

On a Numerical Instability of Davidon-Like Methods

By Yonathan Bard

The Davidon-Fletcher-Powell method of function minimization [2], [3] has attained widespread popularity. Yet it goes wrong from time to time. Among the conditions reported are:

1. Broyden [1] states that negative steps had to be taken occasionally.
2. McCormick [5] noted that reinitialization of the matrix every now and then improved the method's performance.
3. Wolfe [6] has reported cases of convergence to nonstationary points.

The author has encountered similar behavior in his own work, and has found it invariably the result of the matrix turning singular, due to the cause detailed below. The author believes other workers' difficulties probably originate in the same cause.

We wish to find the minimum of a continuously differentiable function $F(\mathbf{x})$ (boldface lower case letters denote vectors). In Davidon's method we proceed iteratively: If \mathbf{x}_i is the value of \mathbf{x} at the i th iteration, then

$$(1) \quad \mathbf{x}_{i+1} = \mathbf{x}_i - \lambda_i \mathbf{H}_i \mathbf{g}_i$$

where \mathbf{g}_i is the gradient of F at $\mathbf{x} = \mathbf{x}_i$, \mathbf{H}_i is a matrix to be defined below, and λ_i is a scalar chosen so as to minimize $F(\mathbf{x})$ along the chosen direction. \mathbf{H}_i is defined iteratively by:

$$(2) \quad \mathbf{H}_{i+1} = \mathbf{H}_i - \frac{\mathbf{H}_i \mathbf{y}_i \mathbf{y}_i^T \mathbf{H}_i}{\mathbf{y}_i^T \mathbf{H}_i \mathbf{y}_i} + \frac{\mathbf{p}_i \mathbf{p}_i^T}{\mathbf{p}_i^T \mathbf{y}_i} = \mathbf{H}_i - \mathbf{A}_i + \mathbf{B}_i \text{ with}$$

$$(3) \quad \mathbf{y}_i = \mathbf{g}_{i+1} - \mathbf{g}_i,$$

$$(4) \quad \mathbf{p}_i = \mathbf{x}_{i+1} - \mathbf{x}_i.$$

The process is started at some arbitrary initial point $\mathbf{x} = \mathbf{x}_0$, and with some arbitrary symmetric positive-definite matrix \mathbf{H}_0 . A common choice is $\mathbf{H}_0 = \mathbf{I}$ (the identity matrix). Fletcher and Powell [3] prove that under these conditions, all \mathbf{H}_i are positive definite, and hence, as long as $\mathbf{g}_i \neq \mathbf{0}$, a point \mathbf{x}_{i+1} can always be found such that $F(\mathbf{x}_{i+1}) < F(\mathbf{x}_i)$. This property is termed "stability" by Fletcher and Powell, as well as by Broyden [1].

Suppose we indeed choose $\mathbf{H}_0 = \mathbf{I}$. The elements of \mathbf{H}_0 will be of the order of magnitude of unity, as will be the elements of \mathbf{A}_0 . The elements of \mathbf{B}_0 , on the other hand, are of the order of magnitude of $\|\mathbf{p}_0\|/\|\mathbf{y}_0\|$, which may be anything: suppose we seek to minimize bF (b a positive constant) in place of F . This leaves \mathbf{x} and \mathbf{p} unchanged, but \mathbf{g} and \mathbf{y} are multiplied by b . Hence, all elements of \mathbf{B}_0 will be scaled by $1/b$. Or, suppose we leave the value of F unchanged, but rescale \mathbf{x} by a factor a . Then \mathbf{p} will also be scaled by a , whereas \mathbf{g} and \mathbf{y} are scaled by $1/a$, and the elements of \mathbf{B}_0 are scaled by a^2 . The magnitude of these elements, thus, depends on the scales

Received December 14, 1967.

chosen for F and \mathbf{x} . In particular, if one was unfortunate enough to choose a scaling for which $\|\mathbf{y}_0\| \gg \|\mathbf{p}_0\|$, the elements of \mathbf{B}_0 will be very small compared to those of $\mathbf{H}_0 - \mathbf{A}_0$, and we will have, approximately:

$$(5) \quad \mathbf{H}_1 \approx \mathbf{H}_0 - \mathbf{A}_0 = \mathbf{H}_0 - \frac{\mathbf{H}_0 \mathbf{y}_0 \mathbf{y}_0^T \mathbf{H}_0}{\mathbf{y}_0^T \mathbf{H}_0 \mathbf{y}_0}.$$

This is a singular matrix, having the null vector \mathbf{y}_0 . Conversely, if $\|\mathbf{y}_0\| \ll \|\mathbf{p}_0\|$, the matrix \mathbf{B}_0 will dominate $\mathbf{H}_0 - \mathbf{A}_0$, and we will have:

$$(6) \quad \mathbf{H}_1 \approx \mathbf{B}_0 = \frac{\mathbf{p}_0 \mathbf{p}_0^T}{\mathbf{p}_0^T \mathbf{y}_0}.$$

Again, \mathbf{H}_1 is singular, being of rank one.

Once an \mathbf{H}_i has turned singular, there is virtually no hope of recovery. Let \mathbf{z} be a null vector of \mathbf{H}_i . Then both \mathbf{z} and \mathbf{y}_i will be null vectors of $\mathbf{H}_i - \mathbf{A}_i$, which (except in the improbable case $\mathbf{z} = k\mathbf{y}_i$) will be at most of rank $n - 2$ (n being the dimensionality of \mathbf{x}). Now \mathbf{B}_i has rank 1, and the rank of $\mathbf{H}_{i+1} = \mathbf{H}_i - \mathbf{A}_i + \mathbf{B}_i$ cannot exceed $n - 1$. Thus, if \mathbf{H}_1 is singular, all subsequent \mathbf{H}_i are likely to be so.

It must be observed that the singularity is only approximate. However, if we carry m significant digits, and if $\|\mathbf{p}_0\|/\|\mathbf{y}_0\| \approx 10^{-m}$ or 10^m , the matrices will be singular to the precision of the calculations.

The problem may be overcome in at least two ways:

1. Increase the number of digits carried, e.g. by using double precision.
2. Scale the variables appropriately. The initial scaling should make the diagonal elements of \mathbf{B}_0 approximately unity. Sometimes the character of the function changes drastically from one region to another: the author has encountered a case where $\|\mathbf{g}_0\| \approx 10^7$, but after taking a step of approximately unit length (i.e. $\|\mathbf{p}_0\| \approx 1$), $\|\mathbf{g}_1\| \approx 1$. In such a case, the initial scaling is wrong for subsequent iterations. The remedy seems to be a rescaling of \mathbf{x} and reinitialization of \mathbf{H} whenever the minimization process seems to bog down at a nonstationary point.

The same difficulties appear in other Davidon-like methods, such as those proposed by Broyden [1] and Greenstadt [4].

Acknowledgement. The author wishes to thank the reviewer for pointing out that scaling F (and not only \mathbf{x}) affects the matrix \mathbf{B}_0 .

International Business Machines, Inc.
590 Madison Avenue
New York, New York 10022

1. C. G. BROYDEN, "Quasi-Newton methods and their application to function minimization," *Math. Comp.*, v. 21, 1967, pp. 368-381.
2. W. C. DAVIDON, *Variable Metric Methods for Minimization*, A.E.C. Research and Development Report ANL-5990 (Rev. TID-4500, 14th ed.), 1959.
3. R. FLETCHER & M. J. D. POWELL, "A rapidly convergent descent method for minimization," *Comput. J.*, v. 6, 1963/1964, pp. 163-168. MR 27 #2096.
4. J. GREENSTADT, *Variations on Variable-Metric Methods*, IBM New York Scientific Center Technical Report No. 320-2901, 1967.
5. G. McCORMICK, Informal remarks on Unconstrained Optimization, SIGMAP Workshop, New York, 1967.
6. P. WOLFE, Informal remarks on Unconstrained Optimization, SIGMAP Workshop, New York, 1967.