MASTER COPY KEEP THIS COPY FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A244 404

...n is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, ...ting and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this ...ucing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson ...nd to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

2. REPORT DATE
08/29/91

3. REPORT TYPE AND DATES COVERED
Final 15 May 88 - 14 May 91

4. TITLE AND SUBTITLE
"New Methods for Nonlinear Optimization"
Final Report for ARO Contract DAAL03-88-K-0086

5. FUNDING NUMBERS
DAAL03-88-K-0086

(2)

6. AUTHOR(S)
Robert B. Schnabel and Richard H. Byrd

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Office of Contracts and Grants
Armory 206, Campus Box B-19
University of Colorado at Boulder
Boulder, CO 80309

92-00772

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 24923.11-MA

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

DTIC
ELECTE
JAN 14 1992
S D
D

13. ABSTRACT *(Maximum 200 words)*

This research project has investigated topics in unconstrained and constrained optimization, solving systems of nonlinear equations, nonlinear least squares, and parallel optimization. Over the course of this research contract, considerable progress was made in all the areas discussed in the proposal, namely tensor methods for nonlinear equations and optimization, trust regions methods for nonlinearly constrained optimization, orthogonal distance regression, semilocal analysis of quasi-Newton methods for nonlinearly constrained optimization, and parallel unconstrained optimization. In addition, we have worked on several other topics, including a new modified Cholesky factorization, analysis and performance of the symmetric rank one update for unconstrained optimization, secant methods for constrained optimization, the behavior of Broyden class methods for unconstrained optimization, and parallel and sequential methods for global optimization.

14. SUBJECT TERMS
Unconstrained and constrained optimization, nonlinear equations, nonlinear least squares, parallel optimization.

15. NUMBER OF PAGES
10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

New Methods for Nonlinear Optimization

Final Report

Robert B. Schnabel and Richard H. Byrd

08/29/91

U.S. Army Research Office

Contract DAAL03-88-K-0086

University of Colorado at Boulder
Boulder, CO 80309-0430

Approved for Public Release;
Distribution Unlimited

DTIC QUALITY INSPECTED

Accesion For
NTIS CRA&I ✓
DTIC TAB
Unannounced
Justification
By
Distribution /
Availability Codes
Dist | Avail and/or Special
A-1

FINAL REPORT for

ARO Contract DAAL03-88-K-0086
New Methods for Nonlinear Optimization
May 15, 1988 - May 14, 1991

Principal Investigator: Robert B. Schnabel
Co-Principal Investigator: Richard H. Byrd

Over the course of this research contract, considerable progress was made in all the areas discussed in the proposal, namely tensor methods for nonlinear equations and optimization, trust regions methods for nonlinearly constrained optimization, orthogonal distance regression, semilocal analysis of quasi-Newton methods for nonlinearly constrained optimization, and parallel unconstrained optimization. We summarize the work in those areas in Sections 1-5. In addition, we have worked on several other topics, including a new modified Cholesky factorization, analysis and performance of the symmetric rank one update for unconstrained optimization, secant methods for constrained optimization, the behavior of Broyden class methods for unconstrained optimization, and parallel and sequential methods for global optimization. This work is summarized in Sections 6-10. Section 11 contains a listing of publications and reports supported by this contract, and Section 12 contains a list of research personnel supported by this contract.

1. Tensor Methods for Nonlinear Problems.

During this research period, we have developed tensor methods for the unconstrained optimization problem. Tensor methods, first developed by Schnabel and Frank in the context of nonlinear equations, augment the standard model (linear for nonlinear equations, quadratic for optimization) by low rank, higher order terms, in a way that makes forming, storing, and solving the higher order model hardly more expensive than the standard model. The development of tensor methods for unconstrained optimization required a significantly different approach than for nonlinear equations, involving the use of a fourth order model. We developed procedures for forming and solving this model, and have shown that the extra costs of forming, storing, and solving this model can be made insignificant in comparison to the underlying costs of forming, storing, and solving the standard quadratic model. In extensive computational tests of methods that use second derivatives, our tensor method using a trust region strategy was about 40-50% more efficient than an analogous standard method on problems where the Hessian at the solution is nonsingular, and 50-60% more efficient on problems where the Hessian at the solution is singular. In cases where secant approximations to the Hessian are made, good improvements were obtained by the tensor method on problems where the Hessian at the solution is singular. This research was performed jointly with a student, Ta-Tung Chow, who received his Ph.D. in Aug. 1989, and is described in a paper by Schnabel in the book of plenary and survey lectures from the 13th International Conference on Mathematical Programming, and in a paper by Schnabel and Chow that just appeared in the *SIAM Journal on Optimization*. Recently we completed a software package that implements the line search version of this tensor method; it is described in a report by Chow, Eskow, and Schnabel that has been submitted for publication.

During this period, we also extended our tensor method for nonlinear equations to apply to the nonlinear least squares problem. In conjunction, we developed trust region versions of these methods to augment the original line search methods. In our computational tests, the tensor nonlinear least squares method is significantly more reliable and efficient than Gauss-Newton and Levenberg-Marquardt methods, to which it is most closely related. It is also considerably more efficient than the NL2SOL method, but not quite as reliable. We have developed a software package that incorporates all these tensor methods for nonlinear equations and least squares. This work has been performed along with a Ph.D. student, A. Bouaricha, and is described in two papers by Schnabel and Bouaricha that are currently in draft form.

During this research period we also began the development of efficient tensor methods for solving large sparse systems of nonlinear equations and nonlinear least squares. We devised an efficient method for finding a root or minimizer of the tensor model for nonlinear equations or least squares when the Jacobian is sparse, and began to consider implementation details connected with this approach. This method is entirely different than the one previously used by us for small dense problems, and can be incorporated with any direct or iterative sparse solver. This research also is being conducted with Ali Bouaricha, and will form the heart of his Ph.D. dissertation. We anticipate writing the first paper on this topic shortly.

Finally, together with another Ph.D. student, Dan Feng, we have recently developed some new techniques for analyzing the convergence rate of tensor methods for nonlinear equations on problems where the Jacobian at the solution has a null space of dimension one. A paper describing this research is currently in draft form.

2. Trust Region Methods for Nonlinearly Constrained Optimization.

We have developed a very robust algorithm for nonlinear equality constrained optimization. It computes each step by a two-phase strategy involving a Levenberg-Marquardt step on the constraints and the solution of a relaxed quadratic program with a trust region restriction. We use a nondifferentiable merit function involving the Euclidean norm of the constraints, along with a second order correction step when necessary.

The global convergence theory for the algorithm is fairly strong. Without any assumptions of linear independence of the constraint gradients or of positive definiteness of the Hessian approximation we can show that a subsequence of the iterates converges to either a Kuhn-Tucker point, a local minimizer of the norm of the constraints, or a point violating a constraint qualification. In addition, if the exact Hessian of the Lagrangian is used we can show that convergence to a strong local minimizer is quadratic.

We have performed computational testing on a variety of test problems, and our algorithm has proved very robust. It has rarely failed to converge, and solves many problems where SQP methods involving a line search fail. In comparisons with a standard line search SQP algorithm it has almost always taken fewer iterations, although the line search method used a quasi-Newton approximation while our method used the exact Hessian. This research is described in the Ph.D. dissertation of one of our students, Emmanuel Omojokun, and in a paper by Byrd and Omojokun that is currently in preparation.

3. Orthogonal Distance Regression.

Previously we developed an efficient method for solving the orthogonal distance regression (ODR) problem, and a software package, ODRPACK, that implements both ODR and ordinary nonlinear least squares (OLS) methods. ODR is used for problems when there are errors in both the independent and the dependent variables. During this research period, we conducted a Monte Carlo study to determine whether it was advantageous to use ODR or OLS in cases when there are errors in both variables. In these tests the parameter estimates produced by ODR were virtually never worse than those produced by OLS, and sometimes were significantly better. This work was performed in conjunction with P. Boggs, J. Donaldson, and C. Spiegelman of the National Bureau of Standards, and is described in a paper by Boggs, Donaldson, Schnabel, and Spiegelman in a special issue of *The Journal of Econometrics* on computational statistics. More recently, Boggs and Donaldson have generalized the ODRPACK code so that it handles certain classes of implicit problems, which was one of the objectives of our proposal.

4. Analysis of Quasi-Newton Methods for Nonlinearly Constrained Optimization.

As a preliminary step in analyzing quasi-Newton methods for constrained optimization, we developed some new analytical tools for doing global analysis of quasi-Newton methods with line searches in regions where the objective function is convex. This has allowed us to extend the results of Powell on convergence of the BFGS with arbitrary positive definite starting Hessian, to allow for a backtracking line search. This work is discussed in a paper by Byrd and Jorge Nocedal of Northwestern University in *SIAM Journal of Numerical Analysis*. This tool, involving the trace of a matrix minus the logarithm of its determinant, has also been useful in analyzing some new parallel methods for optimization, and in analyzing other members of the Broyden class, as well as in our research on constrained optimization.

Our analysis of methods for constrained optimization has been confined mainly to successive quadratic programming methods that use an approximation of the reduced Hessian of the Lagrangian. We have considered such methods used with a line search on a merit function and the updating procedure proposed by Coleman and Conn. For such methods we have shown that if we start close enough to the solution with an arbitrary positive definite initial Hessian, then the iterates will converge r-linearly to the solution. This applies to the l_1 merit function and, with some additional assumptions, to the exact penalty function of Fletcher used as a merit function. We have also shown such methods to be q-superlinearly convergent when used with the Fletcher function, or with the l_1 merit function in conjunction with either a second order correction step or a watchdog strategy. Interestingly, although the algorithm depends on choice of null space basis, the R-linear convergence result is valid for any choice of null space basis. Superlinear convergence occurs when the null space basis is chosen by either the strategy of Coleman and Sorensen or that of Gill, Murray, Saunders, Stewart, and Wright. This also is joint work with Jorge Nocedal and is described in a paper by Byrd and Nocedal in *Mathematical Programming*.

5. Parallel Unconstrained Optimization.

During this research period we completed an extensive study of parallel methods for general unconstrained optimization problems. One aspect we investigated was strategies for utilizing concurrent function evaluations. We introduced the idea of speculative finite difference gradient evaluations, that is the evaluation of the gradient at the same time the function is evaluated and before it is known whether that gradient value will be needed. We showed it to be very effective. We also developed a number of new optimization algorithms for the fairly common situation in parallel computing where there are enough processors to evaluate part, but not all, of the Hessian matrix at each iteration. This work included new theoretical analysis and extensive computational tests that showed that these strategies are worthwhile. We also carefully considered various methods for implementing the linear algebra in the BFGS method, and showed that inverse updates are likely to be most effective on parallel computers, and that the method of Han is equivalent to an inverse, factored update. Much of this research was conducted in conjunction with G. Shultz of Metropolitan State University. It is described in an earlier paper by Schnabel in *Computational Methods in Applied Mechanics and Engineering*, and two papers by Byrd, Schnabel, and Shultz, in *Annals of Operations Research* and in *Mathematical Programming*.

6. Modified Cholesky Factorization for Unconstrained Optimization.

We have developed a new modified Cholesky factorization and shown it to have advantageous theoretical and computational properties. This type of factorization, introduced by Gill and Murray, is central to many optimization algorithms. It takes as input a symmetric matrix A, and produces a Cholesky factorization of $A+D$, where $A+D$ is positive definite, and D is a non-negative diagonal matrix which is zero if A is safely positive definite. The goal is that D be not much larger than necessary, and that the cost be only a small multiple of n^2 operations more than the standard Cholesky factorization. Our method uses an entirely new technique for choosing D. Its cost is still low, and in our tests it often produces a D that is much smaller than the standard method, and usually close to the optimal size. This work has been conducted in conjunction with E. Eskow, a research associate in our department, and is described in a paper by Eskow and Schnabel in the *SIAM Journal on Scientific and Statistical Computing*. We also completed a software package for this new method; it is described in a paper by Schnabel and Eskow that has been accepted for publication in the *ACM Transactions on Mathematical Software*. The new factorization has also been incorporated into several pieces of optimization software by other researchers.

7. The Symmetric Rank One Update for Unconstrained Optimization.

We have conducted a comprehensive study of the use of the symmetric rank one (SR1) update for unconstrained optimization. This topic formed part of the Ph.D. thesis of our student H. Khalfan, completed in Oct. 1989, and Khalfan also spent summer 1990 with us continuing this research. A paper by Khalfan, Byrd, and Schnabel on this topic has been submitted for publication. It describes our computational experiments comparing the SR1 and BFGS updates in line search and trust region methods for unconstrained optimization, which show that the SR1 method is at least as robust and efficient as the BFGS method for this test set. It then shows that the condition of uniform linear independence of the steps, that is assumed in Conn, Gould, and Toint's theoretical analysis of the SR1 method, is often violated in our tests, and gives a new convergence result that does not assume uniform linear independence. This result assumes instead that the Hessian approximations are bounded and positive definite, conditions that appear almost always to be satisfied in practice.

8. Secant Methods for Constrained Optimization.

In collaboration with Richard Tapia and Yin Zhang of Rice University, we completed the development and analysis of an augmented Lagrangian based method for nonlinearly constrained optimization. In the past augmented Lagrangian approaches have been unreliable because of lack of knowledge of the value of the augmentation parameter required by theory. We developed an adaptive method for choosing this parameter which dovetails well with the SALSA version of the augmented Lagrangian first proposed by Tapia. Our implementation of this method slightly outperforms Powell's widely used SQP method using a damped update. We were able to prove that whenever this method converges it converges R-superlinearly, without making any assumptions about the augmentation parameter or the Hessian approximations. A corollary of this result in the unconstrained case is the new result that whenever the standard BFGS method converges it converges Q-superlinearly. A paper describing this work has been accepted by

SIAM Journal on Optimization.

9. Behavior of Broyden Class Methods for Unconstrained Optimization.

In collaboration with Jorge Nocedal of Northwestern University we have completed an investigation of quasi-Newton methods from the Broyden class with negative values of the Broyden parameter. Interestingly, we were able to establish necessary and sufficient conditions on the Broyden parameter for Q-superlinear convergence which were quite close together. We also performed experiments indicating that using negative values of the Broyden parameter has the potential to yield significant improvement. A paper describing this work has been submitted to *SIAM Journal on Optimization.*

10. Parallel and Sequential Methods for Global Optimization.

We have designed, implemented, and tested new adaptive stochastic methods for solving the global optimization problem. These methods are based on the stochastic methods of Rinnooy Kan and Timmer, but decide adaptively how much work to perform in various portions of the feasible region as the computation proceeds. On test problems, this approach has reduced the cost of solving global optimization problems on sequential computers by up to 72%. This approach is particularly conducive to implementation on parallel computers, for it is natural to integrate the adaptive strategy into an algorithm that dynamically assigns work to processors, and eliminates major synchronization points. We have implemented such a method on a network of Sun workstations, and have achieved reductions of up to 85% over our previous parallel implementation of the static, synchronous method of Rinnooy Kan and Timmer. This research is being conducted with a Ph.D. student, Sharon Smith, and is described in a paper by Smith, Eskow, and Schnabel in a book edited by Coleman and Li. We also have extensively examined scheduling strategies for such dynamic parallel algorithms; a preliminary account of this research is given in a report by Smith and Schnabel that will appear in the proceedings of the ICASE Conference on Unstructured Scientific Computations on Scalable Multiprocessors. Recently, we have begun applying these algorithms to global optimization problems arising in molecular conformation.

11. Publications and Technical Reports Associated with Contract DAAG29-84-K-0140.

P.T. Boggs, J.R. Donaldson, R.B. Schnabel, and C.H. Spiegelman, "A computational examination of orthogonal distance regression," *Journal of Econometrics* 38 (1988), pp. 169-201.

R.H. Byrd, "On the convergence of constrained optimization methods with accurate Hessian information on a subspace," *SIAM Journal on Numerical Analysis* 27 (1990), pp. 141-153.

R.H. Byrd, D. Liu and J. Nocedal, "On the behavior of Broyden's class of quasi-Newton methods", Tecnical report NAM 01, Dept. of E.E. and C.S., Northwestern University, to appear in *SIAM Journal on Optimization.*

R.H. Byrd and J. Nocedal, "A Tool for the Analysis of Quasi-Newton Methods with Application to Unconstrained Optimization", *SIAM Journal on Numerical Analysis* 26 (1989), pp. 727-739

R.H. Byrd and J. Nocedal, "An Analysis of Reduced Hessian Methods for Constrained Optimization", *Mathematical Programming,* 49 (1991), pp. 285-325.

R.H. Byrd and E. Omojokun, "A Robust Trust Region Algorithm for Nonlinear Equality Constrained Optimization", (in preparation).

R.H. Byrd, R.B. Schnabel and G.A. Shultz, "Using Parallel Function Evaluations to Improve Hessian Approximations in Unconstrained Optimization", *Annals of Operations Research,* 14 (1988), pp. 167-193.

R.H. Byrd, R.B. Schnabel and G.A. Shultz, "Parallel Quasi-Newton Methods for Unconstrained Optimization", *Mathematical Programming,* 42 (1988), pp. 273-306.

R.H. Byrd, R. A. Tapia and Y. Zhang, "An SQP augmented Lagrangian BFGS algorithm for constrained optimization," Technical Report 89-4, Dept. of Mathematical Sciences, Rice University, to appear in *SIAM Journal on Optimization.*

T. Chow, E. Eskow, and R. B. Schnabel, "A software package for tensor methods for unconstrained optimization," Technical Report CU-CS-491-90, Department of Computer Science, University of Colorado at Boulder, submitted to *ACM Transactions on Mathematical Software.*

J.E. Dennis and R.B. Schnabel, "A view of unconstrained optimization," *Handbooks in Operations Research and Management Science, Volume 1: Optimization,* G.L. Nemhauser, A.H.G. Rinnooy Kan, and M.J. Todd, eds., North-Holland (1989), pp. 1-72.

E Eskow and R.B. Schnabel, "A new modified Cholesky factorization," *SIAM Journal on Scientific and Statistical Computing,* 11(1990), pp. 1136-1158.

E. Eskow and R. B. Schnabel, "Software for a new modified Cholesky factorization," Technical Report CU-CS-443-89, Department of Computer Science, University of Colorado at Boulder, to appear in *ACM Transactions on Mathematical Software.*

H. Khalfan, R. H. Byrd, and R. B. Schnabel, "A theoretical and experimental study of the symmetric rank one update", Technical Report CU-CS-489-90, Department of Computer Science, University of Colorado at Boulder, submitted to *SIAM Journal on Optimization.*

R.B. Schnabel, "Sequential and parallel methods for unconstrained optimization," *Mathematical Programming, Recent Developments and Applications*, M. Iri and K. Tanabe, eds., Kluwer, Tokyo (1989), pp. 227-261.

R. B. Schnabel and T. Chow, "Tensor methods for unconstrained optimization using second derivatives," *SIAM Journal on Optimization* 1 (1991), pp. 293 315.

S. Smith, E. Eskow, and R. B. Schnabel, "Adaptive, asynchronous stochastic global optimization algorithms for

S. Smith, E. Eskow, and R. B. Schnabel, "Adaptive, asynchronous stochastic global optimization algorithms for sequential and parallel computation," *Large Scale Numerical Optimization*, T. Coleman and Y. Li, eds., SIAM, Philadelphia (1990), pp. 207-227.

S. Smith, E. Eskow, and R. B. Schnabel, "An adaptive, asynchronous parallel global optimization algorithm," *Parallel Processing for Scientific Computing*, J. Dongarra, P. Messina, D. C. Sorenson, and R. G. Voigt, eds., SIAM (1990), pp. 192-197.

S. Smith and R. B. Schnabel, "Centralized and distributed dynamic scheduling for adaptive, parallel algorithms", Technical Report CU-CS-516-91, Department of Computer Science, University of Colorado at Boulder, to appear in proceedings of ICASE Conference on Unstructured Scientific Computations on Scalable Multiprocessors.

12. Students Supported (in part) by Contract DAAL03-88-K-0086.

Ali Bouaricha, Ph.D. in progress (expected Dec. 1991)

Ta-Tung Chow, Ph.D. granted 1989

Dan Feng, Ph.D. in progress (expected 1992)

Sharon Smith, Ph.D. granted 1991

Yuanfu Xie, Ph.D. granted 1991

In addition, the following students were supported by their governments but conducted research associated with this contract :

Humaid Khalfan, Ph.D. granted 1989

Emmanuel Omojokun, Ph.D. granted 1989