A NUMERICAL STUDY OF ITERATIVE METHODS FOR LINEARAND NONLINEAR EQUATIONS IN REAL LIFE

Suman[1], Dr.Yogesh Kumar[2]
Department of Mathematics
[1,2]OPJS University, Churu (Rajasthan)

ABSTRACT

This paper constitutes an effort towards the generalization of the most common classical iterative methods used for the solution of linear systems (like Gauss–Seidel, SOR, Jacobi, and others) to the solution of systems of nonlinear algebraic and/or transcendental equations, as well as to unconstrained optimization of nonlinear functions. Convergence and experimental results are presented. The proposed algorithms have also been implemented and tested on classical test problems and on real-life artificial neural network applications and the results to date appear to be very promising.

1. INTRODUCTION

An iterative strategy to settle the straight framework Hatchet = b begins with an underlying estimate x0 to the arrangement x and creates a succession of vectors {xk }∞ k=0 that focalizes to x. Iterative techniques include a procedure that changes over the framework Hatchet = b into an identical arrangement of the shape x = Mx + v, for some settled network M and vector v. After the underlying vector, x0, is chosen, the arrangement of estimated arrangements is produced by registering

xk+1 = Mxk + v,

for every k = 0, 1, 2,....

For huge frameworks containing a large number of conditions, iterative techniques often have unequivocal focal points over direct strategies as far as speed and requests on PC memory. Once in a while, if the precision necessities are not stringent, a humble number of cycles will suffice to create a satisfactory arrangement. Additionally, iterative strategies are often extremely effective for inadequate frameworks issues. In meager issues, the nonzero components of an once in a while put away in a scanty stockpiling group. In other case, it is not important to store A by any means; for instance, in issues including the numerical arrangement of fractional differential conditions as, for this situation, each column of A may be produced as required yet not held after utilize [1]. Another critical favorable position of iterative strategies is that they are typically steady, and they will really hose

mistakes, because of round off or minor bungles, as the procedure proceeds.

2. UNCONSTRAINED OPTIMIZATION OF NONLINEAR FUNCTIONS

It is well-known that a minimize x∗ of a continuous differentiable function f should satisfy the necessary conditions:

∇f x∗ = Θn = (0, 0,..., 0). (3) Eq. (3)

Represents a set of n nonlinear equations which must be solved to obtain x∗. Therefore, one approach to the minimization of the function f is to seek the solutions of the set of Eq. (3) by including a provision to ensure that the solution found does, indeed, correspond to a local minimizer. This is equivalent to solving the following system of equations:

∂1f (x1, x2,...,xn) = 0,

∂2f (x1, x2,...,xn) = 0, . . .

∂nf (x1, x2,...,xn) = 0,

Where ∂if (x1,...,xi,...,xn) means the halfway subsidiary of f regarding the ith parameter. Next, we consider the classes of nonlinear Jacobi and nonlinear SOR techniques connected to framework [2].

The composite nonlinear Jacobi technique and its merging the class of nonlinear Jacobi strategies is broadly utilized for the numerical arrangement of framework [3]. The fundamental element of the nonlinear Jacobi process is that it is a parallel calculation, i.e., it applies a parallel refresh of the factors. Beginning from a self-assertive introductory vector x0 ∈ D, one can subminimize at the kth emphasis the capacity:

f xk 1 ,...,xk i−1, xi, xk i+1,...,xk n , (5) along the ith heading and get the comparing subminimizerxˆi. Clearly for the subminimizerxˆi ∂if xk1,...,xk i−1, xˆi, xk i+1,...,xk n = 0.

This is a one-dimensional subminimization since every one of the segments of the vector xk , with the exception of from the ith part, are kept consistent. At that point the ithpart is refreshed by the condition: xk+1 i = xk i + τkxˆi − xki ,for some unwinding component τk . The target work in [4] is subminimized in parallel for all i. Different composite nonlinear Jacobi preparing calculations can be gotten relying upon the onedimensional minimization strategy connected. It merits seeing that the quantity of the emphasess of the subminimization technique is identified with the asked for exactness in getting the subminimizer approximations. In this manner, critical computational exertion is required with a specific end goal to discover extremely precise

approximations of the subminimizer along every variable heading at every cycle [5].

In addition, this computational exertion is expanded for issues with a high number of factors, as, when preparing neural systems with a few hundred system parameters (additionally called weights). Then again, it is not sure that this substantial computational exertion accelerates the minimization procedure for nonconvex capacities when a long way from a minimizer x∗. In this manner, we propose to get xˆi by limiting the capacity (5) with one emphasis of a subminimization strategy. Take note of that this practice is likewise recommended for the iterative arrangement of nonlinear conditions.

By legitimately tuning the unwinding component τk , we can get better parameter repeats since this element characterizes the length of the minimization venture along the resultant hunt heading. In this manner, we can stay away from transitory motions and additionally to improve the rate of meeting when the present parameter vector is a long way from a minimizer. Next, the joining of the composite nonlinear Jacobi strategy is talked about. The joining investigation is created under suitable suspicions and gives helpful understanding into this class of strategies. The goal is to demonstrate that there is an area of a minimizer of the target work for which union to the minimizer can be ensured [6].

Hypothesis 1. Let f :D⊂Rn → R be twice constantly differentiable in an open neighborhood S0 ⊂ D of a point x∗∈ D for which ∇f (x∗) = Θn and the Hessian, H (x∗) is certain clear with the property Aπ . At that point there exists an open ball

S = S(x∗, r) in S0

(where S(x∗, r) signifies the open ball focused at x∗ with sweep r), to such an extent that any grouping {xk}∞ k=0 produced by the nonlinear Jacobi handle meets to x∗ which limits f . Proof. Consider the decay of H (x∗) into its corner to corner, entirely bring down triangular and entirely upper-triangular parts:

H x∗ = D x∗ − L x∗ − L x∗ . (8)

Since H (x∗) has the property Aπ , the eigenvalues of

Φ x∗ = D x∗−1 L x∗L x∗

Beneath we incorporate three calculations of this class. These calculations utilize an alternate stepsize for every parameter in view of customary one-dimensional minimization techniques. The first requires just the indication of the angle values, while the other two adventure both the capacity and slope values.

The multi-step Jacobi-altered division technique keeping in mind the end goal to figure a minimizer estimate xˆi in the interim [ai, bi] we utilize our adjustment of the cut strategy which is quickly portrayed underneath. An answer of the condition φ(x) = 0, where the capacity φ :

[a, b] ⊂ R → R is nonstop is ensured to exist in the interim (a, b) if the accompanying standard is satisfied:

φ(a) φ(b) < 0, or sgn φ(a)sgn φ(b) = −1

Where sign is the outstanding three esteemed sign capacity. This paradigm is known as Bolzano's presence foundation (for a speculation of this measure to higher measurements see [7]). In view of this foundation different root finding techniques, as, the separation strategy, were made. Here we might utilize the cut strategy which has been adjusted to the accompanying improved form depicted in [8]. There is accounted for that, keeping in mind the end goal to figure a foundation of φ(x) = 0 where φ :[a, b] ⊂ R → R is consistent, an improved adaptation of the separation technique prompts to the accompanying iterative equation:

rp+1 = rp + c • sgn φ rp 2p+1 , p = 0, 1,..., log2(b − a)ε−1 , (9) with c = sgn φ(r0)(b − a), r0 = a,

where ε is the required precision, and •defines the roof work. Obviously the emphases[9] merge to a root r∗∈ (a, b) with the end goal that |rp+1 − r∗| ε if for some rp, p = 1, 2,... , the accompanying holds: sgn φ r0 sgn φ rp = −1. Besides, the quantity of cycles ν, which are required in getting a rough root r to such an extent that |r − r∗| ε for some ε ∈ (0, 1) is given by:

ν = log2(b − a)ε−1 .

Rather than the iterative recipe we can likewise utilize the accompanying one:

rp+1 = rp − c • sgn φ rp 2p+1 , p = 0, 1,..., log2(b − a)ε−1 , (11) where r0 = b.

The composite nonlinear SOR plan and its merging Beginning from a discretionary introductory emphasize x0 ∈ D, the nonlinear SOR plot subminimizes at the kth cycle the capacity: f xk+1 1 ,...,xk+1 i−1 , xi, xk i+1,...,xk n , along the ith heading and acquire the comparing subminimizer xˆi. Again for this situation, the ith segment is refreshed by Eq. The principle distinction from the Jacobi plan is that the adjustment of the xi at the kth emphasis mulls over all the beforehand refreshed factors of a similar cycle. The joining result for the nonlinear SOR plan is as per the following:

Theorem 3: If a quadratic function f (x) of dimension n is minimized sequentially, once along each direction of a set of n linearly

independent, conjugate directions, the global minimum of f will be located in n or less cycles independent of the starting point as well as the order in which the minimization directions are used.

Theorem 4: The directions generated in Powell's method are conjugate

Theorem 5: The proposed method locates the minimum of an n-dimensional quadratic function f (x), in n or less iterations, utilizing only the relative size of the function values of f , independent of the starting point as well as the order in which the minimization directions are used.

A strategy for developing globally convergent algorithms

In this area we introduce a system for growing internationally merged calculations, i.e., calculations with the property that beginning from any beginning stage the succession of the repeats will join to a nearby minimizer of the goal work. This technique is like the nonlinear Jacobi approach, since it uses approximations of the subminimizers in each organize course, and is a parallel calculation. The hypothetical outcome introduced underneath, permits us to furnish the calculations with a system for adjusting the bearing of pursuit to a drop one. Along these lines, a diminishing of the capacity values at every emphasis is guaranteed, and merging to a nearby minimizer of the target capacity is acquired from remote beginning focuses [10].

3. NUMERICAL RESULTS

the proposed algorithms have been tested on various problems of different dimensions and their performance has been compared with several well-known and widely used unconstrained minimization methods. The numerical applications studied here include classical test cases as well as real-life applications such as artificial neural network training.

Classical test problems: The procedures described in Section 2.2.2, have been implemented and tested in two test functions. Our modified version of Powell's method (SIGNOPT) has been compared with two other well-known minimization methods, namely Powell's and Rosenbrock's methods. To study the influence of imprecise information (regarding the values of the objective function), we simulate imprecisions with the following approach: information about f (x) is obtained in the form of f σ (x), where f σ (x) is an approximation to the true function value f (x), contaminated by a small amount of noise. For

the test problems, the reported parameters are: n, the dimension of the objective function; σ , the value of the standard deviation of the simulated noise; and x0 = (x1, x2,...,xn), the classical starting point for each function [11].

Table 1. Function evolutions (Broyden banded function)

n	σ	Powell	Rosenbrock	SIGNOPT
2	0	246	599	536
	0.01	1549	12437	964
	0.10	9369	20524	1289
	0.20	-	15819	1180
	0.30	-	30756	1824
3	0	2182	1426	1002
	0.01	-	43654	3429
	0.10	-	42231	2574
	0.20	-	90159	4000
	0.30	-	-	9289

where $J_i = \{j:\ j \neq i,\ \max(1, i - m_l) \leqslant j \leqslant \min(n, i + m_u)\}$ and $m_l = 5,\ m_u = 1$. For $n = 2$, we have used the starting values $x^0 = (1, 1)$, while for $n = 3$ we have started the methods from the point $x^0 = (1, 1, 1)$. In Table 1 we exhibit the obtained results. SIGNOPT converged in all cases and had predictable performance, while Powell's method diverged as we increased σ and Rosenbrock's method exhibited slow convergence.

$$f(x) = \sum_{p=1}^{P} \sum_{j=1}^{J} \left[\left\{ 1 + \exp\left(\sum_{i=1}^{I} w_{ij} y_{i,p} + \tau_j \right) \right\}^{-1} - t_{j,p} \right]^2,$$

where

$$y_{i,p} = \left\{ 1 + \exp\left(\sum_{k=1}^{K} v_{ki} u_{k,p} + b_i \right) \right\}^{-1}, \quad \text{and}$$

$$x = (v_{11}, \ldots, v_{ki}, \ldots, v_{KI}, b_1, \ldots, b_i, \ldots, b_I, w_{11}, \ldots, w_{ij}, \ldots, w_{IJ}, \tau_1, \ldots, \tau_j, \ldots, \tau_J).$$

Neural networks training In order to train the network we have to find parameter values that minimize the following objective function:

Table 2. Function evolutions (Hilbert function)

n	σ	Powell	Rosenbrock	SIGNOPT
2	0	106	205	322
	0.01	1151	1213	432
	0.10	–	2411	538
	0.20	–	38882	324
	0.30	–	65861	1825
3	0	439	507	428
	0.01	2523	3553	716
	0.02	3370	3814	716
	0.03	11370	3514	716
	0.04	15221	3623	858
	0.05	15598	4343	858
	0.10	24533	5040	4225
4	0	882	717	533
	0.01	6160	3810	359
	0.02	7174	4426	359
	0.03	4297	7136	359
	0.04	9757	6659	359
	0.05	1624	9373	538
	0.10	4505	11332	4824

The parameters vki, bi, wij and τj can be subjective genuine numbers. Eq. (43) gives a distorted depiction of an organic neuron and it is generally used to develop manufactured neural systems. Illustrations utilized for preparing the system are displayed in a limited arrangement C = (c1, c2,...,cp) of

input–output sets cp = (up, tp) where up can either be genuine or twofold esteemed info vectors in RK and tp are genuine or double yield vectors in RJ , for p = 1,...,P, deciding the comparing preparing design. Next, we [11]

give quantitative outcomes applying different techniques in four neural system applications:

- a variation of the Steepest Drop with consistent stepsize (SD);
- the Steepest Plummet with Line Seek (SDLS) [14, p. 30];
- an alteration of the Steepest Plummet with steady stepsize and Force (SDM);
- a Versatile Steepest Plunge with heuristics for tuning the stepsize (ASD);
- the Fletcher–Reeves (FR) technique;
- the Polak–Ribiere (PR) strategy;
- the Polak–Ribiere (PR) strategy obliged by the FR technique (PR–FR);
- the Heuristic Jacobi–Newton strategy (HJN) of Eq. (16); (ix) the one-stage Jacobi with Netwon–Update strategy (JNU) of Eq. (25); and
- the Multi-step SOR-adjusted separation strategy (m-SOR) of Eq. (30).

Take note of that in the execution of FR, PR, PR–FR, the line hunt of has been utilized. Strategies testing has been directed utilizing an arrangement of 1000 haphazardly picked introductory focuses.

Table 3. Result For the XOR problem (N=9)

Algorithm	μ_{GRD}	μ_{FE}	μ_{ASE}	Success
SD	549	549	*n/a*	810/1000
SDLS	64	371	*n/a*	810/1000
SDM	803	803	*n/a*	810/1000
ASD	157	157	*n/a*	810/1000
FR	84	198	*n/a*	130/1000
PR	21	148	*n/a*	380/1000
PR–FR	22	149	*n/a*	410/1000
HJN	52	182	*n/a*	810/1000
m-SOR	*n/a*	*n/a*	193	440/1000

The XOR problem [18] The classification of the four XOR patterns in two classes is an interesting problem because it is sensitive to initial conditions as well as to stepsize variations, and presents a multitude of local minima. The binary patterns are presented to the network in a finite sequence C = (c1, c2,...,cp) of input–output pairs cp = (up, tp) where up are the binary input vectors in R2 determining the binary input pattern and tp are binary output vectors in R1, for p = 1,..., 4, determining the corresponding number of patterns. A neural network with 9 variables is used for this classification task. The termination condition for all algorithms tested is to find a local minimizer with function value f 0.04. The results are summarized in Table 3, where µGRD denotes the mean number of gradient evaluations, µFE denotes the mean number of objective function evaluations required to obtain convergence, Success shows the number of successful simulations out of 1000 runs, i.e., in the successful runs the iterates converge to a minimizer with function value less than or equal to 0.04, and µASE is the mean number of algebraic sing evaluations required by the m-SOR. In this case the number of successful runs is related to the local minima problem. Thus FR, PR and PR–FR usually converge to an undesired local minimum, i.e., a minimizer with function value f > 0.04 which means that some of the patterns are not correctly classified. HJN exhibits better performance than FR, PR and PR–FR with regards to the number of successful runs. HJN also outperforms SD, SDLS, SDM and FR in training speed, measured by the mean number of function and gradient evaluations needed to successfully classify the patterns. Note that PR and PR–FR require less function evaluations than HJN but they reveal a smaller number of successful runs. It is worth noticing that the m-SOR compares favorably to the

conjugate gradient methods in terms of successes. In addition, m-SOR does not require gradient evaluations.

4. CONCLUSION

In this paper, an investigation on the generalization of the most common classical iterative methods used for the solution of linear systems (like Gauss–Seidel, SOR, Jacobi, and others) to the unconstrained optimization of nonlinear functions has been conducted. Although the nonlinear iterative rootfinding methods have been extensively studied, the unconstrained optimization case has not been thoroughly analyzed. Thus, in this work unconstrained optimization algorithms for nonlinear functions based on generalizations of iterative linear methods were introduced. Theoretical convergence results for the proposed algorithms have been derived for computing a local minimizer of a function. A strategy for developing globally convergent modifications of these algorithms has also been proposed.

5. SCOPE OF FUTURE RESEARCH

The new algorithms have been implemented and tested on classical test problems and on real-life artificial neural network applications and the results to date appear to be very promising. In a subsequent communication we intend to implement in parallel the methods of the Jacobi class, using the Parallel Virtual Machine (PVM) Preliminary results indicate that utilizing PVM, the speed up achieved is analogous to the number of the processors used [13], thus considerably shorten the minimization process time

References

1. O. Axelsson, Iterative Solution Methods, Cambridge University Press, New York, 1996.
2. R.P. Brent, Algorithms for Minimization Without Derivatives, Prentice-Hall, Englewood Cliffs, NJ, 1973.
3. P. Brodatz, Textures—A Photographic Album for Artists and Designer, Dover, New York, 1966.
4. J.C. Gilbert, J. Nocedal, Global convergence properties of conjugate gradient methods for optimization, SIAM J. Optimization 2 (1992) 21–42.
5. R. Haralick, K. Shanmugan, I. Dinstein, Textural features for image classification, IEEE Trans. System Man. Cybernetics 3 (1973) 610–621.
6. S. Haykin, Neural Networks: A Comprehensive Foundation, Macmillan, New York, 1994.
7. A. Geist, A. Beguelin, J. Dongarra, W. Jiang, R. Manchek, V. Sunderam, PVM:

Parallel Virtual Machine. A User's Guide and Tutorial for Networked Parallel Computing, MIT Press, Cambridge, MA, 1994.

8. C.T. Kelley, Iterative Methods for Linear and Nonlinear Equations, SIAM, Philadelphia, PA, 1995.
9. G.D. Magoulas, M.N. Vrahatis, G.S. Androulakis, A new method in neural network supervised training with imprecision, in: Proceedings of the IEEE 3rd International Conference on Electronics, Circuits and Systems, IEEE Press, Piscataway, NJ, 1996, pp. 287–290.
10. G.D. Magoulas, M.N. Vrahatis, G.S. Androulakis, Improving the convergence of the back-propagation algorithm using learning rate adaptation methods, Neural Comput. 11 (1999) 1769–1796.
11. B.J. Moré, B.S. Garbow, K.E. Hillstrom, Testing unconstrained optimization, ACM Trans. Math. Software 7 (1981) 17–41. [12] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.