RANK MODIFICATIONS OF SEMI-DEFINITE MATRICES WITH APPLICATIONS TO SECANT UPDATES

MOODY T. CHU*, R. E. FUNDERLIC† AND GENE H. GOLUB‡

Abstract. The BFGS and DFP updates are perhaps the most successful Hessian and inverse Hessian approximations respectively for unconstrained minimization problems. This paper describes these methods in terms of two successive steps: rank reduction and rank restoration. From rank subtractivity and a powerful spectral result, the first step must necessarily result in a positive semi-definite matrix; and the second step is designed to restore positive definiteness. The goal of the research is to better understand the workings of the BFGS and DFP updates to see how they may be modified and yet retain their basic rank and spectral characteristics. The class of BFGS and DFP updates is generalized both in terms of choices for update vectors and rank of the modifications in the formulas. The rank restoration step generalizes naturally to rectangular matrices.

Key words. Rank-one reduction, Wedderburn theorem, BFGS update, DFP update, quasi-Newton methods, rank subtractivity, rank additivity,

AMS(MOS) subject classifications. 65F30,65K05,49D15

1. Introduction.

A typical way of solving the problem

$$\min_{x \in R^n} f(x) \tag{1}$$

where $f : R^n \longrightarrow R$ is twice continuously differentiable is to apply Newton's method to the system of equations

$$g(x) := \nabla f(x) = 0. \tag{2}$$

Even when modified to ensure global convergence, however, Newton's method suffers from a main disadvantage in that the Jacobian $G(x)$ of $g(x)$ (i.e., the Hessian of $f(x)$) is often unavailable or very expensive to compute. For this reason, considerable efforts have been made to develop cheap and reasonable approximations to either $G(x)$ or its inverse, hence the introduction of the so called quasi-Newton methods.

Since the Hessian is always symmetric and often positive definite (especially near the optimal solution), it is important that an approximate Hessian should also possess these properties. Among existing secant methods, perhaps the two most successful Hessian and inverse Hessian approximations in the literature that preserve these properties are the BFGS update and the DFP update [5, 8]. In this paper we consider these methods as consisting of two successive steps—rank reduction and rank restoration. From this viewpoint, we develop general conditions on how the rank modification should be carried

* Department of Mathematics, North Carolina State University, Raleigh, North Carolina 27695-8205. This research was supported in part by National Science Foundation under grant DMS-9422280.

† Department of Computer Science, North Carolina State University, Raleigh, North Carolina 27695-8206.

‡ Department of Computer Science, Stanford University, Stanford, CA 94305. This research was supported in part by National Science Foundation under grant CCR-9505393.

out to maintain the positive semi-definiteness. This development generalizes the class of BFGS and DFP updates both in terms of choices for update vectors and rank of the modifications.

For the purposes of this paper we quickly recapitulate the updates involved in the BFGS and DFP formulas: Let x_c and x_+ denote, respectively, the current and the next approximate of the optimal solution. The secant equation for the approximation H_+ of $G(x_+)$ is

$$H_+ s_c = y_c \tag{3}$$

where

$$\begin{aligned} s_c &:= x_+ - x_c \\ y_c &:= g(x_+) - g(x_c). \end{aligned}$$

The BFGS update H_+ from H_c is given by

$$H_+ := H_c - \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c} + \frac{y_c y_c^T}{y_c^T s_c}. \tag{4}$$

On the other hand, the secant equation for the approximation K_+ of $G(x_+)^{-1}$ is

$$K_+ y_c = s_c, \tag{5}$$

and the corresponding DFP update K_+ from K_c is given by

$$K_+ := K_c - \frac{K_c y_c y_c^T K_c}{y_c^T K_c y_c} + \frac{s_c s_c^T}{s_c^T y_c}. \tag{6}$$

It is interesting to note the dual relationship between (4) and (6) with the interchanges $H \leftrightarrow K$ and $y_c \leftrightarrow s_c$. Since only the secant equation and the format of the update matter in our discussion, the duality enables us not to differentiate between BFGS and DFP; however, in practice, a general consensus is that the BFGS update performs better than the DFP update. Because of the duality we will refer only to BFGS in the sequel. Equation (4) is the focus of this paper.

We use the terminology that the *rank of* a matrix H *is reduced* or that a matrix M *reduces the rank of* H to mean that $H - M$ has rank less than that of H. One of our primary themes will be to show that the (subtractive) second term of the right side of (4) precisely reduces the rank of H_c by one while the (additive) third term of the right side precisely restores the rank by one. In order to better understand the BFGS update formula in terms of these two steps of reduction and restoration we use the notation

$$H - M + P,$$

where the subtraction $H - M$ denotes the *rank reduction* first *step* of the BFGS update, and the addition $(H - M) + P$ denotes the *rank restoration* second *step* of the update. We will see that any matrix of the form

$$M = Hs(s^T H s)^{-1} s^T H \tag{7}$$

when subtracted from H reduces the rank of H by one. Conversely, if the rank of H is symmetrically reduced by one, then the reducing matrix must be of the form (7). Moreover, we show that $H - M$ is positive semi-definite if H is positive semi-definite. The rank restoration step of adding a matrix P of the form $\lambda^{-1}yy^T$ to $H - M$ is characterized and analyzed for the conditions on λ and y to be effective. For this rank restoration we show exactly how the scalar λ and the vector y must relate to M. Also necessary and sufficient conditions are given for maintaining positive definiteness. The underlying rationale for the rank one steps of rank reduction and restoration generalizes naturally to matrices M and P of rank higher than one. For the generalization, rather than y and s vectors, we use upper case letters Y and S for matrices of rank(M) columns. The BFGS and DFP formulas motivated the research of this paper. It is our goal always to relate our results to these formulas even though the setting is generalized to rank greater than one.

Any symmetric update with rank at most two of a symmetric positive definite matrix H_c is of the form:

$$H_+ = H_c + \begin{bmatrix} y & z \end{bmatrix} \begin{bmatrix} p & q \\ q & r \end{bmatrix} \begin{bmatrix} y^T \\ z^T \end{bmatrix} \tag{8}$$

where y, z are vectors and p, q, r are scalars. In an earlier paper, Brodlie et al. [2] studied a special case of (8), i.e.,

$$H_+ = H_c + \begin{bmatrix} u & H_c v \end{bmatrix} \begin{bmatrix} v^T H_c v & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} u^T \\ (H_c v)^T \end{bmatrix}. \tag{9}$$

Their strategy expresses the update in product form

$$H_+ = (I + uv^T) H_c (I + uv^T)^T$$

thereby maintaining the positive semi-definiteness. Although we also specialize our rank two update, we do so by constraining one part of the update to come from a general rank one subtractivity update. From our general, but constrained, setting we answer how to complete the rank two update to arrive again at a positive definite matrix. We show how to do this in complete generality with the BFGS update inferred as a special case. In addition (apparently in contrast to the approach of (8)), the dissection into rank reduction and restoration naturally generalizes to higher rank cases.

We begin in §2 with a brief discussion of several known results used as tools to help understand the phenomena of rank reduction and restoration observed in the BFGS update formula. In particular, we review the Wedderburn rank reduction formula and its generalization. These rank subtractivity characterizations, i.e., rank($H - M$) = rank(H) − rank(M), are important for both the rank reduction and rank restoration steps. The rank subtractivity characterizations are followed by a rank addition characterization, i.e., rank($A + P$) = rank(A) + rank(P), that is key to understanding the restoration step. Finally a spectral result due to Weyl [13] and a simple consequence show that the reduction step results in a positive semi-definite matrix. In §3 we present our main results. We begin with a general rank reduction result for symmetric positive definite matrices followed by a general rank-restoration theorem for rectangular matrices and a symmetric rank-restoration theorem directly applicable to the BFGS update. We conclude with a characterization of maintaining positive definiteness with its application to the BFGS update. All our higher rank generalizations are applicable to BFGS as a special case. Finally, in §4 we summarize our results and point to future work that would include the simultaneous use of multiple vectors for quasi-Newton methods.

2. Preliminaries. We begin with a simple, but far reaching, result first proved by Wedderburn [12, p.69].

LEMMA 2.1. *If $x \in R^n$ and $y \in R^m$ are vectors such that $\omega = y^T Ax \neq 0$, then the matrix*

$$B := A - \omega^{-1} Axy^T A \tag{10}$$

has rank exactly one less than the rank of A.

Egerváry [7], though apparently unaware of Wedderburn's result, proved the complete characterization of rank one subtractivity:

LEMMA 2.2. *Let $u \in R^m$ and $v \in R^n$. Then the rank of the matrix $B = A - \sigma^{-1}uv^T$ is less than that of A if and only if there are vectors $x \in R^n$ and $y \in R^m$ such that $u = Ax$, $v = A^T y$ and $\sigma = y^T Ax$, in which case* $\operatorname{rank}(B) = \operatorname{rank}(A) - 1$.

The Wedderburn rank-one reduction formula (10) has led to a general matrix factorization process (e.g., the LDU and QR decompositions, the Lanczos algorithm and the singular value decomposition are special cases [3]). The following generalization [4] of Lemma 2.2 and its symmetric analogue are key in generalizing the BFGS method to higher rank modifications. We use the terminology *full rank factorization* [1] of a matrix A to mean any factorization $X\Lambda Y^T$ of A such that X and Y each have $\operatorname{rank}(A)$ linearly independent columns and Λ is nonsingular. One way to show such factorizations always exist is to consider an ordered singular value factorization. The diagonal matrix can be reduced to having only nonzero singular values, and the appropriate last rows or columns can be deleted from the other factors.

LEMMA 2.3. (General Rank-Subtractivity Lemma) *Let $UR^{-1}V^T$ be a given full rank factorization of a given matrix M. Then* $\operatorname{rank}(A - M) = \operatorname{rank}(A) - \operatorname{rank}(M)$ *if and only if $U = AX$, $V^T = Y^T A$ and $Y^T AX = R$ for some matrices X and Y.*

Though Lemma 2.3 is a straightforward generalization of Lemma 2.2, both results include somewhat subtle implicit considerations of size and rank. For example, X, Y, U and V each have full column rank equal to $\operatorname{rank}(M) = \operatorname{rank}(R)$ since $Y^T AX = R$ and $UR^{-1}V^T$ is a full rank decomposition of M.

The symmetric analogue, Lemma 2.4, of Lemma 2.3 seems to have been overlooked in the various papers on rank subtractivity. This symmetric version is used for both steps of the BFGS method and their generalizations. The rank one case of Lemma 2.4 not only shows that the first step of BFGS always reduces the rank of H_c by one, but the converse shows that if the first step reduces the rank of H_c by one, then the matrix that effected the reduction must be of the form $M = Hs(s^T Hs)^{-1}s^T H$.

LEMMA 2.4. (Symmetric Rank-Subtractivity Lemma) *Let H and M be symmetric matrices. Then* $\mathrm{rank}(H - M) = \mathrm{rank}(H) - \mathrm{rank}(M)$ *if and only if there is a matrix S such that $M = HS(S^THS)^{-1}S^TH$*

Proof. We first prove the *if* part. By the assumption that S^THS is invertible, the form of M is a full column rank factorization. Lemma 2.3 immediately gives rank subtractivity.

Now we prove the *only if* part. Due to its symmetry, a full rank decomposition of M of the form

$$M = UR^{-1}U^T \tag{11}$$

always exists. By Lemma 2.3, there are matrices S and Y such that $U = HS$, $U^T = Y^TH = S^TH^T = S^TH$, and $Y^THS = R$. But $Y^TH = S^TH$ so that $R = S^THS$. Substituting in (11) gives

$$M = HS(S^THS)^{-1}S^TH.$$

The assertion is proved. $\square$

The following result from [4, Theorem 3.1] is the key to understanding and generalizing the rank restoration step of the BFGS method and its resulting nonsingular matrix.

LEMMA 2.5. (General Rank-Additivity Lemma) *Let A and P be any matrices that are additively compatible. Then rank is additive,*

$$\mathrm{rank}(A + P) = \mathrm{rank}(A) + \mathrm{rank}(P),$$

if and only if there is a matrix B such that $AB = 0$, $BA = 0$ and $PBP = P$.

This rank additivity result can be understood quite intuitively and reasonably, e.g., let $A := diag(1, 0, 0)$ and $P := diag(0, 0, 2)$. Then a sufficient matrix B is $diag(0, 0, \frac{1}{2})$ as B annihilates A on both sides and is an appropriate generalized inverse for P. If P were rather $diag(1, 0, 2)$, then there is no matrix B that both annihilates A and (generalized) inverts P

The next result appears in [10, p. 70], and Stewart [11, p. 203] traced the result back to a 1912 paper by Weyl [13] and showed its relation to the Wielandt-Hoffman Theorem.

LEMMA 2.6. (Weyl) *Let H and B be symmetric and the eigenvalues of H, B and $H + B$ be denoted by η_i, β_i and λ_i respectively with orderings*

$$\begin{aligned} \eta_n &\le \eta_{n-1} \le \ldots \le \eta_1, \\ \beta_n &\le \beta_{n-1} \le \ldots \le \beta_1, \\ \lambda_n &\le \lambda_{n-1} \le \ldots \le \lambda_1. \end{aligned}$$

Then

$$\eta_n \le \lambda_i - \beta_i \le \eta_1 \tag{12}$$

for $i = 1, \ldots, n$.

We now give a simple consequence of Weyl's lemma, Lemma 2.6, that is important for the rank reduction step of the BFGS method.

LEMMA 2.7. *If $H \in R^{n\times n}$ is symmetric positive definite and B is symmetric, then $H + B$ has at least $n - \operatorname{rank}(B)$ positive eigenvalues.*

Proof. Arrange all the eigenvalues as in Weyl's lemma, Lemma 2.6. Observe that the smallest eigenvalue η_n of H is positive. Also exactly $n - \operatorname{rank}(B)$ eigenvalues β_i, say, $i = r, r-1, \ldots, r-(n-\operatorname{rank}(B)-1)$ for a certain r, of B are zero. It follows from (12) that at least these eigenvalues λ_i, $i = r, r-1, \ldots, r-(n-\operatorname{rank}(B)-1)$, of $H + B$ must be positive. □

3. Main Results. In this section we discuss four results generalizing the two steps of of the BFGS formula. The first theorem with $\operatorname{rank}(M) = 1$ along with the Wedderburn rank-one reduction formula (10) guarantees that if H_c is positive definite, then the first step of BFGS results in a positive semi-definite matrix with one zero eigenvalue.

THEOREM 3.1. (Rank Reduction Theorem) *Suppose that H is symmetric positive semi-definite, M is symmetric and* $\operatorname{rank}(H - M) = \operatorname{rank}(H) - \operatorname{rank}(M)$. *Then $H - M$ is positive semi-definite.*

Proof. We first prove the theorem when H is positive definite. By letting the matrix $-M$ take the role of B in Lemma 2.7, $H - M$ must have at least $n - \operatorname{rank}(M)$ positive eigenvalues. From the hypothesis that $H - M$ is rank subtractive, $H - M$ has $\operatorname{rank}(M)$ zero eigenvalues. Thus we have accounted for all the eigenvalues of $H - M$, and they are all either positive or zero. Thus $H - M$ is positive semi-definite.

Suppose now H is only semi-definite. We can diagonalize H so that any rank deficiency is depicted by trailing zeros in the lower-right block of the diagonal matrix $D = U^T HU$. The rank subtractivity hypothesis and Lemma 2.4 ensure that

$$M = HS(S^T HS)^{-1}S^T H$$

for some matrix S. The inertia of

$$H - M = H - HS(S^T HS)^{-1}S^T H$$

is equivalent to that of

$$D - DZ(Z^T DZ)^{-1}Z^T D$$

where $U^T S = Z$. Note that the zero structure in D eliminates any effect from Z in the lower-right corner of D. Thus it suffices to consider the leading nonzero portion of D which is positive definite. □

We can consider the rank subtractivity, Lemma 2.3, and rank additivity, Lemma 2.5, lemmas as characterizations of rank reduction and restoration for an arbitrary matrix A. We now characterize rank restoration of any matrix $A := H - M$ such that A's rank is subtractive, i.e., $\operatorname{rank}(H - M) = \operatorname{rank}(H) - \operatorname{rank}(M)$. Though this general result is not necessary for BFGS, we would be remiss to omit it because it is almost identical to the BFGS symmetric restoration result that is given immediately after the general result. In addition, it is possible the future may bring application for nonsymmetric or the rectangular matrices, (e.g., Gerber and Luk [9] make use of rectangular matrices for Broyden methods and Deuflhard and Freund [6] develop secant methods with nonsymmetric matrices).

THEOREM 3.2. (General Rank-Restoration Theorem) *Let* $\operatorname{rank}(H-M) = \operatorname{rank}(H) - \operatorname{rank}(M)$*; therefore, by the general rank-subtractivity lemma, Lemma 2.3,*

$$M = HS_2(S_1^T HS_2)^{-1} S_1^T H$$

for some matrices S_1 *and* S_2*. Let* P *be a matrix of* $\operatorname{rank}(M)$*, and let* $P = Y_1 \Lambda^{-1} Y_2^T$ *be a full rank decomposition. Then the following results hold:*

1. *If* $Y_2^T S_2$ *and* $Y_1^T S_1$ *are nonsingular, then the rank is restored, i.e.,*

$$\operatorname{rank}[(H-M)+P] = \operatorname{rank}(H). \tag{13}$$

2. *Conversely, if* H *has full column or row rank and (13) holds, then* $Y_2^T S_2$ *or* $Y_1^T S_1$ *are nonsingular respectively.*

Proof. To prove the first part of the theorem, define

$$A := H - M = H - HS_2(S_1^T HS_2)^{-1} S_1^T H,$$

and

$$B := S_2(Y_2^T S_2)^{-1} \Lambda (S_1^T Y_1)^{-1} S_1^T.$$

Clearly both $AS_2 = 0$ and $S_1^T A = 0$; therefore $AB = 0$ and $BA = 0$. Furthermore, by direct verification, $PBP = P$. Thus from the general rank additivity lemma, Lemma 2.5 we have

$$\operatorname{rank}[(H-M)+P] = \operatorname{rank}(H-M) + \operatorname{rank}(P) = \operatorname{rank}(H) - \operatorname{rank}(M) + \operatorname{rank}(P).$$

Note that

$$\operatorname{rank}(P) = \operatorname{rank}(Y_i) = \operatorname{rank}(Y_i^T S_i) = \operatorname{rank}(S_1^T HS_2) = \operatorname{rank}(M), i = 1, 2,$$

where the first and the last equalities are due to the full rank decomposition, while the second and the third equalities follow from the non-singularity of $Y_i^T S_i$ and $S_1^T HS_2$. The assertion (13) therefore is proved.

We prove the second part of the theorem by contradiction. Suppose H has full column rank and $Y_2^T S_2$ is singular. Then there would be a nonzero vector x such that $S_2 x \neq 0$ and $Y_2^T S_2 x = 0$. But then $(H - M + P)S_2 x = 0$, a contradiction to $H - M + P$ having full column rank since $\operatorname{rank}(H - M + P) = \operatorname{rank}(H)$. The proof of the full row case is similar. □

We now give the symmetric rank restoration theorem which for $\operatorname{rank}(M) = 1$ provides considerable understanding of the second step of BFGS. Recall that the first step, the reduction step, involves a rank one matrix M which reduces the rank of H

by one. The second step, the restoration step of BFGS, restores the rank of $H - M$ to that of H. In the BFGS case, the restoration results in a nonsingular matrix. Along with the maintenance of positive definiteness we will discuss this further at the end of the section specifically using the notation of BFGS (4).

Theorem 3.3. (Symmetric Rank-Restoration Theorem) *Suppose H and M are symmetric matrices satisfying* $\operatorname{rank}(H - M) = \operatorname{rank}(H) - \operatorname{rank}(M)$*; therefore, by the symmetric rank subtractivity lemma, Lemma 2.4,*

$$M = HS(S^T HS)^{-1} S^T H$$

for some matrix S. Let P be a symmetric matrix of $\operatorname{rank}(S)$*, and let $P = Y\Lambda^{-1}Y^T$ be a full rank decomposition. Then the following results hold:*

1. *If $Y^T S$ is nonsingular, then the rank is restored, i.e.,*

$$\operatorname{rank}[(H - M) + P] = \operatorname{rank}(H). \tag{14}$$

2. *Conversely, if H is nonsingular and (14) holds, then $Y^T S$ is nonsingular.*

Proof. The proof is immediate from the general rank-restoration theorem, Theorem 3.2, with the choice $S := S_1 = S_2$. □

Notice in Lemma 3.3 that the restoration step (adding a matrix P) specifically presupposes a reduction step (subtracting a matrix M) that has a necessary structure. The actual relationship of the restoration step to the reduction step is particularly important. Just as M has a factor S, P has a factor Y. For rank to be restored, Y and S must be related in the fundamental way of the symmetric restoration theorem.

It is important to note that the assumption of H being nonsingular (or having full column or row rank in the more general restoration theorem) cannot be weakened in the second part of the above theorem. To see this, use the notation of e_i to denote principal axis vectors, and let $H = diag(1, 1, 0) = e_1e_1^T + e_2e_2^T$, $M = diag(1, 0, 0) = e_1e_1^T$ and $P = diag(0, 0, 1) = e_3e_3^T$. These choices result in a successful rank restoration to $H - M + P = diag(1, 0, 1)$, i.e., $\operatorname{rank}(H - M + P) = \operatorname{rank}(H) = 2$; however, $y^T s = e_3^T e_1 = 0$, a singular matrix of order one.

The main concern for application of the BFGS process is not only the rank restoration, but also the maintenance of positive definiteness. Toward this end, we provide the next result which completely characterizes the *generalization* of the BFGS process, the main goal of this paper. After this corollary we discuss its application to the basic BFGS process.

Corollary 3.4. *Let H, M, P, S and Y be the same as in Theorem 3.3. Suppose further that H is positive definite. Then $H - M + P$ is positive definite if and only if P is positive semi-definite and $Y^T S$ is nonsingular.*

Proof. We first prove the *if* part. Note that the non-singularity of $Y^T S$ already implies the non-singularity of $H - M + P$ from the first part of the symmetric restoration theorem, Theorem 3.3. To infer that $H - M + P$ is positive definite we need only determine that it is positive semi-definite. We know from the reduction theorem, Theorem 3.1, that $H - M$ is positive semi-definite. Together with the assumption that P is positive semi-definite, $H - M + P$ is positive definite.

We now prove the *only if* part. First, $Y^T S$ is nonsingular by the second part of the symmetric restoration theorem, Theorem 3.3, since $H - M + P$ has full rank; therefore

$$\operatorname{rank}(P) = \operatorname{rank}(H - M + P) - \operatorname{rank}(H - M).$$

By Theorem 3.1 where the role of H is played by $H - M + P$ and M by $H - M$, we conclude that P is positive semi-definite. $\square$

We now consider the ramifications of the symmetric restoration theorem for the BFGS method. Recall that the symmetric Wedderburn rank-subtractive result, Lemma 2.4, guarantees that the rank of H_c is reduced by one in the reduction step of the BFGS formula (4) (repeated here):

$$H_+ := H_c - \frac{H_c s_c s_c^T H_c}{s_c^T H_c s_c} + \frac{y_c y_c^T}{y_c^T s_c}.$$

Thus for the BFGS formula to be well-defined $s_c^T H_c s_c \neq 0$ and $y_c^T s_c \neq 0$ and the latter inequality means that $Y^T S$ is nonsingular in the symmetric restoration theorem. Thus, the BFGS formula, composed of its reduction and restoration steps yields a symmetric nonsingular matrix. Moreover, H_+ is positive definite if and only if $(y_c^T s_c)^{-1} y_c y_c^T$ is positive semi-definite which is equivalent to $y_c^T s_c > 0$.

In summary, if H_c is positive definite and s_c and y_c are any vectors such that $y_c^T s_c > 0$, then $s_c^T H_c s_c > 0$, and P is positive semi-definite so that H_+ is positive definite. Conversely, if there are vectors s_c and y_c for which H_+ is positive definite, then necessarily $y_c^T s_c > 0$. Though this converse and the higher rank characterizations given in the general and symmetric restoration theorems are apparently new, there has been previous work on maintenance of positive definiteness associated with the additive term of H_+. R. Fletcher ([8] , Theorem 3.2.2, page 54) used a different approach to show that $y_c^T s_c > 0$ implies the updated Hessian is positive definite.

4. Summary and Conclusions. Upon examining the BFGS and the DFP update formulas (4) and (6), we realize that the two rank-one updates involved have the function of, in addition to satisfying the secant equation, first reducing and then restoring the rank of the original matrix by one while maintaining positive definiteness. These notions of rank reduction and rank restoration, along with maintaining positive definiteness, have been generalized in this paper to include higher rank modifications in the BFGS formula. Specifically, the Wedderburn result and its generalization, Lemma 2.3, provide an exact prescription for the form of the rank update; whereas the rank restoration results, Theorem 3.3 and Corollary 3.4, specify how such an update will maintain positive definiteness.

In addition to the fundamental BFGS and DFP methods, the theory developed here should help to understand higher rank modifications of the Hessian matrix. One possible application of this multi-vector update would be the case where the heuristic secant equation is replaced by more sophisticated multi-step formulas, a subject that will be further studied.

Acknowledgment. Professor R. E. Hartwig has kindly pointed out a quicker proof of Theorem 3.1, not directly related to eigenvalues. Let $H = LL^T$ denote the Cholesky decomposition of H. Then

$$H - M = L\left(I - L^T S(S^T LL^T S)^{-1} S^T L\right) L^T. \tag{15}$$

Note that the middle factor $I - L^T S(S^T LL^T S)^{-1} S^T L$ in (15) is a projection matrix. It follows that $H - M$ is positive semi-definite. The proof based on Weyl's lemma, on the other hand, gives additional insight into the spectrum of $H - M$.

REFERENCES

[1] A. Ben-Israel and T. N. E. Greville, Generalized Inverses: Theory and Applications, Wiley, New York, 1974.
[2] K. W. Brodlie, A. R. Gourlay and J. Greenstadt, Rank-one and rank-two corrections to positive definite matrices expressed in product form, J. Inst. Math. Appli., 11(1973), 73-82.
[3] M. T. Chu, R. E. Funderlic and G. H. Golub, A rank-one reduction formula and its applications to matrix factorizations, SIAM Review, 37(1995), 512-530.
[4] R. E. Cline and R. E. Funderlic, The rank of a difference of matrices and associated generalized inverses, Linear Alg. Appl., 24(1979), 185-215.
[5] J. E. Dennis, Jr. and R. B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, Prentice Hall, Englewood Cliffs, 1983.
[6] A. Deuflhard and R. W. Freund, Fast secant methods for the iterative solution of nonsymmetric linear equations, Science and Engineering 2(1990), 244-276.
[7] E. Egerváry, On rank-diminishing operators and their applications to the solution of linear equations, Z. Angew. Math. Phys., 11(1960), 376-386.
[8] R. Fletcher, Practical Methods of Optimization, 2nd ed., Wiley, Chichester, 1990.
[9] R. R. Gerber and F. T. Luk, A generalized Broyden's method for solving simultaneous linear equations, SIAM J. Numer. Anal., 18(1981), 882-890.
[10] A. S. Householder, The Theory of Matrices in Numerical Analysis, Blaisdell, New York, 1964; and Dover, New York, 1975.
[11] G. W. Stewart and J. G. Sun, Matrix Perturbation Theory, Academic Press, Boston, 1990.
[12] J. H. M. Wedderburn, Lectures on Matrices, Colloquium Publications, vol. XVII, American Mathematical Society, New York, 1934; and Dover, New York 1964.
[13] H. Weyl, Das asymptotische Verteilungsgesetz der Eigenwerte linearer partieller Differentialgleichungen, Math. Ann., 71(1912), 441-479.