BEHAVIOR OF BFGS WITH AN EXACT LINE SEARCH ON NONSMOOTH EXAMPLES

ADRIAN S. LEWIS* AND MICHAEL L. OVERTON†

Abstract. We investigate the behavior of the BFGS algorithm with an exact line search on nonsmooth functions. We show that it may fail on a simple polyhedral example, but that it apparently always succeeds on the Euclidean norm function, spiraling into the origin with a Q-linear rate of convergence; we prove this in the case of two variables. Dixon's theorem implies that the result for the norm holds for all methods in the Broyden class of variable metric methods; we investigate how the limiting behavior of the steplengths depends on the Broyden parameter. Numerical experiments indicate that the convergence properties for $\|x\|$ extend to $\|Ax\|$, where A is an $n \times n$ nonsingular matrix, and that the rate of convergence is independent of A for fixed n. Finally, we show that steepest descent with an exact line search converges linearly for any positively homogeneous function that is C^2 everywhere except at the origin, but its rate of convergence for $\|Ax\|$ depends on the condition number of A, in contrast to BFGS.

Key words. BFGS, quasi-Newton, nonsmooth, exact line search, Broyden class, Q-linear convergence

AMS subject classifications. 90C30, 65K05

1. Introduction.

The analysis of variable metric methods with an exact line search was pioneered by Powell, who showed in [Pow71] that the DFP (Davidon-Fletcher-Powell) method converges on strongly convex C^2 functions. That this analysis applies also to the BFGS (Broyden-Fletcher-Goldfarb-Shanno) method follows from Dixon's result [Dix72] that all methods in the Broyden class are equivalent on smooth functions when an exact line search is used. Powell studied the nonconvex smooth case in [Pow72], again assuming the use of an exact line search. He summarized this work as follows: "The lemmas and theorems of this paper are the result of about 18 months of intermittent work to try to explain and understand the behaviour of the variable metric algorithm when $f(x)$ is not convex. They make only a small contribution to this problem, because all the lemmas depend on a condition which may never be satisfied." Several decades later, the convergence of variable metric methods on nonconvex smooth functions remains poorly understood [LF01].

There has been little study of the behavior of variable metric methods on nonsmooth functions. This paper studies the BFGS method using an exact line search on some convex nonsmooth examples. After defining the algorithm in the next section, we show in Section 3 that it can fail on a simple polyhedral function. By contrast, in Section 4, we show that BFGS with an exact line search always succeeds on the Euclidean norm function $f(x) = \|x\|$ in two variables, spiraling into the origin with a Q-linear rate $1/2$ with respect to the number of line searches, independent of the initial Hessian approximation. In Section 5, we give numerical evidence indicating that this fact extends to $f(x) = \|Ax\|$, where A is a nonsingular $n \times n$ matrix, with a convergence rate that grows closer to one as n increases but that, remarkably, is independent of A for fixed n.

*School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853, U.S.A. aslewis@orie.cornell.edu. Research supported in part by National Science Foundation Grant DMS-0806057.

† Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, U.S.A. overton@cs.nyu.edu. Research supported in part by National Science Foundation Grant DMS-0714321.

Dixon's theorem applies to the Euclidean norm function, since the only point where it is nonsmooth is the origin. In the strongly convex smooth case, it is known that regardless of the Broyden parameter, the minimizing steplength converges to one, and the convergence of the function values is superlinear. Neither property holds for the norm, but experiments show that the steplengths converge. We investigate their limiting value with respect to n and the Broyden parameter in Section 6.

The observed property that variable metric methods with an exact line search, when applied to $\|Ax\|$, generate a sequence of function values whose limiting behavior is independent of A is in stark contrast with the method of steepest descent. Indeed, steepest descent with an exact line search is *equivalent* on $\|Ax\|$ and its square, $x^T A^T A x$: convergence of the function values is linear with a rate equal to approximately $1-2/\kappa$ (equivalently, $1-4/\kappa$ for the squared function), where κ is the condition number of $A^T A$. In fact, steepest descent with an exact line search converges linearly on *any* positively homogeneous function that is C^2 everywhere except the origin: we prove this in Section 7.

We make some concluding remarks in Section 8.

2. The Algorithm. The BFGS iteration for minimizing a function $f : \mathbf{R}^n \to \mathbf{R}$ can be written as follows. As is standard, x_k denotes the current point at iteration $k = 0, 1, \ldots$, the positive definite matrix H_k is the current estimate of the inverse Hessian $\nabla^2 f(x_k)^{-1}$, and we abbreviate the gradient $\nabla f(x_k)$ to ∇f_k.

Search direction: $p_k = -H_k \nabla f_k$;
Step length: $x_{k+1} = x_k + \alpha_k p_k$, where $\alpha_k \geq 0$ is chosen by a line search;
Gradient increment: $y_k = \nabla f_{k+1} - \nabla f_k$;
Inverse Hessian factor: $V_k = I - (p_k^T y_k)^{-1} p_k y_k^T$;
Inverse Hessian update: $H_{k+1} = V_k H_k V_k^T + \alpha_k (p_k^T y_k)^{-1} p_k p_k^T$;
Iteration count: $k = k + 1$.

In practice, the update to H does not use matrix-matrix multiplication, but exploits the fact that V_k has rank one and therefore the update can be computed in $O(n^2)$ operations.

Consider a conceptual version of this iteration for a possibly nonsmooth function f, with the step length α_k chosen *by exact line search*: in other words, α_k exactly minimizes the function $\alpha \mapsto f(x_k + \alpha p_k)$. In this case, the BFGS method breaks down if (as will often be the case) the function f is not differentiable at x_{k+1}, since then ∇f_{k+1} is undefined.

To avoid this difficulty we proceed as follows. Suppose the function f is locally Lipschitz and Clarke regular[1] at the iterate x_{k+1}. Since the line search stopped at x_{k+1}, the directional derivative $f'(x_{k+1}; p_k)$ is nonnegative. But Clarke regularity implies there is a Clarke subgradient g (a limit of convex combinations of gradients at points near x_k) such that $g^T p_k = f'(x_{k+1}; p_k)$: we then choose ∇f_{k+1} arbitrarily from the set of such subgradients. If f is smooth at x_{k+1}, then as usual ∇f_{k+1} is simply $\nabla f(x_{k+1})$. More generally, if f is a pointwise maximum of smooth functions $f_1, f_2, \ldots, f_m$, and

$$f_j(x_{k+1} + \epsilon p_k) > f_i(x_{k+1} + \epsilon p_k) \quad \text{for all small } \epsilon > 0 \text{ and } i \neq j,$$

then this definition gives $\nabla f_{k+1} = \nabla f_j(x_{k+1})$.

[1] Assuming f is locally Lipschitz and directionally differentiable everywhere, Clarke regularity simply amounts to upper semicontinuity of the directional derivative $x \mapsto f'(x; p)$ for every fixed direction p. All convex functions are Clarke regular.

3. Failure on a Convex Polyhedral Example. An important difference from the smooth case is that, as in bundle methods, the conceptual nonsmooth BFGS iteration described above can take "null steps". Consider the following example.

EXAMPLE 3.1. We consider the convex polyhedral function

$$f(z) = \max\{2|z_1| + z_2, 3z_2\},$$

which is clearly unbounded below. We begin with initial data

$$x_0 = \begin{bmatrix} 9 \\ 7 \end{bmatrix}, \quad H_0 = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \quad k = 0.$$

Easy calculations give

$$\nabla f_0 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \quad p_0 = -\begin{bmatrix} 5 \\ 3 \end{bmatrix}, \quad \alpha_0 = 2, \quad x_1 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \nabla f_1 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

We then obtain

$$y_0 = \begin{bmatrix} -4 \\ 0 \end{bmatrix}, \quad p_0^T y_0 = 20, \quad V_0 = \begin{bmatrix} 0 & 0 \\ -0.6 & 1 \end{bmatrix}, \quad H_0 = \begin{bmatrix} 2.5 & 1.5 \\ 1.5 & 1.42 \end{bmatrix}.$$

Moving to the next iteration, $k = 1$, we find

$$p_1 = \begin{bmatrix} 3.5 \\ 1.58 \end{bmatrix}, \quad \alpha_1 = 0, \quad x_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

This is a null step. □

Clearly, the occurrence of the null step at iteration $k = 1$ in the above example is insensitive to small changes in the definitions of the linear functions of which f is a pointwise maximum, or in the initial data.

Unfortunately, as the following result makes plain, each time the BFGS iteration takes a null step in a search direction linearly independent of all previous search directions corresponding to null steps, the space of possible future search directions contracts.

PROPOSITION 3.2 (null steps). *If the BFGS iteration takes a null step at iteration k (that is, $\alpha_k = 0$), then the corresponding gradient increment y_k lies in the null space of all future inverse Hessian approximations H_j (for $j > k$) and is orthogonal to all future search directions p_j.*

Proof We proceed by induction on the index j. First notice

$$V_k^T y_k \;=\; \Big(I - (p_k^T y_k)^{-1} y_k p_k^T\Big) y_k \;=\; 0,$$

Since $\alpha_k = 0$, we deduce $H_{k+1} y_k = 0$, and furthermore

$$y_k^T p_{k+1} = -y_k^T H_{k+1} \nabla f_{k+1} = 0.$$

Our claimed result therefore holds when $j = k + 1$.

Suppose, inductively, that the result holds for some value of the index $j > k$: that is, $H_j y_k = 0$ and $p_j^T y_k = 0$. We then deduce

$$V_j^T y_k \;=\; \Big(I - (p_j^T y_j)^{-1} y_j p_j^T\Big) y_k \;=\; y_k,$$

so

$$H_{j+1}y_k \;=\; V_jH_jV_j^Ty_k + \alpha_j(p_j^Ty_j)^{-1}p_jp_j^Ty_k \;=\; 0,$$

and furthermore

$$y_k^Tp_{j+1} = -y_k^TH_{j+1}\nabla f_{j+1} = 0.$$

The result now follows by induction. □

As a consequence of this property, the BFGS iteration with exact line search may jam at a nonoptimal point. Our previous example, which we continue next, is a case in point.

EXAMPLE 3.3. Picking up Example 3.1 where we left it, we next find

$$\nabla f_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \quad y_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad p_1^Ty_1 = 10.16, \quad V_1 = \begin{bmatrix} 0.31\ldots & -0.68\ldots \\ -0.31\ldots & 0.68\ldots \end{bmatrix}.$$

Hence we obtain

$$H_2 \;=\; (0.27\ldots)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad p_2 \;=\; (0.27\ldots)\begin{bmatrix} 3 \\ -3 \end{bmatrix},$$

so, as we expect by the previous result, $H_2y_1 = 0$ and $p_2^Ty_1 = 0$. It is easy to check that the next iterate x_3 is zero, so we have made progress reducing the value of the function f. However, at this point, the iteration jams. Any future search direction p_j (for $j > 2$) must satisfy $p_j^Ty_1 = 0$, so must be a multiple of the vector $[1, -1]^T$. However, at zero, neither of the vectors $\pm[1, -1]^T$ is a descent direction, so the iteration makes no further progress. Nonetheless, zero is not optimal: indeed, the vector $[0, -1]^T$ is a descent direction there. □

In conclusion, applying the BFGS iteration with an exact line search to a polyhedral convex function may fail.

4. Success on the Euclidean Norm. We now turn to the analysis of BFGS with an exact line search on the Euclidean norm $f = \|\cdot\|$. For this function the only point of nonsmoothness is the origin, so breakdown cannot take place unless the method generates the exact solution, as it will on the first iteration if H_0 is a multiple of the identity.

We have

$$\nabla f_k = \frac{x_k}{\|x_k\|} \quad \text{and} \quad p_k = -\frac{H_kx_k}{\|x_k\|}.$$

Using an exact line search, the step α_k is characterized by the relationship

$$0 = p_k^Tx_{k+1} = p_k^T(x_k + \alpha_kp_k),$$

so

$$\alpha_k = -\frac{p_k^Tx_k}{\|p_k\|^2} = \frac{x_k^TH_kx_k}{x_k^TH_k^2x_k}\|x_k\|,$$

and

$$x_{k+1} = x_k - \frac{x_k^T H_k x_k}{x_k^T H_k^2 x_k} H_k x_k.$$

Notice

$$\|x_{k+1}\|^2 = \|x_k\|^2 - \frac{(x_k^T H_k x_k)^2}{x_k^T H_k^2 x_k}.$$

Furthermore,

$$y_k = \frac{x_{k+1}}{\|x_{k+1}\|} - \frac{x_k}{\|x_k\|},$$

so

$$p_k^T y_k = -\frac{p_k^T x_k}{\|x_k\|} = \frac{x_k^T H_k x_k}{\|x_k\|^2}.$$

We next introduce some notation for the normalized iterates:

$$u_k = \frac{x_k}{\|x_k\|}, \quad \beta_k = \|x_k\|, \quad q_k = \frac{p_k}{\|p_k\|}, \quad \text{and} \quad \sigma_k = q_k^T u_k.$$

With this notation we have

$$p_k = -H_k u_k, \quad \alpha_k = -\frac{\sigma_k \beta_k}{\|p_k\|} \quad \text{and} \quad \frac{x_{k+1}}{\|x_k\|} = u_k - \sigma_k q_k.$$

Hence,

$$u_{k+1} = \gamma_k^{-1}(u_k - \sigma_k q_k), \quad \text{where} \quad \gamma_k = \beta_{k+1}\beta_k^{-1} = \sqrt{1 - \sigma_k^2}.$$

Furthermore,

$$y_k = u_{k+1} - u_k \quad \text{and} \quad p_k^T y_k = -p_k^T u_k.$$

Thus we have the updates

$$V_k = I + \sigma_k^{-1} q_k (u_{k+1} - u_k)^T \quad \text{and} \quad H_{k+1} = V_k H_k V_k^T + \beta_k q_k q_k^T$$

As a final simplification, define

$$W_k = \frac{H_k}{\beta_k} \quad \text{and} \quad r_k = \frac{p_k}{\beta_k}. \tag{4.1}$$

The iteration now becomes the following.

ALGORITHM 4.2. *At iteration $k = 0, 1, 2, \ldots$, given the unit vector $u_k \in \mathbf{R}^n$ and the positive-definite matrix W_k, we compute*

$$\begin{aligned}
r_k &= -W_k u_k \\
q_k &= \frac{r_k}{\|r_k\|} \\
\sigma_k &= q_k^T u_k \\
\gamma_k &= \sqrt{1 - \sigma_k^2} \\
u_{k+1} &= \frac{1}{\gamma_k}(u_k - \sigma_k q_k) \\
V_k &= I + \frac{q_k (u_{k+1} - u_k)^T}{\sigma_k} \\
W_{k+1} &= \frac{1}{\gamma_k}\Big(V_k W_k V_k^T + q_k q_k^T\Big).
\end{aligned}$$

(If $\gamma_k = 0$, the algorithm terminates.)

Define two orthogonal matrices

$$R = \begin{bmatrix} \frac{1}{2} & \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} & \frac{1}{2} \end{bmatrix} \quad \text{and} \quad S = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}.$$

Then for any vector $u \in \mathbf{R}^2$, the vectors Ru and Su are obtained by rotating u clockwise through angles $\frac{\pi}{3}$ and $\frac{\pi}{2}$ respectively. Notice that R and S commute.

THEOREM 4.3. *Consider BFGS with exact line search applied to the Euclidean norm in $\mathbf{R}^2$. Suppose the algorithm does not terminate. Then the sequence of iterates $\{x_k\}$ converge to zero at Q-linear rate $\frac{1}{2}$. More precisely, for some strictly positive constant τ, we have*

$$\|x_k\| \sim \frac{\tau}{2^k} \quad \text{as } k \to \infty.$$

The iterates eventually rotate around zero with consistent orientation, either clockwise or counterclockwise, through an angle of magnitude approaching $\frac{\pi}{3}$. The step α_k satisfies

$$\alpha_k \to \frac{1}{4} \quad \text{as } k \to \infty.$$

Furthermore, the inverse Hessian approximation H_k satisfies

$$\text{spectrum}(H_k) \ \sim \ \frac{1}{2^k}\{3 + \sqrt{3}, 3 - \sqrt{3}\}.$$

In terms of Algorithm 4.2 (in the case $n = 2$), if the algorithm does not terminate, then we have the following properties. As $k \to \infty$,

$$\sigma_k \to -\frac{\sqrt{3}}{2}, \quad \gamma_k \to \frac{1}{2}, \quad \|r_k\| \to 2\sqrt{3}.$$

Furthermore, there exists a unit vector $u \in \mathbf{R}^2$ such that exactly one of the following two cases holds.

(i) **(Clockwise case)** *As $k \to \infty$,*

$$R^{-k}u_k \to u, \quad q_k = Su_{k+1},$$
$$R^{-k}W_kR^k \to [u \ \ Su] \begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix} [u \ \ Su]^T.$$

In particular, the vectors u_k eventually rotate clockwise through an angle approaching $\frac{\pi}{3}$.

(i) **(Counterclockwise case)** *As $k \to \infty$,*

$$R^{k}u_k \to u, \quad q_k = S^{-1}u_{k+1},$$
$$R^{k}W_kR^{-k} \to [u \ \ S^{-1}u] \begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix} [u \ \ S^{-1}u]^T.$$

In particular, the vectors u_k eventually rotate counterclockwise through an angle approaching $\frac{\pi}{3}$.

Since both the matrices $[u \quad Su]$ *and* $[u \quad S^{-1}u]$ *are orthogonal, in either case the eigenvalues of the matrix* W_k *approach* $3 \pm \sqrt{3}$.

Proof We first make some observations independent of the dimension n, specifically

$$\sigma_k < 0, \quad \gamma_k \geq 0, \quad q_k^T u_{k+1} = 0, \quad u_k^T u_{k+1} = \gamma_k.$$

Furthermore, we have

$$\begin{aligned} V_k q_k &= 0 \\ V_k u_{k+1} &= \frac{1-\gamma_k}{\sigma_k}\, q_k + u_{k+1} \\ V_k^T u_{k+1} &= u_{k+1} \\ V_k^T q_k &= \frac{1-\gamma_k}{\sigma_k}\, u_{k+1}. \end{aligned}$$

Henceforth we assume $n = 2$. The set $\{q_k, u_{k+1}\}$ is then an orthonormal basis for $\mathbf{R}^2$, so using the above relationships gives

$$V_k = \Big(\frac{1-\gamma_k}{\sigma_k}\, q_k + u_{k+1}\Big) u_{k+1}^T. \tag{4.4}$$

Our initial aim is to find a recurrence relationship for the sequence $\{\gamma_k\}$. To this end, assume the iteration does not terminate, and define another sequence

$$\mu_k = u_{k+1}^T W_k u_{k+1} > 0.$$

Then

$$r_{k+1} = -W_{k+1}u_{k+1} = -\frac{1}{\gamma_k} V_k W_k u_{k+1} = -\frac{\mu_k}{\gamma_k}\Big(\frac{1-\gamma_k}{\sigma_k}\, q_k + u_{k+1}\Big), \tag{4.5}$$

so the vector q_{k+1} is the vector

$$-\frac{1-\gamma_k}{\sigma_k}\, q_k - u_{k+1} \tag{4.6}$$

normalized. The above vector has norm

$$\sqrt{\Big(\frac{1-\gamma_k}{\sigma_k}\Big)^2 + 1} \;=\; \sqrt{\frac{2}{1+\gamma_k}},$$

so

$$\sigma_{k+1} = u_{k+1}^T q_{k+1} = -\sqrt{\frac{1+\gamma_k}{2}}.$$

We deduce our desired recurrence relationship:

$$\gamma_{k+1} = \sqrt{\frac{1-\gamma_k}{2}}. \tag{4.7}$$

The continuous function $f : [0,1] \to \mathbf{R}$ defined by

$$f(\gamma) = \sqrt{\frac{1-\gamma}{2}}$$

has derivative on the interval $[0,1)$ given by

$$f'(\gamma) = -\frac{1}{4f(\gamma)} < 0.$$

Hence f is decreasing on this interval, and so is f'. Clearly f maps the interval $[0,1]$ onto the interval $I = [0, \frac{1}{\sqrt{2}}]$. For $\gamma \in I$,

$$|f'(\gamma)| = -f'(\gamma) \leq -f'\Big(\frac{1}{\sqrt{2}}\Big) < -f'\Big(\frac{7}{8}\Big) = 1,$$

Thus f is a contraction mapping on I, so for any initial point $\gamma_0 \in [0,1]$, the iteration(4.7) converges to the unique fixed point of f, namely $\frac{1}{2}$. We can even specify the rate of convergence:

$$\frac{\gamma_{k+1} - \frac{1}{2}}{\gamma_k - \frac{1}{2}} = \frac{\gamma_{k+1} - f(\frac{1}{2})}{\gamma_k - \frac{1}{2}} \to f'\Big(\frac{1}{2}\Big) = -\frac{1}{2}. \tag{4.8}$$

Next, consider the sequence $\beta_k = \|x_k\|$. From the relationship $\beta_{k+1} = \gamma_k \beta_k$, we deduce

$$\log(2^{k+1}\beta_{k+1}) - \log(2^k \beta_k) = \log(2\gamma_k).,$$

Hence, by induction, we obtain

$$2^{k+1}\beta_{k+1} = \beta_0 + \sum_{j=0}^{k} \log(2\gamma_k).$$

Since $\log(2\gamma_k) \sim 2\gamma_k - 1$ as $k \to \infty$, we deduce from the limit (4.8) that the series $\sum \log(2\gamma_k)$ is eventually alternating, with terms of decreasing magnitude, so converges. Since $\beta_k > 0$ for all k, we deduce that the sequence $\{2^k \beta_k\}$ converges to some limit $\tau \in \mathbf{R}_{++}$.

Since $u_{k+2}^T u_{k+1} = \gamma_{k+1}$, we deduce

$$u_{k+2} = \gamma_{k+1} u_{k+1} \pm \sigma_{k+1} q_k.$$

To decide which case is true, we simply need to check the sign of the quantity

$$q_k^T u_{k+2} \;=\; q_k^T \frac{1}{\gamma_{k+1}} (u_{k+1} - \sigma_{k+1} q_{k+1}),$$

which equals the sign of $q_k^T q_{k+1}$, or equivalently, using the expression (4.6), of the quantity

$$q_k^T \Big(-\frac{1-\gamma_k}{\sigma_k}\, q_k - u_{k+1} \Big) \;=\; -\frac{1-\gamma_k}{\sigma_k} \;>\; 0.$$

We thus deduce

$$u_{k+2} = \gamma_{k+1} u_{k+1} - \sigma_{k+1} q_k. \tag{4.9}$$

We next derive a recurrence relationship for the sequence $\{\mu_k\}$. From equation (4.9), we observe

$$V_k^T u_{k+2} \;=\; \gamma_{k+1} V_k^T u_{k+1} - \sigma_{k+1} V_k^T q_k \;=\; \Big(\gamma_{k+1} - \sigma_{k+1} \frac{1-\gamma_k}{\sigma_k}\Big) u_{k+1}$$

and

$$q_k^T u_{k+2} = -\sigma_{k+1}.$$

We now obtain the recurrence relationship that we need for the sequence $\{\mu_k\}$:

$$\begin{aligned}
\gamma_k \mu_{k+1} &= \gamma_k u_{k+2}^T W_{k+1} u_{k+2} \\
&= u_{k+2}^T \Big(V_k W_k V_k^T + q_k q_k^T \Big) u_{k+2} \\
&= (V_k^T u_{k+2})^T W_k (V_k^T u_{k+2}) + (q_k^T u_{k+2})^2 \\
&= \mu_k \Big(\gamma_{k+1} - \sigma_{k+1} \frac{1-\gamma_k}{\sigma_k} \Big)^2 + \sigma_{k+1}^2.
\end{aligned}$$

Since we know $\gamma_k \to \frac{1}{2}$ and $\sigma_k \to -\frac{\sqrt{3}}{2}$, we deduce

$$\mu_{k+1} - \lambda_k \mu_k \to \beta \quad \text{for some sequence } \lambda_k \to 0, \tag{4.10}$$

where $\beta = \frac{3}{2}$. We now make the following assertion.

Claim: Any sequence of numbers $\{\mu_k\}$ satisfying the property (4.10) must converge to β.

Proof of claim. First note that, by making the change of variables $\hat{\mu}_k = \mu_k - \beta$, we can without loss of generality assume $\beta = 0$. Fix any $\epsilon > 0$. There exists an integer $\bar{k} > 0$ such that, for all integers $k \geq \bar{k}$, we have $|\mu_{k+1} - \lambda_k \mu_k| < \epsilon$ and $|\lambda_k| < \epsilon$, and hence $|\mu_{k+1}| < \epsilon + \epsilon |\mu_k|$. By induction, we deduce

$$|\mu_{\bar{k}+m}| < \epsilon^m |\mu_{\bar{k}}| + \frac{\epsilon(1-\epsilon^m)}{1-\epsilon}$$

for all integers $m > 0$. Letting $m \to \infty$ shows

$$\limsup_{k \to \infty} |\mu_k| \leq \frac{\epsilon}{1-\epsilon}.$$

The claim now follows, since ϵ is arbitrary.

We have therefore shown $\mu_k \to \frac{3}{2}$. From equation (4.5), we see

$$\|r_{k+1}\| = \frac{\mu_k}{\gamma_k} \sqrt{\frac{2}{1+\gamma_k}} \to 2\sqrt{3}.$$

The definition (4.1) implies $\|r_k\| = \|p_k\|/\beta_k$. Hence the step α_k satisfies

$$\alpha_k = -\frac{\sigma_k \beta_k}{\|p_k\|} = -\frac{\sigma_k}{\|r_k\|} \to \frac{1}{4}.$$

The next part of the argument is most easily viewed in the complex plane, so let us define three sequences of complex numbers of modulus one corresponding to the unit vectors u_k:

$$\begin{aligned}
\hat{u}_k &= (u_k)_1 + (u_k)_2 \sqrt{-1} \\
\hat{q}_k &= (q_k)_1 + (q_k)_2 \sqrt{-1} \\
w_k &= \frac{\hat{u}_{k+1}}{\hat{u}_k}.
\end{aligned}$$

We know $u_k^T u_{k+1} \to \frac{1}{2}$, and furthermore, by equation (4.9),

$$u_k^T u_{k+2} = \gamma_{k+1}\gamma_k - \sigma_{k+1}\sigma_k \to -\frac{1}{2}.$$

We therefore obtain

$$\begin{aligned} \operatorname{Re} w_k = \operatorname{Re}\frac{\hat{u}_{k+1}}{\hat{u}_k} &= u_k^T u_{k+1} = \gamma_k \to \frac{1}{2} \\ \operatorname{Re}(w_k w_{k+1}) = \operatorname{Re}\frac{\hat{u}_{k+2}}{\hat{u}_k} &= u_k^T u_{k+2} \to -\frac{1}{2}. \end{aligned}$$

The first limit implies

$$|\operatorname{Im} w_k| \to \frac{\sqrt{3}}{2}.$$

But notice

$$(\operatorname{Im} w_{k+1})(\operatorname{Im} w_k) = (\operatorname{Re} w_{k+1})(\operatorname{Re} w_k) - \operatorname{Re}(w_k w_{k+1}) \to \frac{3}{4}$$

Hence in fact

$$\text{either } \operatorname{Im} w_k \to \frac{\sqrt{3}}{2} \quad \text{or } \operatorname{Im} w_k \to -\frac{\sqrt{3}}{2}.$$

In the first case, the points $\hat{u}_k$ (or equivalently, the unit vectors u_k) eventually always rotate counterclockwise, by an angle of magnitude approaching $\frac{\pi}{3}$. The second case is analogous, but with clockwise orientation.

Consider the counterclockwise case first. Write

$$w_k = e^{\theta_k \sqrt{-1}}, \quad \text{where } \theta_k \in [0, 2\pi).$$

We therefore have $w_k = \cos^{-1}\gamma_k$. If we now fix any number ρ in the interval $(\frac{1}{2}, 1)$, then from the linear convergence property (4.8) we know

$$\left|\gamma_k - \frac{1}{2}\right| < \rho^k, \quad \text{for all sufficiently large } k.$$

Since the function $\cos^{-1}$ has Lipschitz constant less than 2 around the point $\frac{1}{2}$, we deduce

$$\left|\theta_k - \frac{\pi}{3}\right| < 2\rho^k, \quad \text{for all sufficiently large } k. \tag{4.11}$$

Now consider the sequence of complex unit vectors $y_j = \hat{u}_{6j}$. Notice, for any positive integers $j \le l$ we have

$$|y_l - y_j| = \left|\frac{y_l}{y_j} - 1\right| = \left|\frac{\hat{u}_{6l}}{\hat{u}_{6j}} - 1\right| = \left|\prod_{k=6j}^{6l-1} w_k - 1\right|.$$

However, providing j is sufficiently large, inequality (4.11) implies

$$\left|\sum_{k=6j}^{6l-1} \theta_k - 2(l-j)\pi\right| < 2\sum_{k=6j}^{6l-1} \rho^k = 2\rho^{6j}\frac{1-\rho^{6(l-j)}}{1-\rho} \to 0,$$

as $j \to \infty$, and hence

$$\prod_{k=6j}^{6l-1} w_k = e^{\sqrt{-1}\sum_{k=6j}^{6l-1}\theta_k} \to 1.$$

Thus the sequence $\{y_j\}$ is Cauchy, so converges. Consequently, for some complex number $\hat{u}$ of modulus one, we have $\hat{u}_{6j} \to \hat{u}$ as $j \to \infty$. More generally, for any positive integer m we have, as $j \to \infty$,

$$\hat{u}_{6j+m} = \hat{u}_{6j} \prod_{k=6j}^{6j+m-1} w_k \to \hat{u}e^{m\frac{\pi}{3}\sqrt{-1}},$$

so

$$e^{-(6j+m)\frac{\pi}{3}\sqrt{-1}}\hat{u}_{6j+m} \to \hat{u}.$$

Putting together the cases $m = 0, 1, 2, 3, 4, 5$ gives

$$e^{-k\frac{\pi}{3}\sqrt{-1}}\hat{u}_k \to \hat{u}, \quad \text{as } k \to \infty.$$

The clockwise case is almost identical, with $\frac{\pi}{3}$ replaced by $-\frac{\pi}{3}$.

Since the unit vectors q_k and u_{k+1} are orthogonal, we know that one of the two cases $\hat{q}_k = \pm\hat{u}_{k+1}\sqrt{-1}$ holds. By equation (4.9), $\hat{q}_k = \hat{u}_{k+1}\sqrt{-1}$ holds if and only if

$$\hat{u}_{k+2} = \gamma_{k+1}\hat{u}_{k+1} - \sigma_{k+1}\hat{u}_{k+1}\sqrt{-1},$$

or equivalently

$$w_{k+1} = \gamma_{k+1} - \sigma_{k+1}\sqrt{-1}.$$

In that case, the vector $\hat{u}_{k+2}$ is obtained from $\hat{u}_{k+1}$ by counterclockwise rotation. To summarize, in the counterclockwise case we have $\hat{q}_k = \hat{u}_{k+1}\sqrt{-1}$ for all sufficiently large k; in the clockwise case, we have instead $\hat{q}_k = -\hat{u}_{k+1}\sqrt{-1}$

Converting our conclusions back to $\mathbf{R}^2$, we deduce the existence of a unit vector $u \in \mathbf{R}^2$ with the following property: either

$$R^{-k}u_k \to u \quad \text{and} \quad q_k = Su_{k+1}, \quad \text{as } k \to \infty$$

(the clockwise case), or

$$R^k u_k \to u \quad \text{and} \quad q_k = S^{-1}u_{k+1}, \quad \text{as } k \to \infty$$

(the counterclockwise case).

Equation (4.4) shows

$$\begin{aligned} W_{k+1} &= \frac{\mu_k}{\gamma_k}\Big(\frac{1-\gamma_k}{\sigma_k}\,q_k + u_{k+1}\Big)\Big(\frac{1-\gamma_k}{\sigma_k}\,q_k + u_{k+1}\Big)^T + \frac{1}{\gamma_k}q_k q_k^T \\ &= [u_{k+1}\ q_k]\begin{bmatrix} \frac{\mu_k}{\gamma_k}(\frac{1-\gamma_k}{\sigma_k})^2 + \frac{1}{\gamma_k} & \frac{\mu_k}{\gamma_k}(\frac{1-\gamma_k}{\sigma_k}) \\ \frac{\mu_k}{\gamma_k}(\frac{1-\gamma_k}{\sigma_k}) & \frac{\mu_k}{\gamma_k} \end{bmatrix}[u_{k+1}\ q_k]^T. \end{aligned}$$

Using the fact that the matrices R and S commute, we obtain, in the clockwise case

$$R^{-(k+1)}W_{k+1}R^{k+1} \to [u\ \ Su]\begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix}[u\ \ Su]^T,$$

and in the counterclockwise case

$$R^{k+1}W_{k+1}R^{-(k+1)} \to [u \;\; S^{-1}u] \begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix} [u \;\; S^{-1}u]^T,$$

as $k \to \infty$. Our claims about W_k now follow, and those about the inverse Hessian approximation H_k follow from the definition (4.1). □

The following example illustrates well a typical limit cycle.

THEOREM 4.12. *(Spiraling Iterates) Suppose Algorithm 4.2 is initiated with the values*

$$u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad W_0 = \begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix}.$$

Then in each iteration $k = 0, 1, 2, \ldots$, we have

$$u_k = R^k u_0, \quad q_k = R^k \begin{bmatrix} -\frac{\sqrt{3}}{2} \\ \frac{1}{2} \end{bmatrix}$$

$$\sigma_k = -\frac{\sqrt{3}}{2}, \quad \gamma_k = \frac{1}{2}$$

$$W_k = R^k W_0 R^{-k}.$$

The proof is a routine but lengthy induction on k, and is omitted.

COROLLARY 4.13 (spiral behavior). *Consider BFGS with exact line search applied to the Euclidean norm in $\mathbf{R}^2$, initialized by*

$$x_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \textit{and} \quad H_0 = \begin{bmatrix} 3 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{bmatrix}$$

The method generates a sequence of vectors x_k that rotate clockwise through an angle of $\frac{\pi}{3}$ and shrink by a factor $\frac{1}{2}$ at each iteration.

5. Numerical Experiments for the Norm. We do not know how to extend the analysis given in the previous section to $n > 2$. However, numerical experiments implementing the BFGS iteration using the known minimizing steplength α_k indicate that similar results surely hold, not only for $f(x) = \|x\|$, but also for $f(x) = \|Ax\|$, where A is a nonsingular matrix. Remarkably, the convergence of the function values $f_k = \|Ax_k\|$ is observed to be asymptotically Q-linear with rates that are closer to 1 for larger n but independent of A for fixed n. In Figure 5.1, the top left and top right panels show the evolution of f_k for the cases $A = I$ and $A = \mathrm{diag}(1, \ldots, 1/n)$ respectively, for typical runs for $n = 2, 4, 8$ and 16, with both x and H initialized randomly. The bottom two panels display estimated Q-linear convergence rates for the sequence $\{f_k\}$ for varying n, again for $A = I$ and $A = \mathrm{diag}(1, \ldots, 1/n)$, respectively. Each asterisk plots the mean of 10 observed convergence rates, each computed by a least squares fit to a different randomly initialized sequence. Each approximation to a sequence $\{f_k\}$ was made using 40% of the iterates, excluding the first half to avoid the transient initial behavior, and excluding the final 10% to avoid contamination from rounding errors. Since the convergence rates are close to 1 for large n, we plot

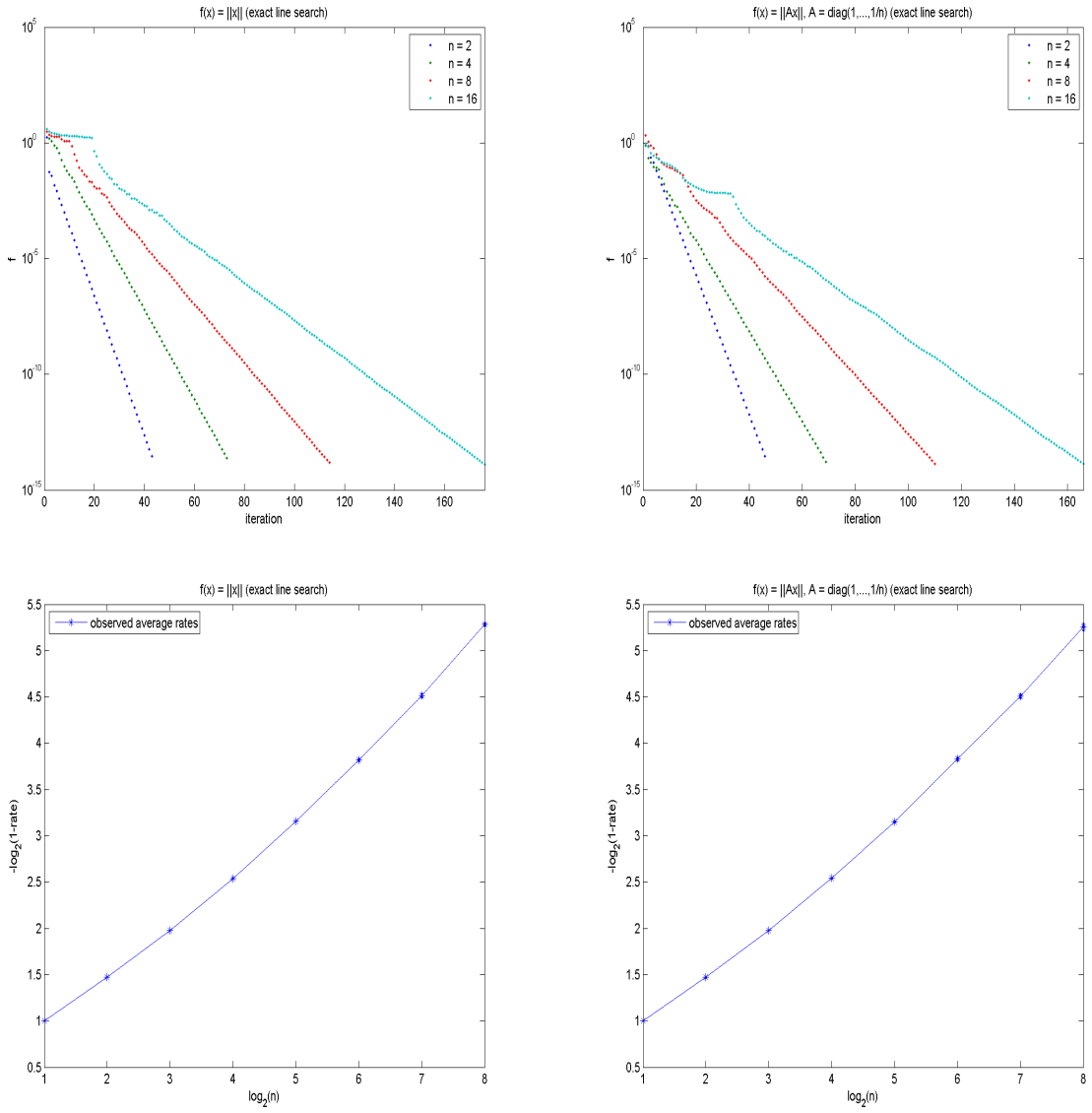

FIG. 5.1. *Convergence of BFGS with an exact line search on* $f(x) = \|Ax\|$. *Top left: plots function values for typical runs for* $n = 2, 4, 8$ *and* 16 *and* $A = I$. *Bottom left: plots* $-\log_2(1-r)$ *against* $\log_2(n)$ *where* r *is the estimated Q-linear convergence rate for the sequence of function values, averaged over 10 runs, for* $A = I$. *Top and bottom right: same for* $A = diag(1, \ldots, 1/n)$.

$-\log_2(1-r)$ against $\log_2(n)$, where r is the average estimated convergence rate. The observed rates grow surprisingly consistently with n, somewhat faster than $1-1/\sqrt{2n}$. Furthermore, the rate of convergence is apparently independent of H_0 unless the method terminates at the origin.

The evolution of the exact steplengths α_k, for the same runs with $n = 2, 4, 8$ and 16, is shown in the top left and right panels of Figure 5.2. Again, there is essentially no difference in the limiting behavior between the cases $A = I$ and $A = \text{diag}(1, \ldots, 1/n)$.

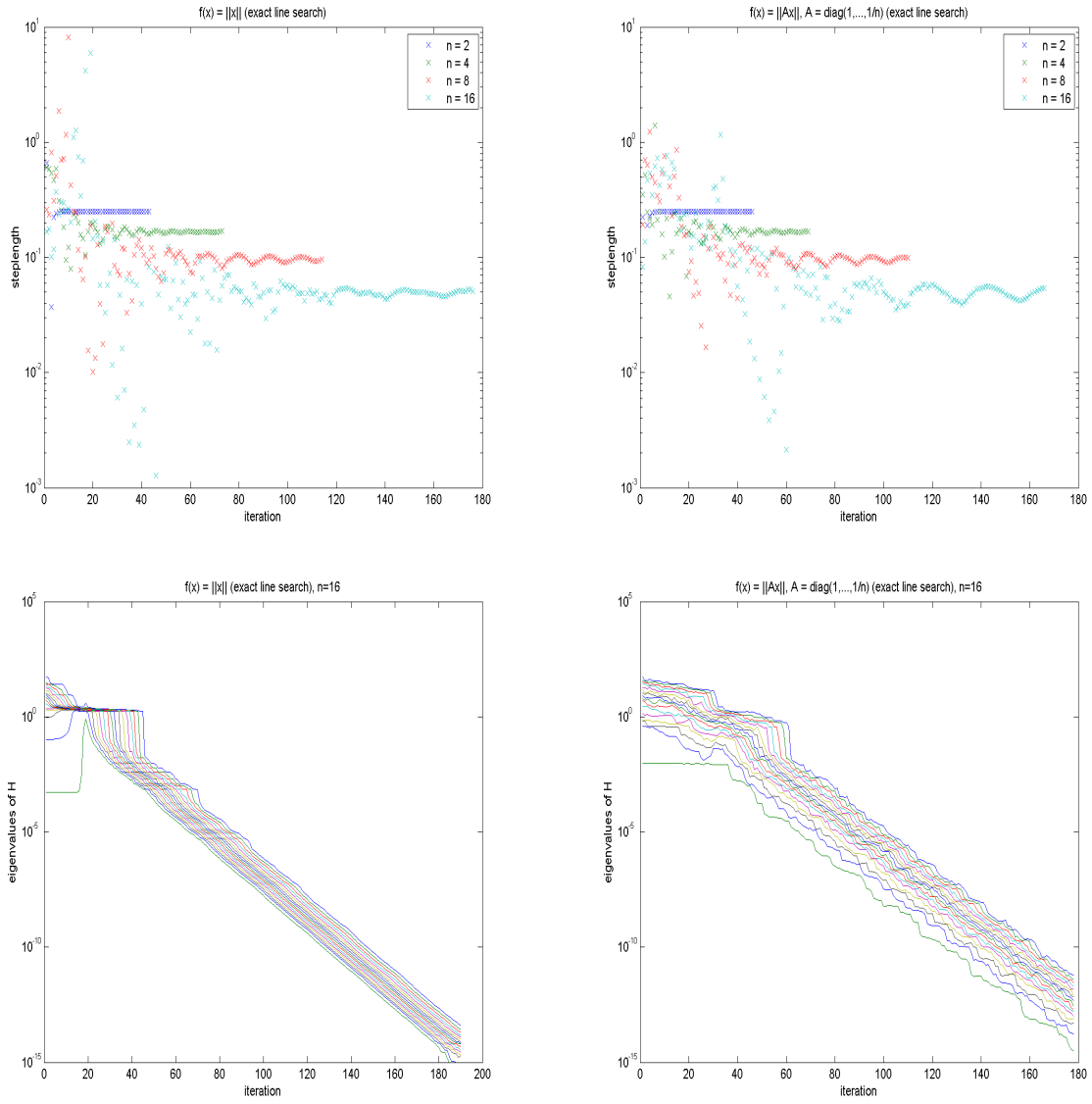

FIG. 5.2. *Convergence of BFGS with an exact line search on* $f(x) = \|Ax\|$. *Top left: plots the exact steplengths for* $n = 2, 4, 8$ *and* 16 *and* $A = I$. *Bottom: eigenvalues of* H_k *for the case* $n = 16$, $A = I$. *Top and bottom right: same for* $A = diag(1, \ldots, 1/n)$.

In both cases α_k converges to the same limits, and these limits decrease as n increases. Note that the limiting behavior of the eigenvalues of H_k *does* depend on A, as is seen by comparing the lower left and right panels of Figure 5.2, which plot the eigenvalues of H_k for the case $n = 16$. In both cases, the eigenvalues of H_k converge to zero, but in the case $A = I$ (bottom left), the convergence is Q-linear and, when the eigenvalues are scaled so the largest is fixed to one, the scaled eigenvalues converge to quantities that are independent of the initial x and H (not shown). For $A = \text{diag}(1, \ldots, 1/n)$ this is not the case; the eigenvalues converge to zero but not Q-linearly (bottom right),

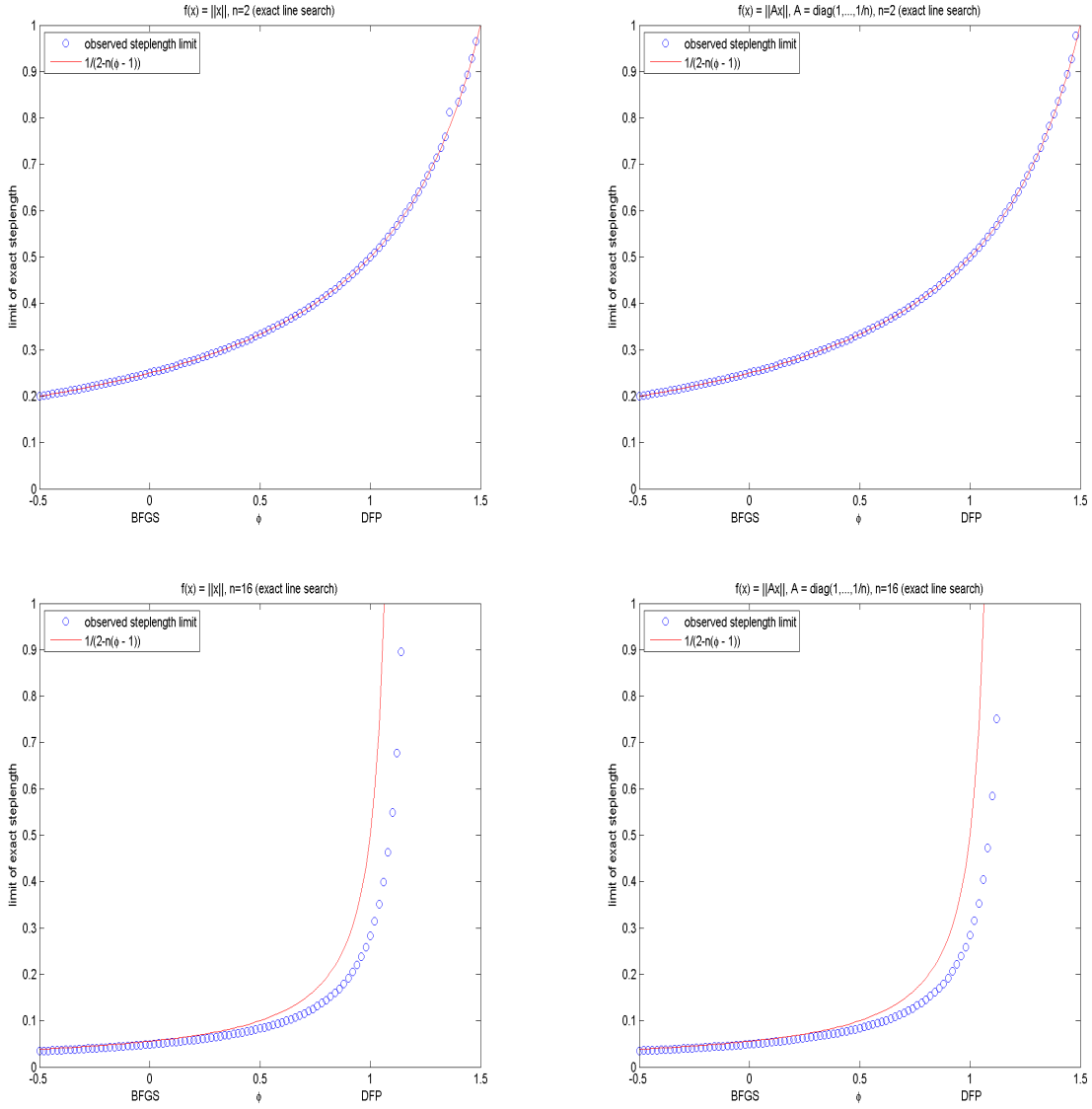

FIG. 6.1. *Limiting steplengths for the Broyden family using an exact line search on* $f(x) = \|Ax\|$. *Top and bottom left: plots the limiting steplengths as a function of the Broyden parameter* ϕ *for* $A = I$ *with* $n = 2$ *and* $n = 16$ *respectively. Top and bottom right: same for* $A = diag(1, \ldots, 1/n)$.

and the scaled eigenvalues of H_k do not converge. It is quite remarkable that BFGS automatically produces a sequence H_k whose limiting behavior *is* dependent on A but that results in a sequence f_k that converges Q-linearly to zero at a rate independent of A, for fixed n. Of course, this independence is familiar for smooth functions, for which convergence is superlinear in nondegenerate cases.

6. Limiting Steplengths for the Broyden Family on the Norm. Dixon's theorem [Dix72], that all variable metric methods in the Brodyen family generate

the same sequence of iterates $\{x_k\}$ when an exact line search is used, applies to the Euclidean norm function without modification. As is well known [NW06, Chap. 6], the Broyden family of updates is defined by a parameter ϕ: when $\phi = 0$, the Broyden update reduces to BFGS, while for $\phi = 1$, we obtain DFP. The sequence of inverse Hessian approximations $\{H_k\}$ *does* depend on ϕ, as does the sequence of steplengths $\{\alpha_k\}$.

Numerical experiments on $f(x) = \|Ax\|$, where A is an $n \times n$ nonsingular matrix, show that the steplengths α_k converge for all $\phi \in [0, 1]$, and Figure 6.1 shows their limiting values as a function of ϕ. The top and bottom left panels show results for $A = I$ with $n = 2$ and $n = 16$ respectively, and the top and bottom right panels show the same for $A = \mathrm{diag}(1, \ldots, 1/n)$. As previously, the results are apparently independent of A. Each blue circle shows the experimentally determined limiting steplength, averaged over 10 randomly initialized runs. Experiments were carried out for ϕ ranging from -0.5 to 1.5. When $\phi < 0$, the updated matrix H_k may not be positive definite, and hence α_k may be negative; nonetheless, as long as H_k is never exactly singular, the steplengths converge to a positive value. For values of ϕ that are sufficiently large, the steplengths diverge.

The solid red curve plots the function $1/(2 - n(\phi - 1)$, which approximates the limiting steplength well for $n = 2$ and seems to be a reasonably good upper bound when $n > 2$. This implies, in the case $\phi = 0$ (BFGS), that $1/(2 + n)$ is an upper bound on the limiting steplength, which is consistent with the results reported in Section 5. For the case $\phi = 1$ (DFP), the upper bound is $1/2$. The results might suggest that DFP is more favorable for use with an inexact line search as fewer steps would be needed, at least on this example. However, this conclusion overlooks the fact that the limiting steplength diverges when ϕ is not much greater than 1, specifically somewhat more than the pole in the upper bound formula, $\phi = 1+2/n$. This indicates a possible instability for DFP, which is perhaps not surprising, given its well known poor performance, with respect to BFGS, for smooth functions [NW06].

7. Steepest Descent on Homogeneous Functions. The apparent property that the limiting behavior of variable metric methods with an exact line search on $\|Ax\|$ is independent of A is in stark contrast with the method of steepest descent on the same function, as we now show.

THEOREM 7.1. *Consider a continuous function $f\colon \mathbf{R}^n \to \mathbf{R}$ that is positively homogeneous. Suppose f is strictly positive and twice continuously differentiable on $\mathbf{R}^n \setminus \{0\}$. Define numbers*

$$M = \max\{\|\nabla^2 f(x)\| : f(x) = 1\},$$
$$m = \min\{\|\nabla f(x)\| : f(x) = 1\}.$$

Then, from any initial point, the method of steepest descent with exact line search either terminates or converges with Q-linear rate at least

$$\rho = \frac{1}{1 + \frac{m^2}{2M}}.$$

Proof The set $\{x \in \mathbf{R}^n : f(x) = 1\}$ is compact, since otherwise it contains a sequence (x_k) satisfying $\|x_k\| \to \infty$, and then, for any cluster point $\bar{x}$ of the normalized sequence $\|x_k\|^{-1} x_k$, continuity and positive homogeneity would imply the contradiction $f(\bar{x}) = 0$. Hence the number M is finite. Positive homogeneity implies the relationships

$$\nabla f(\lambda x) = \nabla f(x) \quad \text{and} \quad \nabla^2 f(\lambda x) = \frac{1}{\lambda} \nabla^2 f(x) \quad \text{for all } 0 \neq x \in \mathbf{R}^n,\ \lambda > 0.$$

In particular, we also see $m > 0$: otherwise, ∇f must vanish at some nonzero point x, and hence throughout the open line segment $(0, x)$, implying that f must be constantly zero on this line segment.

Consider any point x satisfying $f(x) = 1$, and define a vector $g = \nabla f(x)$ and a function $h\colon \mathbf{R} \to \mathbf{R}$ by $h(t) = f(x - tg)$. It then suffices to prove $\min h \le \rho$. By way of contradiction, suppose $h(t) > \rho$ for all numbers $t \in \mathbf{R}$. Then $h(0) = 1$ and $h'(t) = -g^T \nabla f(x - tg)$, and hence $h'(0) = -\|g\|^2$. Furthermore, for all t we have

$$h''(t) = g^T \nabla^2 f(x - tg) g \le \frac{M}{f(x - tg)} \|g\|^2 < \frac{M\|g\|^2}{\rho},$$

For all $t > 0$ we deduce

$$h'(t) < h'(0) + \frac{M\|g\|^2}{\rho} t = \Big(\frac{M}{\rho} t - 1\Big)\|g\|^2,$$

and hence

$$h(t) < 1 + \Big(\frac{M}{2\rho} t^2 - t\Big)\|g\|^2.$$

Putting $t = \rho/M$ then gives

$$\min h < 1 - \frac{\rho}{2M}\|g\|^2 \le 1 - \frac{\rho m^2}{2M} = \rho,$$

which is a contradiction. □

In the case $f(x) = \|Ax\|$, we know from the equivalence of steepest descent with an exact line search to the squared function $x^T A^T A x$ that $(\kappa - 1)/(\kappa + 1)$ is an upper bound on the convergence rate, where κ is the condition number of $A^T A$ [Lue84]. The theorem gives the somewhat weaker upper bound $1/(1 + 1/(4\kappa))$.

8. Conclusions. The polyhedral example makes it clear that BFGS with an exact line search cannot be recommended for general use in the nonsmooth case. On the other hand, the Euclidean norm example demonstrates a rather surprising and consistent behavior of a remarkable algorithm that remains poorly understood several decades after its introduction. Our interest in the algorithm is largely motivated by our conclusion that BFGS with an *inexact* line search *does* have a useful role to play in the minimization of nonsmooth functions, a position we explore in detail in [LO08].

Acknowledgment. Mille grazie a F. Facchinei e agli altri membri del Dipartimento di Informatica e Sistemistica dell' Università di Roma "La Sapienza", dove gran parte di questo lavoro è stato eseguito, per avermi fornito un ambiente piacevole e stimolante.

REFERENCES

[Dix72] L.C.W. Dixon. Quasi-Newton techniques generate identical points. II. The proofs of four new theorems. *Mathematical Programming*, 3:345–358, 1972.

[LF01] D.-H. Li and M. Fukushima. On the global convergence of the BFGS method for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 11:1054–1064, 2001.

[LO08] A.S. Lewis and M.L. Overton. Nonsmooth optimization via BFGS. 2008. Submitted.

[Lue84] D.G. Luenberger. *Linear and Nonlinear Programming*. Addison-Wesley, Reading, MA, 1984.

[NW06] J. Nocedal and S. Wright. *Nonlinear Optimization*. Springer, New York, second edition, 2006.

[Pow71] M.J.D. Powell. On the convergence of the variable metric algorithm. *Journal of the Institute of Mathematics and its Applications*, 7:21–36, 1971.

[Pow72] M.J.D. Powell. Some properties of the variable metric algorithm. In F.A. Lootsma, editor, *Numerical Methods for Nonlinear Optimization*, pages 1–17. Academic Press, 1972.