ON NEWTON-LIKE METHODS

by

James W. Daniel

Computer Sciences Technical Report #21

May 1968

ON NEWTON-LIKE METHODS

by

James W. Daniel

1. Introduction.

Let X and Y be Banach spaces, P a twice continuously Frechet differentiable mapping of some open convex subset Ω of X into Y, and $\Gamma(x)$ a linear mapping of Y into X for each $x \in \Omega$, depending continuously on x. Many authors [1,3; see 2 for extensive bibliography] have considered Newton-like iterations of the form

$$(1) \qquad x_{n+1} = x_n - \Gamma(x_n) P(x_n) \;, \quad x_0 \text{ given} \;,$$

to compute a solution to the equation $P(x) = 0$. Generally speaking, one needs to have $\Gamma(x)$ sufficiently near to $[P'(x)]^{-1}$ to assure convergence; $\Gamma(x) \equiv [P'(x)]^{-1}$ leads to the usual Newton method. To imitate the pure Newton method well, it would appear that $\Gamma(x)$ should be near to a left inverse of $P'(x)$; known theorems however show that in fact the crucial requirement is for $\Gamma(x)$ to be nearly a right inverse of $P'(x)$, a fact clearly noted in [3]. In this short note we wish to point out that a type of convergence theorem can be stated even for $\Gamma(x)$ nearly a left inverse, although the theorems do not contain the usual nearly computable error bounds and existence proofs one is accumstomed to seeing in this field.

Prepared under Contract Number N00014-67-A-0128-0004 at the University of Wisconsin. Reporduction in whole or in part is permitted for any purpose of the United States Government.

2. Results.

Theorem 1. Let $x^* \in \Omega$ satisfy $P(x^*) = 0$. Suppose $\| I - \Gamma(x^*)P'(x^*) \| \leq \delta < 1$. Then given any $r \in (0, 1-\delta)$, there exists a neighborhood U of x^* in Ω such that if x'_n, $n = 0, 1, \ldots$ and x_0 are arbitrary points in U, the extended Newton-like iteration

(2) $$x_{n+1} = x_n - \Gamma(x'_n)P(x_n), \quad n = 0, \ldots$$ converges to the solution

x^* with rate of convergence given by $\| x_n - x^* \| \leq (\delta + r)^n \| x_0 - x^* \|$.

Proof: Given $r \in (0, 1-\delta)$, there exists a spherical neighborhood U of x^* in Ω such that $\| I - \Gamma(u) P'(v) \| \leq \delta + r$ for $u, v \in U$, since $\Gamma(x)$ and $P'(x)$ are continuous. If $x_0, x_1, \ldots x_R \in U$, then x_{R+1} is defined and $x_{R+1} - x^* = x_R - \Gamma(x'_R)P(x_R) - (x^* - \Gamma(x'_R)P(x^*))$, so $\| x_{R+1} - x^* \| \leq$ $\| x_R - x^* \| \ \| I - \Gamma(x'_R)P'(x^* + \vartheta(x_R - x^*)) \|$, $0 < \vartheta < 1$, and hence $\| x_{R+1} - x^* \| \leq (\delta + r) \| x_R - x^* \| < \| x_R - x^* \|$, implying $x_{R+1} \in U$ and also giving the error bound. Q.E.D.

The two special cases $x'_n = x_n$, or $x'_n = x_0$, yield the standard Newton-like and so called modified Newton-like methods. Newton's method itself is known to converge quadratically; following [3], it is possible to state sufficient conditions for the iteration (2) to converge more rapidly than given by the preceding theorem.

Theorem 2. Let $\|P''(x)\| \leq K_1$, $\|\Gamma(x)\| \leq K_2$ for $x \in \Omega$, and let $x^* \in \Omega$ satisfy $P(x^*) = 0$. Suppose $\|I - \Gamma(x^*)P'(x^*)\| \leq \delta < 1$; let $\delta_n \geq \|I - \Gamma(x'_n)P'(x_n)\|$ and suppose that δ_n satisfies $\delta_{n+1} \leq \delta_n^p$, $1 < p \leq 2$. Then for a sufficiently small spherical neighborhood U of x^*, the Newton-like sequence (2) and δ_n are well defined, and x_n converges to x^* with order of convergence at least equal to p, i.e., there exists an $A_0 < 1$ such that $\|x_n - x^*\| \leq \|x_0 - x^*\| \; A_0^{\frac{p^{n+1}-1}{p-1}}$.

Proof: For a small U, the sequences exist by Theorem 1.

$$\|x_{n+1} - x^*\| = \|x_n - x^* - \Gamma(x'_n)[P(x_n) - P(x^*)]\|$$

$$= \|x_n - x^* - \Gamma(x'_n)P'(x_n)(x_n - x^*) + \frac{1}{2}\Gamma(x'_n)P''(x_n + \vartheta(x^* - x_n))(x_n - x^*)^2\|$$

for some $\vartheta \in (0, 1)$. Hence $\|x_{n+1} - x^*\| \leq \delta_n \|x_n - x^*\| + M\|x_n - x^*\|^2$ where $M = \frac{1}{2} K_1 K_2$; for convenience write $e_n = \|x_n - x^*\|$, so $e_{n+1} \leq \delta_n e_n + Me_n^2$. By choosing U smaller if necessary, we can assume $e_0 < 1$, $\delta_0 + Me_0 < 1$; hence $e_n < 1$, $\delta_n + Me_n < 1$ for all n. Therefore $e_{n+1} \leq \delta_n e_n + Me_n^p \equiv A_n e_n$, $A_n \equiv \delta_n + Me_n^{p-1}$. $A_{n+1} = \delta_{n+1} + Me_{n+1}^{p-1} \leq \delta_n^p + M[A_n e_n]^{p-1} = \delta_n^p + A_n^{p-1} Me_n^{p-1} = \delta_n^p + A_n^{p-1}(A_n - \delta_n) = \delta_n^p - A_n^{p-1}\delta_n + A_n^p$. Now $\delta_n \leq A_n \Longrightarrow -A_n^{p-1} \leq -\delta_n^{p-1} \Longrightarrow A_{n+1} \leq \delta_n^p - \delta_n^{p-1}\delta_n + A_n^p = A_n^p$, and hence $A_n \leq A_0^{p^n}$. It follows that $e_n \leq e_0 \, A_0^{\frac{p^{n+1}-1}{p-1}}$. Q.E.D.

If $\Gamma(x)$ is a left inverse of $P'(x)$ and $x'_n = x_n$, then $\delta_n = 0$ and the above theorem states that, asymptotically, quadratic convergence occurs.

3. References

[1] Bartle, R. G., "Newton's method in Banach spaces," Proc. Amer. Math. Soc. 6 (1955), 827-831.

[2] Collatz, L., Functional Analysis and Numerical Mathematics, Academic Press, New York, 1966.

[3] Dennis, J. E., Jr., "On Newton-like methods," Numer. Math. (1968), to appear.