Two-Step Skipping Techniques For Solution of Nonlinear Unconstrained Optimization Problems

Nudrat Aamir1,*, John Ford2

1CECOS University of IT & Emerging Sciences, Peshawar Pakistan
2Department of Mathematical Sciences, University of Essex, Wivenhoe Park, Colchester, Essex, CO4 3SQ, United Kingdom

Received: 26 May, Revised: 07 June, Accepted: 13 June

Abstract—The development of recent engineering systems has introduced increasing levels of complexity and uncertainty over time. Combined with the planning philosophy of engineering itself, this has given rise to several studies addressing the straightforward or multi-objective optimization problems present in these complex systems. Although conventional approaches are often applied to engineering optimization, it depends largely on the character of problem, but they suffered to supply some quick and reasonable feedback to designers and can't be challenging to further possible problems. However, researchers prefer quasi-Newton methods to solve the unconstrained non-linear optimization problems, using updating the approximation to the inverse Hessian. This technique is a computationally expensive operation and, therefore, in this paper we investigate the possibility of skipping update of Hessian approximation on every second step. The experimental results show that the new methods (i.e. with skipping) give better performance in general than existing two-step methods, particularly as the dimension of the test problem increases.

Keywords—unconstrained non-linear optimization, quasi-Newton methods, approximation to the inverse Hessian, skipping updates, two-step methods.

I. INTRODUCTION

Optimization theory is one of the fundamental theory examined in various mathematical branches such as control theory, linear algebra, calculus of variation, linear programming and game theory etc. Optimization problems arise in various fields of science, engineering, economics, business management and space technology etc. Taking simple examples such as, reducing of cost and minimizing profit, finding elevated point of throw for known initial speed.

For the solution of such problems, different optimization techniques are available. To deal such problems, the iterative methods are best in solution. The improvement of efficiency and convergence of iterative method always motivated researchers to work on this method to get best result, further see [3]. Non Linear optimization problem can be solved using various techniques such as, Newton’s method, Conjugate gradient method, Golden section method and many other. Among these methods Newton’s method is of the most interest to get the desired optimal solution. In case of high dimensions, the Hessian calculation becomes expensive. To overcome this difficulty, researchers introduced various techniques such as, Newton’s method with sparse matrix estimation techniques, Quasi-Newton method and partitioned Quasi-Newton method, further we refer [4, 5, 6, 7, 8, 9].

The fact that optimization theory is now used in nearly all branches of life motivates researchers to either develop new improved methods or to further develop the performance of existing methods. This fact also motivated us to further explore the existing single/multi-step quasi-Newton methods for unconstrained non-linear optimization. As it is evident from literature that in many situations, one of the more expensive operations in the code for a quasi-Newton method is the updating of the inverse Hessian approximation $H_i$ to produce $H_{i+1}$. Thus this issue is strongly needed to be addressed to reduce the computational expense and time.

In order to address the said issue a new technique, of skipping the Hessian update in every second iteration, is introduced. In order to get the Current Hessian approximation, two-step methods, employ data from more than one previous step. BFGS formula is also utilised by the two-step method but replacing the vectors normally employed in the BFGS formula by the vectors determined by the two-step version of the quasi-Newton equation. Hence to explain, in-detail, the process of two-step methods, we consider the problem of the unconstrained minimization of a twice differentiable objective function $f(x)$ (where $x \in \mathbb{R}^n$), using two-step quasi-Newton methods. The gradient and Hessian of $f$ are denoted by $g$ and $G$, respectively. In standard quasi-Newton methods the approximation $B_{i+1}$ to the Hessian $G(x_{i+1})$ is required to satisfy the secant equation

$$B_{i+1} s_i = y_i$$ (1)

Authors retain all © copyrights 2021 IJEW. This is an open access article distributed under the CC-BY License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
where $s_i$ is the step in the variable space and $y_i$ is the corresponding step in the gradient space, defined by

$$s_i = x_{i+1} - x_i$$

And

$$y_i = g(x_{i+1}) - g(x_i)$$

Clearly, in performing the update of $B_i$ to produce $B_{i+1}$ these methods employ data from just one step, and so we refer to them as single-step methods. In the case of two-step quasi-Newton methods, the approximation $B_{i+1}$ is required, instead, to satisfy a condition of the form

$$B_{i+1}(s_i - y_is_{i-1}) = y_i - y_i'y_{i-1}$$

(2)

Or

$$B_{i+1}(s_i - y_is_{i-1}) = y_i - y_i'y_{i-1}$$  \(\text{say}\) \(\text{(3)}\)

Ford and Moghrabi noted that the derivation of this relation depends on the construction of interpolating quadratic curves \(\{x(\tau)\}\) and \(\{h(\tau)\}\), where $\tau \in \mathbb{R}$:

$$x(\tau_j) = x_{i+j-1}, \quad j = 0, 1, 2,$$

$$h(\tau_j) = g(x_{i+j-1}), \quad j = 0, 1, 2,$$

Thus $G(x_{i+1})$ will satisfy the relation

$$G(x_{i+1})x' = G(x_{i+2})x' = g'(x_{i+2})$$ \(\text{(4)}\)

where derivatives with respect to $\tau$ are denoted by primes. If we consider

$$r_i \equiv x' \quad \text{and}$$

$$w_i \equiv h' \quad \text{and}$$

and substitute these relations in equation (4), we obtain a relation of the form

$$B_{i+1} = B_i - \frac{g_i' r_i^T}{r_i^T g_i} + \frac{w_i w_i^T}{w_i^T r_i} \quad \text{(5)}$$

The value of the scalar $i$ in equation (2) is given by

$$\delta = \frac{\tau_2 - \tau_1}{\tau_2 - \tau_0}$$

where

$$\gamma_i = \frac{\delta^2 - 1}{2\delta + 1}$$

Extensive numerical experiments have confirmed the need for care in choosing three values $\{\tau_k\}_{k=0}^2$. Ford and Moghrabi [2] described one successful method for choosing suitable values, involving distances computed by using a metric

$$\varphi_M(x_1, z_2) = \sqrt{(z_1 - z_2)M(z_1 - z_2)'} \quad \forall z_1, z_2$$

in $\mathbb{R}^n$. This leads to the adoption of the following two approaches for definition of the set $\{\tau\}_{k=0}^2$ [2].

A. Accumulative Approach

Taking $\tau_2 = 0$, the values $\tau_1$ and $\tau_0$ are computed by accumulating distances between consecutive pairs of points [2]. Hence

$$\tau_2 = 0; \tau_k = \tau_{k+1} - \varphi_M(x_{i+k}, x_{i+k-1}), \quad \text{for} \quad k = 0, 1.$$  \(\text{B. Fixed Point Approach}\)

Again we take $\tau_2 = 0$, but this time find the values of $\tau_1$ and $\tau_2$ by measuring the distances from $x_{i+1}$ to $x_i$ and $x_{i-1}$ directly [2]. It follows that $\tau_2 = 0$; $\tau_0 = -\varphi_M(x_{i+k}, x_{i+k-1})$, for $k = 0, 1$.

For both approaches, possible choices of the weighting matrix $M$ are

- $M = I$;
- $M = B_i$;
- $M = B_{i+1}$.

Each choice gives rise to a corresponding algorithm, which we represent by A1, A2, A3 (accumulative methods) and F1, F2, F3 (fixed-point methods), respectively.

Numerical experiments with these methods showed significant improvements over the standard single-step BFGS method [2].

II. SKIPPING TECHNIQUE

In many situations, one of the more expensive operations in the code for a quasi-Newton method is the updating of the inverse Hessian approximation $H_i$ \(= B^{-1}\) to produce $H_{i+1}$. This is an $O(n^2)$ operation. Following the idea of Kolda et al. [4] and Fukushima et al. [5], we propose to investigate ‘skipping’ certain of these update operations in order to reduce the computational burden. We recognize that omitting to update $H_i$ on every iteration may lead to an increase in function evaluations and/or iterations (because we will be using an ‘older’ approximation to generate some of the search directions), but any such increase will be offset (in computational terms) by the reduced number of update operations. In this paper, we intend to investigate the simplest implementation of this proposal: namely, that of skipping the update on every second iteration. This leads to the following general algorithmic form:

1. Choose $x_0$ and $H_0$; set $i = 1$.  
2. Compute $p_{i-1} = H_{i-1}g_{i-1}$, perform a line search along \(\{x_{i-1} + t p_{i-1}\}\), giving a value of $t_i$ for $t$, and set $x_i = x_{i-1} + t p_{i-1}$.  
3. Compute $p_i = -H_{i}g_i$, perform a line search along \(\{x_i + t p_{i}\}\), giving a value of $t_i$ for $t$, and set $x_{i+1} = x_i + t p_i$.  
4. Update $H_{i-1}$ to produce $H_{i+1}$.  
5. Check for convergence; if unconverged, then $i := i + 2$, and go to step 2.
III. NUMERICAL RESULTS

The new two-step skipping methods were compared with the existing two-step methods and the standard one-step BFGS method (i.e. both non-skipping and skipping modes) also. All the algorithms in these experiments employed the BFGS formula to update the inverse Hessian approximation

\[ H_{i+1} = H_{i-1} + \left( 1 + \frac{w_i^T H_{i-1} w_i}{r_i^T w_i} \right) \frac{r_i r_i^T}{r_i^T w_i} - \left( \frac{H_{i-1} w_i^T r_i}{r_i^T w_i} \right) \frac{r_i r_i^T}{r_i^T w_i} \]

but (in case of multi-step methods) with the usual vectors \( s_i \) and \( y_i \) replaced by \( r_i \) and \( w_i \) in the case of multi-step methods

\[ H_i \approx B_i^{-1} \]

A total of 15 test functions were employed in the experiments with dimensions ranging from 2 to 200, raising the test function set to 93. These were chosen from standard problems described in the literature [10]. For each function, four different starting points were used, giving a total of 372 test problems.

**TABLE I. TEST PROBLEMS AND DIMENSION**

| No | Problem                                    | Dimension |
|----|--------------------------------------------|-----------|
| 1  | Extended Rosenbrock                        | 2, 20, 26, 40, 60, 80, 100, 120 |
| 2  | Extended Wood                              | 4, 12, 24, 48, 68, 92, 112, 140 |
| 3  | Extended Powell Singular                   | 4, 8, 10, 100, 140 |
| 4  | Penalty 1                                  | 10, 14, 20, 30 |
| 5  | Penalty 2                                  | 10, 16, 24, 30 |
| 6  | Modified Trigonometric Function            | 16, 32, 64, 95, 128, 150 |
| 7  | Broyden Triadialal                         | 18, 36, 72, 90, 108, 144, 186 |
| 8  | Discrete Boundary Value                    | 20, 38, 60, 90, 120, 136, 188 |
| 9  | Discrete Integral Equation                 | 20, 84, 100, 150, 175, 200 |
| 10 | Freudenstein and Roth [11]                 | 28, 52, 85, 118, 190 |
| 11 | Variably Dimensioned                      | 30, 55, 75, 100, 130, 150 |
| 12 | Merged Quadratic [11]                      | 30, 50, 70, 110, 136, 180 |
| 13 | Discrete ODE II [12]                       | 33, 44, 66, 88, 110, 176 |
| 14 | Discrete ODE I [13]                        | 42, 58, 78, 96, 114, 160 |

These test functions were classified into the following subsets (where \( n \) is the dimension of the vector \( x \)):

1. Low: \( 2 \leq n \leq 20 \),
2. Medium: \( 20 \leq n \leq 60 \),
3. High: \( 61 \leq n \leq 200 \),
4. Combined: \( 2 \leq n \leq 200 \).

(7) In total there were 15, 24 and 54 functions in the ‘low’, ‘medium’ and ‘high’ subsets, respectively. These gave 60, 96 and 216 test problems in the three subsets. All functions are from J.J. More, B.S. Garbow and K.E. Hillstrom [10], unless otherwise indicated. Summaries of the results from this set of experiments are presented in Table II, III and IV. For each method the total number of function/gradient evaluations, iterations and time to solve all problems in the given test set is stated. The best result, in case of function evaluation, for each set is highlighted with bold text and italic.

Focusing attention on the results given in Table II and III, we observe that nearly all the two-step skipping methods are compatible, and in some cases performing better, with Single-step BFGS skipping method. Following conclusions can be drawn while evaluating the numerical performance of BFGS skipping method and two-step skipping methods in Table II and III [1]:

- It is proved that in low and medium dimension, in case of function evaluation, **F1 Skip** is outperforming the BFGS Skip method. On the other hand **BFGS Skip** is winning on execution time over all the two-step skipping methods.
- The results of single and two step skipping methods in high dimension illustrates that **F1 Skip** is outperforming the BFGS Skip method in both, function evaluation and execution time. Though, it can be noted that in case of computational time A1 Skip is also compatible with BFGS Skip and F1 Skip.
- It is evident from the results, provided in Table III, that **F1 Skip** is outperforming BFGS Skip in, both, function evaluation and computational time.

**TABLE II. RESULTS FOR SKIPPING CASE**

| Methods | Evaluations | Iterations | Time (sec) | Problem set |
|---------|-------------|------------|------------|-------------|
| BFGS     | 16142       | 10805      | 2.39       |             |
| A1 Skip  | 20276       | 1279       | 2.69       |             |
| **F1 Skip** | **15507** | **10199** | **2.46** | **Low**     |
| A2 Skip  | 16206       | 10640      | 2.68       |             |
| F2 Skip  | 16307       | 10845      | 2.84       |             |
| A3 Skip  | 16321       | 10869      | 2.58       |             |
| F3 Skip  | 14753       | 9981       | 3.05       |             |

| Methods | Evaluations | Iterations | Time (sec) | Problem set |
|---------|-------------|------------|------------|-------------|
| BFGS     | 17788       | 10722      | **3.19**   |             |
| A1 Skip  | 17607       | 10649      | 3.20       |             |
| **F1 Skip** | **17419** | **10559** | **3.21** |             |

International Journal of Engineering Works Vol. 8, Issue 06, PP. 170-174, June 2021

www.ijew.io
### TABLE III.  
**TABLE 3: RESULTS FOR SKIPPING CASE**

| Methods | Evaluations | Iterations | Time (sec) | Problem set |
|---------|-------------|------------|------------|-------------|
| BFGS    | 76717       | 46949      | 24.63      | Medium      |
| A1      | 79928       | 48069      | 24.32      |             |
| **F1 Skip** | **74970** | **46084** | **24.39** |             |
| A2      | 76045       | 47271      | 26.24      | Combined    |
| F2      | 76656       | 46997      | 26.35      |             |
| A3      | 76129       | 46661      | 25.17      |             |
| F3      | 76622       | 47176      | 26.05      |             |

### TABLE IV.  
**RESULTS FOR NON SKIPPING CASE**

| Methods | Evaluations | Iterations | Time (sec) | Problem set |
|---------|-------------|------------|------------|-------------|
| BFGS    | 13095       | 10665      | 2.55       | Low         |
| A1      | 12098       | 9692       | 2.42       |             |
| F1      | 12123       | 9489       | 2.39       |             |
| A2      | 12468       | 9652       | 2.80       |             |
| F2      | 13099       | 10357      | 3.04       |             |
| A3      | 12115       | 9226       | 2.52       |             |
| F3      | 13551       | 9976       | 2.15       |             |
| BFGS    | 16664       | 15518      | 4.73       | Medium      |
| A1      | 13848       | 12397      | 3.98       |             |
| F1      | 14098       | 12366      | 4.00       |             |
| A2      | 15130       | 15241      | 4.02       |             |
| F2      | 13372       | 12405      | 4.56       |             |
| A3      | 13959       | 12535      | 4.29       |             |
| F3      | 15106       | 12873      | 4.70       |             |
| BFGS    | 44251       | 42320      | 40.64      | High        |
| A1      | 34204       | 31915      | 29.64      |             |
| F1      | 34757       | 31806      | 29.87      |             |
| A2      | 35876       | 32145      | 31.17      |             |
| F2      | 33421       | 31342      | 30.14      |             |
| A3      | 35856       | 33768      | 32.20      |             |
| F3      | 33425       | 33425      | 31.84      |             |
TABLE V. RESULTS FOR NON SKIPPING CA

| Methods | Evaluations | Iterations | Time (sec) | Problem set |
|---------|-------------|------------|------------|-------------|
| BFGS    | 74010       | 68503      | 48.02      |             |
| A1      | 60150       | 54032      | 36.13      |             |
| F1      | 60978       | 53662      | 36.35      |             |
| A2      | 63474       | 55038      | 38.76      |             |
| F2      | 60243       | 54104      | 37.83      |             |
| A3      | 61930       | 55529      | 39.08      |             |
| F3      | 66058       | 56279      | 39.68      |             |

CONCLUSION

Therefore, by above discussion it can be conclude that two Step Skipping methods (specially F1 Skip) is performing better than the single step skipping method (BFGS).

On the other hand, following conclusion can be drawn by comparing the results of skipping methods (Table II, III) and non skipping methods (Table IV, V):

- It is evident that all the skipping methods (single and two step) are wining on iteration and computational time in comparison with corresponding non skipping method. Though, an increase in number of evaluation can also be noted in both single and two step skipping methods.
- Over all F1 Skip is winning over skipping and non skipping versions of single and two step methods.

In future we intend to focus on a technique where decrease in number of evaluation in the case of skipping methods can be addressed without losing the property of reduction in computational time.

REFERENCES

[1] Nudrat Aamir, Multi-Step Skipping Methods for Unconstrained Non-Linear Optimization, Ph. D Thesis, University of Essex, United Kingdom, 2014.
[2] J.A. Ford and I.A. Moghrabi, Alternative parametric choices for Multi-Step Quasi-Newton Methods, Optimization Methods and Software 50,305-323, 1993.
[3] J.A. Ford and Nudrat Aamir, Multi-Step Skipping Methods for Unconstrained Optimization, AIP conf Proc 1389, 1806, 2011.
[4] Tamara G. Kolda, Dianne P. O'Leary and Larry Nazareth, BFGS with Update Skipping and Varying Memory, J. Optim. 8, 1060-1083,1998.
[5] D. Li, M. Fukushima, On the Global Convergence of BFGS Method for Nonconvex Unconstrained Optimization, SIAM J. Optim. 11(4), 1054-1064, 2001.
[6] J.-L. Hu and Z. Wu and H. McCann and L.E. Devis and C.-G. Xie, BFGS Quasi-Newton Method for Solving Electromagnetic Inverse Problems, IEEE Proc-Micro: Antennas Propag. 153(2), 199-204, 2006.
[7] C.T. Kelley and E.W. Sachs, Quasi-Newton Methods and Unconstrained Optimal Control Problems, IEEE Conference on decision and Control. 25, 1829-1831, 1986.
[8] M.H. Loke and T. Duhlin, A Comparison of the Gauss-newton and Quasi-Newton Methods in Resistivity Imaging Inversion, Journal of Applied Geophysics. 49, 149-162, 2002.
[9] J.A. Ford and I.A. Moghrabi, Multi-step Quasi-Newton Methods for Optimization, Journal Of Computational And Applied Mathematics. 50,305-323, 1994.
[10] J.J. Mor'e, B.S. Garbow, K.E. Hillstrom,Testing Unconstrained Optimization Software, ACM Trans. Math. Software. 7, 17-41,.1981
[11] Ph.L. Toint, On Large Scale Nonlinear Least Squares Calculations, SIAM J. Sci. Statist. Comput. 8, 416-435, 1987.
[12] G. Alefeld and L. Platzeder, A Quadratically Convergent Krawczyklike Algorithm, SIAM J. Numer. Anal. 20, 210-219, 1983.
[13] W. Burmeister (1972), Inversionfreie Verfahren zur Lö sung Nichtlinearen Operatorgleichungen, Z. Angew. Math. Mech. 52, 101-110, 1972.

How to cite this article:

Nudrat Aamir, John Ford “Two-Step Skipping Techniques For Solution of Nonlinear Unconstrained Optimization Problems”, International Journal of Engineering Works, Vol. 8, Issue 06, PP. 170-174, June 2021, https://doi.org/10.34259/ijew.21.806170174.