Math 128A, Summer 2019
Lecture 7, Wednesday 7/3/2019

CLASS ANNOUNCEMENTS: As tomorrow, Thursday, is July 4th Holiday, no lecture.
Plans for today:
- Newton's Method
- Convergence Theorem
- Multiple Roots
  - Double Stepping
  - Line Search
  - Schroder

1 Newton's Method

View Newton's Method as the inverse of first order approximation (via Taylor). That is,

$$f(x) = 0 = f(x_n) + f'(x_n)(x - x_n) + \left[\frac{1}{2} f''(\xi_n)(x - x_n)^2\right],$$

where we drop the bracketed term, and $\min\{x, x_n\} \leq \xi_n \leq \max\{x, x_n\}$. Solving for x, we have:

$$\begin{aligned} -f(x_n) &= f'(x_n)(x_n - x) \\ -\frac{f(x_n)}{f'(x_n)} &= x_{n+1} - x_n \end{aligned}$$

We usually prefer to write Newton's method as above, as opposed to how we write fixed point iteration:

$$x_{n+1} = g(x_n) = x_n - \frac{f(x_n)}{f'(x_n)}$$

Remark: For error analysis (in numerical analysis in general), we always try to subtract similar equations:

$$\begin{aligned} & 0 = f(x_n) + f'(x_n)(x_{n+1} - x_n) \\ - \qquad & 0 = f(x_n) + f'(x_n)(x - x_n) + \frac{1}{2} f''(\xi_n)(x - x_n)^2 \end{aligned}$$

Subtracting these give:

$$0 = f'(x_n)(x_n)\underbrace{(x_{n+1} - x)}_{e_{n+1}} - \frac{1}{2} f''(\xi_n)\underbrace{(x - x_n)^2}_{e_n^2}$$

Where on the left we have the new error, and on the right we have the old error, squared.

Thus we have:

$$e_{n+1} = \left(\frac{1}{2}\frac{f''(\xi_n)}{f'(x_n)}e_n\right)e_n$$

As mentioned before, either this gets accurate very quickly or goes bad quickly (which would feed us NaN early, rather than later).
Dividing by the solution, then we have the **relative error**.

$$\frac{e_{n+1}}{x} = \frac{1}{2}\left[\frac{xf''(\xi_n)}{f'(x_n)}\right]\left(\frac{e_n}{x}\right)^2$$

Recall that relative is dimensionless, so we can express these as percentages (whereas for absolute error, we have to specify units like inch). Everyone knows what we're talking about when we say 0.7% error.

Definition: Dimension -

We define brackets around something to denote its dimension, as such.

$$\left[\frac{d^2f}{dx^2}\right] =: [f''] := \frac{F}{x^2}$$
$$\frac{xF/x^2}{F/x} = 1$$

For example, let $f(x)$ have units $\underset{F}{\text{therbligs}}$ $(\underset{x}{\text{lightyears}})$.
We have $f(x) \iff 10^6 f(10^{-6}x) = 0$.
If $f : \mathbb{R}^n \to \mathbb{R}^n$, then $Df(x)$ is the $n \times n$ matrix,

$$\left[\frac{\partial f_i}{\partial x_j}\right].$$

So if $f(x,y) = \left(\overset{f_1(x,y)}{x^2 - y^2}, \overset{f_2(x,y)}{2xy}\right)$, then

$$[Df] = \begin{bmatrix} 2x & -2y \\ 2y & 2x \end{bmatrix}$$

1.1 2 Dimensional Newton's Method (Aside)

We won't use this in this course, but just so we know it's the same as one dimention:

$$f(x) \cong f(x_n) + Df(x_n)(x_n - x) = 0$$
$$-f(x_n) = Df(x_n)(x_{n+1} - x_n)$$

On the left and on the right, we have column vectors, where $Df(x_n)$ is a square matrix.
Then we have:

$$x_{n+1} = x_n - Df(x_n)^{-1}f(x_n)$$

Usually we can't invert the jacobian because it depends on x, but we can get close to it by using (backslash) in Matlab.

1.2 Back to 1 Dimension

$$\frac{e_{n+1}}{x} = \frac{1}{2}\left(\frac{xf''(\xi_n)}{f'(x_n)}\right)\left(\frac{e_n}{x}\right)^2$$

Theorem 1.1. Convergence Theorem for Newton's Method: Suppose $f(x) = 0$, and $\frac{|x-x_0|}{|x|} \leq c$ (the relative error is invariant), where

$$\max_s f''(s) \cdot \max_s \frac{s}{f'(s)} \leq \frac{1}{c}$$
$$\max_{|x-s|\leq c|x|} [f''(s)] \cdot \max_{|x-s|\leq c|x|} \left[\frac{s}{f'(s)}\right] \leq \frac{1}{c}.$$

If these are satisfied, then Newton's method x_n converges quadratically to x.

1.3 Example

Recall our example of the square root.

$$\begin{aligned} f(x) &= x^2 - a \\ f'(x) &= 2x \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \frac{x_n^2 - a}{2x_n} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \end{aligned}$$

Using the theorem above,

$$x^2 - a = 0, \qquad \frac{|x - x_0|}{|x|} \leq C = 1$$
$$\max_{|x_s|\leq C|x|} 2 \cdot \max_{|x_s|\leq C|x|} \frac{s}{2s} \leq \frac{1}{c} = 1$$

and thus we see we have quadratic convergence from any x_0 with:

$$\begin{aligned} |x - x_0| &\leq |x| \\ |\sqrt{a} - x_0| &\leq \sqrt{a} \\ -\sqrt{a} \leq \sqrt{a} - x_0 &\leq \sqrt{a} \\ -2\sqrt{a} &\leq -x_0 \\ 2\sqrt{a} &\geq x_0 \\ \implies x_0 &\leq 2\sqrt{a} \end{aligned}$$

So if $1 \leq a \leq 2$, then $1 \leq \sqrt{a} \leq \sqrt{2}$, which implies ($x_0 \leq 2 \implies$ quadratic convergence).

1.4 Newton's Method for Multiplicity > 1

In this case with multiple case, we have problems because:

$$f(x) = 0 = f'(x) = f''(x) = \cdots = 0, \quad f^{(m)}(x) \neq 0$$

$$\begin{aligned} f(x) &= x^3 \\ f'(x) &= 3x^2 \\ f''(x) &= 6x \\ f^{(3)} &= 6 \end{aligned}$$

This causes issues because Newton's method says:

$$x_{n+1} = x_n - \frac{f(x_n)}{\underbrace{f'(x_n)}_{?}}$$

We can be more specific, via Taylor:
Let $g(y) := \frac{f^{(m)}(x)}{m!} + \frac{f^{(m+1)}(x)(y-x)}{(m+1)!} + \cdots$.

$$\begin{aligned} f(y) &= f(x) + f'(x)(y-x) + \cdots + \frac{f^{(m-1)}(x)(y-x)^{m-1}}{(m-1)!} + \frac{f^{(m)}(x)(y-x)^m}{m!} + \cdots \\ &= g'(y)(y-x)^m + g(y)m(y-x)^{m-1} \\ &= \frac{g(y)\cdot(y-x)^m}{g'(y)\cdot(y-x)^m + m\cdot g(y)\cdot(y-x)^{m-1}} \\ &= \frac{g(y)\cdot(y-x)}{g'(y)\cdot(y-x) + m\cdot g(y)} \to \frac{0}{m\cdot g(y)} \end{aligned}$$

Then,

$$\begin{aligned} x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \frac{g(x_n)\cdot\overbrace{(x_n - x)}^{e_n}}{g'(x_n)\cdot\underbrace{(x_n - x)}_{e_n} + m\cdot g(x_n)} \\ x &= x \\ e_{n+1} &= e_n - \frac{g(x_n)e_n}{g'(x_n)\cdot e_n + m\cdot g(x_n)} \\ &= e_n - \frac{1}{m}e_n + O\left(e_n^2\right) \end{aligned}$$

To see the first line for e_{n+1}, consider $m\cdot g(x_n)$ is nonzero, so

$$\begin{aligned} \frac{g(x_n)}{m\cdot g(x_n) + g'(x_n)\cdot e_n} &= \frac{g(x_n)}{g(x_n)\cdot g\left(1 + \frac{g'(x_n)}{m\cdot g(x_n)}e_n\right)} \\ &= \frac{1}{m}\left(1 - \frac{g'(x_n)}{m\cdot g(x_n)}e_n + O\left(e_n\right)^2\right) \end{aligned}$$

Where we have a geometric series, hence the last line.
WHen $m = 3$, then

$$|e_{n+1}| \leq \frac{2}{3}|e_n| + O\left(e_n^2\right)$$

So Newton fails to have quadratic convergence, if $m \geq 2$. Using the secant method never gets (as good as) quadratic convergence, but it can be safe. Or, we can try to fix Newton's method somehow, using **double stepping**. That is,

$$x_{n+1} = x_n - \underbrace{m}_{\neq 1} \frac{f(x_n)}{f'(x_n)}$$

This would cancel the $\frac{1}{m}$ from above, where m is the multiplicity of the root. The downside, is practically we don't know m (but we need m), but we know that Newton's method is just not working. The upside, is we can find m by:

1.5 Option 1:

- Take 1 step, with $m = 1$, which gives $x_{n+1}^1 = x_n - \underbrace{\frac{f(x_n)}{f'(x_n)}}_{=:\Delta_n}$
- Try $m = 2$, which gives $x_{n+1}^2 = x_{n+1}^1 - \Delta_n$.
- $\vdots$
- $x_{n+1}^k = x_{n+1}^{k-1} - \Delta_n$.

Choose $i \leq j \leq k$, so

$$|f(x_{n+1}^2)| = \min_{1 \leq j \leq k} |f(x_{n+1}^j)|$$

This determines m automatically, which costs a few more f evaluation, but k is rarely larger than 3.

1.6 Option 2:

This is to generalize this a bit more, which is that we are checking intergral values:

$$x_{n+1} = x_n - t\frac{f(x_n)}{f'(x_n)},$$

We choose t with $0 \leq t < \infty$ to minimize $|f(x_{n+1}^t)| \leq \frac{1}{2}|f(x_n)|$. We take the Newton direction, then we find a better approach.

1.7 Option 3:

Estimate m from $f(x_n), f'(x_n), f''(x_n)$. So,

$$\begin{aligned}
f(y) &= (y-x)^m \cdot g(y) \\
f'(y) &= (y-x)^m \cdot g'(y) + m(y-x)^{m-1} g(y) \\
f''(y) &= (y-x)^m \cdot g''(y) + 2m(y-x)^{m-1} g'(y) + m(m-1)(y-x)^{m-2} \cdot g(y)
\end{aligned}$$

We want to choose the largest (right-most) terms of the sequence (as we let $y \to x$). So consider:

$$\begin{aligned}
f \cdot f''(y) &= m(m-1)(y-x)^{2m-2} g(y)^2 + O(y-x)^{2m-1} \\
(f')^2 &= m^2 (y-x)^{2m-2} g(y)^2 + \cdots
\end{aligned}$$

Subtracting these two equations, we have:

$$(f')^2 = m(y-x)^{2m-2} g(y)^2 + \cdots$$

Hence, as $y \to x$,

$$\frac{(f')^2}{(f')^2 - f \cdot f''} \to m$$

2 Newton-Schroder

$$x_{n+1} = x_n - \frac{[f'(x_n)]^2}{[f'(x_n)]^2 - f(x_n)\underbrace{f''(x_n)}} \cdot \frac{f(x_n)}{f'(x_n)}$$

The only cost here is to compute $f''(x_n)$. Of course, it's harder to explain what this means in multi-dimensional settings.
This is quadratically convergent, for simple roots **and** roots with multiplicity. So this is a pretty fail-safe way of conducting Newton's method.

$$\begin{aligned}
\left(\frac{f(x)}{f'(x)}\right)' &= 1 - \frac{f(x) f''(x)}{[f'(x)]^2} \\
&= \frac{f'(x)^2 - f(x) \cdot f''(x)}{[f'(x)]^2}
\end{aligned}$$

This would suggest that Newton-Schroder's method is

$$\begin{aligned}
x_{n+1} &= x_n - \left[\frac{f(x_n)}{f'(x_n)}\right] \Big/ \left[\frac{f(x_n)}{f'(x_n)}\right]' \\
&= x_n - \frac{F(x_n)}{F'(x_n)},
\end{aligned}$$

where we let

$$F(y) = \frac{f(y)}{f'(y)} = \frac{(y-x)^m g(y)}{(y-x)^m \cdot g'(y) + m \cdot (y-x)^{m-1} \cdot g(y)}$$

Remark: This is interesting because F no longer has a zero with multiplicity more than one, at $y = x$. The numerator $f(y)$ has a zero of multiplicity m, but the denominator has a zero with multiplicity $(m-1)$.