MAT360 Section Summary: 2.4

Error Analysis for Iterative Methods

Summary

Is there a good way of getting a handle on the number of terms in Newton's method? That's essentially the subject of this section.

We learned a bit previously in section 2.2: in 2.2 we obtained useful bounds for fixed-point methods, e.g.

$$|p_n - p| \leq \frac{k^n}{1-k}|p_1 - p_0| \tag{1}$$

where $g(x) \in [a,b]$, $\forall x \in [a,b]$, and $|g'(x)| \leq k < 1$ on $[a,b]$, which brackets the fixed point p. You can use this for Newton's method, but perhaps we can do better, since the convergence is better (quadratic, rather than linear).

Theorem 2.5 (from section 2.3): Let $f \in C^2[a,b]$. If $p \in [a,b]$ is such that $f(p) = 0$ and $f'(p) \neq 0$, then $\exists \delta > 0$ such that Newton's method generates a sequence $\{p_n\}_{n=1}^{\infty}$ converging to p for any initial approximation $p_0 \in [p-\delta, p+\delta]$.

This result is "obvious" (I claimed, in 2.2), since

$$|p_{n+1} - p| \approx \frac{1}{2}|g^{''}(p)||p_n - p|^2$$

when p_n gets into close proximity (i.e. a δ-neighborhood) of p. We can be assured of "contracting" as long as the magnitude of $g^{''}(x)$ is bounded (e.g. $|g^{''}(x)| < M$) in that neighborhood, so long as

$$\frac{1}{2}M|p_n - p| < 1$$

It's obviously true when $p_n = p$, and we simply choose $|p_n - p| < \frac{2}{M}$ to be assured that we'll converge by the Fixed-Point Theorem (2.3).

Definition 2.6: Suppose that $\{p_n\}_{n=0}^{\infty}$ is a sequence that converges to p, with $p_n \neq p$ for all n. If positive constants λ and α exist with

$$\lim_{n\to\infty} \frac{|p_{n+1} - p|}{|p_n - p|^{\alpha}} = \lambda$$

then the sequence **converges to p of order α, with asymptotic error constant λ**.

1. If $\alpha = 1$, the sequence is **linearly convergent** *(e.g. standard convergent fixed point function, with $g'(p) \neq 0$)*, whereas

2. if $\alpha = 2$, the sequence is **quadratically convergent** *(e.g. Newton's method, with $g'(p) \neq 0$)*.

Q: What does *asymptotic* mean?
Q: Is bisection linearly convergent?[1] Contrast this with Exercise #11, for your homework.

Theorem 2.7: Let $g \in C[a,b]$ be such that $g(x) \in [a,b]$, $\forall x \in [a,b]$. Suppose, in addition, that g' is continuous on (a,b) and a positive constant $k < 1$ exists with

$$|g'(x)| \leq k$$

$\forall x \in (a,b)$. If $g'(p) \neq 0$, then for any number p_0 in $[a,b]$, the sequence of iterates

$$p_n = g(p_{n-1})$$

for $n \geq 1$ converges only linearly to the unique fixed point $p \in [a,b]$.

Proof (by the MVT)

[1] The Bisection Algorithm is Not Linearly Convergent. Sui-Sun Cheng and Tzon-Tzer Lu, College Math Journal: Volume 16, Number 1, (1985), Pages: 56-57.

Theorem 2.8: Let p be a solution of the equation $x = g(x)$. Suppose that $g'(p) = 0$ and g'' is continuous and strictly bounded by M on an open interval I containing p. Then $\exists \delta > 0$ such that, for $p_0 \in [p - \delta, p + \delta]$, the sequence $\{p_n = g(p_{n-1})\}_{n=1}^{\infty}$ converges at least quadratically to p. Moreover, for sufficiently large values of n,

$$|p_{n+1} - p| < \frac{M}{2}|p_n - p|^2$$

(Hence, Newton's method is quadratic.)

Proof (by Taylor series, and Fixed-Point theorem)

Example: Here's where we can make use of the quadratic convergence to address our opening question about the number of iterates of Newton's method: For problem #5b, for example, with

$$f(x) = x^3 + 3x^2 - 1$$

$p_0 = 3$ and a solution $p_3 = -2.87939$, we use

$$g(x) = x - \frac{x^3 + 3x^2 - 1}{3x^2 + 6x}$$

and then compute the first and second derivatives of g. We note that by theorem 2.2 there is a unique fixed point in the interval $[-3, -2.74]$; also we see that g has a maximum value of $|g''(x)| < 2.5$ on the interval $[-3, -2.74]$. g has a maximum value of $< -.27$ on the interval, so we could use Equation (1) above to make our estimate (it gives 8 iterations).

We can do better, of course!

Theorem: the secant method is of order the golden mean.

Motivation: #14

It's possible to create methods that are of higher order than Newton's, but one does so at the expense of more constraints on f (e.g. $f \in C^3[a,b]$, and greater computational complexity:

Example: #13, p. 83

Definition 2.9: A solution p of $f(x) = 0$ is a **zero of multiplicity** m of f if, for $x \neq p$, we can write $f(x) = (x-p)^m q(x)$, where $\lim_{x \to p} q(x) \neq 0$.

Theorem 2.10: $f \in C^1[a,b]$ has a simple zero at $p \in (a,b) \iff f(p) = 0$, but $f'(p) \neq 0$.

Theorem 2.11: $f \in C^m[a,b]$ has a zero of multiplicity m at $p \in (a,b) \iff 0 = f(p) = f'(p) = \ldots = f^{(m-1)}(p)$, but $f^{(m)}(p) \neq 0$.

To handle roots p of f multiplicity $m > 1$, we use a trick called "deflation". Consider

$$\mu(x) = \frac{f(x)}{f'(x)}$$

Claim: μ has a simple zero at p (and hence we can use straightforward Newton's method on μ to find the root p).