OPTIMIZATION AND EIGENVALUE COMPUTATION
WITH APPLICATION TO METEOROLOGY AND OCEANOGRAPHY

RECEIVED
JUN 16 1993
OSTI

Progress Report

for Period August 1, 1992 - July 31, 1993

Jorge Nocedal

Northwestern University
Evanston, IL 60208

May 1993

Prepared for

THE U.S. DEPARTMENT OF ENERGY
AGREEMENT NO. DE-FG02-87ER25047

MASTER

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

PROGRESS REPORT

on the Project

“Optimization and Eigenvalue Computations with Application to Meteorology and Oceanography”

Principal Investigator: Jorge Nocedal

Grant No. DE-FG02-87ER25047-A002

This is a request for the continuation of grant DE-FG02-87ER25047-A002, which now enters its second funding period. The project is proceeding as planned: there is no change in the scope of the project nor in the proposed budget for the coming year.

1. Work Completed.

Our computational experience has shown that dynamic scaling techniques can have a dramatic effect in the efficiency of optimization algorithms, especially when the number of variables is very large. Conjugate gradient and limited memory methods do not have adequate information on the Hessian matrix and dynamic scaling of the variables can supply some of this missing information. In [1] we developed a full-rank scaling technique for variable metric methods that is globally and superlinearly convergent on general problems. The numerical results indicate that the new scaling improves the performance of the BFGS method significantly, particularly on ill-conditioned problems. The idea is to update a factorization of the Hessian approximation, $B_k = W_k W_k^T$, and control the size of the columns of W . The first column of W is guaranteed to have the correct size, because an orthogonal rotation is used to pre-process the Hessian approximation, and this pre-processing ensures that the size of the first column equals the average curvature detected during the iteration. For the rest of the columns, the strategy is to increase them in size (scale them up) if they are small compared to the first column. This research project is completed and our paper has been accepted for publication [1].

We have also proposed a new technique for accelerating the optimization process in the solution of data assimilation problems. This is an area where optimization plays a crucial role, since the four dimensional assimilation approach consists of using a nonlinear optimization code to adjust the initial conditions of the model so that measurements distributed in space and time can be adequately predicted. Each iteration of the optimization method requires a full simulation of the model - a process that typically takes 1 hour on the CRAY 90. We have found that by computing a few extreme eigenvalues of the Hessian matrix, and by incorporating this information in a limited memory code, the number of simulations can be reduced by at least 50%. This is highly significant and represents a substantial contribution to the area of variational assimilation of data. We have also developed an extension

of the limited memory BFGS method that is capable of handling simple bounds on the variables, providing more freedom to the modeler [2]. This work was possible by the development of compact matrix representations of limited memory matrices [3], which keep the computation cost of the iteration to a minimum. Indeed we have shown that the operations with limited memory BFGS matrices are optimal in terms of arithmetic operations, since they can be rewritten as the reverse automatic differentiation of an auxiliary function [4].

We have also developed a theoretical and algorithmic framework for combining trust region and line search approaches [5]. When the number of variables is large, and when the step generated by a trust region method does not result in a decrease in the objective function, it is not desirable to resolve the trust region problem from scratch. We have shown that it is possible to take the failed point and back-track, provided the solution of the trust region problems satisfies two conditions that we specify. These ideas will be particularly useful in our LP method for solving large systems of nonlinear equations.

2. Research Plans for the Coming Year.

Most of our effort will focus on the development of a new algorithm for solving large constrained optimization problems. During the next year we will concentrate in the case when only nonlinear equality constraints are present, but will allow the number of variables and constraints to be arbitrary. We will study how to approximately compute the vertical and horizontal steps required by our algorithm, while maintaining a fast rate of convergence. One of the most difficult aspects of this work is how to detect and cope with ill-conditioning in the Jacobian matrix of constraint gradients. So far we have developed a set of subroutines, based on the Harwell routine MA28, which find a new basis if the algorithm appears to be performing poorly. However it is necessary to make this identification earlier – before difficulties arise. We plan to develop a new algorithm and implement it using the Harwell data structures.

Another aspect of this research that requires careful consideration is the solution of the horizontal step. One needs to minimize an approximation to the Lagrangian on the tangent space of the constraints. We have found that the Conjugate Gradient method proposed by Steihaug performs extremely well in practice and requires a low computational cost. However when the representation of the tangent space of the constraints is ill-conditioned, this approach can become quite expensive and erratic. We propose to develop a new version of the conjugate gradient method that is effective in this setting.

The robustness and efficiency of our method also requires that the choice of the penalty parameter be made with great precision. The standard approach is to keep increasing the parameter as the iteration proceeds to ensure that it is large

enough when the iterates approach the solution. It is also customary to never let the penalty parameter decrease, since this greatly simplifies the design and analysis of the method. In practice, however, the penalty parameter may become large too soon, and this forces the iterates to follow the constraints too closely. In some case, a large penalty parameter results in a failure to converge. Therefore we need to develop a strategy that allows re-adjustments of the size of the penalty parameter. We intend to follow a novel approach in which information from the whole run is used to construct a model of the penalty function and its evolution is tracked adaptively. Such an approach will allow drastic changes in some circumstances, and will ensure that smooth changes in the parameter are made for well-conditioned problems. We intend to release the first version of our code in the summer of 1994.

References:

- [1] An Automatic Scaling Algorithm for Quasi-Newton Methods, (1991), with Marucha Lalee, to appear in *Mathematical Programming*
- [2] A limited memory method for bound constrained optimization, Northwestern University, EECS Dept., Rep. NAM 03, 1992, with R. Byrd and P. Lu.
- [3] Representations of Quasi-Newton Matrices and their use in Limited Memory Methods, (1992), Northwestern University, EECS Dept., Rep. NAM 06, to appear in *Mathematical Programming*, with R. Byrd and R. Schnabel.
- [4] The Limited Memory Step Computation and Automatic Differentiation, (1992), Northwestern University, EECS Dept., Rep. NAM 02, to appear in *Applied Math Letters*, with J. C. Gilbert.
- [5] Combining Trust Region and Line Search Techniques, (1992), Northwestern University, EECS Dept., Rep. NAM 05, submitted for publication in *SIAM Journal on Optimization*, with Y. Yuan.

END

**DATE
FILMED**

7 / 13 / 93