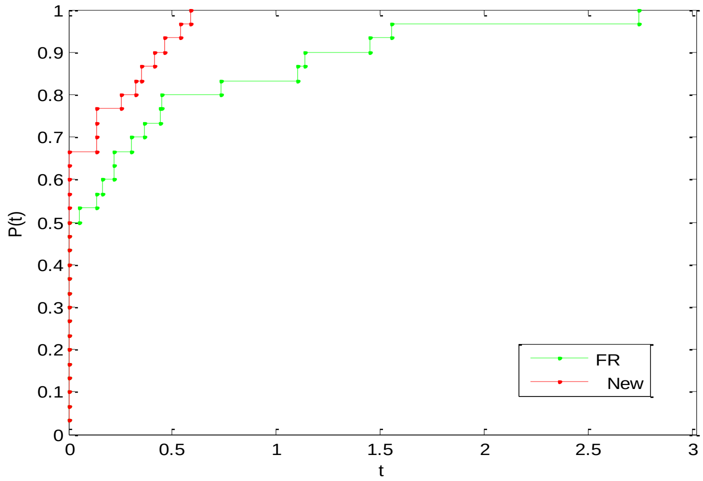

Figure 2. Performance measure based on the NR

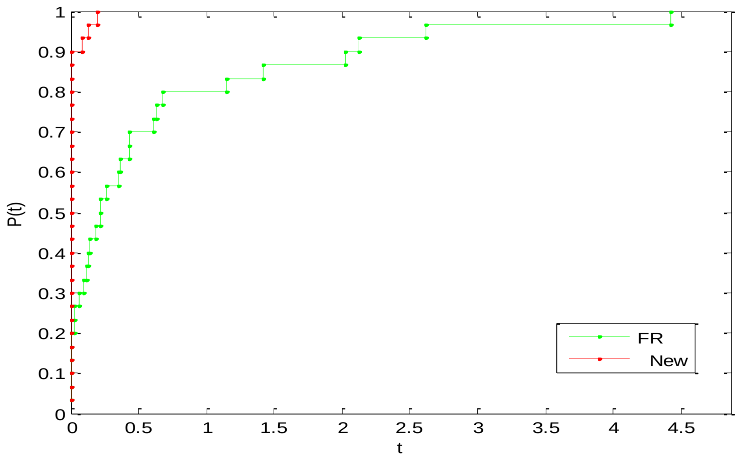

Figure 3. Performance measure based on the NF

5. CONCLUSION

On the basis of the quadratic model, we proposed a unique conjugate gradient technique. The proposed technique satisfies both the descent and convergence requirements. According to the numerical data, the new strategy outperformed the FR-method in terms of the number of iterations, restart, and evaluation functions. In addition, the new method surpasses the previous FR.

REFERENCES

[1] Y. H Dai and L. Z Liao, "New conjugacy conditions and related nonlinear conjugate gradient methods," *Applied Mathematics and Optimization*, vol. 43, no. 1, pp. 87–101, 2001, doi: 10.1007/s002450010019.

[2] J. Nocedal and S. J. Wright, "Sequential quadratic programming,"in *Numerical optimization*. New York: Springer, 1999.

[3] Z. Dai and F. Wen, "Some improved sparse and stable portfolio optimization problems," *Finance Research Letters*, vol. 27, pp. 46–52, 2018, doi: 10.1016/j.frl.2018.02.026.

[4] R. Fletcher and C. M. Reeves, "Function minimization by conjugate gradients," *The Computer Journal*, vol. 7, no. 2, pp. 149–154, 1964, doi: 10.1093/comjnl/7.2.149.

[5] Y.-H. Dai and Y. Yuan, "A nonlinear conjugate gradient method with a strong global convergence property," *SIAM Journal on Optimization*, vol. 10, no. 1, pp. 177–182, 1999, doi: 10.1137/S1052623497318992.

[6] R. F. RF, *Practical Method of Optimization, Unconstrained Optimization*, New York: Wiley, 1987.

[7] M. R. Hestenes and E. Stiefel, "Methods of Conjugate Gradients for Solving," *Journal of Research of the Nat ional Bureau of Standards*, vol. 49, no. 6, pp. 409-436, 1952.

[8] M. R. Hajmohammadi, "Design and analysis of multi-scale annular fins attached to a pin fin," *International Journal of Refrigeration*, vol. 88, pp. 16-23, 2018, doi: 10.1016/j.ijrefrig.2017.11.032.

[9] E. Polak and C. Ribiere, "Note surla convergence de directions conjugates," *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse*, vol. 3, no. 1, pp. 35–43, 1969.

[10] M. J. D. Powell, "Restart procedures for the conjugate gradient method," *Mathematical Programming volume*, vol. 12, no. 1, pp. 241-254, 1977, doi: 10.1007/BF01593790.
[11] B. A. Hassan, H. O. Dahawi, and A. S. Younus, "A new kind of parameter conjugate gradient for unconstrained optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 17, no. 1, pp. 404-411, 2020, doi: 10.11591/ijeecs.v17.i1.pp404-411.
[12] H. N. Jabbar and B. A. Hassan, "Two-versions of descent conjugate gradient methods for large-scale unconstrained optimization," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, no. 3, pp. 1643–1649, 2021, doi: 10.11591/ijeecs.v22.i3.pp1643-1649.
[13] Y. Dai, J. Han, G. Liu, D. Sun, H. Yin, and Y. Yuan, "Convergence properties of nonlinear conjugate gradient methods," *SIAM Journal on Optimization*, vol. 10, no. 2, pp. 345–358, 2000, doi: 10.1137/S1052623494268443.
[14] B. A. Hassan, F. Alfarag, A. Ibrahim, and A. Abubakar, "An improved quasi-Newton equation on the quasi-Newton methods for unconstrained optimizations," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 22, pp. 389-397, doi: 10.11591/ijeecs.v22.i2.pp997-1005.
[15] W.W Hager.and H. Zhang, "A survey of nonlinear conjugate gradient methods," *Pacific journal of Optimization*, vol. 2, no. 1, pp. 35-58, 2006.
[16] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 147-161, 2008.
[17] J. C. Gilbert and J. Nocedal, "Global convergence properties of conjugate gradient methods for optimization," *SIAM Journal on Optimization*, vol. 2, no. 1, pp. 21-42, 1992, doi: 10.1137/0802003.
[18] W. R. Boland, E. R. Kamgnia, and J. S. Kowalik, "A conjugate-gradient optimization method invariant to nonlinear scaling," J *Journal of Optimization Theory and Applications*, vol. 27, no. 2, pp. 221-230, 1979, doi: 10.1007/BF00933228.
[19] J.-F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal, *Numerical optimization: theoretical and practical aspects*, Berlin: Springer, 2006.
[20] Y. Liu and C. Storey, "Efficient generalized conjugate gradient algorithms, part 1: theory," *Journal of Optimization Theory and Applications*, vol. 69, no. 1, pp. 129-137, 1991, doi: 10.1007/BF00940464.
[21] I. M. Sulaiman, N. A. Bakar, M. Mamat, M. Malik, and A. M. Ahmed, "A new hybrid conjugate gradient algorithm for optimization models and its application to regression analysis," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 23, no. 2, pp. 1100-1109, 2021, doi: 10.11591/ijeecs.v23.i2.pp1100-1109.
[22] M. Malik, A. Abubakar, S. Ibrahim, M. Mamat, S. Abas, and S. Firman, "A new three-term conjugate gradient method for unconstrained optimization with applications in portfolio selection and robotic motion control," *IAENG International Journal of Applied Mathematics*, vol. 51, no. 3, pp. 471–486, 2021.
[23] B. A. Hassan and E. S. Al-Rawi, "A modified Newton's method for solving functions of one variable," *Italian Journal of Pure and Applied Mathematics*, pp. 577–582, 2021.
[24] L. Ladislav and V. Jan, "Nonlinear conjugate gradient methods," *Programs and Algorithms of Numerical Mathematics, Proceedings of Seminar*. Prague: Institute of Mathematics AS CR, 2015.
[25] B. A. Hassan and M. A.aq Kahya, "A new class of quasi-Newton updating formulas for unconstrained optimization," *Journal of Interdisciplinary Mathematics*, vol. 24, no. 8, pp. 2355–2366, 2022, doi: 10.1080/09720502.2021.1961980.
[26] E. D. Dolan and J. J. Moré, "Benchmarking optimization software with performance profiles," *Mathematical Programming volume*, vol. 91, no. 2, pp. 201–213, 2002, doi: 10.1007/s101070100263.

BIOGRAPHIES OF AUTHORS

Ranen M. Sulaiman is currently a Professor in Department of Mathematics, College of Computer Science and Mathematics, University of Mosul. He obtained his M.Sc degrees in Mathematics from the University of Mosul, in 2020 and 2021, respectively with specialization in optimization. To date, he has published more than 4 research paper in various international journals and conferences. He currently works on iterative methods. His research interest in applied mathematics, with a field of concentration of optimization include conjugate gradient, steepest descent methods, Broyden's family and quasi- Newton methods. He can be contacted at email: ranen.sulyman@uomosul.edu.iq.

Basim A. Hassan is currently a Professor in Department of Mathematics, College of Computer Science and Mathematics, University of Mosul. He obtained his M.Sc and Ph.D degrees in Mathematics from the University of Mosul, in 2000 and 2010, respectively with specialization in optimization. To date, he has published more than 80 research paper in various international journals and conferences. He currently works on iterative methods. His research interest in applied mathematics, with a field of concentration of optimization include conjugate gradient, steepest descent methods, Broyden's family and quasi- Newton methods with application in signal recovery and image restoration. He can be contacted at email: basimah@uomosul.edu.iq.