Stellarator Optimization Using the Levenberg-Marquardt Method with Broyden Update

M. A. Kraemer, A. S. Ware

University of Montana - Missoula

S. P. Hirshman

Oak Ridge National Laboratory

Abstract

Optimization of quasi-omnigeneous stellarator (QOS) configurations has led to the discovery of a new class of compact configurations with high beta ballooning stability limits. The design of the stellarator leads to nonlinear least squares problems with typically n=40 parameters and several hundreds of target functions. The Levenberg-Marquardt iterative method is used to solve these problems. In each iteration the Jacobi matrix is approximated by a difference scheme that requires n evaluations of the set of target functions. Each function evaluation uses several minutes of supercomputer time to solve the MHD equations and determine ballooning stability.

By the utilization of a Broyden update - a generalization of the secant method in one dimension - only one function evaluation is required to update the approximate Jacobi matrix in each Levenberg-Marquardt iteration. Although the number of iterations necessary to satisfy the convergence criteria may be increased, the total number of function evaluations is reduced. This is demonstrated by examples.

Levenberg-Marquardt Method for Nonlinear Least Squares

- An iterative method to minimize the sum of squares ("Chi-square") of m functions in n variables.
- It requires the evaluation or approximation of the Jacobian matrix in every iteration.

Finite Difference Approximation of the Jacobian Matrix

- In stellarator optimizations the Jacobian is not available & must be approximated.
- The gradients of m functions of n variables are approximated by finite differences.
- This requires n evaluations of the set of functions, one in each of n directions, using a fixed stepsize.

Broyden Update of the Jacobian

In each iteration the approximate Jacobian is obtained from the previous one by the update formula [3,4]:

$$B_i = B_{i-1} + (\Delta y_{i-1} - B_{i-1}\Delta x_{i-1})\frac{(\Delta x_{i-1})^T}{(\Delta x_{i-1})^T \Delta x_{i-1}}$$

The Broyden update requires no(!) function evaluations.

Properties of Broyden Update

- Satisfies secant equation:

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

- Updates approximate Jacobian only in the direction Δx_i :

$$B_{i+1}s = B_i s \quad \text{for all s orthogonal to } \Delta x_i$$

Loss of Gradient Information

- Since the Broyden update only changes the Jacobian in the direction of the x step, in stellarator optimizations reduction of Chi-square slows down after typically 3-4 successful iterations.
- Eventually, the method no longer provides a descent direction.

Resetting the Jacobian

- Whenever Chi-square cannot be further reduced with Broyden updates of the Jacobian, reset the Jacobian using finite differences.
- Check stopping conditions for the optimization after such Jacobian resets. [6,11]

The LM Parameter

- In each LM iteration, the LM parameter must be chosen to reduce Chi-square.
- For large parameter values, the step to the next x iterate is small, guarantees descent direction.
- For small parameter values, the step to the next x iterate is large, might not provide descent.

Optimizing the LM Parameter

- Outer loop: performs Broyden update or resets the Jacobian with finite differences.
- Inner loop: rescales the range for LM parameters, until Chi-square is reduced.
- Search loop: determines optimal LM parameter in range given in inner loop.

Options for Search Loop

- Evaluate Chi-square for a set of fixed LM parameters, and choose the one with lowest Chi-square. Multi-processing is possible.
- Use a line search algorithm to minimize Chi-square within given LM parameter range, e.g. golden section line search [10].

Results for 3 Test Problems [9]

	WATSON 20 var., 31 func. iterations / eval. min. chi-square	OSBORNE 1 5 var., 33 func. iterations / eval. min. chi-square	OSBORNE 2 11 var., 65 func. iterations / eval. min. chi-square
Broyden's update	13 / **41** 2.213 E-10	29 / **40** 5.465 E-5	13 / **35** 4.066 E-2
Finite diff. Jacobian	5 / **85** 3.334 E-10	11 / **75** 5.619 E-5	13 / **159** 4.014 E-2
Broyden's update with line search	14 / **58** 2.250 E-11	8 / **37** 6.407 E-5	10 / **49** 6.124 E-2

➔LM optimization requires fewer function evaluations with Broyden's update than with finite difference Jacobian.

Stellarator Optimization with Broyden Update

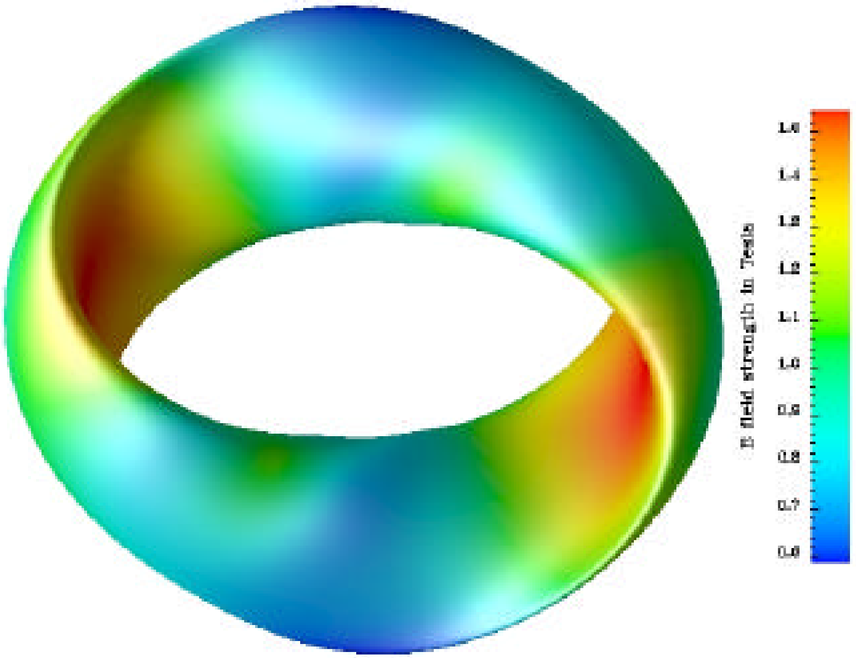
- The Broyden update method has been tested on stellarator optimization as part of the QOS program
  (see Poster 3B51, this meeting)
- Optimizing a compact stellarator for good confinement and stability properties

Reference Configuration

- Properties of the reference case:
  A=2.6 β=2.5% $<|B|>$=1.0 T
  Tor. Current = 68 kA

- Outer surface:
  (color indicates |B|)

- Large χ^2 due to being ballooning unstable
  (χ^2 =1.90E+06)



Controlling the Number of Parameters

- Four cases were tested:
  - p12: Vary only the current profile
  - p21: Vary the current & pressure profiles
  - p30: Vary the boundary Fourier coefficients
  - p39: Vary the boundary Fourier coefficients and the pressure profile

Results for Stellarator Optimization

Case	p12	p21	p30	p39
Init. Chi-sq	1.41E+6	2.63E+6	1.50E+6	2.71E+6
End Chi-sq	2.70E+3	2.44E+3	7.54E+4	7.82E+4
Function eval., FD Jacobian	105	187	421	335
Function eval., Broyden update	80	71	86	83

Conclusions

- Number of function evaluations required for typical stellarator optimization is reduced, when Broyden update is used in combination with Jacobian resets.
- Resetting the Jacobian with finite difference approximation is necessary whenever progress criteria are no longer met.

Future Goals

- Replace golden section line search by parabolic (Brent's method [2,10]) interpolation to reduce function evaluations in search loop.
- Introduce multi-processing within the function evaluations (VMEC software), while using single-processing in LM procedure.

References

- 1. Bittner, L. (1959): Eine Verallgemeinerung des Sekantenverfahrens (regula falsi) zur naeherungsweisen Berechnung der Nullstellen eines nichtlinearen Gleichungssystems. Wiss. Zeit. der Technischen Hochschule Dresden **9**, 325.
- 2. Brent, R.P. (1973): Algorithms For Minimization Without Derivatives. Prentice-Hall, Englewood Cliffs, NJ.
- 3. Broyden, C.G. (1965): A class of methods for solving nonlinear simultaneous equations. Math. Comp. **19**, 577-593.
- 4. Broyden, C.G. (2000): On the discovery of the “good Broyden” method. Math. Program., Ser. B **87**: 209-213.
- 5. Broyden, C.G. (2001): private communication.
- 6. Dennis, J.E., Jr., Schnabel, R.B. (1983): Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Englewood Cliffs, NJ.

References, Continued

- 7. Levenberg, K. (1944): A method for the solution of certain problems in least squares. SIAM J. Numer. Anal. **16**, 588-604.
- 8. Marquardt, D. (1963): An algorithm for least squares estimation of nonlinear parameters. SIAM J. Appl. Math. **11**, 431-441.
- 9. More, J.J., Garbow, B.S., Hillstrom, K.E. (1981): Testing unconstrained optimization software. TOMS **7**, 17-41.
- 10. Press, W.H., Teukolsky, S.A., Vetterling, W.T., Flannery, B.P. (1992): Numerical Recipes in FORTRAN. The Art of Scientific Computing, Second Edition. Cambridge University Press, Cambridge.
- 11. Schnabel, R.B. (2001): private communication.
- 12. Wolfe, P. (1959): The secant method for simultaneous nonlinear equations. Comm. Assoc. Comp. Mach. **2** (12), 12-13.