Sizing and Least Change
Secant Methods

J. E. Dennis
H. Wolkowicz

Center for Research on Parallel Computation
Rice University
P.O. Box 1892
Houston, TX 77251-1892

SIZING AND LEAST CHANGE SECANT METHODS

by J.E. Dennis, Jr.[1] and H. Wolkowicz[2]

November, 1989

Abstract. The function $\omega(A) = \frac{trace(A)}{n \, det \, (A)^{\frac{1}{n}}}$ is introduced as a measure of deviation of a positive-definite matrix from the identity. This appears to be a more uniform measure than the standard ℓ_2 condition number since it takes all the eigenvalues of A into account. Optimal quasi-Newton updates are given with respect to various applications of this measure. This yields the inverse-sized BFGS and sized DFP updates suggested by Oren and Luenberger, and it gives rise to a new one-parameter class of updates based on these two updates just as the Broyden class is based on the BFGS and DFP updates. Also considered are alternatives to sizing after the first step. This leads to some interesting weighted Frobenius norm problems for weak forms of the secant condition and a particular Fletcher dual pair in the Broyden class.

[1] Professor, Department of Mathematical Sciences, Rice University, Houston, Texas, U.S.A. and Adjunct Professor, Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada. Research partially supported by Air Force Grant AFOSR-89-0363 and National Science Foundation Grant DMS-8903751.
[2] Professor, Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada.

This report is issued simultaneously at the Department of Combinatorics and Optimization, University of Waterloo, Ontario, Canada and the Department of Mathematical Sciences, Rice University, Houston, Texas, U.S.A.

1 Introduction

We consider quasi-Newton methods for the unconstrained optimization problem

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

where f is twice continuously differentiable. The methods use a local quadratic model of the form

$$f(x_c + s) \approx f(x_c) + g_c^t s + \frac{1}{2} s^t B_c s \tag{2}$$

where x_c is the current approximation to a minimizer x^*, B_c is the current approximation to the true Hessian G at x_c, and g_c is the gradient at x_c. We will use the notation that $B^{-1} = H$. Our particular interest is in the secant-type methods based upon approximating Newton's method by accumulating Hessian approximations using gradient differences. These methods have the property that the next Hessian approximation $B = B_+$ satisfies the secant condition

$$Bs = y \equiv g_+ - g_c \quad \text{or} \quad Hy = s \equiv x_+ - x_c.$$

The best known class of such approximations is the Broyden family of updates. The best known members of this family are the *BFGS* and the *DFP* methods. We will not restrict ourselves to the Broyden family.

Using the standard ℓ_2 measure of conditioning κ, optimal updates in the Broyden family of rank-two updates have been found by Davidon [1]. Specifically, Davidon chooses a member $B_+(\phi)$ of the Broyden class that minimizes $\kappa(H_c B_+(\phi))$. Although it is common to call these methods *optimally conditioned*, we think it is more salient, as well as more in the mainstream of the subject, to view $\kappa(H_c B_+(\phi))$ as a measure of the change made in the Hessian approximation by the update. This is because we like to view κ as a measure of deviation of a matrix from a multiple of the identity.

In this paper we will use a different measure ω of the deviation of a matrix from the identity. To us, this measure seems more relevant to the updating context than κ. Furthermore, it is very similar to the measure used in the proofs in [2]. For some interesting results on that specific measure, see [3]. We give some properties of ω and relate it to κ in Section 2. In Section 3, we find least-change secant updates from the Broyden class using $\omega(H_c B_+)$ and $\omega(B_c H_+)$. These results are interesting, but they mainly serve as lemmas for our main results in Section 5 where we give very interesting connections between updates that minimize the measures $\omega(H_c B_+)$ and $\omega(B_c H_+)$ and the so-called Oren-Luenberger [5] *scaling*.

In order to interpret the results of Section 3, and to prepare for the results of Section 5, we give some results on Oren-Luenberger scaling in Section 4. Also, since Oren-Luenberger scaling, which we prefer to call *sizing*, is generally regarded as useful only in the first step of an iteration [6], we look for an alternative for subsequent iterations. This leads us to some interesting weighted Frobenius norm problems for weak forms of the secant condition

$$s^T B s = s^T y \quad \text{or} \quad y^T H y = y^T s.$$

2

These problems are solved with some surprises in Section 4 by rank-one updates, which we call *weak secant updates*.

In Section 5, we bring together sizing, and weak and strong least-change secant updating in ω and in the traditional weighted Frobenius norms associated with the DFP and BFGS methods. This leads to even stronger connections between sizing, ω-least-change secant updates, and the DFP and BFGS methods. Weak updating followed by weighted Frobenius updating leads to a pair of *Fletcher-dual* members of the Broyden class which we have not seen identified before, but which resemble the Hoshino update [7].

In Section 6, we consider the special two-dimensional example of Powell in [8] used there to illustrate that the BFGS behaves better than the DFP. We show that the ω-least-change secant methods for this special case is a sized DFP, or equivalently an inverse-sized BFGS, and we give the corresponding numerical results for the sized DFP. These results are better than for the BFGS. We also include some numerical tests for other problems.

2 Preliminaries

Let A be an $n \times n$ symmetric positive definite matrix (denoted s.p.d.), with eigenvalues

$$\lambda_1 \geq \ldots \geq \lambda_n > 0, \tag{3}$$

and corresponding eigenvectors $u_1, \ldots, u_n$. Then the usual ℓ_2 condition number of A is given by

$$\kappa(A) = \lambda_1 / \lambda_n. \tag{4}$$

The condition number $\kappa(A)$ is used in perturbation analysis for matrix inversion, e.g. for the systems of linear equations $Ax = b$ and $A\bar{x} = \bar{b}$, we obtain bounds on the relative differences

$$\frac{1}{\kappa(A)} \frac{||b - \bar{b}||}{||b||} \leq \frac{||x - \bar{x}||}{||x||} \leq \kappa(A) \frac{||b - \bar{b}||}{||b||}, \tag{5}$$

see e.g. [9]. This condition number acts as an upper bound on the amplification factor of the relative change in the right hand side in bounding the relative change in the solution

It has often been noted that κ depends only on the largest and smallest eigenvalues We propose using the following measure which depends more uniformly on all the eigenvalues.

$$\omega(A) \equiv \frac{trace(A)/n}{det(A)^{\frac{1}{n}}}. \tag{6}$$

Just as the usual condition number $\kappa(A)$ is a measure of how close A is to the pencil αI, where I is the identity matrix and $\alpha > 0$, the function $\omega(A)$ similarly measures the 'distance' of A from αI. This measure takes all the eigenvalues of A into account and so is a more uniform or average indicator of the distance of A.

Moreover, $\omega(A)$ can be calculated in terms of the actual data of A rather than its spectrum and can be more easily differentiated and manipulated.

We now give some useful properties of ω and address some related issues. Here we restrict ourselves to the s.p.d. matrices and use the Lowner order, i.e. $A \geq B$ means $A - B$ is positive semidefinite. See [10, pg. 475]. Remember that a function is pseudoconvex means:

$$(y - x)^t \nabla f(x) \geq 0 \Longrightarrow f(y) \geq f(x). \tag{7}$$

See [11].

Proposition 2.1 *The measure $\omega(A)$ satisfies*

(i) $1 \leq \omega(A) \leq \kappa(A) \leq \frac{(\kappa(A)+1)^2}{\kappa(A)} \leq 4\omega^n(A),$
   with equality in the first and second inequality if and only if A is a multiple of the identity and equality in the last inequality if and only if

$$\lambda_2 = \ldots = \lambda_{n-1} = \frac{\lambda_1 + \lambda_n}{2}; \tag{8}$$

(ii) $\omega(\alpha A) = \omega(A)$, *for all $\alpha > 0$;*

(iii) if $n = 2$, $\omega(A)$ is isotonic with $\kappa(A)$.

(iv) The measure ω is pseudoconvex on the set of s.p.d. matrices, and thus any stationary point is a global minimizer of ω.

(v) Let V be a full rank $m \times n$ matrix, $n \leq m$. Then the optimal column scaling that minimizes the measure ω, i.e.

$$min \; \omega((VD)^t(VD)), \tag{9}$$

over D positive, diagonal, is given by

$$D_{ii} = \frac{1}{\|V_{\cdot i}\|}, i = 1, ..., n, \tag{10}$$

where V_{ii} is the i-th column of V.

Proof. That $1 \leq \omega(A)$ follows from the arithmetic-geometric mean inequality, while $\omega(A) \leq \kappa(A) \leq \frac{(\kappa(A)+1)^2}{\kappa(A)}$ follows from the definitions. The equality conditions also follow directly from the definitions. To prove the last inequality in (i), we fix λ_1 and λ_n and thus also $\kappa(A)$. We now minimize $\omega(A)$ by differentiating. (Note (iv) in the Proposition.) This yields the equality conditions (8). Substitution shows that

$$min \; \omega^n(A) = \frac{(\kappa(A) + 1)^2}{\kappa(A)}. \tag{11}$$

If $n = 2$

$$2\omega(A) = \frac{\kappa(A) + 1}{\kappa(A)^{\frac{1}{2}}} = \kappa(A)^{\frac{1}{2}} + \kappa(A)^{-\frac{1}{2}}. \tag{12}$$

The derivative of $\omega(A)$ with respect to $\kappa(A)$ can now be seen to be positive since $\kappa(A) \geq 1$.

The function $det(A)$ is log concave and strictly increasing and the function $det(A)^{\frac{1}{n}}$ is concave and increasing. See e.g. [10]. The trace function is linear and so convex. Thus ω is the quotient of a convex function by a concave function and so it is pseudoconvex. Pseudoconvex functions have the property that every stationary point is a global minimizer. See [11].

To prove *(v)*, let V be given. Then the arithmetic-geometric mean inequality yields

$$
\begin{aligned}
\omega((VD)^t(VD)) &= \frac{trD^tV^tVD/n}{(detD^tV^tVD)^{\frac{1}{n}}} \\
&= \frac{trV^tVD^2/n}{(detD^tD)^{\frac{1}{n}}(detV^tV)^{\frac{1}{n}}} \\
&= \frac{\Sigma\|V_{\cdot i}\|^2D_{ii}^2/n}{(\Pi\|V_{\cdot i}\|^2D_{ii}^2)^{\frac{1}{n}}}\left(\frac{\Pi\|V_{\cdot i}\|^2}{detV^tV}\right)^{\frac{1}{n}} \\
&\geq \left(\frac{\Pi\|V_{\cdot i}\|^2}{detV^tV}\right)^{\frac{1}{n}}
\end{aligned}
\tag{13}
$$

with equality if and only if $\|V_{\cdot i}\|^2 D_{ii}^2 = constant, i = 1, \ldots, n.$ □

Property *(i)* above shows that $\omega(A)$ is a valid condition number, e.g. we can replace $\kappa(A)$ by $4\omega^n(A)$ in (5). Property *(v)* shows that the measure ω predicts the best column scaling. Note that this is the best scaling used in practice, see e.g. [9], and it is not the one found by minimizing the measure κ. Moreover, the proof of *(v)* is particularly simple.

We use ω as a measure of 'best' in determining some quasi-Newton updates. This leads to minimizing this measure subject to constraints. The following lemma shows, under very mild assumptions, that we do not have to worry about maintaining positive definiteness in our updates since they will solve problems like the one posed here.

Lemma 2.1 *Given the s.p.d. matrix C, consider the quantity*

$$
\begin{aligned}
\mu^* \quad &= \quad \inf_B \omega(BC) \\
\text{subject to} \quad &v^t Bv = \gamma \\
&B \text{ s.p.d.}, \ B \in \Omega,
\end{aligned}
$$

where $v \neq 0$ and $\gamma > 0$ are given, and Ω is a closed set of symmetric matrices. Assume that a feasible B exists. Then the finite value μ^* is <u>attained</u> at some B^* s.p.d.

Proof. First note that

$$1 \le \mu^* \le \alpha, \tag{14}$$

where $\omega(\frac{\gamma}{\|v\|^2}C) = \alpha < \infty$, since $\frac{\gamma}{\|v\|^2}I$ satisfies the equality constraint. Choose $\{B_k\}$ such that each B_k is feasible and

$$\lim_{k \to \infty} \omega(B_k C) = \mu^*. \tag{15}$$

If either $\lambda_1(B_k)$ is unbounded above, or $\lambda_n(B_k)$ is not bounded away from zero, then $\sup \kappa(B_k) = \sup \omega(B_k) = \infty$, by Prop. 2.1*(i)*. But this implies $\omega(B_k C) \to \infty$, a contradiction. Thus we can assume that $B_k \to \tilde{B}$ for some $\tilde{B} \in \Omega$ which is s.p.d. $\square$

In the sequel we consider the space of s.p.d. matrices as a subspace of the $n \times n$ matrices with the inner product

$$\langle A, B \rangle = trace(AB).$$

The gradients of functionals restricted to this subspace are symmetric matrices.

3 ω-Optimal Rank-2 Updates

We now consider the Broyden family of updates

$$B_\phi = B_c - \frac{1}{s^t B_c s} B_c s s^t \, B_c + \frac{1}{y^t s} y y^t + (1 - \phi) s^t B_c s \, w w^t, \tag{16}$$

where $s = x_+ - x_c$ is the step taken, $y = g_+ - g_c$, the current approximation B_c to the Hessian is s.p.d., and

$$w = \frac{1}{y^t s} y - \frac{1}{s^t B_c s} B_c s. \tag{17}$$

If $\phi = 1$ we get the $BFGS$ update and $\phi = 0$ yields the DFP update. The updates for $\phi \in [0, 1]$ are called the convex class. This is not the most common parameterization, but it is well known and it allows us to use results directly from [12] without reproving them here and uselessly lengthening the paper.

If we form the Fletcher dual updates, i.e. we exchange the roles of y and s and let $H_c = B_c^{-1}$, then we get the inverse updates

$$H_{\hat\phi} = H_c - \frac{1}{y^t H_c y} H_c y y^t H_c + \frac{1}{y^t s} s s^t + (1 - \hat\phi) y^t H_c y v v^t, \tag{18}$$

where

$$v = \frac{1}{y^t s} s - \frac{1}{y^t H_c y} H_c y. \tag{19}$$

We now have that $\hat\phi = 1$ and $\hat\phi = 0$ yield the DFP and $BFGS$ updates, respectively.

Every member of the Broyden family of updates satisfies the secant condition

$$B_\phi s = y \,. \tag{20}$$

Furthermore, let

$$a = y^t H_c y, \quad b = y^t s, \quad c = s^t B_c s. \tag{21}$$

Note that $b^2 \le ac$ with equality if and only if $B_c^{-\frac{1}{2}} y$ and $B_c^{\frac{1}{2}} s$ are collinear, which is true if and only if y and $B_c s$ are collinear, which is true if and only if $H_c y$ and s are collinear. From [12], B_ϕ is s.p.d. if $b^2 < ac$, $b > 0$ and $\phi < \frac{ac}{ac-b^2}$.

A κ-optimal rank-two update is found in [1] by minimizing the measure $\kappa(H_c B_\phi)$ over the Broyden family of rank-two updates. Note that the spectrum of a matrix product $C_1 C_2$ is equal to the spectrum of $C_2 C_1$ and $\kappa(C) = \kappa(C^{-1})$. Consequently, we can replace $H_c B_\phi$ by any of $B_\phi H_c$, $B_\phi^{-1} B_c$, $B_c B_\phi^{-1}$, $B_\phi^{-\frac{1}{2}} B_c B_\phi^{-\frac{1}{2}}$ etc. (See also, [12, Chapter VII].) Related work is found in [13].

We now consider the problem of finding those updates in the Broyden family that minimize the measure ω. In the following, we assume $n \ge 2$, B_c is an $n \times n$ s.p.d. matrix, $s, y \in \mathbf{R}^n$ such that $s^t y > 0$, and $y, B_c s$ are linearly independent. If y and $B_c s$ are linearly dependent, then the entire Broyden family of updates (as well as the sized updates) reduce to the symmetric rank-one

or SR1, update. Otherwise, $\phi_{SR1} = -\frac{c}{b-c}$ and $\hat{\phi}_{SR1} = -\frac{a}{b-a}$. Moreover, the inverse of the DFP update B_0 is H_1; the inverse of the $BFGS$ update B_1 is H_0. In general,

$$\hat{\phi} = \iota(\phi) = \frac{1 - \phi}{1 + \phi[\frac{b^2}{ac} - 1]} \tag{22}$$

is a 1-1 and onto mapping (c.f. [14]) that relates ϕ and $\hat{\phi}$ for which $B_\phi^{-1} = H_{\hat{\phi}}$. (This formula corrects a typographical error in [12] and [14].) The mapping satisfies $\iota^2(\phi) = \phi$ and the convex class $\phi \in [0,1] \mapsto \hat{\phi} \in [0,1]$.

Lemma 3.1 *[12, pg. 111] The matrix $B_c^{-\frac{1}{2}} B_\phi B_c^{-\frac{1}{2}}$ has $n - 2$ unit eigenvalues and the two remaining eigenvalues are*

$$\lambda_\pm(\phi) = f_1(\phi) \pm (f_1(\phi)^2 - f_2(\phi))^{\frac{1}{2}}, \tag{23}$$

where

$$f_1(\phi) = \frac{a(b + c) - \phi(ac - b^2)}{2b^2}, \tag{24}$$

We now present the ω-optimal update from the Broyden family. Notice that for large n, the ω-optimal update looks more and more like the BFGS update but that it is in the convex class only when $a \geq b$.

Theorem 3.1 *The minimum over ϕ of $\omega(B_c^{-\frac{1}{2}} B_\phi B_c^{-\frac{1}{2}})$ is attained at*

$$\phi_* = 1 + \frac{(a - b)b}{(1 - n)(ac - b^2)}. \tag{25}$$

Furthermore, B_{ϕ_*} is s.p.d. and

$$\iota(\phi_*) = \frac{\frac{-(a-b)b}{(1-n)(ac-b^2)}}{1 + (1 + \frac{(a-b)b}{(1-n)(ac-b^2)})(\frac{b^2}{ac} - 1)}. \tag{26}$$

Proof. From Lemma 3.1,

$$\omega(B_c^{-\frac{1}{2}} B_\phi B_c^{-\frac{1}{2}}) = \frac{(2f_1(\phi) + n - 2)/n}{(f_2(\phi))^{1/n}}. \tag{27}$$

We can now substitute using (24), differentiate and solve for ϕ_*, the critical point. This yields (25). Since ω is pseudoconvex and both f_1 and f_2 are linear in ϕ, ϕ_* is the global minimizer. The ω-optimal update is s.p.d. because it is easy to show that for $n \geq 2$, $\phi_* < \frac{ac}{ac-b^2}$. (This also follows from Lemma 2.1 with the appropriate choice of the closed set Ω.) □

Corollary 3.1 *The minimum over $\hat{\phi}$ of $\omega(B_c^{\frac{1}{2}} H_{\hat{\phi}} B_c^{\frac{1}{2}})$ is attained at*

$$\hat{\phi}_* = 1 + \frac{(c - b)b}{(1 - n)(ac - b^2)}.$$

Furthermore, $H_{\hat{\phi}_*}$ is s.p.d. and

$$\iota(\hat{\phi}_*) = \frac{\frac{-(c-b)b}{(1-n)(ac-b^2)}}{1 + (1 + \frac{(c-b)b}{(1-n)(ac-b^2)})(\frac{b^2}{ac} - 1)}. \tag{28}$$

Proof. The update formulas (16) and (18) are obtained by exchanging the roles of y and s, i.e., the secant condition can be expressed as $B_\phi s = y$ or $H_{\hat\phi} y = s$, where ϕ and $\hat\phi$ are related by (22). This is equivalent to exchanging the roles of a and c in Theorem 3.1. $\qquad\square$

Note that although $\kappa(A) = \kappa(A^{-1})$, whenever A^{-1} is defined, this is not the case for the measure ω. Therefore, the optimal update obtained in Theorem 3.1 is not, in general, the same as the update obtained in Corollary 3.1. Nor are the values of the measure ω equal. The following table summarizes the results.

measure	ϕ	$\hat\phi$
$\omega(H_c^{\frac{1}{2}} B_\phi H_c^{\frac{1}{2}})$ (*for optimal* ϕ)	$\phi_* = 1 + \frac{(a-b)b}{(1-n)(ac-b^2)}$	$\iota(\phi_*) = \hat\phi = \dfrac{\frac{-(a-b)b}{(1-n)(ac-b^2)}}{1+(1+\frac{(a-b)b}{(1-n)(ac-b^2)})(\frac{b^2}{ac}-1)}$
$\omega(B_c^{\frac{1}{2}} H_{\hat\phi} B_c^{\frac{1}{2}})$ (*for optimal* $\hat\phi$)	$\iota(\hat\phi_*) = \phi = \dfrac{\frac{-(c-b)b}{(1-n)(ac-b^2)}}{1+(1+\frac{(c-b)b}{(1-n)(ac-b^2)})(\frac{b^2}{ac}-1)}$	$\hat\phi_* = 1 + \frac{(c-b)b}{(1-n)(ac-b^2)}$

ω-optimal rank-two updates

In general, to find an optimal ϕ_*, we minimize the measure $\omega(H_c^{\frac{1}{2}} B_\phi H_c^{\frac{1}{2}})$ over B_ϕ, and then use Stoer's formula to find the corresponding $\hat\phi = \iota(\phi_*)$. Conversely, an optimal $\hat\phi$ refers to the measure $\omega(B_c^{\frac{1}{2}} H_{\hat\phi} B_c^{\frac{1}{2}})$. Notice that for large n, ϕ_* is near 1, which corresponds to the DFP, and $\hat\phi_*$ is in the convex class only when $c \geq b$.

9

4 Sizing and Weak Secant Updating

The Broyden family, B_ϕ, of rank-two updates satisfies the secant condition and preserves positive definiteness whenever $y^t s > 0$ for $\phi < \frac{ac}{ac-b^2}$. Now let B_+ be any symmetric matrix that satisfies the secant condition $B_+ s = y$. This guarantees that the curvature information in B_+ along the step s is approximately correct, i.e.,

$$\nabla f(x + s) = g_+ = g + Gs + \bigcirc(\|s\|^2) \tag{29}$$

and

$$s^t G s \approx y^t s = s^t B_+ s, \tag{30}$$

$$y^t G^{-1} y \approx y^t s = y^t H_+ y. \tag{31}$$

We refer to the above two equations as the direct and inverse *weak secant conditions*. We use B_+ rather than B_ϕ since we will no longer restrict ourselves to updates from the Broyden family. By choosing updates that minimize a measure such as $\omega(B_+), \kappa(B_+)$ or $\|B_+ - B_c\|_F$, we attempt to guarantee that the new directional information does not destroy too much information already built up in B_c. Note that the measure ω does this uniformly over all the eigenvalues while κ only deals with the two extreme eigenvalues. However, the low-rank updates bring one eigenvalue of $H_+ B_c$ (or $H_c B_+$) to 1 at each iteration. If the eigenvalues of $H_+ B_c$ are large, this results in ill-conditioning (c.f. [15, pg. 275]). This can be corrected by sizing B_c. More precisely, B_c is to be replaced by $\frac{y^t s}{s^t B_c s} B_c$ before updating (c.f. [16]). Conversely, if H_c is replaced by $\frac{y^t s}{y^t H_c y} H_c$ before updating, then we are sizing H_c. We now find the ω-optimal ϕ and $\hat{\phi}$ to determine a member of the Broyden family that is obtained after sizing. For each sizing, we obtain two 'optimal' matrices and their inverses.

Theorem 4.1 *If $B_c \leftarrow \frac{b}{c} B_c$, then the optimal ϕ_* and corresponding $\hat{\phi} = \iota(\phi_*)$ are given by*

$$\phi_* = 1 + \frac{1}{1 - n}, \tag{32}$$

and

$$\hat{\phi} = \iota(\phi_*) = \frac{1}{(n - 1) + (n - 2)(\frac{b^2}{ac} - 1)}, \tag{33}$$

respectively. Similarly, if $H_c \leftarrow \frac{c}{b} H_c$, then the optimal $\hat{\phi}$ and corresponding $\phi = \iota(\hat{\phi}_*)$ are

$$\hat{\phi}_* = 1, \tag{34}$$

and

$$\iota(\hat{\phi}_*) = \phi = 0. \tag{35}$$

All values are in the convex class. The optimal ϕ_* gives the DFP update if $n = 2$ and approaches the BFGS update as n grows. For every n, the optimal $\hat{\phi}_*$ gives the DFP update. If $B_c \leftarrow \frac{a}{b} B_c$,

then the optimal $\phi_* = 1$ and the BFGS update is optimal. Similarly, if $H_c \leftarrow \frac{b}{a}H_c$, then the optimal $\hat{\phi}_*$ and corresponding $\phi = \iota(\hat{\phi}_*)$ are given by

$$\hat{\phi}_* = 1 + \frac{1}{1-n}, \tag{36}$$

and

$$\iota(\hat{\phi}_*) = \phi = \frac{1}{(n-1) + (n-2)(\frac{b^2}{ac}-1)}, \tag{37}$$

respectively. This is always in the convex class. The optimal ϕ_* is the BFGS update. The optimal $\hat{\phi}_*$ is also the BFGS update, if $n = 2$, but it approaches the DFP update as n grows.

Before giving the proof, we summarize the results in some tables. We prematurely include the result of Theorem 4.4 that the ϕ values are the same after weak updating.

measure	ϕ	$\hat{\phi}$
$\omega(H_c^{\frac{1}{2}} B_\phi H_c^{\frac{1}{2}})$ (*for optimal* ϕ)	$\phi_* = 1 + \frac{1}{1-n}$	$\iota(\phi_*) = \hat{\phi} = \frac{1}{(n-1)+(n-2)(\frac{b^2}{ac}-1)}$
$\omega(B_c^{\frac{1}{2}} H_{\hat{\phi}} B_c^{\frac{1}{2}})$ (*for optimal* $\hat{\phi}$)	$\iota(\hat{\phi}_*) = \phi = 0$	$\hat{\phi}_* = 1$

ω-optimal rank-two updates after sizing $B_c \leftarrow \frac{b}{c}B_c$ (or direct shift)

measure	ϕ	$\hat{\phi}$
$\omega(H_c^{\frac{1}{2}} B_\phi H_c^{\frac{1}{2}})$ (*for optimal* ϕ)	$\phi_* = 1$	$\iota(\phi_*) = \hat{\phi} = 0$
$\omega(B_c^{\frac{1}{2}} H_{\hat{\phi}} B_c^{\frac{1}{2}})$ (*for optimal* $\hat{\phi}$)	$\iota(\hat{\phi}_*) = \phi = \frac{1}{(n-1)+(n-2)(\frac{b^2}{ac}-1)}$	$\hat{\phi}_* = 1 + \frac{1}{1-n}$

ω-optimal rank-two updates after inverse sizing $H_c \leftarrow \frac{b}{a}H_c$ (or weak inverse update)

Proof. Since $\tilde{B} \leftarrow \frac{b}{c}B$, we see that

$$\tilde{b} \leftarrow b, \ \tilde{c} \leftarrow \frac{b}{c}c = b, \ \tilde{a} \leftarrow \frac{ca}{b}$$

Thus Theorem 3.1 yields

$$\phi_* = 1 + \frac{(\frac{ca}{b} - b)b}{(1-n)(ac-b^2)} = 1 + \frac{1}{1-n}. \tag{38}$$

Applying (26) then yields (33). If $\tilde{B}^{-1} \leftarrow \frac{b}{a}\tilde{B}^{-1}$ or $\tilde{B} \leftarrow \frac{a}{b}B$, we have $\tilde{b} \leftarrow b, \tilde{a} \leftarrow b, \ \tilde{c} \leftarrow \frac{a}{b}c$, and (25) yields

$$\tilde{\phi} = 1 + \frac{(b-b)b}{(1-n)(ac-b^2)} = 1. \tag{39}$$

11

Using Corollary (3.1) instead of Theorem 3.1 completes the proof. □

Notice that $n = 2$ is a very special case.

Sizing B_c (or H_c) can cause a drastic change in all the eigenvalues of B_c so that though G and B_c may have had a relatively good overlapping of their spectra, they no longer do. After the first sizing step, a better strategy might be to *shift* the spectrum by a rank-one update. If we want to size both B_c and H_c simultaneously, we can use a rank-two update. Consequently, we will consider finding the 'closest' matrix to B_c that satisfies the weak secant conditions.

The next few results hold some surprises for experts in the field.

Theorem 4.2 *Let u, v be nonzero vectors in $\mathbb{R}^n$, and let A, M be symmetric matrices with M s.p.d. Then,*

$$\tilde{A} = A + \frac{(u^T v - v^T A v) M v v^T M}{(v^T M v)^2}$$

uniquely solves

$$\min \quad \|W^T(\tilde{A} - A)W\|_F$$
$$subject\ to \quad v^T \tilde{A} v = u^T v$$

independent of W such that $M^{-1} = (WW^T)$. Moreover, if A is s.p.d. and $M = A$ then $\tilde{A}$ is s.p.d. if and only if $u^t v > 0$.

Proof. First note that $\tilde{A}$ is feasible, and let C be any other feasible matrix. Set $x = W^{-1}v$. Then $W^T(\tilde{A} - A)W = \frac{v^T(C-A)v}{x^T x} \cdot \frac{x x^T}{x^T x}$. Thus, $\|W^T(\tilde{A} - A)W\|_F = \left|\frac{x^T(W^T(C-A)W)x}{x^T x}\right| \leq \|W^T(C-A)W\|_2$. Now if $A = M$ is s.p.d. then $\tilde{A}$ is s.p.d. if and only if the rank-1 update $W^T \tilde{A} W = I + \frac{u^T v - v^T A v}{(v^T M v)^2}W^{-1}vv^T W^{-T}$ is s.p.d. The latter is s.p.d. if and only if

$$0 < 1 + trace\frac{(u^T v - v^T A v)}{(v^T A v)^2}W^{-1}vv^T W^{-T},$$

which is equivalent to $\frac{v^T A v - u^T v}{(v^T A v)^2}v^T A v < 1$, which is true if and only if $u^T v > 0$. □

Now we apply this theorem to direct shifting and then to weak inverse updating, i.e. to shifting B and then to shifting H. We use the terms Greenstadt and DFP to refer to the choice of W. In fact, Greenstadt never considered least changes to B_c, only to H_c.

Corollary 4.1 *The direct weak Greenstadt update*

$$\tilde{B} = B_c + \frac{1}{(s^T B_c s)^2}(y^T s - s^T B_c s)B_c s s^T B_c \tag{40}$$

uniquely solves

$$\min \quad \|W^T(\tilde{B} - B_c)W\|_F$$
$$subject\ to \quad s^T \tilde{B} s = y^T s$$

for any square W such that $H_c = WW^T$. Moreover, $\tilde{B}$ is s.p.d. if and only if $y^T s > 0$. Also, the direct weak DFP update

$$\tilde{B} = B_c + \frac{1}{(y^T s)^2}(y^T s - s^T B_c s)yy^T \tag{41}$$

12

uniquely solves

$$\min \quad \|W^T(\tilde{B} - B_c)W\|_F$$
$$\text{subject to} \quad s^t\tilde{B}s = y^Ts \tag{42}$$

for any square W such that $(WW^T)^{-1}s = y$.

It is surprising that (40) is a hereditarily positive definite update, but (41) is not. This is directly opposite the case for strong secant methods with the same weighted Frobenius norms since (41) corresponds to the DFP secant update. Now we will see in the next corollary that the same twist holds for weak inverse updating; the Greenstadt inverse update is hereditarily positive definite, but the BFGS is not, as we found with the first two randomly generated examples in MATLAB that we tried.

Corollary 4.2 *The inverse weak Greenstadt update*

$$\tilde{H} = H_c + \frac{1}{(y^T H_c y)^2}(y^Ts - y^T H_c y)H_c yy^T H_c$$

uniquely solves

$$\min \quad \|W^T(\tilde{H} - H_c)W\|_F$$
$$\text{subject to} \quad y^T\tilde{H}y = y^Ts \tag{43}$$

for any square W such that $B_c = WW^T$. *Moreover,* $\tilde{H}$ *is s.p.d. if and only if* $y^Ts > 0$. *Also the weak BFGS update*

$$\tilde{H} = H_c + \frac{1}{(y^Ts)^2}(y^Ts - y^T H_c y)ss^T \tag{44}$$

uniquely solves

$$\min \quad \|W^T(\tilde{H} - H_c)W\|_F$$
$$\text{subject to} \quad y^T\tilde{H}y = y^Ts \tag{45}$$

for any square W such that $(WW^T)^{-1}y = s$.

If we try to find the the 'best' update with respect to our measure ω that satisfies the spectral conditions (30) and (31), then we just scale B_c or H_c since the value of $\omega(\tilde{B}H_c)$ or $\omega(\tilde{B}^{-1}B_c)$ will be one. It is interesting that if we let $\sigma > 0$ and apply the following theorem to σB_c, we get $(\sigma B_c)_+ = \sigma B_+$.

Theorem 4.3 *The update* $\tilde{B}$ *that solves*

$$\min \quad \omega(H_c^{-\frac{1}{2}}\tilde{B}H_c^{-\frac{1}{2}}) \quad (or \ \omega(B_c^{\frac{1}{2}}\tilde{B}^{-1}B_c^{\frac{1}{2}}))$$
$$\text{subject to} \quad s^t\tilde{B}s = y^ts \tag{46}$$

is

$$\tilde{B} \leftarrow \frac{y^ts}{s^t B_c s}B_c. \tag{47}$$

13

Proof.

Since $H_c^{-\frac{1}{2}} \tilde{B} H_c^{-\frac{1}{2}} = \frac{y^t s}{s^t B_s} I$, we have obtained the minimum of κ and ω, i.e., $\omega = 1 = \kappa$. □

Corollary 4.3 *The update $\tilde{B}$ that solves*

$$\min \quad \omega(B_c^{\frac{1}{2}} \tilde{B}^{-1} B_c^{\frac{1}{2}}) \text{ or } \omega(B_c^{\frac{1}{2}} \tilde{B}^{-1} B_c^{\frac{1}{2}}))$$

$$\text{subject to} \quad y^t \tilde{B}^{-1} y = y^t s \tag{48}$$

is obtained from

$$\tilde{B} \leftarrow \frac{y^t H_c y}{y^t s} B_c. \tag{49}$$

measure	optimal	optimal inverse
$\|H_c^{\frac{1}{2}}(\tilde{B} - B)H_c^{\frac{1}{2}}\|_F$ (*constraint* $s^t B s = y^t s$)	$\tilde{B} = B_c + \frac{1}{c^2}(b - c)B_c s s^t B_c$	$\tilde{B}^{-1} = H_c + \frac{c-b}{bc} s s^t$
$\|B_c^{\frac{1}{2}}(\tilde{B}^{-1} - H_c)B_c^{\frac{1}{2}}\|_F$ (*constraint* $y^t H y = y^t s$)	$\tilde{B} = B_c + \frac{a-b}{ba} yy^t$	$\tilde{B}^{-1} = H_c + \frac{1}{a^2}(b - a)H_c y y^t H_c$
$\omega(H_c^{\frac{1}{2}} \tilde{B} H_c^{\frac{1}{2}})$ (*or* $\omega(B_c^{\frac{1}{2}} \tilde{B}^{-1} B_c^{\frac{1}{2}})$) (*constraint* $s^t B s = y^t s$)	$\tilde{B} = \frac{b}{c} B_c$	$\tilde{B}^{-1} = \frac{c}{b} B_c^{-1}$
$\omega(B_c^{\frac{1}{2}} \tilde{B}^{-1} B_c^{\frac{1}{2}}$ (*or* $\omega(H_c^{\frac{1}{2}} \tilde{B} H_c^{\frac{1}{2}})$) (*constraint* $y^t H y = y^t s$)	$\tilde{B} = \frac{a}{b} B_c$	$\tilde{B}^{-1} = \frac{b}{a} B_c^{-1}$

Optimal updates for weak secant equations.

In Theorem 4.1 we presented the optimal rank-2 updates in the Broyden family obtained after sizing and using the measure ω. We now show that we obtain the same optimal ϕ_* (and $\hat{\phi}_*$) to strongly update the weakly updated matrix. This does not mean that the corresponding B_+ matrices will be the same, it just means they are obtained from the sized or weakly updated B_c (or H_c) using the same formula from the Broyden class.

Theorem 4.4 *The optimal ϕ_* and $\hat{\phi}_*$ expressions in Theorem 4.1 are unchanged if we replace sizing B_c ($B_c \leftarrow \frac{b}{c} B$ and $H_c \leftarrow \frac{c}{b} H_c$) with the direct weak Greenstadt update (4.1) of Corollary 4.1 and we replace inverse sizing H_c($H_c \leftarrow \frac{b}{a} H_c$ and $B_c \leftarrow \frac{a}{b} B_c$) with the inverse weak Greenstadt update (4.2) of Corollary (4.2).*

Proof.

Suppose that we apply the direct weak update. Then the Sherman-Morrison formula yields

$$\tilde{H} = H_c + \frac{c - b}{bc} s s^t,$$

14

which implies

$$\tilde{a} \leftarrow a + \frac{b}{c}(c - b).$$

We also have $\tilde{b} \leftarrow b$ and $\tilde{c} \leftarrow b$. Therefore, Theorem 3.1 yields

$$
\begin{aligned}
\phi_* &= 1 + \frac{(\tilde{a} - b)b}{(1 - n)(\tilde{a}\tilde{c} - b^2)} \\
&= 1 + \frac{(a + b - \frac{b^2}{c} - b)b}{(1 - n)((a + \frac{b(c-b)}{c})b - b^2)} \\
&= 1 + \frac{1}{1 - n}.
\end{aligned}
$$

Similarly, the optimal $\hat{\phi}_* = 1$, since $\tilde{b} = \tilde{c}$. By exchanging the roles of H_c and B_c, we see that the weak inverse update yields

$$\hat{\phi}_* = 1 + \frac{1}{1 - n}, \quad \phi_* = 1 .$$

$\square$

We now apply the weighted Frobenius norm measures after weak updating. Although we do not restrict the updates to the Broyden class, we get a new Fletcher-dual pair of updates in the Broyden class. These symmetric updates are hereditarily positive definite but not always in the convex class. The new updates are *(iii)* and *(iv)* in the following theorem.

Theorem 4.5 *The following are equivalent updating sequences. The first four are hereditarily positive definite.*

(i) *The result of a direct weak Greenstadt update or a weak BFGS update followed in either case by a BFGS update is a BFGS update.*

(ii) *The result of an inverse weak Greenstadt update or a weak DFP update followed in either case by a DFP update is a DFP update.*

(iii) *The result of a direct weak Greenstadt update followed by a DFP update is the $\phi = 1 - \frac{b}{c}, \hat{\phi} = \frac{1}{\frac{b}{a} - (\frac{b^2}{ac} - 1)}$ update from the Broyden class.*

(iv) *The result of an inverse weak Greenstadt update followed by a BFGS update is the $\hat{\phi} = 1 - \frac{b}{a}$, $\phi = \frac{1}{\frac{b}{c} - (\frac{b^2}{ac} - 1)}$ update from the Broyden class.*

The following sequences may not be hereditarily positive definite

(v) *The result of a weak DFP update followed by a BFGS update is the $\phi = \frac{c}{b}$, $\hat{\phi} = \frac{ab - ac}{ab - ac + b^2}$ member of the Broyden class.*

(vi) *The result of a weak BFGS update followed by a DFP update is the $\hat{\phi} = \frac{a}{b}$, $\phi = \frac{cb - ac}{cb - ac + b^2}$ member of the Broyden class.*

Proof.

The proofs are much the same, so we will do only *(i)* and *(iv)* since they seem to be the most interesting updates.

The direct weak Greenstadt update is

$$\bar{B} = B_c + \frac{b-c}{c^2}B_c ss^T B_c \text{ and } \bar{B}s = (1 + \frac{b-c}{c})B_c s.$$

The weak BFGS is

$$\tilde{H} = H_c + \frac{b-c}{b^2}ss^T \text{ and } s - \tilde{H}y = s - H_c y - (\frac{b-a}{b})s = (1 - \frac{b-a}{b})s - H_c y$$

In the first case,

$$
\begin{aligned}
B_+ &= \bar{B} - \frac{\bar{B}ss^T\bar{B}}{c} + \frac{yy^T}{b} \\
&= B_c + \frac{b-c}{c^2}B_c ss^T B_c - (1 + \frac{b-c}{c})^2\frac{B_c ss^T B_c}{c} + \frac{yy^T}{b} \\
&= B_c - \frac{B_c ss^T B_c}{c} + \frac{yy^T}{b}.
\end{aligned}
$$

In the second,

$$
\begin{aligned}
H_+ &= \tilde{H} + \frac{(s - \tilde{H}y)s^T + s(s - \tilde{H}y)^T}{b} - \frac{y^T(s - \tilde{H}y)ss^T}{b^2} \\
&= H_c + \frac{b-a}{b}\frac{ss^T}{b} + \frac{((1 - \frac{b-a}{b})s - H_c t)s^T + s((1 - \frac{b-a}{b})s - H_c y)^T}{b} - 0 \\
&= H_c + \frac{(s - H_c y)s^T + s(s - H_c y)^T}{b} + (\frac{b-a}{b^2} - 2\frac{b-a}{b^2})ss^T,
\end{aligned}
$$

which is the BFGS update of H_c. The proof of *(iv)* is as direct.

The inverse weak Greenstadt update of H_c is

$$\tilde{H} = H_c + \frac{(b-a)}{a^2}H_c yy^T H_c$$

and $\tilde{H}y = (1 + \frac{b-a}{a})H_c y$, $s - \tilde{H}y = s - H_c y - \frac{b-a}{a}H_c y$. So, following with a BFGS update,

$$
\begin{aligned}
H_+ &= \tilde{H} + \frac{(s - \tilde{H}y)s^T + s(s - \tilde{H}y)^T}{b} - \frac{y^T(s - \tilde{H}y)ss^T}{b^2} \\
&= H_c + \frac{b-a}{a^2}H_c yy^T H_c + \frac{(s - H_c y)s^T + s(s - H_c y)^T}{b} \\
&\quad - (\frac{b-a}{ab})(H_c ys^T + sy^T H_c) \pm \frac{b-a}{b^2}ss^T \\
&= H_c + \frac{(s - H_c y)s^T + s(s - H_c y)^T}{b} - \frac{b-a}{b^2}ss^T \\
&\quad + (b-a)(\frac{H_c y}{a}\frac{y^T H_c}{a} - \frac{H_c y}{a}\frac{s^T}{b} - \frac{s}{b}\frac{y^T H_c}{a} + \frac{s}{b}\frac{s^T}{b}).
\end{aligned}
$$

The first three terms are the BFGS and the last is

$$(b-a)(\frac{s}{b} - \frac{H_c y}{a})(\frac{s}{b} - \frac{H_c y}{a})^T = (\frac{b}{a} - 1)a\ vv^T,$$

where v is given by (19). Now use (18) with $\hat{\phi} = 0$ for the BFGS portion and we get

$$H_+ = H_c - \frac{1}{a}H_c yy^T H_c + \frac{1}{b}ss^T + avv^T + (\frac{b}{a} - 1)avv^T$$

which is (18) with $1 - \hat{\phi} = 1 + \frac{b}{a} - 1 = \frac{b}{a}$ and *(iv)* is proven. □

5 The ω-Optimal Update

In Section 3 we found the 'best' rank-two updates in the Broyden class, i.e., the rank-two updates parametrized by ϕ and $\hat{\phi}$ in (16) and (18) that minimize the measure ω. However, if we do not restrict ourselves to rank-two updates but only to maintaining positive definiteness and the secant equation, then we obtain a different result. We see that the best update of B_c in the case $ac \neq b^2$ is obtained by inverse sizing B_c and then applying the $BFGS$ update. Similarly, the best possible inverse update is obtained by sizing and operating on H_c with DFP. The updates are different although the optimal values of ω are ultimately equal.

We continue to assume that $y^t s > 0$ and that y and Bs are linearly independent.

Theorem 5.1 *Assume B_c is s.p.d. and $ac \neq b^2$. Then for*

$$\alpha = \frac{b}{a} = \frac{n}{trace(H_c B_+)},$$

the BFGS update of $\frac{1}{\alpha} B_c$,

$$H_+ = \alpha H_c + us^T + su^T, \quad u = \frac{s - \alpha H_c y}{b}$$

is the unique solution of

$$\min \qquad w(H_c B_+)$$
$$subject\ to \quad B_+ s = y, \quad B_+\ s.p.d.$$

In addition, the Lagrange multiplier for the secant equation is uniquely

$$\frac{2(s - \alpha H_c y)}{ban(det(H_c B_+))^{\frac{1}{n}}}$$

and the optimal value of the measure is

$$\omega(H_c B_+) = (\frac{ac}{b^2})^{\frac{1}{n}}.$$

Proof. First, we note that an optimal B_+ s.p.d. exists from Lemma 2.1. The Lagrangian for our problem is

$$L(\lambda, B_+) = g(B_+) + \lambda^t(B_+ s - y),$$

where

$$g(B_+) = \frac{trace(H_c B_+)}{n(det(H_c))^{\frac{1}{n}}(det(B_+))^{\frac{1}{n}}}$$

and $\lambda \in \mathbf{R}^n$ is the Lagrange multiplier. For simplicity of notation, we multiply the Lagrangian by the constant $n(det(H_c))^{\frac{1}{n}}$ and remove this constant from λ at the end. Now

$$
\begin{aligned}
\lambda^t B_+ s &= trace(\lambda^t B_+ s) \\
&= trace(B_+ s \lambda^t) \\
&= trace(\lambda s^t B_+) \\
&= trace(s \lambda^t B_+) \\
&= trace(\frac{s\lambda^t + \lambda s^t}{2} B_+)
\end{aligned}
$$

by adding the previous two equivalences and dividing by 2. Therefore the gradient of the linear functional $\lambda^t B_+ s = \langle \frac{s\lambda^t + \lambda s^t}{2}, B_+ \rangle$ on the subspace of symmetric matrices is

$$\frac{s\lambda^t + \lambda s^t}{2}.$$

Using the cofactor expansion of the determinant along a row of the matrix, i.e. $det(A) = \sum_j a_{kj}(-1)^{k+j} det(A(k,j))$ where $A(k,j)$ denotes the submatrix of A obtained by deleting row k and column j, we see that the gradient of $det(B_+)$ is $adj\, B_+$, the (symmetric) matrix of signed cofactors. We will also need Cramer's rule, i.e. $B_+^{-1} = \frac{1}{det(B_+)} adj\, B_+$.

We can now differentiate the Lagrangian with respect to B_+ and equate the derivative to 0;

$$0 = \frac{1}{(det\ B_+)^{\frac{2}{n}}}\{(det\ B_+)^{\frac{1}{n}} H_c - \frac{trace(\ H_c B_+)}{n}(det\ B_+)^{\frac{1}{n}-1}(adj\ B_+)\} + \frac{s\lambda^t + \lambda s^t}{2}$$

or

$$0 = \frac{n}{trace(\ H_c B_+)} H_c - B_+^{-1} + \frac{n(det\ B_+)^{\frac{1}{n}}}{trace(\ H_c B_+)} \frac{s\lambda^t + \lambda s^t}{2}. \tag{50}$$

Let

$$\alpha = \frac{n}{trace(\ H_c B_+)}, u = \frac{n(det\ B_+)^{\frac{1}{n}}}{2\ trace(\ H_c B_+)}\lambda. \tag{51}$$

Then (50) becomes

$$B_+^{-1} = \alpha H_c + su^t + us^t. \tag{52}$$

Since $B_+^{-1} y = s$, we obtain

$$\alpha H_c y + su^t y + us^t y = s \tag{53}$$

or

$$u = \frac{s}{s^t y} - \frac{\alpha H_c y}{s^t y} - s(\frac{u^t y}{s^t y}). \tag{54}$$

Let

$$\beta = \frac{u^t y}{s^t y} \tag{55}$$

and substitute (54) in (53). We get

$$\alpha H_c y + s\frac{s^t y}{s^t y} - \frac{sy^t \alpha H_c y}{s^t y} - ss^t \beta y + \frac{ss^t y}{s^t y} - \frac{\alpha H_c y}{s^t y}s^t y - ss^t \beta y = s$$

or

$$-s(2s^t \beta y) + s(1 - \frac{y^t \alpha H_c y}{s^t y}) = 0. \tag{56}$$

Therefore,

$$\beta = \frac{1}{2s^t y}(1 - \frac{y^t \alpha H_c y}{s^t y}). \tag{57}$$

¿From (54), (55), and (57), we conclude

$$u = \frac{1}{s^t y}(s - \alpha H_c y) - (\frac{s^t y - y^t \alpha H_c y}{2(s^t y)^2})s. \tag{58}$$

19

We can now substitute for u in (52) and obtain

$$
\begin{aligned}
B_+^{-1} &= \alpha H_c + us^t + su^t \\
&= \alpha H_c + \frac{ss^t}{s^t y} - \frac{\alpha H_c y s^t}{s^t y} - \frac{ss^t}{2s^t y} + \frac{ss^t}{2s^t y}\frac{y^t \alpha H_c y}{s^t y} \\
&\quad + \frac{ss^t}{s^t y} - \frac{sy^t \alpha H_c}{s^t y} - \frac{ss^t}{2s^t y} + \frac{ss^t}{2s^t y}\frac{y^t \alpha H_c y}{s^t y} \\
&= \alpha H_c + \frac{ss^t}{s^t y} - \frac{\alpha H_c y s^t + sy^t \alpha H_c}{s^t y} + \frac{ss^t y^t \alpha H_c y}{(s^t y)^2}.
\end{aligned}
\tag{59}
$$

Note that (59) is the *BFGS* update of (αH_c) and is equivalent to (18) with $\hat\phi = 0$ (c.f. [15, pg. 269]. Since the update B_+ is the best possible for our measure and since it is a rank-two update of $\frac{1}{\alpha}B_c$, we conclude that $\frac{1}{\alpha}$ is the constant that makes the *BFGS* update the best among all rank-two updates which includes the Broyden class. We can now apply Theorem 3.1, i.e. we want ϕ in (25) to be one in order to get the *BFGS* update. Since $b > 0$, this is equivalent to scaling B_c so that the new a equals b, i.e., $y^t \alpha H_c y = b$ or

$$
\alpha = \frac{b}{a}.
\tag{60}
$$

The values for the Lagrange multiplier λ and for α are given in (51). The optimal value

$$
\omega(H_c B_+) = \omega((\alpha H_c)B_+),
$$

by Proposition 2.1(i). Therefore, B_+ is the BFGS update of the sized B_c, i.e. $\phi = 1$, $\tilde a = b$, $\tilde b = b$, $\tilde c = ac/b$, and Lemma 3.1 yields

$$
\begin{aligned}
\omega(\alpha H_c B_+) &= \frac{(2f_1(1) + n - 2)/n}{(f_2(1))^{\frac{1}{n}}} \\
&= \frac{(2 + n - 2)/n}{(b^2/ac)^{\frac{1}{n}}}.
\end{aligned}
$$

$\square$

Corollary 1 *Assume B_c s.p.d. and $ac \neq b^2$. Then*

$$
\alpha = \frac{b}{c} = \frac{n}{trace(H_+ B_c)}
$$

and the DFP update of αB_c,

$$
B_+ = \alpha B_c + uy^t + yu^t, \quad u = \frac{y - \alpha B_c s}{b},
$$

is the unique solution of

$$
\min \qquad \omega(B_c H_+)
$$
$$
subject\ to \quad H_+ y = s, \quad H_+ s.p.d.
$$

20

In addition, the Lagrange multiplier for the secant equation is uniquely

$$\frac{2(y - \alpha B_c y)}{b\alpha n(det(B_c B_+^{-1}))^{\frac{1}{n}}}$$

and the optimal value of the measure is equal to the optimal value in Theorem 5.1.

Proof.

We need only exchange the roles of B_c and H_c in the theorem and note that the optimal value does not change. $\qquad\square$

This theorem and corollary state that we obtain the same updated matrix B_+ whether we apply the optimal ϕ_* formula to B_c or the sized B_c, or indeed any $\sigma B_c, \sigma > 0$. Therefore, the ω-optimal update of σB_c is also the sized BFGS update of B_c. Similarly, the ω-optimal inverse update of σH_c is the sized DFP update of H_c.

6 Concluding Remarks

In this paper we have studied the measure ω as it relates to the derivation of updates of the Hessian in least change secant methods. We have seen, in Section 5, that sizing of the Hessian arises naturally from this measure. In particular, the inverse-sized BFGS and sized DFP are the optimal ω updates. This further illustrates the central role of the BFGS and DFP updates. We have also considered weak secant updating in place of sizing.

Powell [8] shows that the DFP update performs far worse than the BFGS update when applied with direct prediction steps to the simple quadratic function

$$f(x) = \frac{1}{2}(x_1^2 + x_2^2).$$

He shows that the DFP was far less effective than the BFGS at reducing large eigenvalues. Of course, sizing can reduce large eigenvalues immediately. In this respect, sizing can be considered a 'fix' for the DFP. This is corroborated in the following numerical data which compares Powell's data with sized DFP and sized BFGS. The initial Hessian approximation is the diagonal matrix $diag(1, \lambda_1)$, while the initial point is $x_1 = (\cos \psi_1 \ \sin \psi_1)^t$. The numbers represent the number of iterations needed to obtain the condition $\|x_{k+1}\| < \epsilon \|x_1\|$. The numbers in brackets are for the sized updates.

λ_1 \ ψ_1	20 deg	40 deg	60 deg	70 deg	80 deg	85 deg	87 deg	88 deg
10	5 (9)	6 (10)	7 (7)	8 (6)	7 (3)	6 (9)	5 (9)	4 (9)
100	5 (10)	7 (14)	8 (16)	9 (10)	10 (7)	10 (6)	9 (7)	9 (6)
10^4	5 (14)	7 (27)	8 (30)	9 (20)	11 (12)	12 (8)	13 (11)	14 (13)
10^6	5 (19)	7 (15)	8 (15)	9 (34)	11 (14)	12 (10)	13 (8)	14 (6)
10^9	5 (12)	7 (14)	8 (17)	9 (10)	11 (8)	12 (21)	13 (10)	14 (8)

Table 5.1. Number of iterations for the BFGS (sized BFGS) when $\epsilon = 10^{-4}$

λ_1 \ ψ_1	20 deg	40 deg	60 deg	70 deg	80 deg	85 deg	87 deg	88 deg
10	6 (8)	10 (5)	14 (5)	16 (5)	14 (5)	9 (4)	7 (6)	6 (7)
100	8 (8)	15 (5)	29 (6)	47 (6)	89 (8)	106 (8)	84 (7)	59 (6)
1000	10 (8)	19 (5)	45 (6)	83 (7)	230 (8)	549 (10)	855 (10)	1000 (10)
10^4	12 (8)	24 (5)	60 (6)	119 (7)	380 (9)	1141 (10)	2420 (11)	4102 (12)
10^6	15 (8)	34 (5)	92 (6)	181 (7)	752 (9)	3482 (10)	5162 (11)	9194 (11)

Table 5.2. Number of iterations for the DFP (sized DFP) when $\epsilon = 10^{-4}$

From Theorems 4.1 and 5.1 and Corollary 5.1, the sized DFP, inverse-sized BFGS, optimal ϕ and inverse optimal $\hat{\phi}$, are all equal in the case n=2. (In fact, Proposition 2.1 (iii) implies that they are also equal to the optimally conditioned sized symmetric rank-one update.) This clearly

appears to be the best update in the above tables. We now decrease ϵ. The following results show that we do not appear to lose asymptotically when we use the ω optimal updates in the case $n = 2$.

λ_1 \ ψ_1	20 deg	40 deg	60 deg	70 deg	80 deg	85 deg	87 deg	88 deg
10	6 (9)	7 (6)	9 (6)	9 (5)	8 (5)	7 (7)	6 (7)	5 (8)
100	6 (9)	8 (6)	9 (7)	10 (8)	11 (8)	11 (7)	11 (7)	10 (7)
10^4	6 (9)	8 (6)	10 (7)	11 (9)	12 (10)	14 (10)	15 (10)	15 (11)
10^6	6 (9)	8 (6)	10 (7)	11 (9)	12 (10)	14 (11)	15 (11)	16 (12)
10^9	6 (9)	8 (6)	10 (7)	11 (9)	12 (10)	14 (11)	15 (11)	16 (12)

Table 5.3. Number of iterations for the BFGS (sized DFP) when $\epsilon = 10^{-6}$

λ_1 \ ψ_1	20 deg	40 deg	60 deg	70 deg	80 deg	85 deg	87 deg	88 deg
10	7 (10)	9 (7)	10 (7)	10 (7)	10 (6)	8 (5)	7 (8)	6 (9)
100	7 (10)	9 (7)	11 (7)	12 (7)	13 (9)	13 (9)	12 (8)	11 (8)
10^4	7 (10)	9 (7)	11 (7)	12 (8)	14 (10)	15 (11)	16 (12)	17 (13)
10^6	7 (10)	9 (7)	11 (7)	12 (8)	14 (10)	15 (11)	16 (12)	17 (12)
10^9	7 (10)	9 (7)	11 (7)	12 (8)	14 (10)	15 (11)	16 (12)	17 (12)

Table 5.4. Number of iterations for the BFGS (sized DFP) when $\epsilon = 10^{-9}$

We have also tested out 22 methods on the standard set of 19 test problems from [4]. We include some of the results below. The methods are:

1. BFGS

2. optimal ϕ

3. optimal $\hat{\phi}$

4. size at first step only and use optimal ϕ

5. size at first step only and use optimal $\hat{\phi}$

6. inverse size at first step only and use optimal ϕ

7. inverse size at first step only and use optimal $\hat{\phi}$

8. size at first step only, direct shift subsequently and use optimal ϕ

9. size at first step only, direct shift subsequently and use optimal $\hat{\phi}$

10. inverse size at first step only, weak inverse update subsequently and use optimal ϕ

23

11. inverse size at first step only, weak inverse update subsequently and use optimal $\hat{\phi}$

(The numbers in the following two tables include iterations/function-evaluations. *** indicates an error or lack of convergence.)

Methods Problems	1	2	3	4	5	6	7
1	9/11	6/8	5/8	6/8	46/48	6/8	46/48
2	29/38	36/43	40/47	29/33	40/44	29/34	50/55
3	37/48	41/45	39/47	45/48	53/59	45/49	56/66
4	3/5	4/8	3/5	4/9	5/9	4/9	5/9
5	149/191	***	107/754	209/273	***	208/275	***
6	29/37	27/33	49/59	30/37	73/80	34/40	77/83
7	20/30	20/30	18/27	18/27	18/27	18/27	18/27
8	37/47	35/41	42/52	53/60	66/71	54/59	70/75
9	138/184	201/259	167/215	148/191	145/186	144/187	148/189
10	871/1131	865/1077	838/1064	253/292	233/263	276/323	237/273
11	***	22/26	***	22/26	***	22/26	***
12	23/48	***	***	31/40	***	***	56/65
13	30/42	***	71/88	54/72	105/120	57/74	105/120
14	28/30	28/30	27/28	29/31	31/33	30/32	31/32
15	52/75	55/73	72/95	38/52	46/58	39/54	47/61
16	109/128	125/142	119/139	40/45	72/76	44/48	71/75
17	14/19	17/22	16/46	17/22	167/174	17/22	150/158
18	53/81	50/67	51/70	41/50	72/84	41/50	74/85
19	21/35	22/32	21/33	27/32	31/38	29/35	33/40
average	91.8/121.1	97.1/121.0	99.1/163.4	57.7/70.9	75.1/85.6	60.9/75.1	74.9/85.9

Table 5.5. Methods 1 through 7

Methods	8	9	10	11
Problems				
1	9/11	9/11	10/12	9/11
2	30/38	31/40	35/42	27/35
3	48/51	109/111	47/50	44/48
4	3/7	3/7	3/7	3/7
5	164/214	162/210	***	150/200
6	34/40	44/51	44/50	33/40
7	18/27	18/27	18/27	18/27
8	53/58	129/133	58/63	52/57
9	148/193	***	151/200	147/198
10	266/308	984/1030	270/310	271/325
11	12/16	12/16	13/17	***
12	32/41	***	***	***
13	52/64	63/73	65/77	46/64
14	30/32	54/55	30/31	29/31
15	38/56	40/56	***	37/55
16	45/49	121/130	50/54	45/49
17	15/20	15/20	17/22	14/19
18	40/47	63/70	46/54	40/48
19	29/37	68/64	30/34	29/34
average	56.1/68.9	113.2/123.8	55.4/65.6	58.5/73.4

Table 5.6. Methods 8 through 11

References

[1] W.C. DAVIDON. Optimally conditioned optimization algorithms without line searches. *Math. Prog.*, 9:1–30, 1975.

[2] R.H. BYRD, R.H. NOCEDAL, and Y.YUAN. Convergence analysis of a class of quasi-newton methods on convex problems. *SIAM J. Numer. Anal.*, 24:1171–1191, 1987.

[3] R. FLETCHER. A new variational result for quasi-newton formulae. Technical Report NA/119, Department of Mathematical Sciences, University of Dundee, 1989.

[4] GARBOW, B.S., HILLSTROM, K.E. and MORÉ, J.J. (1980). User guide for MINPACK-1. Technical Report ANL-80-74, Argonne National Labs, Argonne, IL. Available from National Technical Information Service, Springfield, Virginia.

[5] S.S. OREN and D.G. LUENBERGER. Self-scaling variable metric (SSVM) algorithms, part I. Criteria and sufficient conditions for scaling a class of algorithms. *Manage. Sci.*, 20:845–862, 1974.

[6] K.H. PHUA and D.F. SHANNO. Numerical comparison of several variable metric algorithms. *J. Optim. Theory Appl.*, 25:507–518, 1978.

[7] S. HOSHINO. A formulation of variable metric methods. *Journal Instit. Math. Appl.*, 10:394–403, 1972.

[8] M.J.D. POWELL. How bad are the BFGS and DFP methods when the objective function is quadratic. *Math. Prog.*, 34:34–47, 1986.

[9] G.H. GOLUB and C.F. VAN LOAN. *Matrix Computations.* Johns Hopkins University Press, Baltimore, Maryland, 1983.

[10] A.W. MARSHALL and I. OLKIN. *Inequalities: Theory of Majorization and its Applications.* Academic Press, New York, NY, 1979.

[11] O.L. MANGASARIAN. *Nonlinear Programming.* McGraw-Hill, New York, NY, 1969.

[12] R.B. SCHNABEL. *Analysing and improving quasi-Newton methods for unconstrained optimization.* PhD thesis, Department of Computer Science, Cornell University, Ithaca, NY, 1977. Also available as TR-77-320.

[13] S.S. OREN and E. SPEDICATO. Optimal conditioning of self-scaling variable metric algorithms. *Math. Prog.*, 10:70–90, 1976.

[14] J. STOER. On the convergence rate of imperfect minimization algorithms in Broyden's β class. *Math. Prog.*, 9:313–335, 1975.

[15] D.G. LUENBERGER. *Linear and Nonlinear Programming*. Reading. Addison-Wesley, Massachusetts, second edition, 1984.

[16] J.E. DENNIS Jr., D.M. GAY, and R.E. WELSCH. An adaptive nonlinear least-squares algorithm. *TOMS*, 7:348–368, 1981.