Math. Program., Ser. B 87: 209–213 (2000)

Digital Object Identifier (DOI) 10.1007/s101070000155

C.G. Broyden

On the discovery of the "good Broyden" method

Dedicated to Dr. William C. Davidon to commemorate his **70**th birthday

Received: February 5, 2000 / Accepted: February 19, 2000
Published online March 15, 2000 –

Abstract. This short note traces the events that led to the unsymmetric rank one formula known as the "good Broyden" update [5, 6], which is widely used within derivative-free mathematical software for solving a system of nonlinear equations.

I left University in 1956 with an indifferent degree in Physics and took up a post with the English Electric Company in Leicester. The company was involved in the design and construction of nuclear reactors and I was employed as a computer programmer.

One of the problems with which I was concerned was the solution of systems of differential equations. These equations were used in the performance calculations of the reactors and modelled their behaviour as time passed, and some of the favoured methods of solution were the "predictor-corrector" methods. They worked roughly as follows. Given a temperature distribution at time t they would calculate the equivalent distribution at time $t + \delta t$ where δt was a small time increment. This calculation would then be repeated many times in the hope that the computed temperature would subside before the Magnox cladding of the fuel rods melted.

The new temperature distribution would be obtained by getting an approximate solution of the "corrector" equations, which were essentially of the form

$$\mathbf{x} = \mathbf{g}(\mathbf{x}) \tag{1}$$

where $\mathbf{x} \in \mathbb{R}^n$ represented unknown temperatures and $\mathbf{g}(\mathbf{x})$ was some non-linear function. Since it was nonlinear, equation (1) could only be solved iteratively and the job of the "predictor" equations was to give an explicit (and usually not too accurate) initial approximation to be subsequently refined by the correctors. Equation (1) suggests the iterative procedure

$$\mathbf{x}_{i+1} = \mathbf{g}(\mathbf{x}_i),$$

but often only one correction step was applied so that the solution obtained was none too accurate either. However, as computers and software improved and people became more ambitious, more and more correction steps were taken and soon we were thinking in

C.G. Broyden: Universitá di Bologna, Facoltá di Scienze MM.FF.NN., Via Sacchi N.3, 47023 Cesena (FO), Italy. e-mail: broyden@CsR.UniBo.It

terms of solving (1) exactly, or at least as exactly as one can solve nonlinear equations. The chosen method was, of course, to set $\mathbf{f}(\mathbf{x}) \equiv \mathbf{x} - \mathbf{g}(\mathbf{x})$ and then to solve $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ using Newton's method.

Now the problem with Newton's method is that it needs the Jacobian of $\mathbf{f}(\mathbf{x})$ to be calculated for every iteration, but in this case the inefficiency was worse because of the way it was to be used. Not only did the equations have to be solved for t but also for $t + \delta t,\ t + 2\delta t,\ t + 3\delta t$ etc. and the Jacobian did not change all that much when going from one solution to the next. It did seem to be rather a waste to keep recomputing it so I began to look at ways of short-cutting this, perhaps by using an approximation to the Jacobian.

In an earlier draft of this note I wrote at this point "I then came across a paper [1] written in the Computer Journal by one J.G.P. Barnes on the secant method" but I now realise that this was far from the true story. I distinctly remember Barnes's paper but it could not have been an influence at the time as it and my paper [3], in which I presented my update, were both published in 1965. So what were the influences acting on me at that point? Newton's method, certainly. My paper [3] starts off with a description of this method and I had used it before in one of my programs. I also knew the 1963 Fletcher-Powell Computer Journal paper [7] but I don't remember thinking of this at the time as an approximate Newton method. Although I included the DFP update formula in [3] it had been presented in [7] as a type of conjugate gradient method and I probably thought of it as such. I actually wrote "Powell [8] regards it not as a version of Newton's method but as a conjugate gradient method" and I probably went along with that. It was not until later that I wondered how one should think of it when applied to non-quadratic functions and realised its close connection to the master algorithm. I certainly *knew of* Bill Davidon's paper [4] but probably had not seen it at the time. This was because it was published as some arcane report and my industrial-type library found it difficult to obtain such things. I did at some stage read it but at the time I probably went along with the Fletcher-Powell [7,8] view of the algorithm. So the following description of the secant method perhaps played less of a part in the derivation of the "good Broyden" update than my recollection suggests.

The secant method as described in [1] is an approximation to Newton's method, and it uses an *estimate* $\mathbf{B}$ of the Jacobian, the estimate being based on the Taylor series approximation to the vector function $\mathbf{f}(\mathbf{x})$. This is

$$\mathbf{f}(\mathbf{x} + \mathbf{s}) = \mathbf{f}(\mathbf{x}) + \mathbf{J}\mathbf{s} + \text{higher order terms}$$

where $\mathbf{J}$ is the Jacobian of $\mathbf{f}$ evaluated at $\mathbf{x}$ and $\mathbf{s}$ is an arbitrary vector of increments. If we can ignore the higher order terms (and even if we can't) the best thing we can do with our approximation $\mathbf{B}$ is to make it satisfy

$$\mathbf{f}(\mathbf{x} + \mathbf{s}) = \mathbf{f}(\mathbf{x}) + \mathbf{B}\mathbf{s},$$

an equation now generally known as the secant equation following a suggestion by John Dennis (see for example [6]). This, of course, does not define $\mathbf{B}$ precisely, merely imposes conditions upon it, but, in the context of an iteration like that of Newton, this is not too much of a problem. Suppose that the iteration to solve $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ had been going

for some time so that a lot of approximate solutions $\mathbf{x}_j$ and corresponding function values $\mathbf{f}_j$ had been computed. Suppose, also, that we had reached an iteration i, (so that we knew $\mathbf{x}_i$ and $\mathbf{f}_i$) and now wanted to determine an approximation $\mathbf{B}_i$ to the Jacobian so that we could compute the next approximation $\mathbf{x}_{i+1}$ by

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{B}_i^{-1}\mathbf{f}_i. \tag{2}$$

If we define $\mathbf{s}_j$ and $\mathbf{y}_j$ by

$$\mathbf{s}_j = \mathbf{x}_{j+1} - \mathbf{x}_j$$

and

$$\mathbf{y}_j = \mathbf{f}_{j+1} - \mathbf{f}_j$$

and then if we require $\mathbf{B}_i$ to satisfy $\mathbf{B}_i\mathbf{s}_j = \mathbf{y}_j,\ i-n \le j \le i-1$ it is not too difficult to show that provided the sets of vectors $\{\mathbf{s}_j : i-n \le j \le i-1\}$ and $\{\mathbf{y}_j : i-n \le j \le i-1\}$ are linearly independent then $\mathbf{B}_i$ is well-defined and nonsingular and $\mathbf{x}_{i+1}$ can be computed. I might state in parentheses here that the secant method for systems of equations is probably due to Gauss, but that the first computer-era references seem to be the works of Wolfe [10] and Bittner [2].

So far so good! We had an updating procedure that seemed to work provided that certain conditions of linear independence were satisfied, but the problem was that it did not work very well. In fact it proved to be quite numerically unstable. Now at that time I knew nothing of condition numbers – I may not even have known what a norm was – but it did seem to me that if, in the two-dimensional case, the vectors $\mathbf{s}_j,\ j = i-1, i-2$, were nearly parallel then there would be trouble. My unease sprang from my intuition as a physicist. If we have two rods in the horizontal plane, joined to each other by a hinge at their mid-points so that the angle between them can change, and we rest a plate on them then if the angle between them is a right-angle the plate is fairly stable. If the angle is reduced so that looking down on them the rods look like an "X" which gets thinner and thinner then the plate becomes more and more unstable until the rods coincide and the plate topples. I therefore felt that the Jacobian would be well-defined if the two vectors $\mathbf{s}_j$ were orthogonal, but would not be so well defined if the information defining it came from two virtually co-linear steps, and often in the secant method two successive steps can be almost co-linear. So what to do?

One solution canvassed at the time was to take the "best" two or more vectors $\mathbf{s}_j$ from the last m and use these to define the Jacobian, but this did not appeal to me. What is "best" and how does one choose m? Besides, it was inelegant and I thought that mathematics should be elegant, (I still do) so I reasoned thus:

- We have new information, namely $\mathbf{s}_{i-1}$ and $\mathbf{y}_{i-1}$ that we wish to incorporate into the new Jacobian $\mathbf{B}_i$, i.e. we have new information in the direction $\mathbf{s}_{i-1}$
- We can use this by making $\mathbf{B}_i$ satisfy $\mathbf{B}_i\mathbf{s}_{i-1} = \mathbf{y}_{i-1}$
- We have no new information in any direction $\mathbf{s}$ orthogonal to $\mathbf{s}_{i-1}$
- We should therefore require, if possible, that $\mathbf{B}_i\mathbf{s} = \mathbf{B}_{i-1}\mathbf{s}$ for all $\mathbf{s}$ orthogonal to $\mathbf{s}_{i-1}$, i.e. no change to $\mathbf{B}_{i-1}$ in any direction orthogonal to $\mathbf{s}_{i-1}$.

After a certain amount of scratching around I realised that this uniquely defined $\mathbf{B}_i$ to be

$$\mathbf{B}_i = \mathbf{B}_{i-1} + (\mathbf{y}_{i-1} - \mathbf{B}_{i-1}\mathbf{s}_{i-1})\frac{\mathbf{s}_{i-1}^T}{\mathbf{s}_{i-1}^T\mathbf{s}_{i-1}},$$

a formula which appeared in my paper [3] and came to be known as the "good Broyden" update[1] [5,6]. The fact that it turned out to be a rank-1 update was pure serendipity.

When I was at University they did not teach matrices to physicists. As matrices seemed to be important in my new environment, I decided to learn something about them and one of the things that I came across, or more probably derived for myself, was the Sherman-Morrison formula [9]. In those days one wrote computer programs by punching machine-code instructions in binary on Hollerith cards, and the solution of ten linear simultaneous equations was a task not to be undertaken lightly. The Sherman-Morrison formula enabled me to avoid all that by updating the inverse of $\mathbf{B}_i$, $\mathbf{H}_i$ say, by

$$\mathbf{H}_i = \mathbf{H}_{i-1} + (\mathbf{s}_{i-1} - \mathbf{H}_{i-1}\mathbf{y}_{i-1})\frac{\mathbf{s}_{i-1}^T\mathbf{H}_{i-1}}{\mathbf{s}_{i-1}^T\mathbf{H}_{i-1}\mathbf{y}_{i-1}},$$

and the new approximate solution was, from equation (2), given by

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{H}_i\mathbf{f}_i.$$

So that's it! Some good old-fashioned physical intuition allied to a bit of technique. The work appeared in [3]. And English Electric even paid me a page fee for the published paper!

Acknowledgements. The author thanks the Consiglio Nazionale di Ricerca, Italy (CNR), for partial financial support under Il Progetto Finalizzato Trasporti 2, Contract No. 93.01861.PF74. He is also happy to acknowledge the help given by Professor K.A. Ariyawansa in the preparation of this note.

References

1. Barnes, J. (1965): An algorithm for solving nonlinear equations based on the secant method. Comput. J. **8**, 66–72
2. Bittner, L. (1959): Eine Verallgemeinerung des Sekantenverfahrens (regula falsi) zur näherungsweisen Berechnung der Nullstellen eines nichtlinearen Gleichungssystems. Wissen. Zeit. der Technischen Hochschule Dresden **9**, 325
3. Broyden, C.G. (1965): A class of methods for solving nonlinear simultaneous equations. Math. Comp. **19**, 577–593
4. Davidon, W.C. (1959): Variable metric method for minimization. Report ANL-5990 (Rev.), Argonne National Laboratory, Argonne, IL
5. Dennis, J.E., Jr., Schnabel, R.B. (1981): A new derivation of symmetric positive definite secant updates. In: Mangasarian, O.L., Meyer, R.R., Robinson, S.M., eds., Nonlinear Programming 4, 1980. Academic Press, New York, NY, 167–199

[1] The formula is referred to as "good" due to its better numerical performance relative to another formula that I also presented in [3], which has become to be known as the "bad Broyden" update. See [5] for a discussion of these two updates and their relations to the DFP and BFGS updates.

6. Dennis, J.E., Jr., Schnabel, R.B. (1983): Numerical Methods for Unconstrained Optimization and Nonlinear Equations. Prentice-Hall, Inc., Englewood Cliffs, NJ
7. Fletcher, R., Powell, M.J.D. (1963): A rapidly convergent descent method for minimization. Comp. J. **6**, 163–168
8. Powell, M.J.D. (1964): An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J. **7**, 155–162
9. Sherman, J., Morrison, W.J. (1950): Adjustment of an inverse of matrix corresponding to changes in the elements of a given column or a given row of the original matrix. (Abstract of paper presented at the 12th Summer Meeting of the Institute of Mathematical Statistics, Boulder, CO, 1949) Ann. Math. Stat. **20**, 621
10. Wolfe, P. (1959): The secant method for simultaneous nonlinear equations. Comm. Assoc. Comp. Mach. **2**(12), 12–13