- Dennis, J. E., and R. B. Schnabel, 1996, “Numerical methods for unconstrained optimization and nonlinear equations”, SIAM: pp 168–174.

Excerpt (from Google Books)

171

LEMMA 8.1.1 Let $A \in \mathbb{R}^{n \times n}$, $s, y \in \mathbb{R}^n$, $s \neq 0$. Then for any matrix norms $\|\cdot\|$, $|||\cdot|||$ such that

$$\| A \cdot B \| \leq \| A \| \cdot ||| B ||| \tag{8.1.6}$$

and

$$\left|\left|\left| \frac{vv^T}{v^T v} \right|\right|\right| = 1, \tag{8.1.7}$$

the solution to

$$\min_{B \in Q(y, s)} \| B - A \| \tag{8.1.8}$$

is

$$A_+ = A + \frac{(y - As)s^T}{s^T s}. \tag{8.1.9}$$

In particular, (8.1.9) solves (8.1.8) when $\|\cdot\|$ is the l_2 matrix norm, and (8.1.9) solves (8.1.8) uniquely when $\|\cdot\|$ is the Frobenius norm.

172

ALGORITHM 8.1.2 BROYDEN'S METHOD

Given $F : \mathbb{R}^n \longrightarrow \mathbb{R}^n$, $x_0 \in \mathbb{R}^n$, $A_0 \in \mathbb{R}^{n \times n}$

Do for $k = 0, 1, \ldots$:

Solve $A_k s_k = -F(x_k)$ for s_k

$$x_{k+1} := x_k + s_k$$

$$y_k := F(x_{k+1}) - F(x_k)$$

$$A_{k+1} := A_k + \frac{(y_k - A_k s_k)s_k^T}{s_k^T s_k}. \tag{8.1.10}$$

We will also refer to this method as the *secant method*. At this point, the reader may have grave doubts whether it will work. In fact, it works quite well locally, as we suggest below by considering its behavior on the same problem that we solved by Newton's method in Section 5.1. Of course, like Newton's method, it may need to be supplemented by the techniques of Chapter 6 to converge from some starting points.

There is one ambiguity in Algorithm 8.1.1: how do we get the initial approximation A_0 to $J(x_0)$? In practice, we use finite differences this one time to get a good start. This also makes the minimum-change characteristic of Broyden's update more appealing. In Example 8.1.3, we assume for simplicity that $A_0 = J(x_0)$.

EXAMPLE 8.1.3 Let

$$F(x) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1^2 + x_2^2 - 9 \end{bmatrix},$$

which has roots $(0, 3)^T$ and $(3, 0)^T$. Let $x_0 = (1, 5)^T$, and apply Algorithm 8.1.2

with

$$A_0 = J(x_0) = \begin{bmatrix} 1 & 1 \\ 2 & 10 \end{bmatrix}.$$

Then

$$F(x_0) = \begin{bmatrix} 3 \\ 17 \end{bmatrix}, \qquad s_0 = -A_0^{-1}F(x_0) = \begin{bmatrix} -1.625 \\ -1.375 \end{bmatrix},$$

$$x_1 = x_0 + s_0 = \begin{bmatrix} -0.625 \\ 3.625 \end{bmatrix}, \qquad F(x_1) = \begin{bmatrix} 0 \\ 4.53125 \end{bmatrix}.$$

Therefore, (8.1.10) gives

$$A_1 = A_0 + \begin{bmatrix} 0 & 0 \\ -1.625 & -1.375 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0.375 & 8.625 \end{bmatrix}.$$

The reader can confirm that $A_1 s_0 = y_0$. Note that

The reader can confirm that $A_1 s_0 = y_0$. Note that

$$J(x_1) = \begin{bmatrix} 1 & 1 \\ -1.25 & 7.25 \end{bmatrix},$$

so that A_1 is not very close to $J(x_1)$. At the next iteration,

$$s_1 = -A_1^{-1} F(x_1) \cong \begin{bmatrix} 0.549 \\ -0.549 \end{bmatrix}, \qquad x_2 = x_1 + s_1 \cong \begin{bmatrix} -0.076 \\ 3.076 \end{bmatrix},$$

$$F(x_2) \cong \begin{bmatrix} 0 \\ 0.466 \end{bmatrix}, \qquad A_2 \cong \begin{bmatrix} 1 & 1 \\ -0.799 & 8.201 \end{bmatrix}.$$

Again A_2 is not very close to

$$J(x_2) = \begin{bmatrix} 1 & 1 \\ -0.152 & 6.152 \end{bmatrix}.$$

The complete sequences of iterates produced by Broyden's method, and for comparison, Newton's method, are given below. For $k \geq 1$, $(x_k)_1 + (x_k)_2 = 3$ for both methods; so only $(x_k)_2$ is listed below.

for both methods; so only $(x_k)_2$ is listed below.

Broyden's Method		Newton's Method
$(1, 5)^T$	x_0	$(1, 5)^T$
3.625	x_1	3.625
3.0757575757575	x_2	3.0919117647059
3.0127942681679	x_3	3.0026533419372
3.0003138243387	x_4	3.0000023425973
3.0000013325618	x_5	3.0000000000018
3.0000000001394	x_6	3.0
3.0	x_7	

Example 8.1.3 is characteristic of the local behavior of Broyden's method. If any components of $F(x)$ are linear, such as $f_1(x)$ above, then the corresponding rows of the Jacobian approximation will be correct for $k \geq 0$, and the corresponding components of $F(x_k)$ will be zero for $k \geq 1$ (Exercise 4). The rows of A_k corresponding to nonlinear components of $F(x)$ may not be very accurate, but the secant equation still gives enough good information that there is rapid convergence to the root. We show in Section 8.2 that the rate of convergence is q-superlinear, not q-quadratic.

8.2 LOCAL CONVERGENCE ANALYSIS OF BROYDEN'S METHOD

...