Some Notes from the Book: Afternotes on Numerical Analysis by G. W. Stewart

John L. Weatherwax*

October 8, 2004

Nonlinear Equations

By the Dawn's Early Light

In the example considered in this chapter we are trying to find a value for θ that satisfies the equation

$$\frac{2V_0^2\sin(\theta)\cos(\theta)}{g} - d = 0\,.$$

Consider writing this slightly differently using the identity

$$2\sin(\theta)\cos(\theta) = \sin(2\theta)\,,$$

as

$$\sin(2\theta)\left(\frac{V_0^2}{g}\right) = d\,.$$

Since we know that $\sin(2\theta)$ is less than 1. Thus the above product must take the expression $\frac{V_0^2}{g}$ and make it smaller (to equal d) by multiplying by $\sin(2\theta)$. If this fraction is already to small i.e. if

$$\frac{V_0^2}{g} < d\,,$$

then there will be no solution for θ.

*wax@alum.mit.edu

Newton's Method

The book argues using geometry and Taylor's theorem that Newton's method can be expressed as the difference equation

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}\,, \quad \text{for} \quad k = 1, 2, \cdots\,. \tag{1}$$

If we want to use this to calculate the reciprocal of a number a we can look for the root of a function $f(x)$ given by

$$f(x) = \frac{1}{x} - a\,.$$

Then we have the first derivative given by

$$f'(x) = -\frac{1}{x^2}\,,$$

and Newton's iteration given by Equation 1 gives

$$x_{k+1} = x_k - \frac{\left(\frac{1}{x_k} - a\right)}{-\frac{1}{x_k^2}} = x_k + x_k - a x_k^2 = 2x_k - a x_k^2\,. \tag{2}$$

As as an aside we wonder if given the Newton's iteration expression can we determine what function $f(x)$ we are looking for a zero of. Thus given the iterations $x_{k+1} = \phi(x_k)$ if we take

$$\phi(x) \equiv x - \frac{f(x)}{f'(x)}\,, \tag{3}$$

by solving for $f(x)$ we have

$$\frac{f'(x)}{f(x)} = \frac{1}{x - \phi(x)} \quad \text{so} \quad \ln(f(x)) = \int^x \frac{dx'}{x' - \phi(x')}$$

or

$$f(x) = \exp\left(\int^x \frac{dx'}{x' - \phi(x')}\right)\,. \tag{4}$$

We can test this idea on Equation 2 where $\phi(x) = 2x - ax^2$, then the denominator in the above integral is given by $x - (2x - ax^2) = -x + ax^2$ and we need to evaluate the integral of

$$\frac{1}{-x + ax^2} = -\frac{1}{x(1 - ax)} = \frac{A}{x} + \frac{B}{1 - ax}\,.$$

If we multiply by x and let $x = 0$ we see that $A = -1$. If we multiply by $1 - ax$ and let $x = \frac{1}{a}$ we get that $B = -a$ and so have the partial fraction expansion of

$$\frac{1}{-x + ax^2} = -\frac{1}{x} - \frac{a}{1 - ax}\,.$$

Integrating these gives

$$\int^x \frac{1}{-x + ax^2} = -\ln(x) + \ln(1 - ax) = \ln\left(\frac{1 - ax}{x}\right)\,.$$

Then using Equation 4 we see that $f(x)$ is given by

$$f(x) = \frac{1 - ax}{x} = \frac{1}{x} - a\,.$$

Notes on Local convergence analysis

The book derives that the error convergence for a fixed point method $x_{k+1} = \phi(x_k)$ is given by

$$e_{k+1} = \phi'(\xi_k) e_k \,, \tag{5}$$

where $e_k \equiv x_k - x_*$ and ξ_k is a point between x_k and x_*. If $|\phi'(x)| \le C < 1$ in a region around x_* then this method will converge. As we are considering Newton's iterations as the formula used to determine the functional form for $\phi(x)$ via Equation 3 we have that

$$\phi'(x) = 1 - \frac{f'(x)}{f'(x)} + \frac{f(x)f''(x)}{f'(x)^2} = \frac{f(x)f''(x)}{f'(x)^2} \,. \tag{6}$$

Since x_* is a root of $f(x)$ we have $f(x_*) = 0$ and from the above $\phi'(x_*) = 0$. Thus by continuity we expect there to be a region around x_* such that the needed convergence inequality $|\phi'(x)| \le C < 1$ holds.

We can consider how *fast* this iterative algorithm converges to x_* by Taylor expanding $\phi(x)$ about x_*. We have

$$\phi(x_k) - \phi(x_*) = \phi'(x_*)(x_k - x_*) + \frac{1}{2}\phi''(\eta_k)(x_k - x_*)^2 \,,$$

where η_k is a point between x_k and x_*. Since $x_{k+1} = \phi(x_k)$, $x_* = \phi(x_*)$, $\phi'(x_*) = 0$, and using the definition of e_k the above becomes

$$e_{k+1} = \frac{1}{2}\phi''(\eta_k) e_k^2 \,.$$

Thus we need to evaluate ϕ''. Using Equation 6 we have

$$\phi''(x) = \frac{f'(x)f''(x)}{f'(x)^2} + \frac{f(x)f'''(x)}{f'(x)^2} - 2\frac{f(x)f''(x)}{f'(x)^3} \,. \tag{7}$$

When we evaluate this at $x = x_*$ using that x_* is a root we find that

$$\phi''(x_*) = \frac{f''(x_*)}{f'(x_*)} \,.$$

Since as we iterate, assuming that we are converging we will have $x_k \to x_*$ and thus

$$\lim_{k\to\infty} \left(\frac{e_{k+1}}{e_k^2} \right) = \lim_{k\to\infty} \left(\frac{1}{2}\phi''(\eta_k) \right) = \frac{1}{2}\phi''(x_*) = \frac{f''(x_*)}{2f'(x_*)} \,.$$

It is this constant number that determine how our error e_k changes from time step to time step

$$e_{k+1} \approx \frac{f''(x_*)}{2f'(x_*)} e_k^2 \,.$$

As another way to derive this expression, consider Newton's iterations where we assume that $x_k = x_* + \epsilon_k$ and we Taylor expand everything about x_*. Then in that case we have that the

Newton iterations have

$$\begin{aligned}
\epsilon_{k+1} &= x_{k+1} - x_* = x_k - x_* - \frac{f(x_k)}{f'(x_k)} \\
&= \epsilon_k - \frac{f(\epsilon_k + x_*)}{f'(\epsilon_k + x_*)} = \epsilon_k - \frac{f'(x_*)\epsilon_k + \frac{1}{2}f''(x_*)\epsilon_k^2 + \frac{1}{6}f'''(\eta_k)\epsilon_k^3}{f'(x_*) + f''(x_*)\epsilon_k + \frac{1}{2}f'''(\xi_k)\epsilon_k^2} \\
&\approx \epsilon_k - \frac{1}{f'(x_*)}\left(f'(x_*)\epsilon_k + \frac{1}{2}f''(x_*)\epsilon_k^2 + \frac{1}{6}f'''(\eta_k)\epsilon_k^3\right)\left(1 - \frac{f''(x_*)}{f'(x_*)}\epsilon_k - \frac{1}{2}\frac{f'''(\xi_k)}{f'(x_*)}\epsilon_k^3\right) \\
&= \frac{f''(x_*)}{2f'(x_*)}\epsilon_k^2 + O(\epsilon_k^3)\,,
\end{aligned}$$

the same results as before.

As a summary, we recall that **the number of significant digits** n in the approximation x_k to x_* can be given by

$$n = -\frac{\log(|x_* - x_k|)}{|x_*|} \tag{8}$$

Notes on a quasi-Newton method

As another comment, note that we could use *any* method to approximate the derivative $f'(x_k)$. Namely many of the methods presented in [**?**] could be used.

Notes on iterating a fixed point

For the fixed point x_* that satisfies $x_* = \phi(x_*)$ we can derive a recursion relationship for the error $e_k \equiv x_k - x_*$, using Taylor's theorem with a remainder. To do this we expand $\phi(x_k)$ about the point x_* where we have

$$\begin{aligned}
\phi(x_k) &= \phi(x_*) + \phi'(x_*)(x_k - x_*) + \frac{\phi''(x_*)}{2}(x_k - x_*)^2 + \cdots + \frac{\phi^{(p-1)}(x_*)}{(p-1)!}(x_k - x_*)^{p-1} \\
&\quad + \frac{\phi^{(p)}(\xi_k)}{p!}(x_k - x_*)^p\,,
\end{aligned}$$

where ξ_k is a point between x_k and x_*. If the first $p-1$st derivative of ϕ vanish at x_* then the above becomes

$$\phi(x_k) = x_* + \frac{\phi^{(p)}(\xi_k)}{p!}(x_k - x_*)^p\,.$$

Since $x_{k+1} = \phi(x_k)$ this gives

$$e_{k+1} = \frac{\phi^{(p)}(\xi_k)}{p!}e_k^p\,, \tag{9}$$

our desired recurrence relationship for e_k.

Notes on Newton's method with multiple zeros

When x_* is a zero of $f(x)$ of multiplicity m then using Taylor's theorem we can show that $f(x) = (x - x_*)^m g(x)$ with $g(x_*) \neq 0$. For such a function Newton's iteration function ϕ becomes

$$\phi(x) = x - \frac{f(x)}{f'(x)} = x - \frac{(x - x_*)g(x)}{mg(x) - (x - x_*)g'(x)}\,.$$

To study convergence to the root x_* we need $\phi'(x_*)$. We find the first derivative given by

$$\begin{aligned}\phi'(x) = 1 &- \frac{g(x)}{mg(x) - (x - x_*)g'(x)} \\ &- \frac{(x - x_*)g(x)}{(mg(x) - (x - x_*)g'(x))^2}(mg'(x) - g'(x) - (x - x_*)g''(x))\,.\end{aligned}$$

We could simplify that expression but since we only want to evaluate it at x_* we don't need to. At the point x_* we find

$$\phi'(x_*) = 1 - \frac{g(x_*)}{mg(x_*)} = 1 - \frac{1}{m}\,.$$

You might recall Equation 6 and argue that $\phi'(x_*) = 0$ since $f(x_*) = 0$. Showing that Newton's method must have at least quadratic convergence. These statements are true only in the case where $f'(x_*) \neq 0$ which unfortunately when we have multiple roots is not true. Thus using Equation 9 we have that

$$e_{k+1} \approx \phi^{(1)}(\xi_k)e_k = \left(1 - \frac{1}{m}\right)e_k\,.$$

Notes on the secant method: convergence

Consider the two dimensional function $\phi(u, v)$ given by the secant method fixed point mapping

$$\phi(u, v) = u - \frac{f(u)(u - v)}{f(u) - f(v)} = \frac{vf(u) - uf(v)}{f(u) - f(v)}\,. \tag{10}$$

If $v = x_*$ then since x_* is a root of f we have

$$\phi(u, x_*) = \frac{x_* f(u)}{f(u)} = x_*\,, \tag{11}$$

and if $u = x_*$ then in the same way

$$\phi(x_*, v) = \frac{-x_* f(v)}{-f(v)} = x_*\,. \tag{12}$$

To prove convergence we will need the expression for the two-dimensional Taylor series of ϕ with error term, which states that $\phi(x_* + p, x_* + q)$ is equal to

$$\begin{aligned}&\phi(x_*, x_*) + \phi_u(x_*, x_*)p + \phi_v(x_*, x_*)q \\ +\ &\frac{1}{2}\left[\phi_{uu}(x_* + \theta p, x_* + \theta q)p^2 + 2\phi_{uv}(x_* + \theta p, x_* + \theta q)pq + \phi_{vv}(x_* + \theta p, x_* + \theta q)q^2\right]\,,\end{aligned}$$

where $\theta \in [0, 1]$. From Equation 11 we can take the u derivative and we see that $\phi_u(u, x_*) = 0$. In the same way using Equation 12 we can take the v derivative and get $\phi_v(x_*, v) = 0$. Thus the u and v derivatives at the point (x_*, x_*) are zero

$$\phi_u(x_*, x_*) = \phi_v(x_*, x_*) = 0\,,$$

and the second and third terms of the Taylor expansion of $\phi(x_* + p, x_* + q)$ vanish. Now consider the uu derivative of ϕ that is the coefficient of the p^2 term. By a linear Taylor expansion in its *second* argument we have

$$\begin{aligned}\phi_{uu}(x_* + \theta p, x_* + \theta q) &= \phi_{uu}(x_* + \theta p, x_*) + \phi_{uuv}(x_* + \theta p, x_* + \tau_q \theta q)\theta q \\ &= \phi_{uuv}(x_* + \theta p, x_* + \tau_q \theta q)\theta q\,.\end{aligned}$$

Where we have used Equation 11 to argue that $\phi_{uu}(x_* + \theta p, x_*) = 0$. Now consider the vv derivative of ϕ that is the coefficient of the q^2 term. By a linear Taylor expansion in its *first* argument and using Equation 12 we have

$$\begin{aligned}\phi_{vv}(x_* + \theta p, x_* + \theta q) &= \phi_{vv}(x_*, x_* + \theta q) + \phi_{vvu}(x_* + \tau_p \theta p, x_* + \theta q)\theta p \\ &= \phi_{vvu}(x_* + \tau_p \theta p, x_* + \tau_q \theta q)\theta p\,.\end{aligned}$$

In both case τ_p and τ_q are in $[0, 1]$. With these expressions we have that $\phi(x_* + p, x_* + q)$ is equal to x_* plus the expression

$$\frac{pq}{2}\left[\phi_{uuv}(x_* + \theta p, x_* + \tau_q \theta q)\theta p + 2\phi_{uv}(x_* + \theta p, x_* + \theta q) + \phi_{vvu}(x_* + \tau_p \theta p, x_* + \tau_q \theta q)\theta q\right]\,, \tag{13}$$

We now derive the recursion relationship between the errors at various timesteps. Let $e_0 = x_0 - x_*$ and $e_1 = x_1 - x_*$ and take $p = e_1$ and $q = e_0$, then $\phi(x_* + p, x_* + q) = \phi(x_1, x_0) = x_2$, since given x_1 and x_0 the secant fixed point function ϕ is how we get the next iterate. The error in the point x_2 using the Taylor series computed above to evaluate the increment gives

$$\begin{aligned}e_2 &= \phi(x_* + e_1, x_* + e_0) - x_* \\ &= \frac{e_1 e_0}{2}\,[\text{bracketed term in 13 with } q \text{ replaced with } e_0 \text{ and } p \text{ with } e_1] \\ &\equiv \frac{e_1 e_0}{2} r(e_1, e_0)\,.\end{aligned} \tag{14}$$

Where we have defined $r(e_1, e_0)$ in the above expression. We can evaluate r at one point namely $(0, 0)$. Where from Equation 13 we find

$$r(0, 0) = 2\phi_{uv}(x_*, x_*)\,. \tag{15}$$

This expression may or may not be zero but it is a constant number. Thus we can make the *product* $vr(u, v)$ as small as we need if v is taken small by keeping v close to the origin. Thus we can find a region in (u, v) around $(0, 0)$ where $vr(u, v)$ is still "small". What we mean is that we can find a δ such that when $|u|, |v| \le \delta$ we have

$$|vr(u, v)| \le C < 1\,.$$

We start our iterations with e_0 and e_1 such that $|e_0|, |e_1| \le \delta$ and then find

$$|e_2| = \frac{1}{2}|e_1||e_0 r(e_1, e_0)| \le \frac{C}{2}|e_1| < |e_1| < \delta\,.$$

Thus e_2 is inside this δ region as well. Because of this we have that

$$|e_1 r(e_2, e_1)| \leq C < 1\,.$$

So the bound on e_3 is given by

$$|e_3| = \frac{1}{2}|e_2||e_1 r(e_2, e_1)| \leq \frac{C}{2}|e_2| < C|e_2| < C^2|e_1|\,.$$

Continuing this for arbitrary k we have

$$|e_k| < C^{k-1}|e_1|\,,$$

and since $0 < C < 1$ we have $|e_k| \to 0$ and the secant method converges.

We now consider the convergence rate of the secant method. Since the secant method has errors that satisfy Equation 14 or for general k

$$e_{k+1} = \frac{e_k e_{k-1}}{2} r(e_k, e_{k-1})\,. \tag{16}$$

We claim that we have super linear convergence of order p where $p = \frac{1+\sqrt{5}}{2}$. This means that we need to show

$$\lim_{k\to\infty} \frac{|e_{k+1}|}{|e_k|^p} = C\,,$$

for some constant C and the numerical value of p specified. To show this define the sequence s_k as $s_k \equiv \frac{|e_{k+1}|}{|e_k|^p}$. Then solving for $|e_{k+1}|$ in terms of s_k we get

$$|e_{k+1}| = s_k |e_k|^p\,.$$

Decrementing k by one and putting the result into the right-hand-side of this last expression we get

$$|e_{k+1}| = s_k |s_{k-1}|e_{k-1}|^p|^p = s_k s_{k-1}^p |e_{k-1}|^{p^2}\,.$$

We now have expressed $|e_{k+1}|$ and $|e_k|$ in terms of $|e_{k-1}|$ thus

$$|r(e_k, e_{k-1})| = \frac{2|e_{k+1}|}{|e_k||e_{k-1}|} = \frac{2 s_k s_{k-1}^p |e_{k-1}|^{p^2}}{s_{k-1}|e_{k-1}|^p |e_{k-1}|} = 2 s_k s_{k-1}^{p-1} |e_{k-1}|^{p^2-p-1}\,.$$

For the value of p suggested one can show $p^2 - p - 1 = 0$ and we end with

$$\frac{1}{2}|r(e_k, e_{k-1})| = s_k s_{k-1}^{p-1}\,. \tag{17}$$

We still want to prove that the limit of s_k is finite. It seems like if s_k limits to a constant s then it must satisfy

$$\frac{1}{2}|r(0,0)| = s^p \quad \text{so} \quad s = \left(\frac{1}{2}|r(0,0)|\right)^{1/p}\,.$$

Using Equation 15 we can relate this limit to the function ϕ and the fixed point x_*. This seems to be enough to show that a limit to s_k exists. We can still discuss the method presented in the book however. If we take the logarithm of Equation 17 we get

$$\log\left(\frac{1}{2}|r(e_k, e_{k-1})|\right) = \log(s_k) + (p-1)\log(s_{k-1}) \quad \text{or} \quad \rho_k = \sigma_k + (p-1)\sigma_{k-1}\,,$$

using the definitions of the sequence ρ_k and σ_k. Assuming limits as $k \to \infty$ exist we must have $\rho_* = \sigma_* + (p-1)\rho_*$. Subtracting these two equations gives

$$\rho_k - \rho_* = \sigma_k - \sigma_* + (p-1)(\sigma_{k-1} - \sigma_*)\,.$$

or changing the order of the terms

$$\sigma_k - \sigma_* = (\rho_k - \rho_*) - (p-1)(\sigma_{k-1} - \sigma_*)\,.$$

Then to use the theorem in the book we make the association $n = 1$, $a_1 = -(p-1)$, $\epsilon_k = \sigma_k - \sigma_*$, and $\eta_k = \rho_k - \rho_*$. Then the difference equation above will converge to zero $\epsilon_k \to 0$ if the roots of

$$x + (p-1) = 0\,,$$

are inside the unit circle. Since this root is $x = -p + 1 \approx -0.618$ we see that it is inside the unit circle, showing that a limit to σ_k exists.