1. Search Directions

In this chapter we study the choice of search directions used in our basic updating scheme

$$x^{k+1} = x^k + t_k d^k \ .$$

for solving

$$\mathcal{P} \qquad \min_{x \in \mathbb{R}^n} f(x).$$

All of the search directions considered can be classified as *Newton-like* since they are all of the form

$$d^k = -H_k \nabla f(x^k),$$

where H_k is a symmetric $n \times n$ matrix. If $H_k = \mu_k I$ for all k, the resulting search directions a a scaled steepest descent direction with scale factors μ_k. More generally, we choose H_k to approximate $\nabla^2 f(x^k)^{-1}$ in order to approximate Newton's method for optimization. The Newton is important since it possesses rapid local convergence properties, and can be shown to be *scale independent*. We precede our discussion of search directions by making precise a useful notion of speed or *rate of convergence*.

1.1. Rate of Convergence.

We focus on notions of quotient rates convergence, or Q-convergence rates. Let $\{x^\nu\} \subset \mathbb{R}^n$ and $\bar{x} \in \mathbb{R}^n$ be such that $\bar{x}^\nu \to \bar{x}$. We say that $\bar{x}^\nu \to \bar{x}$ at a *linear* rate if

$$\limsup_{\nu \to \infty} \frac{\|x^{\nu+1} - \bar{x}\|}{\|x^\nu - \bar{x}\|} < 1 \ .$$

The convergence is said to be *superlinear* if this limsup is 0. The convergence is said to be *quadratic* if

$$\limsup_{\nu \to \infty} \frac{\|x^{\nu+1} - \bar{x}\|}{\|x^\nu - \bar{x}\|^2} < \infty \ .$$

For example, given $\gamma \in (0,1)$ the sequence $\{\gamma^\nu\}$ converges linearly to zero, but not superlinearly. The sequence $\{\gamma^{\nu^2}\}$ converges superlinearly to 0, but not quadratically. Finally, the sequence $\{\gamma^{2^\nu}\}$ converges quadratically to zero. Superlinear convergence is much faster than linear convergences, but quadratic convergence is much, much faster than superlinear convergence.

1.2. Newton's Method for Solving Equations.

Newton's method is an iterative scheme designed to solve nonlinear equations of the form

$$g(x) = 0, \tag{1.1}$$

where $g : \mathbb{R}^n \to \mathbb{R}^n$ is assumed to be continuously differentiable. Many problems of importance can be posed in this way. In the context of the optimization problem $\mathcal{P}$, we wish to locate critical points, that is, points at which $\nabla f(x) = 0$. We begin our discussion of Newton's method in the usual context of equation solvng.

Assume that the function g in (1.1) is continuously differentiable and that we have an approximate solution $x^0 \in \mathbb{R}^n$. We now wish to improve on this approximation. If $\overline{x}$ is a solution to (1.1), then

$$0 = g(\overline{x}) = g(x^0) + g'(x^0)(\overline{x} - x^0) + o\|\overline{x} - x^0\|.$$

Thus, if x^0 is "close" to $\overline{x}$, it is reasonable to suppose that the solution to the linearized system

$$0 = g(x^0) + g'(x^0)(x - x^0) \tag{1.2}$$

is even closer. This is Newton's method for finding the roots of the equation $g(x) = 0$. It has one obvious pitfall. Equation (1.2) may not be consistent. That is, there may not exist a solution to (1.2).

For the sake of the present argument, we assume that (3) holds, i.e. $g'(x^0)^{-1}$ exists. Under this assumption (1.2) defines the iteration scheme,

$$x^{k+1} := x^k - [g'(x^k)]^{-1} g(x^k), \tag{1.3}$$

called the Newton iteration. The associated direction

$$d^k := -[g'(x^k)]^{-1} g(x^k). \tag{1.4}$$

is called the Newton direction. We analyze the convergence behavior of this scheme under the additional assumption that only an approximation to $g'(x^k)^{-1}$ is available. We denote this approximation by J_k. The resulting iteration scheme is

$$x^{k+1} := x^k - J_k g(x^k). \tag{1.5}$$

Methods of this type are called *Newton-Like methods.*

Theorem 1.1. *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable, $x^0 \in \mathbb{R}^n$, and $J_0 \in \mathbb{R}^{n\times n}$. Suppose that there exists $\bar{x}$, $x_0 \in \mathbb{R}^n$, and $\epsilon > 0$ with $\|x_0 - \bar{x}\| < \epsilon$ such that*

(1) $g(\overline{x}) = 0$,
(2) $g'(x)^{-1}$ *exists for* $x \in B(\overline{x}; \epsilon) := \{x \in \mathbb{R}^n : \|x - \overline{x}\| < \epsilon\}$ *with*
$$\sup\{\|g'(x)^{-1}\| : x \in B(\overline{x}; \epsilon)] \le M_1$$
(3) g' *is Lipschitz continuous on* $c\ell B(\overline{x}; \epsilon)$ *with Lipschitz constant* L, *and*
(4) $\theta_0 := \frac{LM_1}{2}\|x^0 - \overline{x}\| + M_0 K < 1$ *where* $K \ge \|(g'(x^0)^{-1} - J_0)y^0\|$, $y^0 := g(x^0)/\|g(x^0)\|$, *and* $M_0 = \max\{\|g'(x)\| : x \in B(\overline{x}; \epsilon)\}$.

Further suppose that iteration (1.5) is initiated at x^0 where the J_k's are chosen to satisfy one of the following conditions;

(i) $\|(g'(x^k)^{-1} - J_k)y^k\| \le K$,
(ii) $\|(g'(x^k)^{-1} - J_k)y^k\| \le \theta_1^k K$ *for some* $\theta_1 \in (0, 1)$,
(iii) $\|(g'(x^k)^{-1} - J_k)y^k\| \le \min\{M_3\|x^k - x^{k-1}\|, K\}$, *for some* $M_2 > 0$, *or*
(iv) $\|(g'(x^k)^{-1} - J_k)y^k\| \le \min\{M_2\|g(x^k)\|, K\}$, *for some* $M_3 > 0$,

where for each $k = 1, 2, \ldots$, $y^k := g(x^k)/\left\|g(x^k)\right\|$.

These hypotheses on the accuracy of the approximations J_k yield the following conclusions about the rate of convergence of the iterates x^k.

(a) *If (i) holds, then $x^k \to \overline{x}$ linearly.*
(b) *If (ii) holds, then $x^k \to \overline{x}$ superlinearly.*
(c) *If (iii) holds, then $x^k \to \overline{x}$ two step quadratically.*
(d) *If (iv) holds, then $x^k \to \overline{x}$ quadratically.*

Proof. We begin by inductively establishing the basic inequalities

$$\|x^{k+1} - \overline{x}\| \le \frac{LM_1}{2}\|x^k - \overline{x}\|^2 + \|(g'(x^k)^{-1} - J_k)g(x^k)\|, \tag{1.6}$$

and

$$\|x^{k+1} - \overline{x}\| \le \theta_0\|x^k - \overline{x}\| \tag{1.7}$$

as well as the inclusion

$$x^{k+1} \in B(\bar{x};\epsilon) \tag{1.8}$$

for $k = 0, 1, 2, \ldots$. For $k = 0$ we have

$$\begin{aligned} x^1 - \overline{x} &= x^0 - \overline{x} - g'(x^0)^{-1}g(x^0) + \left[g'(x^0)^{-1} - J_0\right] g(x^0) \\ &= g'(x^0)^{-1}\left[g(\overline{x}) - (g(x^0) + g'(x^0)(\overline{x} - x^0))\right] \\ &\quad + \left[g'(x^0)^{-1} - J_0\right] g(x^0), \end{aligned}$$

since $g'(x^0)^{-1}$ exists by the hypotheses. Consequently, the hypothese (1)–(4) plus the quadratic bound lemma imply that

$$\begin{aligned} \left\|x^{k+1} - \overline{x}\right\| &\le \left\|g'(x^0)^{-1}\right\| \left\|g(\overline{x}) - \left(g(x^0) + g'(x^0)(\overline{x} - x^0)\right)\right\| \\ &\qquad + \left\|\left(g'(x^0)^{-1} - J_0\right) g(x^0)\right\| \\ &\le \frac{M_1 L}{2}\left\|x^0 - \overline{x}\right\|^2 + K\left\|g(x^0) - g(\overline{x})\right\| \\ &\le \frac{M_1 L}{2}\left\|x^0 - \overline{x}\right\|^2 + M_0 K\left\|x^0 - \overline{x}\right\| \\ &\le \theta_0\left\|x^0 - \overline{x}\right\| < \epsilon, \end{aligned}$$

whereby (1.6) – (1.8) are established for $k = 0$.

Next suppose that (1.6) – (1.8) hold for $k = 0, 1, \ldots, s - 1$. We show that (1.6) – (1.8) hold at $k = s$. Since $x^s \in B(\overline{x}, \epsilon)$, hypotheses (2)–(4) hold at x^s, one can proceed exactly as in the case $k = 0$ to obtain (1.6). Now if any one of (i)–(iv) holds, then (i) holds. Thus, by (1.6), we find that

$$\begin{aligned} \|x^{s+1} - \overline{x}\| &\le \tfrac{M_1 L}{2}\|x^s - \overline{x}\|^2 + \|(g'(x^s)^{-1} - J_s)g(x^s)\| \\ &\le \left[\tfrac{M_1 L}{2}\theta_0^s\|x^0 - \overline{x}\| + M_0 K\right]\|x^s - \overline{x}\| \\ &\le \left[\tfrac{M_1 L}{2}\|x^0 - \overline{x}\| + M_0 K\right]\|x^s - \overline{x}\| \\ &= \theta_0\|x^s - \overline{x}\|. \end{aligned}$$

Hence $\|x^{s+1} - \overline{x}\| \le \theta_0\|x^s - \overline{x}\| \le \theta_0\epsilon < \epsilon$ and so $x^{s+1} \in B(\overline{x}, \epsilon)$. We now proceed to establish (a)–(d).

(a) This clearly holds since the induction above established that

$$\left\|x^{k+1} - \overline{x}\right\| \le \theta_0\left\|x^k - \overline{x}\right\|.$$

(b) From (1.6), we have

$$\begin{aligned}\left\|x^{k+1}-\overline{x}\right\| &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+\left\|(g'(x^k)^{-1}-J_k)g(x^k)\right\| \\ &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+\theta_1^k K\left\|g(x^k)\right\| \\ &\leq \left[\frac{LM_1}{2}\theta_0^k\left\|x^0-\overline{x}\right\|+\theta_1^k M_0 K\right]\left\|x^k-\overline{x}\right\|\end{aligned}$$

Hence $x^k \to \overline{x}$ superlinearly.
(c) From (1.6) and the fact that $x^k \to \bar{x}$, we eventually have

$$\begin{aligned}\left\|x^{k+1}-\overline{x}\right\| &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+\left\|(g'(x^k)^{-1}-J_k)g(x^k)\right\| \\ &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+M_2\left\|x^k-x^{k-1}\right\|\left\|g(x^k)\right\| \\ &\leq \left[\frac{LM_1}{2}\left\|x^k-\overline{x}\right\|+M_0M_2\left[\left\|x^{k-1}-\overline{x}\right\|+\left\|x^k-\overline{x}\right\|\right]\right]\left\|x^k-\overline{x}\right\| \\ &\leq \left[\frac{LM_1}{2}\theta_0\left\|x^{k-1}-\overline{x}\right\|+M_0M_2(1+\theta_0)\left\|x^{k-1}-\overline{x}\right\|\right] \\ &\qquad \times\theta_0\left\|x^{k-1}-\overline{x}\right\| \\ &= \left[\frac{LM_1}{2}\theta_0+M_0M_2(1+\theta_0)\right]\theta_0\left\|x^{k-1}-\overline{x}\right\|^2.\end{aligned}$$

Hence $x^k \to \overline{x}$ two step quadratically.
(d) Again by (1.6) and the fact that $x^k \to \bar{x}$, we eventually have

$$\begin{aligned}\left\|x^{k+1}-\overline{x}\right\| &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+\left\|(g'(x^k)^{-1}-J_k)g(x^k)\right\| \\ &\leq \frac{LM_1}{2}\left\|x^k-\overline{x}\right\|^2+M_2\left\|g(x^k)\right\|^2 \\ &\leq \left[\frac{LM_1}{2}+M_2M_0^2\right]\left\|x^k-\overline{x}\right\|^2 \ .\end{aligned}$$

□

Note that the conditions required for the approximations to the Jacobian matrices $g'(x^k)^{-1}$ given in (i)–(ii) do not imply that $J_k \to g'(\bar{x})^{-1}$. The stronger conditions

$(i)'$ $\left\|g'(x^k)^{-1}-J_k\right\| \leq \|g'(x^0)^{-1}-J_0\|$,
$(ii)'$ $\left\|g'(x^{k+1})^{-1}-J_{k+1}\right\| \leq \theta_1\left\|g'(x^k)^{-1}-J_k\right\|$ for some $\theta_1 \in (0,1)$,
$(iii)'$ $\left\|g'(x^k)^{-1}-J_k\right\| \leq \min\{M_2\left\|x^{k+1}-x^k\right\|, \|g'(x^0)^{-1}-J_0\|\}$ for some $M_2>0$, or
$(iv)'$ $g'(x^k)^{-1}=J_k$,

which imply the conditions (i) through (iv) of Theorem 1.1 respectively, all imply the convergence of the inverse Jacobian approximates to $g'(\bar{x})^{-1}$. The conditions $(i)'$–$(iv)'$ are less desirable since they require greater expense and care in the construction of the inverse Jacobian approximates.

1.3. **Newton's Method for Minimization.** We now translate the results of previous section to the optimization setting. The underlying problem is

$$\mathcal{P} \qquad \min_{x\in\mathbb{R}^n} f(x)\ .$$

The Newton-like iterations take the form

$$x^{k+1} = x^k - H_k \nabla f(x^k),$$

where H_k is an approximation to the inverse of the Hessian matrix $\nabla^2 f(x^k)$.

Theorem 1.2. *Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice continuously differentiable, $x^0 \in \mathbb{R}^n$, and $H_0 \in \mathbb{R}^{n\times n}$. Suppose that*

(1) *there exists $\overline{x} \in \mathbb{R}^n$ and $\epsilon > \|x^0 - \bar{x}\|$ such that $f(\overline{x}) \le f(x)$ whenever $\|x - \bar{x}\| \le \epsilon$,*
(2) *there is a $\delta > 0$ such that $\delta \|z\|_2^2 \le z^T \nabla^2 f(x) z$ for all $x \in B(\overline{x}, \epsilon)$,*
(3) *$\nabla^2 f$ is Lipschitz continuous on $clB(\overline{x}; \epsilon)$ with Lipschitz constant L, and*
(4) *$\theta_0 := \frac{L}{2\delta}\|x^0 - \overline{x}\| + M_0 K < 1$ where $M_0 > 0$ satisfies $z^T \nabla^2 f(x) z \le M_0 \|z\|_2^2$ for all $x \in B(\overline{x}, \epsilon)$ and $K \ge \|(\nabla^2 f(x^0)^{-1} - H_0) y^0\|$ with $y^0 = \nabla f(x^0) / \|\nabla f(x^0)\|$.*

Further, suppose that the iteration

$$x^{k+1} := x^k - H_k \nabla f(x^k) \tag{1.9}$$

is initiated at x^0 where the H_k's are chosen to satisfy one of the following conditions:

(i) $\left\|(\nabla^2 f(x^k)^{-1} - H_k) y^k\right\| \le K$,
(ii) $\left\|(\nabla^2 f(x^k)^{-1} - H_k) y^k\right\| \le \theta_1^k K$ *for some* $\theta_1 \in (0, 1)$,
(iii) $\left\|(\nabla^2 f(x^k)^{-1} - H_k) y^k\right\| \le \min\{M_2 \left\|x^k - x^{k-1}\right\|, K\}$, *for some* $M_2 > 0$, *or*
(iv) $\left\|(\nabla^2 f(x^k)^{-1} - H_k) y^k\right\| \le \min\{M_3 \left\|\nabla f(x^k)\right\|, K\}$, *for some* $M_3 > 0$,

where for each $k = 1, 2, \ldots$ $y^k := \nabla f(x^k) / \left\|\nabla f(x^k)\right\|$.

These hypotheses on the accuracy of the approximations H_k yield the following conclusions about the rate of convergence of the iterates x^k.

(a) *If (i) holds, then $x^k \to \overline{x}$ linearly.*
(b) *If (ii) holds, then $x^k \to \overline{x}$ superlinearly.*
(c) *If (iii) holds, then $x^k \to \overline{x}$ two step quadratically.*
(d) *If (iv) holds, then $x^k \to \overline{x}$ quadradically.*

To more fully understand the convergence behavior described in this theorem, let us examine the nature of the controling parameters L, M_0, and M_1. Since L is a Lipschitz constant for $\nabla^2 f$ it loosely corresponds to a bound on the third–order behavior of f. Thus the assumptions for convergence make implicit demands on the third derivative. The constant δ is a local lower bound on the eigenvalues of $\nabla^2 f$ near $\bar{x}$. That is, f behaves locally as if it were a *strongly convex function* (see exercises) with modulus δ. Finally, M_0 can be interpreted as a local Lipschitz constant for ∇f and only plays a role when $\nabla^2 f$ is approximated inexactly by H_k's.

We now illustrate the performance differences between the method of steepest descent and Newton's method on a simple one dimensional problem. Let $f(x) = x^2 + e^x$. Clearly, f is a strongly convex function with

$$\begin{aligned} f(x) &= x^2 + e^x \\ f'(x) &= 2x + e^x \\ f''(x) &= 2 + e^x \ > \ 2 \\ f'''(x) &= e^x. \end{aligned}$$

If we apply the steepest descent algorithm with backtracking ($\gamma = 1/2,\ c = 0.01$) initiated at $x^0 = 1$, we get the following table

k	x^k	$f(x^k)$	$f'(x^k)$	s
0	1	.37182818	4.7182818	0
1	0	1	1	0
2	$-.5$	.8565307	-0.3934693	1
3	$-.25$	.8413008	0.2788008	2
4	$-.375$	.8279143	$-.0627107$	3
5	$-.34075$	.8273473	.0297367	5
6	$-.356375$	.8272131	$-.01254$	6
7	$-.3485625$	.8271976	.0085768	7
8	$-.3524688$	.8271848	$-.001987$	8
9	$-.3514922$	.8271841	.0006528	10
10	$-.3517364$	.827184	$-.0000072$	12

If we apply Newton's method from the same starting point and take a unit step at each iteration, we obtain a dramatically different table.

x	$f'(x)$
1	4.7182818
0	1
$-1/3$	.0498646
$-.3516893$	.00012
$-.3517337$	.00000000064

In addition, one more iteration gives $|f'(x^5)| \le 10^{-20}$. This is a stunning improvement in performance and shows why one always uses Newton's method (or an approximation to it) whenever possible.

Our next objective is to develop numerically viable methods for approximating Jacobians and Hessians in Newton-like methods.

1.4. **Matrix Secant Methods.** Let us return to the problem of finding $\overline{x} \in \mathbb{R}^n$ such that $g(\overline{x}) = 0$ where $g : \mathbb{R}^n \to \mathbb{R}^n$ is continuously differentiable. In this section we consider Newton-Like methods of a special type. Recall that in a Newton-Like method the iteration scheme takes the form

$$x^{k+1} := x^k - J_k g(x^k), \tag{1.10}$$

where J_k is meant to approximate the inverse of $g'(x^k)$. In the one dimensional case, a method proposed by the Babylonians 3700 years ago is of particular significance. Today we call it the *secant method*:

$$J_k = \frac{x^k - x^{k-1}}{g(x^k) - g(x^{k-1})}. \tag{1.11}$$

With this approximation one has

$$g'(x^k)^{-1} - J_k = \frac{g(x^{k-1}) - [g(x^k) + g'(x^k)(x^{k-1} - x^k)]}{g'(x^k)[g(x^{k-1}) - g(x^k)]}.$$

Near a point x^* at which $g'(x^*) \neq 0$ one can use the MVT to show there exists an $\alpha > 0$ such that

$$\alpha \|x - y\| \leq \|g(x) - g(y)\|.$$

Consequently, by the Quadratic Bound Lemma,

$$\left\|g'(x^k)^{-1} - J_k\right\| \leq \frac{\frac{L}{2}\left\|x^{k-1} - x^k\right\|^2}{\alpha \left\|g'(x^k)\right\| \left\|x^{k-1} - x^k\right\|} \leq K \left\|x^{k-1} - x^k\right\|$$

for some constant $K > 0$ whenever x^k and x^{k-1} are sufficiently close to x^*. Therefore, by our convergence Theorem for Newton Like methods, the secant method is locally two step quadratically convergent to a non–singular solution of the equation $g(x) = 0$. An additional advantage of this approach is that no extra function evaluations are required to obtain the approximation J_k.

1.4.1. *Matrix Secant Methods for Equations.* Unfortunately, the secant approximation (1.11) is meaningless if the dimension n is greater than 1 since division by vectors is undefined. But this can be rectified by multiplying (1.11) on the right by $(g(x^{k-1}) - g(x^k))$ and writing

$$J_k(g(x^k) - g(x^{k-1})) = x^k - x^{k-1}. \tag{1.12}$$

Equation (1.12) is called the *Quasi-Newton equation* (QNE), or *matrix secant equation* (MSE), at x^k. Here the matrix J_k is unknown, but is required to satisfy the n linear equations of the MSE. These equations determine an n dimensional affine manifold in $\mathbb{R}^{n \times n}$. Since J_k contains n^2 unknowns, the n linear equations in (1.12) are not sufficient to uniquely determine J_k. To nail down a specific J_k further conditions on the update J_k must be given. What conditions should these be?

To develop sensible conditions on J_k, let us consider an overall iteration scheme based on (1.10). For convenience, let us denote J_k^{-1} by B_k (i.e. $B_k = J_k^{-1}$). Using the B_k's, the MSE (1.12) becomes

$$B_k(x^k - x^{k-1}) = g(x^k) - g(x^{k-1}). \tag{1.13}$$

At every iteration we have (x^k, B_k) and compute x^{k+1} by (1.10). Then B_{k+1} is constructed to satisfy (1.13). If B_k is close to $g'(x^k)$ and x^{k+1} is close to x^k, then B_{k+1} should be chosen not only to satisfy (1.13) but also to be as "close" to B_k as possible. With this in mind, we must now decide what we mean by "close". From a computational perspective, we prefer "close" to mean *easy to compute*. That is, B_{k+1} should be *algebraically close* to B_k in the sense that B_{k+1} is only a rank 1 modification of B_k. Since we are assuming that B_{k+1} is a rank 1 modification to B_k, there are vectors $u, v \in \mathbb{R}^n$ such that

$$B_{k+1} = B_k + uv^T. \tag{1.14}$$

We now use the matrix secant equation (1.13) to derive conditions on the choice of u and v. In this setting, the MSE becomes

$$B_{k+1}s^k = y^k,$$

where

$$s^k := x^{k+1} - x^k \qquad \text{and} \qquad y^k := g(x^{k+1}) - g(x^k)\ .$$

Multiplying (1.14) by s^k gives

$$y^k = B_{k+1}s^k = B_k s^k + uv^T s^k\ .$$

Hence, if $v^T s^k \neq 0$, we obtain

$$u = \frac{y^k - B_k s^k}{v^T s^k}$$

and

$$B_{k+1} = B_k + \frac{\left(y^k - B_k s^k\right) v^T}{v^T s^k}. \tag{1.15}$$

Equation (1.15) determines a whole class of rank one updates that satisfy the MSE where one is allowed to choose $v \in \mathbb{R}^n$ as long as $v^T s^k \neq 0$. If $s^k \neq 0$, then an obvious choice for v is s^k yielding the update

$$B_{k+1} = B_k = \frac{\left(y^k - B_k s^k\right) s^{k^T}}{s^{k^T} s^k}. \tag{1.16}$$

This is known as Broyden's update. It turns out that the Broyden update is also analytically close.

Theorem 1.3. *Let $A \in \mathbb{R}^{n\times n}$, s, $y \in \mathbb{R}^n$, $s \neq 0$. Then for any matrix norms $\|\cdot\|$ and $|||\cdot|||$ such that*

$$\|AB\| \leq \|A\|\ |||B|||$$

and

$$\left|\left|\left| \frac{vv^T}{v^T v} \right|\right|\right| \leq 1,$$

the solution to

$$\min\{\|B - A\| : Bs = y\} \tag{1.17}$$

is

$$A_+ = A + \frac{(y - As)s^T}{s^T s}. \tag{1.18}$$

In particular, (1.18) solves (1.17) when $\|\cdot\|$ is the ℓ_2 matrix norm, and (1.18) solves (1.17) uniquely when $\|\cdot\|$ is the Frobenius norm.

Proof. Let $B \in \{B \in \mathbb{R}^{n\times n} : Bs = y\}$, then

$$\begin{aligned}
\|A_+ - A\| &= \left\| \frac{(y-As)s^T}{s^Ts} \right\| = \left\| (B-A)\frac{ss^T}{s^Ts} \right\| \\
&\le \|B-A\| \left\|\left| \frac{ss^T}{s^Ts} \right|\right\| \le \|B-A\| .
\end{aligned}$$

Note that if $|||\cdot||| = \|\cdot\|_2$, then

$$\begin{aligned}
\left\| \frac{vv^T}{v^Tv} \right\|_2 &= \sup\left\{ \left\| \frac{vv^T}{v^Tv} x \right\|_2 \,\middle|\, \|x\|_2 = 1 \right\} \\
&= \sup\left\{ \sqrt{\frac{(v^Tx)^2}{\|v\|^2}} \,\middle|\, \|x\|_2 = 1 \right\} \\
&= 1,
\end{aligned}$$

so that the conclusion of the result is not vacuous. For uniqueness observe that the Frobenius norm is strictly convex and $\|A\cdot B\|_F \le \|A\|_F \|B\|_2$. □

Therefore, the Broyden update (1.16) is both algebraically and analytically close to B_k. These properties indicate that it should perform well in practice and indeed it does.

Algorithm: Broyden's Method
Initialization: $x^0 \in \mathbb{R}^n$, $B_0 \in \mathbb{R}^{n\times n}$
Having (x^k, B_k) compute (x^{k+1}, B_{x+1}) as follows:
Solve $B_k s^k = -g(x^k)$ for s^k and set

$$\begin{aligned}
x^{k+1} : &= x^k + s^k \\
y^k : &= g(x^{k+1}) - g(x^k) \\
B_{k+1} : &= B_k + \frac{(y^k - B_k s^k){s^k}^T}{{s^k}^T s^k}.
\end{aligned}$$

We would prefer to write the Broyden update in terms of the matrices $J_k = B_k^{-1}$ so that we can write the step computation as $s^k = -J_k g(x^k)$ avoiding the need to solve an equation. To obtain the formula for J_k we use the the following important lemma for matrix inversion.

Lemma 1.1. (Sherman-Morrison-Woodbury*) Suppose $A \in \mathbb{R}^{n\times n}$, $U \in \mathbb{R}^{n\times k}$, $V \in \mathbb{R}^{n\times k}$ are such that both A^{-1} and $(I + V^TA^{-1}U)^{-1}$ exist, then*

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^TA^{-1}U)^{-1}V^TA^{-1}$$

The above lemma verifies that if $B_k^{-1} = J_k$ exists and ${s^k}^T J_k y^k = {s^k}^T B_k^{-1} y^k \neq 0$, then (1.19)

$$J_{k+1} = \left[B_k + \frac{(y^k - B_k s^k){s^k}^T}{{s^k}^T s^k} \right]^{-1} = B_k^{-1} + \frac{(s^k - B_k^{-1}y^k){s^k}^T B_k^{-1}}{{s^k}^T B_k^{-1} y} = J_k + \frac{(s^k - J_k y^k){s^k}^T J_k}{{s^k}^T J_k y}.$$

In this case, it is possible to directly update the inverses J_k. It should be cautioned though that this process can become numerically unstable if $|{s^k}^T J_k y^k|$ is small. Therefore, in practise, the value $|{s^k}^T J_k y^k|$ must be monitored to avoid numerical instability.

Although we do not pause to establish the convergence rates here, we do give the following result due to Dennis and Moré (1974).

Theorem 1.4. *Let $g : \mathbb{R}^n \to \mathbb{R}^n$ be continuously differentiable in an open convex set $D \subset \mathbb{R}^n$. Assume that there exists $x^* \in \mathbb{R}^n$ and $r, \beta > 0$ such that $x^* + r\mathbb{B} \subset D$, $g(x^*) = 0$, $g'(x^*)^{-1}$ exists with $\|g'(x^*)^{-1}\| \le \beta$, and g' is Lipschitz continuous on $x^* + r\mathbb{B}$ with Lipschitz constant $\gamma > 0$. Then there exist positive constants ϵ and δ such that if $\|x^0 - x^*\|_2 \le \epsilon$ and $\|B_0 - g'(x^0)\| \le \delta$, then the sequence $\{x^k\}$ generated by the iteration*

$$\left[\begin{array}{lcl} x^{k+1} & := & x^k + s^k \text{ where } s^k \text{ solves } 0 = g(x^k) + B_k s \\ B_{k+1} & := & B_k + \frac{(y^k - B_k s^k) s_k^T}{s_k^T s^k} \text{ where } y^k = g(x^{k+1}) - g(x^k) \end{array}\right.$$

is well-defined with $x^k \to x^*$ superlinearly.

1.4.2. *Matrix Secant Methods for Minimization.* We now extend these matrix secant ideas to optimization, specifically minimization. The underlying problem we consider is

$$\mathcal{P} : \underset{x \in \mathbb{R}^n}{\text{minimize}} \ f(x) \ ,$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is assumed to be twice continuously differentiable. In this setting, we wish to solve the equation $\nabla f(x) = 0$ and the MSE (1.13) becomes

$$H_{k+1} y^k = s^k \ , \tag{1.20}$$

where $s^k := x^{k+1} - x^k$ and

$$y^k := \nabla f(x^{k+1}) - \nabla f(x^k).$$

Here the matrix H_k is intended to be an approximation to the inverse of the hessian matrix $\nabla^2 f(x^k)$. Writing $M_k = H_k^{-1}$, a straightforward application of Broyden's method gives the update

$$M_{k+1} = M_k + \frac{(y^k - M_k s^k){s^k}^T}{{s^k}^T s^k}.$$

However, this is unsatisfactory for two reasons:

1. (1) Since M_k approximates $\nabla^2 f(x^k)$ it must be symmetric.
2. (2) Since we are minimizing, then M_k must be positive definite to insure that $s^k = -M_k^{-1} \nabla f(x^k)$ is a direction of descent for f at x^k.

To address problem 1 above, one could return to equation (1.15) an find an update that preserves symmetry. Such an update is uniquely obtained by setting

$$v = (y^k - M_k s^k).$$

This is called the symmetric rank 1 update or SR1. Although this update can on occasion exhibit problems with numerical stability, it has recently received a great deal of renewed interest. The stability problems occur whenever

$$v^T s^k = (y^k - M_k s^k)^T s^s \tag{1.21}$$

has small magnitude. The inverse SR1 update is given by

$$H_{k+1} = H_k + \frac{(s^k - H_k y^k)(s^k - H_k y^k)^T}{(s^k - H_k y^k)^T y^k}$$

which exists whenever $(s^k - H_k y^k)^T y^k \neq 0$.

We now approach the question of how to update M_k in a way that addresses both the issue of symmetry and positive definiteness while still using the Broyden updating ideas. Given a symmetric positive definite matrix M and two vectors s and y, our goal is to find a symmetric positive definite matrix $\bar{M}$ such that $\bar{M}s = y$. Since M is symmertic and positive definite, there is a non-singular $n \times n$ matrix L such that $M = LL^T$. Indeed, L can be chosen to be the lower triangular Cholesky factor of M. If $\overline{M}$ is also symmetric and positive definite then there is a matrix $J \in \mathbb{R}^{n\times n}$ such that $\overline{M} = JJ^T$. The MSE (**??**) implies that if

$$J^T s = v \tag{1.22}$$

then

$$Jv = y. \tag{1.23}$$

Let us apply the Broyden update technique to (1.23), J, and L. That is, suppose that

$$J = L + \frac{(y - Lv)v^T}{v^T v}. \tag{1.24}$$

Then by (1.22)

$$v = J^T s = L^T s + \frac{v(y - Lv)^T s}{v^T v}. \tag{1.25}$$

This expression implies that v must have the form

$$v = \alpha L^T s$$

for some $\alpha \in \mathbb{R}$. Substituting this back into (1.25) we get

$$\alpha L^T s = L^T s + \frac{\alpha L^T s (y - \alpha L L^T s)^T s}{\alpha^2 s^T L L^T s}.$$

Hence

$$\alpha^2 = \left[\frac{s^T y}{s^T M s}\right]. \tag{1.26}$$

Consequently, such a matrix J satisfying (1.25) exists only if $s^T y > 0$ in which case

$$J = L + \frac{(y - \alpha M s)s^T L}{\alpha s^T M s},$$

with

$$\alpha = \left[\frac{s^T y}{s^T M s}\right]^{1/2},$$

yielding

$$\overline{M} = M + \frac{yy^T}{y^T s} - \frac{Mss^T M}{s^T M s}. \tag{1.27}$$

Moreover, the Cholesky factorization for $\overline{M}$ can be obtained directly from the matrices J. Specifically, if the QR factorization of J^T is $J^T = QR$, we can set $\overline{L} = R$ yielding

$$\overline{M} = JJ^T = R^T Q^T QR = \overline{L}\overline{L}^T.$$

The formula for updating the inverses is again given by applying the Sherman-Morrison-Woodbury formula to obtain

$$\overline{H} = H + \frac{(s+Hy)^T y s s^T}{(s^T y)^2} - \frac{Hys^T + sy^T H}{s^T y}, \tag{1.28}$$

where $H = M^{-1}$. The update (1.27) is called the BFGS update and (1.28) the inverse BFGS update. The letter BFGS stand for Broyden, Flethcher, Goldfarb, and Shanno.

We have shown that beginning with a symmetric positive definite matrix M_k we can obtain a symmetric and positive definite update M_{k+1} that satisfies the MSE $M_{k+1}s_k = y_k$ by applying the formula (1.27) whenever ${s^k}^T y^k > 0$. We must now address the question of how to choose x^{k+1} so that ${s^k}^T y^k > 0$. Recall that

$$y = y^k = \nabla f(x^{k+1}) - \nabla f(x^k)$$

and

$$s^k = x^{k+1} - x^k = t_k d^k ,$$

where

$$d^k = -t_k H_k \nabla f(x^k)$$

is the matrix secant search direction and t_k is the stepsize. Hence

$$\begin{aligned} {y^k}^T s^k &= \nabla f(x^{k+1})^T s^k - \nabla f(x^k)^T s^k \\ &= t_k(\nabla f(x^k + t_k d_k)^T d^k - \nabla f(x^k)^T d^k) , \end{aligned}$$

where $d^k := -H_k \nabla f(x^k)$. Since H_k is positive definite the direction d^k is a descent direction for f at x^k and so $t_k > 0$. Therefore, to insure that ${s^k}^T y^k > 0$ we need only show that $t_k > 0$ can be choosen so that

$$\nabla f(x^k + t_k d^k)^T d^k \geq \beta \nabla f(x^k)^T d^k \tag{1.29}$$

for some $\beta \in (0,1)$ since in this case

$$\nabla f(x^k + t_k d_k)^T d^k - \nabla f(x^k)^T d^k \geq (\beta - 1)\nabla f(x^k)^T d^k > 0.$$

But this precisely the second condition in the weak Wolfe conditions with $\beta = c_2$. Hence a successful BFGS update can always be obtained. The BFGS update and is currently considered the best matrix secant update for minimization.

BFGS Updating

$$\begin{aligned} \sigma &:= \sqrt{{s^k}^T y^k} \\ \hat{s}^k &:= s^k/\sigma \\ \hat{y}^k &:= y^k/\sigma \\ H_{k+1} &:= H_k + (\hat{s}^k - H_k\hat{y}^k)(\hat{s}^k)^T + \hat{s}^k(\hat{s}^k - H_k\hat{y}^k)^T - (\hat{s}^k - H_k\hat{y}^k)^T\hat{y}^k\hat{s}^k(\hat{s}^k)^T \end{aligned}$$