7th International Conference on
Combinatorics, Cryptography, Computer Science and Computing
November. 16-17, 2022

Upscaling Parameters for Conjugate Gradient Method in Unconstrained Optimization

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

Hawraz N. Jabbar , Yoksal A. Laylani
Department of Mathematics,
College of Sciences
University of Kirkuk, Iraq

ABSTRACT

Conjugate gradient algorithm is a powerful iterative method and based on the parameters conjugate gradient. Using second order Taylor series to derive an effective parameters conjugate gradient by belonging to the conjugacy condition. We analyze the convergence properties of the new algorithm, then give some numerical results which show the modified algorithms are robust and efficient.

KEYWORDS: Parameters conjugate gradient, Convergence property, Numerical results

1 INTRODUCTION

The majority of optimization techniques are iterative, [12]. We aim to resolve the following mathematical optimization issue for the continuously differentiable objective function $f : R^n \to R$:

$$Min \quad f(x) \ , \ x \in R^n \tag{1}$$

We shall examine many iterative techniques of the type:

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

where α_k is the step length and d_k is the search direction, [10]. Step length is determined by:

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

Using the Wolfe conditions the step length α_k will be accepted:

$$f(x_k + \alpha_k d_k) \le f(x_k) + \delta \alpha_k g_k^T d_k \tag{4}$$

$$d_k^T g(x_k + \alpha_k d_k) \ge \sigma \, d_k^T g_k \tag{5}$$

where $0 < \delta < \sigma < 1$, see [18-19]. Here, the following give the search direction:

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

where β_k is a scalar that is calculated using a different formula depending on the conjugate gradient technique being employed. This approach was expanded by Fletcher and Reeves for use with broader functions [9]. The following is the formula for calculating β_k using Fletcher-Reeves:

$$\beta_k^{FR} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k}. \tag{7}$$

The stepsize and the CG parameter are the two main variables that have an impact on the numerical performance of the CG technique. Some famous classical choices of β_k can be found in [7,8,11,14,16 and 17].

Numerous CG approaches have been put forth recently that have both strong global convergence features and promising numerical performance. These methods are based on the conjugacy requirement of numerous CG parameters (see [3-6]). The conjugacy requirement was used in previous CG algorithms:

$$d_{k+1}^T Q d_k = 0 \quad . \tag{8}$$

that is based on the precise line search is crucial to the mathematical experiment and convergence analysis [15].

In this study, we provide novel conjugate gradient optimization techniques and analyze their convergence. Numerical tests reveal that our approaches can perform better than those already in use.

2 A NEW FORMULAS FOR β_k .

Now we using the second order Taylor series derive the new formulas, we have:

$$f_k = f_{k+1} - g_{k+1}^T s_k + \frac{1}{2} s_k^T Q s_k \tag{9}$$

Using exact line search in (9), we get:

$$s_k^T Q s_k = 2(f_k - f_{k+1}) \tag{10}$$

The conjugacy condition, we get:

$$d_{k+1}^T Q s_k = 0 \tag{11}$$

Applying (11) on (10) and after some algebra, we obtained:

$$\beta_k = \frac{\frac{2(f_k - f_{k+1})}{s_k^T y_k} g_{k+1}^T y_k}{d_k^T y_k} \tag{12}$$

If exact line search is utilized, then β_k is such that:

$$\beta_k^{BK1} = \frac{\frac{2(f_k - f_{k+1})}{s_k^T y_k} \|g_{k+1}\|^2}{d_k^T y_k} \tag{13}$$

and

$$\beta_k^{BK2} = \frac{\frac{2(f_k - f_{k+1})}{s_k^T g_k} \|g_{k+1}\|^2}{d_k^T g_k} \tag{14}$$

and

$$\beta_k^{BK3} = \frac{\frac{2(f_k - f_{k+1})}{\alpha_k g_k^T g_k} \|g_{k+1}\|^2}{g_k^T g_k} \quad . \tag{15}$$

The ways that the aforementioned parameters produced are known as the BK1, BK2, and BK3, respectively.

Next the algorithm of the proposed method is presented as follows:

I : Given $x_0 \in R^n$. Let $d_0 = -g_0$, $k = 0$.

II : If $\|g_{k+1}\| \le 10^{-6}$ yes, then stop; if not, continue.

III : Compute α_k by using (4)-(5).

IV : Let $x_{k+1} = x_k + \alpha_k d_k$, and compute β_k by (13-15).

V : Compute the search direction $d_{k+1} = -g_{k+1} + \beta_k d_k$.

VI : Set $k = k+1$ and continue with step 1.

3 CONVERGENCE ANALYSIS:

We assume that $f(x)$ meets the following conditions in what follows:

i. $f(x)$ is bounded on the set $\Psi = \{x \in R^n : f(x) \le f(x_0)\}$.

ii. g is Lipschitz continuous, i.e. there exits a constant $L > 0$. That is:

$$\|g(z) - g(u)\| \le L \|z - u\| , \ \forall \ z, u \in R^n . \tag{17}$$

With a constant ¥ ≥ 0 such that $\|\nabla f(x)\| \le$ ¥ , see [2, 13].

Lemma 1.1

Assuming that the assumptions are correct, think about any recurrence expression (2) with search direction (3). The Zoutendijk criterion (16) is satisfied in that case [20].

Theorem 3.1.

Let d_{k+1} be a sequence generated by new Algorithms. Then, it holds that:

$$d_{k+1}^T g_{k+1} < 0 \quad \text{and} \quad d_{k+1}^T g_{k+1} = \beta_k d_k^T g_k . \tag{18}$$

Proof :

Using (3), we get:

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \beta_k d_k^T g_{k+1} \tag{19}$$

Equation (14) indicates that,

$$\|g_{k+1}\|^2 = \frac{\beta_k^{BK1} d_k^T y_k}{\dfrac{2(f_k - f_{k+1})}{s_k^T y_k}} . \tag{20}$$

Adding (20) to (19), we got:

$$d_{k+1}^T g_{k+1} = \beta_k^{BK1} [d_k^T g_{k+1} - d_k^T y_k] = \beta_k^{BK1} d_k^T g_k \ < 0 \tag{21}$$

The proof is so finished.

Theorem 3.2.

Assume that the direction d_{k+1} was produced using new algorithms. Then:

$$\lim_{k \to \infty} \inf \|g_k\| = 0 \ . \tag{22}$$

Proof : Similar to the Theorem 2 proof in [7], the proof is straightforward.

4 NUMERICAL RESULTS

This section includes the outcomes of novel techniques on a number of test problems. The codes are created in Fortran and double precision calculations are used. Computers are used to conduct every exam. Our tests are based on a set of 50 nonlinear situations that a transducer is capable of producing. These problems are discussed in Andrei [1] and are part of the CUTE test. The standard stop, s.t. $\|g_{k+1}\| \le 10^{-6}$, is used in all algorithms. The total number of function evaluations (NOF) and the total number of iterations are both taken into account while evaluating the algorithms' performance (NOI). Tables 1and 2 contains the results.

Table 1 : The numerical results of the FR and New methods with n=100.								
	FR		BK1		BK2		BK3	
P.No.	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
Trigonometric	**19**	**35**	**19**	**36**	**16**	**34**	**18**	**34**
Extended Rosenbrock	**47**	**93**	**44**	**90**	**43**	**88**	**34**	**72**
Extended White & Holst	**43**	**88**	**42**	**85**	**41**	**85**	**38**	**89**
Extended Beale	**32**	**52**	**16**	**30**	**13**	**26**	**17**	**32**
Extended Tridiagonal	**32**	**64**	**9**	**19**	**10**	**21**	**15**	**31**
Extended Three Expo Terms	**15**	**25**	**10**	**15**	**14**	**20**	**17**	**27**
Generalized Tridiagonal 2	**37**	**67**	**37**	**59**	**43**	**68**	**45**	**64**
Extended Himmelblau	**12**	**25**	**10**	**19**	**10**	**19**	**10**	**19**
Extended Powell	**189**	**313**	**87**	**163**	**63**	**118**	**87**	**166**
Quad. Diagonal Perturbed	**124**	**231**	**47**	**85**	**45**	**81**	**36**	**101**
Extended Wood	**71**	**110**	**29**	**54**	**25**	**47**	**33**	**60**
ARWHEAD	**9**	**18**	**8**	**16**	**8**	**16**	**10**	**19**
NONDIA	**13**	**25**	**13**	**26**	**11**	**21**	**11**	**22**
DIXMAANE	**121**	**218**	**82**	**129**	**81**	**124**	**75**	**122**
Partial Perturbed Quadratic	**74**	**123**	**77**	**120**	**74**	**113**	**77**	**126**
LIARWHD	**23**	**45**	**19**	**34**	**20**	**38**	**20**	**36**
ENGVAL1	**34**	**57**	**24**	**46**	**22**	**42**	**24**	**47**
DENSCHNA	**20**	**33**	**11**	**20**	**10**	**19**	**13**	**23**
DENSCHNC	**49**	**80**	**19**	**33**	**13**	**24**	**18**	**31**
Extended Block-Diagonal	**122**	**156**	**11**	**21**	**11**	**20**	**12**	**23**
Generalized quartic GQ1	**11**	**24**	**7**	**16**	**7**	**16**	**8**	**18**
Generalized quartic GQ2	**112**	**147**	**34**	**56**	**35**	**53**	**39**	**61**
STAIRCASE S1	**671**	**1066**	**503**	**816**	**504**	**793**	**462**	**747**
TRIDIA	**398**	**605**	**292**	**468**	**322**	**496**	**345**	**548**
Extended PSC1	**15**	**31**	**8**	**17**	**8**	**17**	**8**	**17**
Total	**2293**	**3731**	**1005**	**2473**	**1449**	**2399**	**1472**	**2535**

Table 2 : The numerical results of the FR and New methods with n=1000.								
	FR		BK1		BK2		BK3	
P.No.	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**	**NI**	**NF**
Trigonometric	**38**	**65**	**37**	**67**	**36**	**66**	**40**	**69**
Extended Rosenbrock	**78**	**171**	**37**	**83**	**40**	**89**	**42**	**92**
Extended White & Holst	**46**	**92**	**41**	**90**	**41**	**84**	**44**	**94**
Extended Beale	**22**	**42**	**12**	**24**	**12**	**23**	**15**	**29**
Extended Tridiagonal	**77**	**129**	**13**	**27**	**13**	**26**	**18**	**36**
Extended Three Expo Terms	**fail**	**fail**	**61**	**1447**	**19**	**32**	**10**	**17**
Generalized Tridiagonal 2	**73**	**115**	**52**	**84**	**49**	**79**	**56**	**89**
Extended Himmelblau	**14**	**29**	**11**	**21**	**10**	**19**	**10**	**19**
Extended Powell	**fail**	**fail**	**97**	**185**	**70**	**133**	**84**	**156**
Quad. Diagonal Perturbed	**445**	**711**	**162**	**288**	**149**	**260**	**186**	**340**
Extended Wood	**47**	**84**	**26**	**52**	**25**	**49**	**26**	**50**
ARWHEAD	**12**	**82**	**7**	**34**	**8**	**55**	**8**	**16**
NONDIA	**15**	**29**	**13**	**25**	**14**	**29**	**16**	**33**
DIXMAANE	**345**	**634**	**237**	**376**	**234**	**357**	**256**	**403**
Partial Perturbed Quadratic	**370**	**616**	**253**	**428**	**274**	**457**	**272**	**463**
LIARWHD	**27**	**55**	**21**	**44**	**22**	**48**	**18**	**39**
ENGVAL1	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**
DENSCHNA	**19**	**35**	**9**	**18**	**18**	**30**	**9**	**18**
DENSCHNC	**129**	**166**	**13**	**29**	**13**	**26**	**12**	**24**
Extended Block-Diagonal	**130**	**166**	**13**	**25**	**12**	**23**	**10**	**20**
Generalized quartic GQ1	**9**	**22**	**7**	**18**	**7**	**18**	**7**	**18**
Generalized quartic GQ2	**110**	**145**	**35**	**54**	**31**	**56**	**34**	**57**
STAIRCASE S1	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**	**fail**
TRIDIA	**1255**	**1991**	**fail**	**fail**	**1279**	**1996**	**1026**	**1657**
Extended PSC1	**8**	**17**	**7**	**15**	**7**	**15**	**7**	**15**
Total	**2014**	**3405**	**792**	**1802**	**1015**	**1809**	**855**	**1924**

According to this data, the BK1, BK2 and BK3 approaches have a clear advantage over the FR technique since they save around (38-42)% in NOF but only (42-59)% in NOI.

Table 3 : The Performance Percentage for the BK1,BK2 and BK3 algorithms compared with FR method				
	FR algorithm	**BK1 (13)**	**BK2 (14)**	**BK3 (15)**
NI	**100 %**	**41.72%**	**58.02%**	**54.02%**
NF	**100 %**	**59.90%**	**58.96%**	**62.48%**

5 CONCLUSION

Finally, we presented the conjugate gradient methods BK1, BK2, and BK3, as well as a novel modified conjugate gradient formula. Wolfe line search settings allowed us to identify its worldwide convergence. Simulations have demonstrated that BK1, BK2, and BK3 may decrease function evaluations and iterations. While the conventional secant relation only employs gradient values, the modified conjugate gradient approach uses both gradient and function values.

REFERENCES

[1] N. Andrei, Open problems in conjugate gradient algorithms for unconstrained optimization, B. Malays. Math. Sci. Soc. 34(2) (2011), pp. 319-330.

[2] Basim A. Hassan, Issam A. R. Moghrabi, (2022), Amodified secant equation quasi-Newtonmethod for unconstrained optimization, Journal of Applied Mathematics and Computing, pp.1-4.

[3] Basim A. Hassan, Hussein O. D., Azzam S. Y., (2020), A new kind of parameter conjugate gradient for unconstrained optimization, Indonesian Journal of Electrical Engineering and Computer Science, pp. 404-411.

[4] Basim A. Hassan1, Zeyad M. Abdullah and Saif A. Husssin, (2022), New Family of Conjugate Gradient Method to Solving Unconstrained Optimization Problems , To be paper.

[5] Basim A. Hassan, Kanikar Muangchoo, Fadhil Alfarag, Abdulkarim Hassan Ibrahim Auwal Bala Abubakar,(2021), An improved quasi-Newton equation on the quasi-Newtonmethods for unconstrained optimizations, Indonesian Journal of Electrical Engineering and Computer Science. Vol. 22, No. 2, pp. 389~397.

[6] Basim A. Hassan, Zeyad M. A., Hawraz N. J. , (2019), A descent extension of the Dai - Yuan conjugate gradient technique, Indonesian Journal of Electrical Engineering and Computer Science,, pp. 661-668.

[7] Dai Y. H. and Yuan Y., (1999), A Nonlinear Conjugate Gradient Method with A Strong Global Convergence Property. SIAM Journal on Optimization 10(1), 177-182.

[8] Daniel J.W., (1967), The conjugate gradient method for linear and nonlinear operator equations. SIAM J. Numer. Anal. 4, 10-26.

[9] Fletcher R. and Reeves C., (1964), Function minimization by conjugate gradients, Comput. J. 7, 149-154.

[10] Gilbert J. C. and Nocedal J.,(1992), Global convergence properties of conjugate gradient methods for optimization. SIAM J. Opt. 2(1), 21-42.

[11] Hestenes M. R. and Stiefel , E. L.,(1952), Methods of conjugate gradients for solving linear systems. J. Res. Nat. Bur. Standards Sect. 5(49), 409-436.

[12] Hager W. and Zhang H.,(2006), “A survey of nonlinear conjugate gradient methods,” Pac. J. Optim., vol. 2, no. 1, pp. 35-58.

[13] Hawraz N. Jabba and Basim A. Hassan, (2021), 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.

[14] Liu Y. and Storey C.,(1991), Efficient Generalized Conjugate Gradient Algorithms, Part 1: Theory. J. Opt. Theory Appli. 69, 129-137.

[15] Nocedal J.and Wright S.J.,(2006), Numerical Optimization, 2nd ed., Springer, New York.

[16] Polak E. and Ribi`ere G., (1969), Note sur la convergence de directions conjug´ees. Rev. Francaise Informat Recherche Opertionelle, 3e Ann´ee 16, 35-43.

[17] Polyak B. T.,(1969),, The conjugate gradient method in extremem problems, USSR Comp. Math. Math. Phys. 9, 94-112.

[18] Wolfe P.,(1969), "Convergence Conditions for Ascent Methods", SIAM Rev., vol. 11, pp. 226-235.

[19] Wolfe, P.: Convergence conditions for ascent methods. II: some corrections. SIAM Rev. 13, 185–188 (1971).

[20] Zoutendijk G.,(1970),Nonlinear programming, computational methods, in Integer and Nonlinear Programming, J. Abadie, ed., North-Holland, Amsterdam, pp. 37-86.